



OPEN Impact of irrelevant speech and non-speech sounds on serial recall of verbal and spatial items in children and adults

Larissa Leist^{1✉}, Thomas Lachmann^{1,2,3} & Maria Klatte¹

Short-term memory for sequences of verbal items such as written words is reliably impaired by task-irrelevant background sounds, a phenomenon known as the “Irrelevant Sound Effect” (ISE). Different theoretical accounts have been proposed to explain the mechanisms underlying the ISE. Some of these assume specific interference between obligatory sound processing and phonological or serial order representations generated during task performance, whereas others posit that background sounds involuntarily divert attention away from the focal task. To explore the roles of phonological processing, serial order retention, and attention control, we analyzed the effects of environmental non-speech sounds and unfamiliar speech on serial recall of verbal items (pictures representing German nouns) and spatial items (dot locations) in children ($n=137$) and adults ($n=98$). In the verbal task, both age groups were equally affected by background sounds, with speech impairing recall more than environmental sounds. In the spatial task, no ISE was found in adults and fourth graders, but third graders exhibited significant performance impairment from both sounds. There was no habituation to the sound effects across the experimental trials. The findings indicate that both specific interference and attention capture may contribute to the ISE, with the impact of attention capture potentially decreasing with age.

In our daily lives, we are frequently exposed to environmental noise, such as traffic noise, music, and other people's conversations. Mental work in noisy environments is becoming increasingly prevalent in our digitalized society as people are taking greater advantage of the opportunity—or the obligation—to work virtually anywhere, e.g., at home, in cafés, on trains and airplanes. The issue of environmental noise is also increasingly important for children. For them, learning in classrooms is often accompanied by task-irrelevant sounds from classmates' activities or noise pollution from outside, which has been shown to have a negative impact on mental performance and well-being^{1–4}. To avoid or reduce such impairments, we need to know more about the mechanisms underlying noise effects in different age groups. Therefore, in the present study we investigated the effects of different types of environmental noise on the short-term memory performance in children and adults.

Research on noise-induced impairments of short-term memory goes back more than 50 years, and has focused on the so-called “Irrelevant Sound Effect” (ISE). The ISE refers to the impairment of short-term memory for verbal items (e.g., digits, words, letters) presented visually due to task-irrelevant background sounds that participants are instructed to ignore. The usual paradigm for studies on the ISE uses a standard short-term memory task which requires immediate serial recall of lists of 5 to 9 items presented in succession for about one second each. The ISE is a robust and reliable phenomenon which has been replicated numerous times using a wide range of materials, designs, and procedures^{5–10}. There is widespread agreement that both speech and non-speech sounds have a detrimental impact on verbal serial recall in adults.

Several theoretical frameworks have been proposed to explain the mechanisms underlying the ISE. Some of these emphasize automatic, content- or process- based interference between sound characteristics and specific task demands, whereas others attribute the ISE to more general diversions of attention away from the focal task (for a comprehensive review, see Leist et al.¹¹).

Early studies performed by Baddeley and his colleagues^{5,7,12,13} focused on speech as irrelevant sound and attributed its detrimental effect to interference between phonological representations in the phonological loop

¹Center for Cognitive Science, Cognitive and Developmental Psychology Unit, University of Kaiserslautern-Landau (RPTU), 67663 Kaiserslautern, Germany. ²Centro de Investigación Nebrija en Cognición (CINC), Universidad Nebrija, 28007 Madrid, Spain. ³Brain and Cognition Research Unit, Faculty of Psychology and Educational Sciences, University of Leuven (KU), 3000 Leuven, Belgium. ✉email: lleist@rptu.de

component of working memory¹⁴. The phonological loop is a storage system for verbal information. It consists of two parts: a phonological store, which holds phonological representations in a passive form, and an articulatory rehearsal process, which prevents trace decay in the store through subvocal repetitions of the to-be-remembered (TBR) list items. According to the phonological loop account, the ISE results from specific (content-based) interference between the phonological codes of the visually presented items and the irrelevant speech that gains automatic access to the phonological store. Consequently, sounds lacking speech-specific properties such as tones or instrumental music should not impair verbal short-term memory, and tasks requiring short-term memory for nonverbal items should be unimpaired by irrelevant speech.

However, a number of later findings did not support these predictions. First, it was shown that the ISE is also evoked by reversed speech^{15–18}, tones^{10,19}, and music^{20–23}, even though the disruption is most pronounced with speech²⁴. Second, the disruptive effect of background speech is independent from its phonological similarity to the list items²⁵, which contradicts the assumption that similarity between the irrelevant stimuli and the TBR items determines the magnitude of the disruption. Third, the phonological loop account was unable to explain findings demonstrating a dominant role of seriation in ISE evocation. Concerning task characteristics, verbal memory tasks without a serial component were shown to be unaffected or only mildly affected by irrelevant sounds²⁶, whereas a nonverbal task requiring the recall of spatial locations in serial order was significantly affected, see Jones et al.²⁷ (Exp. 4). Concerning sound characteristics, it was shown that sounds representing coherent sequences of distinct, different auditory tokens (“changing-state” sounds, e.g., sequences of different syllables produced by a single speaker) produce stronger impairments when compared to continuous, unstructured sounds like broadband noise or sounds consisting of single, repeated tokens (“steady-state” sounds, e.g., repetitions of single syllables or tones). Aiming to integrate these findings, Jones et al.²⁸ provided another explanation of the ISE. According to this “changing-state account”, sounds with a changing-state characteristic are automatically structured and organized into streams of ordered objects via pre-attentive auditory streaming processes. Sequential information generated during this process interferes with the serial rehearsal of the list items that participants rely on during serial recall tasks. Thus, in contrast to “interference-by-content” proposed by the phonological loop account, the changing-state account assumes interference-by-process (seriation) as the mechanism underlying the ISE.

The changing-state account has been integrated into a broader theoretical framework of serial recall and respective impairments, which assumes that serial order retention is not accomplished through specialized memory structures or processes, but through general-purpose motor and perceptual functions (perceptual-gestural view of short-term memory, see^{29,30}). According to this model, ordered recall of unrelated items is achieved through a deliberate rehearsal process that involves mapping the perceived item sequence to a motor output plan, a record of which is then cyclically repeated until recall is completed. This holds for items of any origin or modality, as they all share a common representational space. Therefore, this model predicts equivalent effects of speech and non-speech changing-state sounds, and equivalent effects of changing-state sounds on serial recall of verbal and nonverbal (e.g., spatial) items.

Neither the phonological loop nor the changing-state account incorporates a specified role of attention in ISE evocation. In contrast, Cowan’s Embedded Processes Model^{31,32} emphasizes the role of attention in memory performance and its impairments. Cowan^{31,32} posits that memory representations vary with respect to activation. Working memory is considered as a subset of long-term memory, comprising those representations that are currently in a heightened state of activation. Representations with the strongest activation enter the so-called “focus of attention”, which is limited in capacity to about four items. Maintaining a longer sequence of items is achieved by recirculating them through the focus of attention. This mechanism is termed “attentional refreshing”, and is available for verbal as well as nonverbal items. According to the attentional account, refreshing is impaired through involuntary intrusions of task-irrelevant sounds into the focus of attention, leading to poorer recall.

Aiming to explore the role of attention in the ISE, several studies compared the effects of irrelevant sounds between children and adults. Given their immature attention control, children are likely to be less able than adults to focus attention on the task in the presence of distracting stimuli. Thus, if the ISE were to result from attention capture, children should be more affected than adults. However, the findings on developmental change in ISE magnitude are inconsistent. While some studies^{11,33–35} identified a larger ISE in children when compared to adults, others found no differences between groups^{35–40}.

Another approach to examine the impact of attention capture is to investigate the magnitude of the ISE over the time-course of task completion. According to the attentional account, the ISE should diminish with time, as participants habituate to the irrelevant sound with repeated exposure. Research in this area also revealed mixed results. Some studies provided evidence for habituation (and thereby attention capture)^{41,42}, while others did not^{43,44}. Researchers have attributed the inconsistency in habituation effects to variations in methodologies.

Especially, it has been suggested that habituation effects are more pronounced when the same sound sequence is presented in each trial of the respective condition (specific habituation) when compared to paradigms where a sound condition is represented by different sound sequences, e.g., different sequences of syllables representing a changing-state condition (unspecific habituation)⁴⁵. According to this view, unspecific habituation proceeds more slowly and is thus not always observable over the time course of a typical ISE experiment.

In the present study, we aim to get further insight into the roles of phonological interference, seriation, and attention control in the ISE. For this aim, we investigate the effects of environmental sounds and narrative speech on serial recall of verbal and nonverbal (spatial) items in children and adults. As previously stated, in contrast to the assumption of phonology-based interference, Jones et al.²⁷ reported similar effects of changing-state speech on verbal and spatial serial recall, and attributed the disruption to interference between pre-attentive seriation processes involved in obligatory sound processing, and deliberate serial rehearsal of the list items. However, the finding of an ISE in spatial serial recall could not be reproduced in several subsequent studies^{11,46–50}. One of these studies¹¹ compared performance of children and adults, and found a significant impairment due to

changing-state speech (spoken syllables) in the verbal task, which was more pronounced in the children than in the adults. For the spatial task, no sign of a sound-induced disruption was found in either group. The latter finding is in contrast to both the changing-state and the attention-capture account of the ISE, as both assume domain-general maintenance mechanisms as the locus of the ISE.

In this study, we used narrative unfamiliar speech and a mixture of environmental non-speech sounds such as dogs barking, telephone ringing, and traffic sounds as task-irrelevant background sounds. For both sound conditions, the same sequence was used in each of the respective trials, thereby increasing the probability to detect potential habituation effects⁴⁵. Using nonspeech environmental sounds, we deviate from prior studies, which analyzed the role of attention in the ISE by using auditory deviants (sounds that violate expectations, i.e., a rarely presented “deviant” sound within a sequence of repeated “standard” sounds, with deviation defined by acoustic properties like timbre⁵¹, or a change in semantic category⁵²), taboo words^{53,54}, or self-relevant words (e.g., the name of the participants⁵⁵) as irrelevant sounds, and/or by varying the predictability of the respective sequences^{56,57}. These studies provide evidence that irrelevant sounds may impair visual-verbal serial recall performance in adults through attention capture.

We decided for environmental sounds for several reasons. First, performing mental task in the presence of irrelevant environmental sounds is a daily challenge in working environments such as classrooms and open-plan offices. Second, sequences of different environmental sounds should evoke only a minor or no changing-state effect due to their lack of stream coherence. Especially, following the changing-state account, in order to evoke interference with seriation, an irrelevant sound sequence should represent a coherent stream consisting of different auditory tokens emanating from a single source, e.g., a sequence of different syllables uttered by a single speaker, or a melody played by a single instrument. Detrimental effects evoked by a sequence of unrelated environmental sounds may thus be considered as indicators of attention capture. Third, evidence from field studies^{2,58} and from behavioral and EEG experimental studies^{36,59,60} confirms that environmental sounds capture attention, especially in children. One of the studies³⁶ explored age differences in the ISE and found that non-speech classroom noise impaired visual-verbal serial recall performance in 7-year-old first graders, whereas older children and adults were unaffected.

In light of the existing research, the following predictions arise: if attention capture plays a dominant role in the ISE, the speech and environmental sounds should evoke significant disruption in both tasks, with stronger effects in children when compared to adults. Furthermore, we should observe a decrease of disruption over the time-course of task completion due to specific habituation. Concerning the interference-based accounts, the following predictions can be derived. If the ISE results from phonological interference (Phonological Loop Account), task-irrelevant speech should impair performance in the verbal task. This effect should be equivalent in children and adults, and should remain stable over the course of the task. Environmental sounds should evoke a minor or no ISE in the verbal task as they lack phonological properties. The spatial task should not be impaired by any background sound, neither in the adults nor in the children, because spatial items are not processed in the phonological loop. In contrast, according to the changing-state account, speech should evoke impairments in both tasks, regardless of age. The disruptive effect should remain stable over the course of the tasks. The mixture of environmental sounds should evoke no or minor disruption, as it does not represent a coherent auditory stream, but a sequence of unrelated sounds from different sources.

For the verbal task, items were displayed as colorful drawings of monosyllabic German nouns. Drawings were used instead of written words in order to rule out effects of children's reading abilities on serial recall performance. Pictorial presentation does not alter participants' strategies when compared to printed words^{61,62}, for a review see Leist et al.¹¹. We recruited children aged about 9 years, as we assume that they employ verbal rehearsal strategies when memorizing verbal items^{63,64}. However, it is important to recognize that rehearsal strategies can vary significantly among younger children, influencing their recall performance and susceptibility to distractions^{65,66}.

For the spatial task, items were spatial locations presented as uniform black dots at varying positions on the screen. Sequence length for adults (children) was 8 (5) for the verbal task, and 7 (5) for the spatial task. Items were presented one-by-one on the computer screen with an onset-to-onset rate of 2 seconds. After a 5 seconds rehearsal interval, all items presented in the respective trial were shown in random arrangement on the screen. For serial recall, participants clicked on the items in the order of presentation. The experimental factors Task (verbal, spatial) and Sound condition (silence, environmental sounds, speech) were varied between and within participants, respectively (i.e., each participant performed either the spatial or the verbal task in each of the three sound conditions). A recording of a female speaker reading a newspaper article in a language unfamiliar to all participants was used as irrelevant speech.

Results

An item was scored as correct when it was recalled in the correct serial position^{51,67}. All analyses are based on proportion correct scores. Figure 1 illustrates the proportion of correct answers based on age group and sound condition.

The data were analyzed using a $3 \times 2 \times 2$ mixed-design Analysis of Variance (ANOVA) with Sound condition (silence, environmental sounds, speech) as within-subject factor and Age group (children, adults) and Task (verbal, spatial) as between-subject factors. The analysis revealed significant main effects of Sound condition, $F(2,462) = 24.770$, $p < .001$, $\eta_p^2 = .097$, and Age group, $F(1,231) = 27.871$, $p < .001$, $\eta_p^2 = .108$. The main effect of Task was not significant, $F(1,231) < 1$. Importantly, there were significant two-way interactions between Sound condition \times Task, $F(2,462) = 8.762$, $p < .001$, $\eta_p^2 = .037$, and between Sound Condition \times Age group, $F(2,462) = 3.889$, $p = .021$, $\eta_p^2 = .017$, demonstrating a stronger effect of sound in the verbal when compared to the spatial task, and a stronger effect of sound in children when compared to adults, respectively.

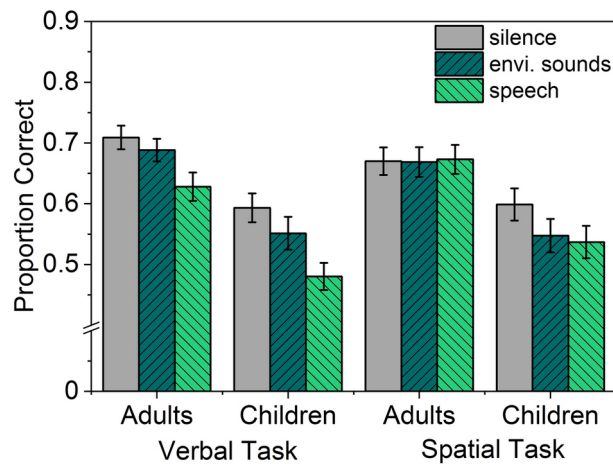


Fig. 1. Mean proportion correct scores of both the verbal and the spatial task with respect to age group and sound condition. Error bars denote standard errors of the mean.

Performance in the silent control condition did not differ between tasks, for either the adults, ($t(96) = 1.308$, $p = .194$, $d = .265$), and the children, ($t(135) = .160$, $p = .873$, $d = .027$). Hence, within age groups, a potential difference in the ISE between tasks cannot be attributed to differences in task difficulty. Across age groups, t -test revealed significantly better performance of the adults in the verbal task ($t(119) = 3.565$, $p < .001$, $d = .656$) and a tendency for better performance in the spatial task ($t(112) = 1.928$, $p = .056$, $d = .367$). The interaction between Task and Age group and the three-way interaction were not significant, $F(1,231) < 1$ and $F(2,462) < 1$, respectively. To evaluate the significant two-way interactions of Sound Condition with Age group and Task, separate analyses were performed for the verbal and spatial tasks.

Verbal task

Data were analyzed using a 3×2 mixed-design ANOVA with Sound condition (silence, environmental sounds, speech) as within-subject factor and Age group (children, adults) as between-subject factor. Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 7.11$, $p = .029$, therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\epsilon = .97$)⁶⁸. There were significant main effects of Sound condition, $F(1.94,230) = 29.274$, $p < .001$,

$\eta_p^2 = .197$, and Age group, $F(1,119) = 19.608$, $p < 0.001$, $\eta_p^2 = .141$. The interaction was not significant, $F(1.94,230) < 1$, indicating comparable sound effects in children and adults. Bonferroni-corrected post hoc tests revealed that performance in the speech condition was lower than the silent and environmental sounds conditions (each $p < .001$). Performance in the environmental sounds condition was lower than in the silent control condition ($p = .019$).

Spatial task

Proportion correct scores were analysed using a 3×2 mixed-design ANOVA with Sound condition (silence, environmental sounds, speech) as within-subject factor and Age group (children, adults) as between-subject factor. There were significant main effects of Sound condition, $F(2,224) = 3.394$, $p = .035$, $\eta_p^2 = .029$, and Age group, $F(1,112) = 9.787$, $p = .002$, $\eta_p^2 = .080$, and a significant interaction, $F(2,224) = 3.394$, $p = .025$, $\eta_p^2 = .032$, indicating that the sound effect differed between age groups.

Separate analyses were performed for each age group in order to analyze the significant interaction. A one-way repeated measures ANOVA of the factor Sound condition (silence, environmental sounds, speech) did not reach significance for the adults, $F(2,92) < 1$, but revealed a significant main effect for the children, $F(2,132) = 6.205$, $p = .003$, $\eta_p^2 = .086$.

Post hoc Bonferroni-corrected tests of the children's spatial task performance revealed lower performance in the speech condition ($p = .006$) and in the environmental sounds condition ($p = .030$) when compared to silence. The sound conditions, environmental sounds and speech, did not differ ($p = 1.00$).

The significant sound effect in children is in contrast to Leist et al.¹¹, who reported no disruptive effect of sounds on children's performance in the spatial task. However, in that study, only third graders were included, whereas in the current study, the spatial task was performed by both third and fourth graders. As attention control develops rapidly during the early years of schooling^{69–73}, we conducted an additional analysis to explore the potential influence of class level on the sound effect. This analysis was performed on 34 third graders (20 female, 14 male, 7;1–9;6 years, $M = 8;4$, $SD = .4$) and 33 fourth graders (20 female, 13 male, 8;1–11;3 years, $M = 9;6$, $SD = .6$). Figure 2 illustrates the proportion of correct answers based on class level and sound condition. A 3×2 mixed-design ANOVA with the within-subject factor Sound condition (silence, environmental sounds, speech) and the between-subjects factor Class level (third grade, fourth grade) revealed significant main effects of Sound condition, $F(2,130) = 6.400$, $p = .002$, $\eta_p^2 = .090$, and Class level, $F(1,65) = 14.141$, $p < .001$, $\eta_p^2 = .179$, and a significant interaction, $F(2,130) = 4.053$, $p = .020$, $\eta_p^2 = .059$. Separate repeated-measures ANOVAs revealed a

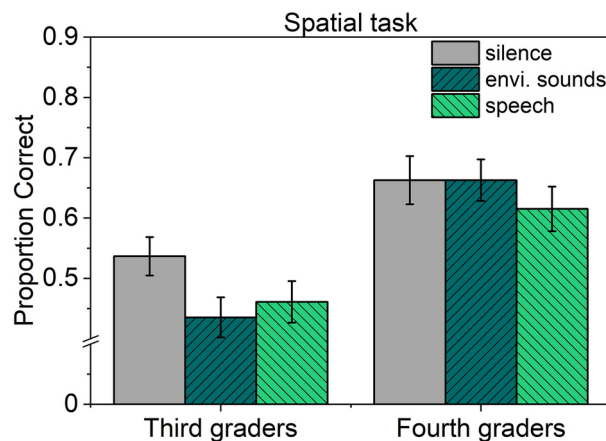


Fig. 2. Mean proportion correct scores of the spatial task with respect to class level and sound condition. Error bars denote standard errors of the mean.

significant effect of Sound condition for the third-graders, $F(2,66) = 8.957$, $p < .001$, $\eta_p^2 = .213$, but not for the fourth-graders, $F(2,64) = 2.061$, $p = .136$, $\eta_p^2 = .060$.

Analyses concerning habituation

To investigate the impact of habituation on the ISE, an additional analysis was conducted to determine whether the sound induced impairments revealed in the preceding analyses diminish over time. If the sound effects are caused by attention capture, one might expect that the first trials with sound would be more disrupted than the last ones, i.e., over time, participants become habituated to the repeated sound sequence (specific habituation, see Röer et al.⁴⁵). Proportion correct scores were calculated for each age group and task for the ordinal trial position, with sixteen trials (adults) and eight trials (children) for each sound condition. Figure 3 depicts the resulting proportion correct scores.

Two-factor repeated measures ANOVAs were performed with Sound condition (silence, environmental sounds, speech) and Trial (1–16) for adults, and Trial (1–8) for children, separately for each task (verbal, spatial).

The analysis for the adults' spatial task performance revealed no significant effects of Sound condition, $F(2,92) < 1$, $p = .939$, Trial, $F(15,690) < 1$, $p = .982$, and no interaction, $F(30,1380) < 1$, $p = .879$. The analysis for the adults' verbal task performance revealed a significant main effect of Sound condition, $F(2,100) = 16.058$, $p < .001$, $\eta_p^2 = .243$, no main effect of Trial, $F(15,750) = 1.24$, $p = .236$, and no interaction, $F(30,1500) = 1.410$, $p = .070$. The same pattern was found for the children's verbal task performance, a significant main effect of Sound condition, $F(2,138) = 17.606$, $p < .001$, $\eta_p^2 = .203$, no main effect of Trial, $F(7,483) < 1$, $p = .434$, and no interaction, $F(14,966) < 1$, $p = .633$. The analysis of the children's spatial task performance revealed significant main effects of Sound condition, $F(2,132) = 6.205$, $p = 0.003$, $\eta_p^2 = 0.086$, and,

Trial, $F(7,462) = 2.317$, $p = 0.025$, $\eta_p^2 = 0.034$, but no interaction, $F(14,924) < 1$, $p = .456$, indicating that, regardless of sound conditions, children's overall performance in the spatial task declined over time. Thus, no evidence for habituation was found.

Discussion

This study provides a follow-up extension of earlier research conducted by Leist et al.¹¹, who showed that children's and adults' serial recall performance of verbal but not spatial items is impaired by task-irrelevant changing-state speech. Here, we investigated the impact of task-irrelevant speech and environmental non-speech sounds on verbal and spatial serial recall in primary school children and adults. The tasks were identical to Leist et al.¹¹, requiring the serial reconstruction of pictures displaying monosyllabic German words (verbal task) and of uniform black dots displayed at various positions on the screen (spatial task).

In the verbal task, children and adults were equally affected by both task-irrelevant background sounds. Speech impaired recall performance significantly more than environmental sounds. There was no evidence of habituation across experimental trials, neither in the children nor in the adults. For the spatial task, performance of adults was unaffected by the irrelevant sounds, which is in contrast to the classical finding of Jones et al.²⁷, but in line with more recent findings^{11,50}. However, for children's performance in the spatial task, background speech and environmental sounds evoked significant disruption of equal magnitude. Further analyses revealed that the significant sound effect in the child sample resulted from a reliable impairment in the third-grade children, whereas fourth graders were unaffected. The pattern of results is discussed with respect to the theoretical accounts outlined in the introduction.

In adults, the ISE was evident for visual-verbal serial recall, but was completely absent for serial recall of spatial items. Given the consistent evidence from the current and other studies with considerable statistical power^{11,50}, we may conclude that serial recall of spatial items is unaffected by irrelevant sounds in adults. This poses severe difficulties for theories positing that the ISE results from impairments of domain-general maintenance mechanisms, i.e., perceptual-motor sequencing (changing-state account) and attentional

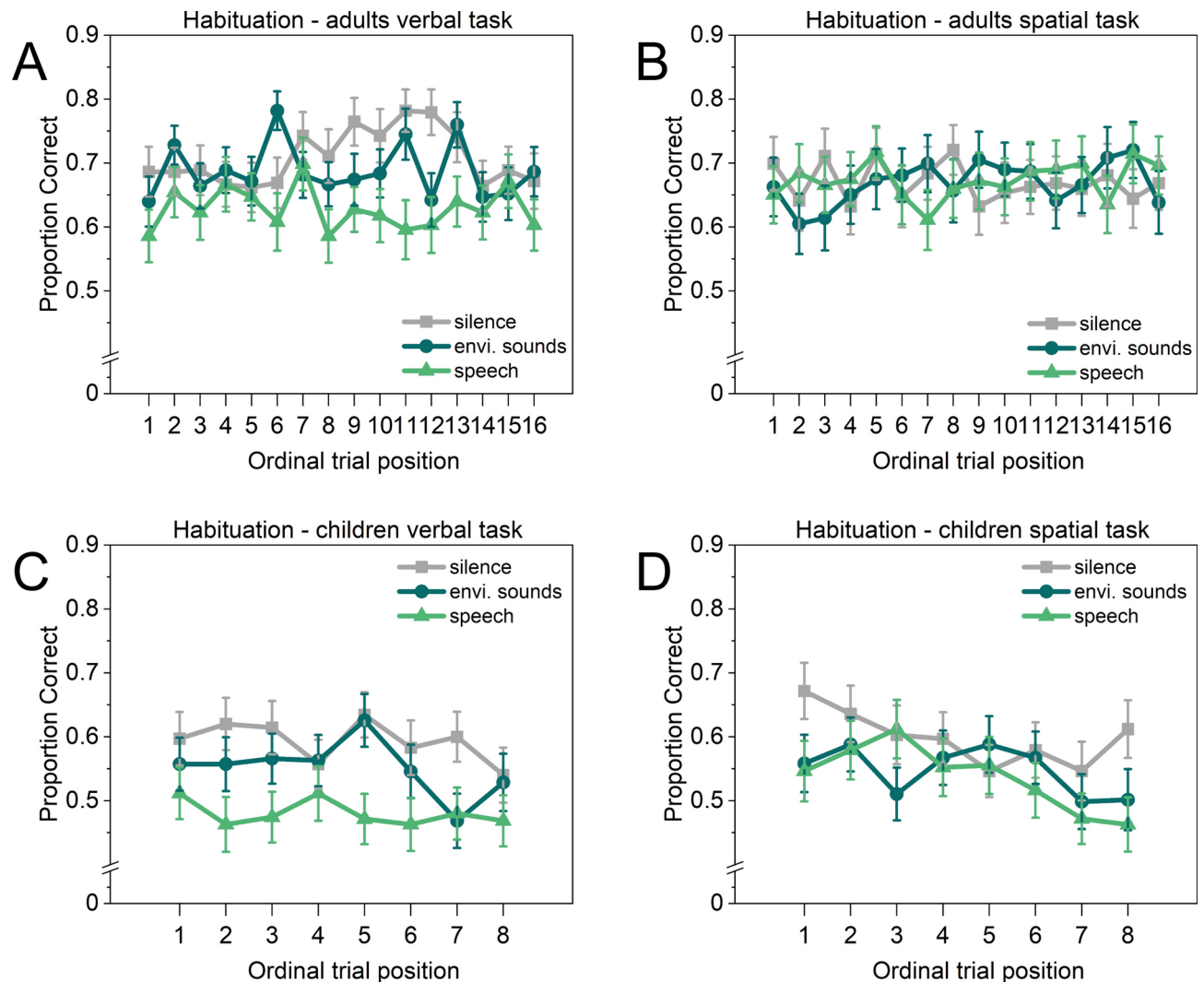


Fig. 3. Mean proportion correct scores as a function of sound condition and ordinal trial position of sixteen (adults) and eight (children) trials for both tasks (verbal, spatial). Error bars denote standard errors of the mean.

refreshing (attention-capture account), as both predict sound-induced performance decrements irrespective of item modality. Importantly, according to the attention capture account, ISE should even be stronger for spatial when compared to verbal serial recall, as it has been shown that visuo-spatial working memory tasks rely much more on domain-general attentional resources^{74,75}. The immunity of spatial serial recall to the ISE is in line with the phonological loop account, which considers the retention of phonological codes in a verbal short-term store as the locus of disruption.

The finding that environmental non-speech sounds also impaired verbal serial recall performance, albeit to a lesser extent than speech, mirrors the results reported in Buchner et al.⁷⁶. These authors explain their findings in the framework of the changing-state account, arguing that speech sounds have a more prominent changing-state characteristic when compared to environmental sounds. In particular, according to the changing-state account, speech is more disruptive than non-speech sounds because it represents a coherent changing-state auditory stream, whereas the sound mix consists of isolated and unrelated auditory tokens. However, the difference in effect magnitude between speech and environmental sounds might also be explained by the phonological loop account, by assuming that complex non-speech sounds are processed in the phonological loop, as was suggested for instrumental music⁷, or that environmental sounds—like spoken words—automatically activate phonological representations which interfere with those of the list items. Recent findings suggest that spoken words and environmental sounds evoke comparable phonological activation⁷⁷, but see Kukona et al.⁷⁸ for contradicting results. For the attention capture account, the reduced effect of environmental sounds is difficult to explain. One would expect that a mixture of heterogeneous, partly alarming sounds has more potential to attract participants' attention when compared to narrative, meaningless speech. In addition, the attention capture account predicts habituation with repeated exposure, which was not found for either sound.

Concerning developmental change, the absence of age-related differences in the ISE on the verbal task is in line with the interference-based accounts, as both predict comparable effects across age groups provided that the same strategies are applied. The lack of an age effect is difficult to handle for the attentional account. Due to the immature attention control in children, they should be more prone to sound-induced attention capture. However, the current results need to be considered in the context of prior studies on developmental change in the ISE on verbal serial recall, which yielded inconsistent results. Some mirrored the current finding of equivalent impairments in children and adults^{35–40,79}, whereas others found stronger impairments in the children^{11,33–35}, thereby confirming a role of attention in the ISE. To account for these discrepancies, following Klatte et al.³⁶ and Elliott et al.³⁵, we suggest that children and adults are equally prone to specific interference through task-irrelevant sounds in visual-verbal serial recall (provided that the children use the same retention strategies), but that the disruptive effect may be superimposed by sound-induced diversions of attention. The contribution of the latter depends on characteristics of the sounds and method, and on the participants' ability to control attention, which increases with age and schooling^{71–73}. Developmental studies on the ISE differ considerably in these respects, for example concerning children's age and class levels, types of background sounds (e.g., meaningless vs. meaningful narrative speech, environmental sounds, tones), sound levels, and methods of sound presentation (blocked vs. trial-by-trial variations of sound conditions, presentation via earphones vs. loudspeakers). All of these aspects have an impact on the sounds' potential to grab attention away from the primary task. In addition, it has to be kept in mind that, even within a restricted age range or class level, there are huge individual differences in children's attention control^{69,70}. This might explain why, in our study, we did not find an age effect on verbal serial recall, whereas a prior study using comparable methods and age groups¹¹ reported greater impairment in children compared to adults.

Concerning the spatial task, the current study revealed disruptive effects of equal magnitude due to speech and environmental sounds in children, but not in adults. We attribute the effect observed in the children to attention capture for several reasons. First, additional analyses revealed that the effect was only evident in the third-grade children, but absent in the fourth graders. In view of the increase in attention control with age and schooling, we may conclude that the fourth-grade children were better able to focus attention on the task in the presence of irrelevant sounds. In line with this argument, Klatte et al.³⁶ found a significant impairment of verbal serial recall due to non-speech classroom noise in 7-year-old first graders, whereas older children were unaffected. Second, the current findings mirror those of a previous study³⁵, which showed that children and adults were equally impaired by irrelevant speech in a visual-verbal serial recall task, but only the children were also affected in a verbal task that has consistently proven immune to the ISE in adults. The authors infer that children and adults are equally prone to sound-induced interference in visual-verbal serial recall, but that, in children, attentional diversions are also at play, making them vulnerable to sound-induced impairments irrespective of task demands and the sounds' changing-state.

To summarize, none of the described theoretical accounts of the ISE is able to integrate the full pattern of results in our study. The findings are best explained by assuming that both specific interference with visual-verbal serial recall and more general effects on attention contribute to the observed effects, as suggested in the “duplex mechanism account” of auditory distraction⁸⁰. Following this account, specific interference and attention capture represent independent mechanisms of sound-induced performance decrements, which may act in isolation or in combination, depending on characteristics of the tasks, the sounds, and the individuals affected. In children, diversions of attention have a stronger impact on the overall disruption, as children are less able to control attention in the presence of task-irrelevant stimuli when compared to adults. However, there is huge interindividual variability in children's attention abilities over the primary school years, which may contribute to the discrepancies in results from studies on developmental change in the ISE.

Several limitations of the present study must be acknowledged. First, the children performed only 8 trials per sound condition, and we thus cannot rule out that habituation effects become evident when the time-course of exposure is extended. Second, with the environmental sounds used here, only limited evidence for performance decrements through attention capture was found in adults. However, as stated in the introduction, a number of studies found clear evidence for task impairments through sound-induced diversions of attention in adults^{51–55}. Obviously, the attention-grabbing potential of the sounds used in those studies (i.e., auditory deviants, taboo-words^{53,54}, self-relevant words⁵⁵) is stronger when compared to the current environmental sounds. Third, as in the majority of respective studies, our child participants aged 8 years or more. As attention control increases strongly during the primary school years and the ability to filter out irrelevant sounds becomes adult-like with about age 10, future studies on age differences in the ISE should include younger children (school starters, aged 6 to 7 years). Fourth, due to the strong correlation between chronological age and class level ($r(67) = .82$), we do not know whether age or schooling is responsible for the stronger sound effect on the spatial task observed in third graders compared to fourth graders.

Experimental research on the ISE in children does not only increase our knowledge on the mechanisms underlying this effect, but may also contribute to practical issues. As verbal short-term memory plays a crucial role in language and reading acquisition⁸¹, sound-induced disruption could result in lasting deficits. Field studies have already demonstrated that noise in classrooms affects reading acquisition and verbal precursors of reading^{1,2,4,82}. Therefore, experimental studies on the ISE in children are critical for creating suitable acoustic environments for learning³⁶. For example, tasks that involve verbal short-term memory, such as reading and spelling in beginning readers, vocabulary learning, or mental arithmetic should not be performed in the presence of background speech or music.

Methods

Participants

An a priori power analysis was conducted using G*Power⁸³ (version 3.1.9.7) to estimate the sample size based on data from Leist et al.¹¹ ($N = 164$). The effect size reported in that study was 0.176. With a significance criterion of $\alpha = .05$ and a power of .80, the minimum sample size required for the mixed ANOVA with this effect size is $N = 108$. A total of 137 children, recruited from two primary schools in the Kaiserslautern region and in North Rhine-Westphalia in Germany (76 female, 61 male, aged between 7 years, 9 months and 11 years, 3 months, mean age of 9;0 years), and 98 students from the University of Kaiserslautern-Landau (45 female, 53 male, aged between 19 and 36 years, mean age of 23.1 years), participated in this study.

The verbal task was performed by 51 adults (22 female, 29 male) aged between 19–26 years ($M = 22.2$ years, $SD = 1.82$ years) and 70 children (36 female, 34 male) aged between 8;0 years to 10 years, 9 months ($M = 8$ years, 6 months, $SD = 7$ months). The spatial task was performed by 47 adults (23 female, 24 male) aged between 19–36 years ($M = 24.0$ years, $SD = 3.48$ years) and 67 children (40 female, 27 male) aged between 7 years, 9 months to 11 years, 3 months ($M = 9$ years, 1 month, $SD = 7$ months).

All participants were native German speakers and had normal or corrected-to-normal vision and normal hearing according to self-reports (adults) or parental reports (children). Adults received either course credit or payment for participation (10 €). Children received writing materials, e.g., pencils, erasers.

Apparatus

A 15.6-inch HP laptop (ProBook 650 G1) running Python 3.7/Psychopy 3.1.5⁸⁴ on Windows 10 was used to control the experimental trial. Display resolution was 1920×1080, with a refresh rate of 60 Hz. The background sounds were played via open headphones (Sennheiser HD650), connected with a Focusrite Scarlett 2i2 2nd generation audio interface to the laptop.

Irrelevant sounds

Task performance was measured during three different sound conditions: silent control, environmental non-speech sounds, and speech. The environmental sounds encompassed a sequence of diverse common sounds (e.g., dog barking, bicycle bell, church bell, approaching car, chainsaw, cackling hens, siren horn). These environmental sounds were played one after the other, without pause, for two to three seconds per sound. The speech consisted of a female Danish speaker reading an article from a newsletter identical to Klatte et al.³⁶. Danish was unfamiliar to all participants. For both sound conditions, the same sequence was played in each of the respective trials.

To-be-remembered items

For the verbal task, for children and adults, items were drawn from a set of ten colorful drawings representing the German monosyllabic words *Bett* (bed), *Bus* (bus), *Eis* (ice cream), *Frosch* (frog), *Kamm* (comb), *Mond* (moon), *Pilz* (mushroom), *Schal* (scarf), *Schiff* (ship) and *Zaun* (fence). For the adults the set included the additional items *Brief* (letter), *Haus* (house), *Herz* (heart), *Hut* (hat), *Nuss* (nut) and *Schwein* (pig). Four lists of five items (drawn out of 10) and eight lists of eight items (drawn out of 16) were created for the children and the adults, respectively. The lists were created quasi-randomly, with each item appearing in one of four lists (children) or eight lists (adults). To generate parallel versions of each list, random permutations of the list items were used. In addition, one practice list for each sound condition was constructed. All drawings were presented in a rectangular black frame (102×73 mm) in the center of a white screen.

For the spatial task, items were presented as black dots (10 mm diameter) on a white background in a 304×183 mm black frame. Dot positions were drawn from a total of 80 possible locations that were arranged in a 8×10 grid (not shown to the participants) in the frame. As shown by Parmentier and colleagues^{85,86}, the difficulty of reproducing a sequence of dot locations depends on the complexity of the spatial path formed by the successive dots. Path complexity is determined by the number of crossings, the path length, and the angular degree. In order to verify equal task difficulty across sound conditions, we first created eight lists of five (children) and sixteen lists of seven dot positions (adults). The resulting path of each list was mirrored horizontally and vertically, thereby varying dot locations while keeping path complexity constant (for further details, see Leist et al.¹¹). With this strategy, three parallel versions of each list were constructed, which were assigned randomly to the three sound conditions. In addition, three practice lists were constructed.

Serial recall tasks

Both tasks are illustrated in Figure 4. Each trial consisted of a presentation phase, a retention interval, and a reconstruction phase. Items were presented one after the other with a 1500 ms presentation time and a 500 ms interstimulus interval. The final list item was followed by a 5000 ms retention interval. In the reconstruction phase, all items were displayed simultaneously on the screen. In the verbal task, drawings were randomly positioned in a predefined array of 5 (children) or 8 (adults) frames. In the spatial task, the dots were shown in the same locations as in the presentation phase. Participants were required to reproduce the serial order of drawings/dots by clicking on the items in the order in which they were presented. When an object was clicked, its shading changed to indicate that it has been selected. There was no time limit for responding and no possibility for error correction. After selecting the final item, the participants were shown a visual cue instructing them to hit the spacebar to initiate the next trial.

Procedure

Adults were tested in two to five-person groups in a sound-attenuated booth at the University of Kaiserslautern-Landau. Children were tested individually or in groups of up to five at their school in a large classroom, where

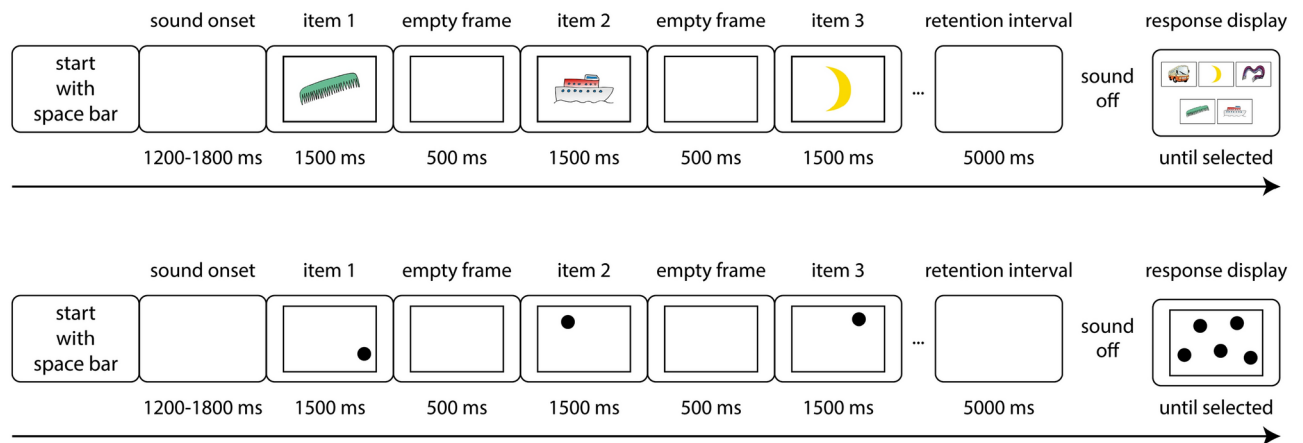


Fig. 4. The experimental procedure for the serial recall tasks. Top: Verbal Task, bottom: Spatial Task. The children were shown five items per trial in both tasks, while the adults were shown eight (verbal task) and seven (spatial task) items per trial. The items in the verbal task were drawings, while the items in the spatial task were uniform black dots. All items shown in the relevant trial were presented again on the response display in the reconstruction phase. The pictures in the verbal task were arranged in an array of 5 (children) and 8 (adults) frames. The spatial task's black dots were displayed in the same location as they had been shown in the trial. Note that the items are enlarged for illustrative purposes.

five computer workstations were set up about ten meters apart. A researcher or trained student assistant instructed each child individually. Adults received written instruction. Participants were instructed to ignore the sounds and focus on the serial recall task. For the verbal task, all pictures used were presented and named by a female speaker. To introduce the background sounds to the participants, both the environmental non-speech sounds and speech were played for 4 seconds each, followed by one practice trial per sound condition. Then, the children completed a total of 24 experimental trials, eight for each sound condition. Adults performed a total of 48 trials, 16 for each sound condition.

For both tasks, the three versions of each list were randomly assigned to the sound conditions. Sound conditions varied from trial to trial and were quasi-randomized: All sound conditions were presented in random order before being randomized again. In the trials with sound presentation, the sound started when the participant initiated the trial with the space bar. A random interval between 1200 and 1800 ms was introduced before the presentation of the first list item, to avoid correlations between sound and item onset⁸⁷. The sound was played throughout both the presentation phase and retention interval and stopped at the beginning of the reconstruction phase. The average sound level (LAeq) of the background sounds was 59 dB, as measured by ITA Artificial Head⁸⁸. The test session lasted about 25 minutes for children and 35 minutes for adults.

Data availability

The datasets generated during and analyzed during the current study as well as the materials are available in the OSF repository, https://osf.io/eayhb/?view_only=441bef7c4f5f487091b7e31f43b684dc.

Received: 26 June 2024; Accepted: 6 January 2025

Published online: 14 January 2025

References

- Klatte, M., Hellbrück, J., Seidel, J. & Leistner, P. Effects of classroom acoustics on performance and well-being in elementary school children: A field study. *Environ. Behav.* **42**, 659–692. <https://doi.org/10.1177/0013916509336813> (2010).
- Klatte, M. et al. Effects of aircraft noise on reading and quality of life in primary school children in Germany: Results from the Norah study. *Environ. Behav.* **49**, 390–424. <https://doi.org/10.1177/0013916516642580> (2017).
- Thompson, R. et al. Noise pollution and human cognition: An updated systematic review and meta-analysis of recent evidence. *Environ. Int.* **158**, 106905. <https://doi.org/10.1016/j.envint.2021.106905> (2022).
- Gheller, F., Spicciarelli, G., Scimemi, P. & Arfé, B. The effects of noise on children's cognitive performance: A systematic review. *Environ. Behav.* <https://doi.org/10.1177/00139165241245823> (2024).
- Salamé, P. & Baddeley, A. D. Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *J. Verbal Learn. Verbal Behav.* **21**, 150–164 (1982).
- Colle, H. A. Auditory encoding in visual short-term recall: effects of noise intensity and spatial location. *J. Verbal Learn. Verbal Behav.* **19**, 722–735. [https://doi.org/10.1016/s0022-5371\(80\)90403-x](https://doi.org/10.1016/s0022-5371(80)90403-x) (1980).
- Salamé, P. & Baddeley, A. D. Effects of background music on phonological short-term memory. *Q. J. Exp. Psychol. Sect. A* **41**, 107–122 (1989).
- Hanley, J. R. & Broadbent, C. The effect of unattended speech on serial recall following auditory presentation. *Br. J. Psychol.* **78**, 287–297. <https://doi.org/10.1111/j.2044-8295.1987.tb02247.x> (1987).
- Ellermeier, W. & Zimmer, K. Individual differences in susceptibility to the “irrelevant speech effect”. *J. Acoust. Soc. Am.* **102**, 2191–2199 (1997).
- Jones, D. M. & Macken, W. J. Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **19**, 369. <https://doi.org/10.1037/0278-7393.19.2.369> (1993).

11. Leist, L., Lachmann, T., Schlittmeier, S. J., Georgi, M. & Klatte, M. Irrelevant speech impairs serial recall of verbal, but not spatial items in children and adults. *Mem. Cogn.* **51**, 307–320. <https://doi.org/10.3758/s13421-022-01359-2> (2023).
12. Salamé, P. & Baddeley, A. D. Phonological factors in STM: Similarity and the unattended speech effect. *Bull. Psychon. Soc.* **24**, 263–265. <https://doi.org/10.3758/bf03330135> (1986).
13. Salamé, P. & Baddeley, A. D. Noise, unattended speech and short-term memory. *Ergonomics* **30**, 1185–1194 (1987).
14. Baddeley, A. D. & Logie, R. H. Working memory: The multiple-component model. In *Models of Working Memory* (eds Baddeley, A. D. & Logie, R. H.) (Cambridge University Press, 1999).
15. Jones, D. M., Miles, C. & Page, J. Disruption of proofreading by irrelevant speech: Effects of attention, arousal or memory?. *Appl. Cogn. Psychol.* **4**, 89–108. <https://doi.org/10.1002/acp.2350040203> (1990).
16. LeCompte, D. C., Neely, C. B. & Wilson, J. R. Irrelevant speech and irrelevant tones: The relative importance of speech to the irrelevant speech effect. *J. Exp. Psychol. Learn. Mem. Cogn.* **23**, 472–483. <https://doi.org/10.1037/0278-7393.23.2.472> (1997).
17. Surprenant, A. M. Effects of noise on identification and serial recall of nonsense syllables in older and younger adults. *Aging Neuropsychol. Cogn.* **14**, 126–143. <https://doi.org/10.1080/13825580701217710> (2007).
18. Viswanathan, N., Dorsi, J. & George, S. The role of speech-specific properties of the background in the irrelevant sound effect. *Q. J. Exp. Psychol.* **67**, 581–589. <https://doi.org/10.1080/17470218.2013.821708> (2014).
19. Divin, W., Coyle, K. & James, D. T. T. The effects of irrelevant speech and articulatory suppression on the serial recall of silently presented lipread digits. *Br. J. Psychol.* **92**, 593–616. <https://doi.org/10.1348/000712601162365> (2001).
20. Klatte, M., Kilcher, H. & Hellbrück, J. Wirkungen der zeitlichen struktur von hintergrundschall auf das arbeitsgedächtnis und ihre theoretischen und praktischen implikationen. *Zeitschrift für Exp. Psychol.* **42**, 517–544 (1995).
21. Ellermeier, W. & Hellbrück, J. Is level irrelevant in “irrelevant speech”? Effects of loudness, signal-to-noise ratio, and binaural unmasking. *J. Exp. Psychol. Hum. Percept. Perform.* **24**, 1406–1414. <https://doi.org/10.1037/0096-1523.24.5.1406> (1998).
22. Schlittmeier, S. J., Hellbrück, J., Thaden, R. & Vorländer, M. The impact of background speech varying in intelligibility: Effects on cognitive performance and perceived disturbance. *Ergonomics* **51**, 719–736. <https://doi.org/10.1080/00140130701745925> (2008).
23. Schweppe, J. & Knigge, J. Irrelevant music: How suprasegmental changes of a melody's tempo and mode affect the disruptive potential of music on serial recall. *Mem. Cogn.* **48**, 982–993. <https://doi.org/10.3758/s13421-020-01037-1> (2020).
24. Schlittmeier, S. J., Weißgerber, T., Kerber, S., Fastl, H. & Hellbrück, J. Algorithmic modeling of the irrelevant sound effect (ISE) by the hearing sensation fluctuation strength. *Atten. Percept. Psychophys.* **74**, 194–203 (2011).
25. Bell, R., Mund, I. & Buchner, A. Disruption of short-term memory by distractor speech: Does content matter?. *Q. J. Exp. Psychol.* **64**, 146–168. <https://doi.org/10.1080/17470218.2010.483769> (2011).
26. Marsh, J. E., Hughes, R. W. & Jones, D. M. Interference by process, not content, determines semantic auditory distraction. *Cognition* **110**, 23–38. <https://doi.org/10.1016/j.cognition.2008.08.003> (2009).
27. Jones, D. M., Farrand, P., Stuart, G. & Morris, N. Functional equivalence of verbal and spatial information in serial short-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **21**, 1008–1018. <https://doi.org/10.1037/0278-7393.21.4.1008> (1995).
28. Jones, D. M., Beaman, C. & Macken, W. J. The object-oriented episodic record model. *Model. short-term memory* 209–238 (1996).
29. Jones, D. M. & Macken, B. In the beginning was the deed: Verbal short-term memory as object-oriented action. *Curr. Dir. Psychol. Sci.* **27**, 351–356. <https://doi.org/10.1177/0963721418765796> (2018).
30. Jones, D., Hughes, R. & Macken, W. Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *J. Mem. Lang.* **54**, 265–281. <https://doi.org/10.1016/j.jml.2005.10.006> (2006).
31. Cowan, N. *Attention and Memory* (Oxford University Press, 1995).
32. Cowan, N., Morey, C. C. & Naveh-Benjamin, M. *An Embedded-Processes Approach to Working Memory: How Is It Distinct From Other Approaches, and to What Ends?* (Oxford University Press, 2020).
33. Elliott, E. M. The irrelevant-speech effect and children: theoretical implications of developmental change. *Mem. Cogn.* **30**, 478–487. <https://doi.org/10.3758/bf03194948> (2002).
34. Elliott, E. M. & Briganti, A. M. Investigating the role of attentional resources in the irrelevant speech effect. *Acta Psychol.* **140**, 64–74. <https://doi.org/10.1016/j.actpsy.2012.02.009> (2012).
35. Elliott, E. M. et al. Distraction in verbal short-term memory: Insights from developmental differences. *J. Mem. Lang.* **88**, 39–50. <https://doi.org/10.1016/j.jml.2015.12.008> (2016).
36. Klatte, M., Lachmann, T., Schlittmeier, S. J. & Hellbrück, J. The irrelevant sound effect in short-term memory: Is there developmental change?. *Eur. J. Cogn. Psychol.* **22**, 1168–1191. <https://doi.org/10.1080/09541440903378250> (2010).
37. Schwarz, H. et al. Age differences in the irrelevant sound effect: A serial recognition paradigm. *psihologija* **48**, 35–43. <https://doi.org/10.2298/PSI1501035S> (2015).
38. Röer, J. P., Bell, R., Körner, U. & Buchner, A. Equivalent auditory distraction in children and adults. *J. Exp. Child Psychol.* **172**, 41–58. <https://doi.org/10.1016/j.jecp.2018.02.005> (2018).
39. Meinhardt-Injac, B., Imhof, M., Wetzel, N., Klatte, M. & Schlittmeier, S. J. The irrelevant sound effect on serial recall is independent of age and inhibitory control. *Auditory Percept. Cogn.* <https://doi.org/10.1080/25742442.2022.2064692> (2022).
40. Joseph, T. N., Hughes, R. W., Sörqvist, P. & Marsh, J. E. Differences in auditory distraction between adults and children: A duplex-mechanism approach. *J. Cogn.* <https://doi.org/10.5334/joc.15> (2018).
41. Bell, R., Röer, J. P., Dentale, S. & Buchner, A. Habituation of the irrelevant sound effect: Evidence for an attentional theory of short-term memory disruption. *J. Exp. Psychol. Learn. Mem. Cogn.* **38**, 1542–1557. <https://doi.org/10.1037/a0028459> (2012).
42. Morris, N. & Jones, D. M. Habituation to irrelevant speech: Effects on a visual short-term memory task. *Percept. Psychophys.* **47**, 291–297. <https://doi.org/10.3758/bf03205003> (1990).
43. Röer, J. P., Bell, R., Dentale, S. & Buchner, A. The role of habituation and attentional orienting in the disruption of short-term memory performance. *Mem. Cogn.* **39**, 839–850. <https://doi.org/10.3758/s13421-010-0070-z> (2011).
44. Jones, D. M., Macken, W. J. & Mosdell, N. A. The role of habituation in the disruption of recall performance by irrelevant sound. *Br. J. Psychol.* **88**, 549–564. <https://doi.org/10.1111/j.2044-8295.1997.tb02657.x> (1997).
45. Röer, J. P., Buchner, A. & Bell, R. Auditory distraction in short-term memory: Stable effects of semantic mismatches on serial recall. *Auditory Percept. Cogn.* **2**, 143–162. <https://doi.org/10.1080/25742442.2020.1722560> (2019).
46. Bergström, K., Lachmann, T. & Klatte, M. Wann stört lärm das geistige arbeitsgedächtnis? einfluss von aufgaben- und geräuschcharakteristiken bei der wirkung moderaten lärms auf arbeitsgedächtnisleistungen. *Fortschritte der Akustik. Beiträge zur 38. Jahrestagung für Akustik DAGA* (2012).
47. Klatte, M. & Hellbrück, J. Effects of irrelevant speech on serial recall of verbal and spatial materials. In Schick, A. & Klatte, M. (eds.) *Contributions to Psychological Acoustics. Results of the 7th Oldenburg Symposium on Psychological Acoustics*, 531–538 (1997).
48. Kvetnaya, T. Registered replication report: Testing disruptive effects of irrelevant speech on visual-spatial working memory. *J. Eur. Psychol. Stud.* **9**, 10–15. <https://doi.org/10.5334/jeps.450> (2018).
49. Tremblay, S., Parmentier, F. B., Hodgetts, H. M., Hughes, R. W. & Jones, D. M. Disruption of verbal-spatial serial memory by extraneous air-traffic speech. *J. Appl. Res. Mem. Cogn.* **1**, 73–79. <https://doi.org/10.1016/j.jarmac.2012.04.004> (2012).
50. Marsh, J. E. et al. Changing-state irrelevant speech disrupts visual-verbal but not visual-spatial serial recall. *J. Exp. Psychol. Learn. Mem. Cogn.* <https://doi.org/10.1037/xlm0001360> (2023).
51. Hughes, R. W., Vachon, F. & Jones, D. M. Disruption of short-term memory by changing and deviant sounds: Support for a duplex-mechanism account of auditory distraction. *J. Exp. Psychol. Learn. Mem. Cogn.* **33**, 1050. <https://doi.org/10.1037/0278-7393.33.6.1050> (2007).

52. Vachon, F., Marsh, J. E. & Labonté, K. The automaticity of semantic processing revisited: Auditory distraction by a categorical deviation. *J. Exp. Psychol. Gen.* **149**, 1360 (2020).
53. Röer, J. P., Körner, U., Buchner, A. & Bell, R. Attentional capture by taboo words: A functional view of auditory distraction. *Emotion* **17**, 740–750. <https://doi.org/10.1037/emo0000274> (2017).
54. Rettie, L. et al. Warning—taboo words ahead! Avoiding attentional capture by spoken taboo distractors. *J. Cogn. Psychol.* <https://doi.org/10.1080/20445911.2023.2285860> (2023).
55. Röer, J. P., Bell, R. & Buchner, A. Self-relevance increases the irrelevant sound effect: Attentional disruption by one's own name. *J. Cogn. Psychol.* **25**, 925–931 (2013).
56. Röer, J. P., Bell, R. & Buchner, A. Specific foreknowledge reduces auditory distraction by irrelevant speech. *J. Exp. Psychol. Hum. Percept. Perform.* **41**, 692–702. <https://doi.org/10.1037/xhp0000028> (2015).
57. Bell, R., Röer, J. P., Marsh, J. E., Storch, D. & Buchner, A. The effect of cognitive control on different types of auditory distraction: A preregistered study. *Exp. Psychol.* **64**, 359–368. <https://doi.org/10.1027/1618-3169/a000372> (2017).
58. Banbury, S. P. & Berry, D. C. Office noise and employee concentration: Identifying causes of disruption and potential improvements. *Ergonomics* **48**, 25–37 (2005).
59. Wetzel, N., Widmann, A. & Schröger, E. The cognitive control of distraction by novelty in children aged 7–8 and adults. *Psychophysiology* **46**, 607–616 (2009).
60. Wetzel, N., Schröger, E. & Widmann, A. Distraction by novel and pitch-deviant sounds in children. *Front. Psychol.* **7**, 1949 (2016).
61. Poloczek, S., Henry, L. A., Messer, D. J. & Büttner, G. Do children use different forms of verbal rehearsal in serial picture recall tasks? A multi-method study. *Memory* **27**, 758–771. <https://doi.org/10.1080/09658211.2018.1563615> (2019).
62. Schiano, D. J. & Watkins, M. J. Speech-like coding of pictures in short-term memory. *Mem. Cogn.* **9**, 110–114 (1981).
63. Elliott, E. M. et al. Multitab direct replication of flavell, beach, and chinsky (1966): Spontaneous verbal rehearsal in a memory task as a function of age. *Adv. Methods Pract. Psychol. Sci.* **4**, 251524592110181. <https://doi.org/10.1177/25152459211018187> (2021).
64. Lehmann, M. & Hasselhorn, M. Rehearsal dynamics in elementary school children. *J. Exp. Child Psychol.* **111**, 552–560. <https://doi.org/10.1016/j.jecp.2011.10.013> (2012).
65. AuBuchon, A. M. et al. Lexical access speed and the development of phonological recoding during immediate serial recall. *J. Cogn. Dev.* **23**, 624–643 (2022).
66. Morey, C. C. et al. Is verbal rehearsal strategic? An investigation into overt rehearsal of nameable pictures in 5- to 10-year-old children. *J. Cogn. Dev.* <https://doi.org/10.1080/15248372.2024.2389123> (2024).
67. Vachon, F., Labonté, K. & Marsh, J. E. Attentional capture by deviant sounds: A noncontingent form of auditory distraction?. *J. Exp. Psychol. Learn. Mem. Cogn.* **43**, 622–634. <https://doi.org/10.1037/xlm0000330> (2017).
68. Girden, E. ANOVA (SAGE Publications Inc., 1992).
69. Romine, C. B. & Reynolds, C. R. A model of the development of frontal lobe functioning: Findings from a meta-analysis. *Appl. Neuropsychol.* **12**, 190–201. https://doi.org/10.1207/s15324826an1204_2 (2005).
70. Burrage, M. S. et al. Age- and schooling-related effects on executive functions in young children: A natural experiment. *Child Neuropsychol.* **14**, 510–524. <https://doi.org/10.1080/09297040701756917> (2008).
71. Brouwers, S. A. et al. Cognitive development through schooling and everyday life: A natural experiment among Kharwar children in India. *Int. J. Behav. Dev.* **41**, 309–319 (2017).
72. Brod, G., Bunge, S. A. & Shing, Y. L. Does one year of schooling improve children's cognitive control and alter associated brain activation?. *Psychol. Sci.* **28**, 967–978 (2017).
73. Morrison, F. J., Kim, M. H., Connor, C. M. & Grammer, J. K. The causal impact of schooling on children's development: Lessons for developmental science. *Curr. Dir. Psychol. Sci.* **28**, 441–449 (2019).
74. Vernucci, S. et al. The structure of working memory during childhood: a systematic review. *J. Cogn. Psychol.* **33**, 103–118 (2021).
75. Morey, C. C. & Miron, M. D. Spatial sequences, but not verbal sequences, are vulnerable to general interference during retention in working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **42**, 1907–1918. <https://doi.org/10.1037/xlm0000280> (2016).
76. Buchner, A., Bell, R., Rothermund, K. & Wentura, D. Sound source location modulates the irrelevant-sound effect. *Mem. Cogn.* **36**, 617–628. <https://doi.org/10.3758/mc.36.3.617> (2008).
77. Bartolotti, J. et al. Listening to speech and non-speech sounds activates phonological and semantic knowledge differently. *Q. J. Exp. Psychol.* **73**, 1135–1149. <https://doi.org/10.1177/1747021820923944> (2020).
78. Kukona, A. & Jordan, A. Online mouse cursor trajectories distinguish phonological activation by linguistic and nonlinguistic sounds. *Psychon. Bull. Rev.* **30**, 362–372. <https://doi.org/10.3758/s13423-022-02153-6> (2022).
79. Leist, L. et al. Differential effects of task-irrelevant monaural and binaural classroom scenarios on children's and adults' speech perception, listening comprehension, and visual-verbal short-term memory. *Int. J. Environ. Res. Public Heal.* **19**, 15998. <https://doi.org/10.3390/ijerph192315998> (2022).
80. Hughes, R. W. Auditory distraction: A duplex-mechanism account. *Psych J.* **3**, 30–41 (2014).
81. Baddeley, A. D., Gathercole, S. & Papagno, C. The phonological loop as a language learning device. *Psychol. Rev.* **105**, 158–173. <https://doi.org/10.1037/0033-295x.105.1.158> (1998).
82. Dockrell, J. E. & Shield, B. The impact of sound-field systems on learning and attention in elementary school classrooms. *J. Speech Lang. Hear. Res.* **55**, 1163–1176 (2012).
83. Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).
84. Peirce, J. W. PsychoPy—psychophysics software in python. *J. Neurosci. Methods* **162**, 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017> (2007).
85. Parmentier, F. B. R., Elford, G. & Maybery, M. Transitional information in spatial serial memory: Path characteristics affect recall performance. *J. Exp. Psychol. Learn. Mem. Cogn.* **31**, 412–427. <https://doi.org/10.1037/0278-7393.31.3.412> (2005).
86. Parmentier, F. *Exploring the determinants of memory for spatial sequences* (Taylor & Francis, 2014).
87. Bell, R., Dentale, S., Buchner, A. & Mayr, S. Erp correlates of the irrelevant sound effect. *Psychophysiology* **47**, 1182–1191 (2010).
88. Schmitz, A. Ein neues digitales kunstkopfmesssystem [a new digital artificial head measuring system]. *Acustica* **81**, 416–420 (1995).

Acknowledgements

The authors thank all children, parents and participants who took part in the study and all research assistants, especially Aarabis Srijevaghavan, Hannah Plückebaum and Charlotte Mock for their help with data collection and Markus Kurtz for his technical assistance. The illustrations (Figure 4) were created by the first author. This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; KL 1305/4).

Author contributions

Conceptualization, M.K. and L.L.; methodology, M.K. and L.L.; software, L.L.; validation, L.L. and M.K.; formal analysis, L.L.; investigation, L.L.; resources, T.L.; data curation, L.L.; writing—original draft preparation, L.L.; writing—review and editing, all; visualization, L.L.; supervision, M.K.; project administration, L.L. and M.K.;

funding acquisition, M.K. and T.L. All authors reviewed the manuscript and agreed to the published version of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Declarations

Competing interests

The authors declare no competing interests.

Ethics declarations

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of the German Psychological Society (Deutsche Gesellschaft für Psychologie, DGPs). For the children, study approval was also obtained from the Rhineland-Palatinate school authority (Aufsichts- und Dienstleistungsdirektion, ADD). Informed written consent was provided by the adult participants and by the legal guardians of the children.

Additional information

Correspondence and requests for materials should be addressed to L.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025