The irrelevant sound effect in short-term memory: Is there developmental change?

Maria Klatte and Thomas Lachmann
Department of Psychology, University of Kaiserslautern, Kaiserslautern, Germany

Sabine Schlittmeier and Jürgen Hellbrück
Department of Psychology, Catholic University of Eichstätt-Ingolstadt, Eichstätt, Germany

With two experiments, effects of irrelevant speech and classroom noise on serial recall of common nouns presented pictorially were investigated in children and adults. Experiment 1 used fixed list lengths for children (first graders) and adults. Experiment 2 used list lengths adjusted to participants' (second–third graders, adults) individual spans. In both experiments, children and adults were equally impaired by irrelevant speech. This contrasts with a related study (differences in methodology) by Elliott (2002), who reported severe increase in the detrimental impact of irrelevant speech with decreasing age. In both experiments, classroom noise had no effect in overall analyses. For Experiment 1, however, separate group analyses revealed impairment in children. Results suggest that effects of irrelevant sounds on serial recall stem from two separate mechanisms: Specific interference due to the sounds’ automatic access to short-term memory, and/or attention capture. Only for the latter there is developmental change.

Keywords: Attention; Children; Development; Irrelevant sound effect; Noise; Serial recall; Short-term memory.
Over the last decades, a number of studies have shown that immediate serial recall of sequences of visually presented verbal items such as digits, syllables, or words is impaired if certain task irrelevant background sounds are presented during the experimental trials. This so-called “irrelevant sound effect” (ISE) was found for irrelevant background speech as well as for nonspeech sounds such as tones (e.g., Divin, Coyle, & James, 2001; Elliott, 2002; Jones & Macken, 1993) and instrumental music (Klatte, Kilcher, & Hellbrück, 1995; Schlittmeier, Hellbrück, & Klatte, 2008). However, the occurrence of the ISE has been shown to depend on the inherent properties of the irrelevant sound. The recall performance is specifically impaired by irrelevant background sounds with a changing state characteristic, i.e., sounds consisting of distinct auditory–perceptive objects that vary consecutively. For example, irrelevant sounds consisting of different consonants or tones evoke an ISE, whereas steady state sounds such as continuous broadband noise or repetitions of single syllables or tones have a minor effect or no effect at all (see for a review, Beaman, 2005a; Jones, 1999; Neath, 2000).

In the present study, we focused on developmental aspects of the ISE by comparing the effects of irrelevant speech and nonspeech sounds in children to those in adults. Hitherto, only few studies have addressed the ISE in children (Elliott, 2002; Elliott, Baghat, & Lynn, 2007; Klatte, Meis, Sukowski, & Schick, 2007). All of these found that the ISE is present in elementary school children.

Investigations on the ISE in children may help to answer practical as well as theoretical questions. Concerning the former, the ISE in children is of interest in the context of noise effects on children’s cognitive development. Verbal short-term memory is of major importance for the acquisition of oral and written language in children (Baddeley, 2003a; Baddeley, Gathercole, & Papagno, 1998; Lachmann, 2002; Steinbrink & Klatte, 2008). The importance of short-term memory in language and reading acquisition on the one hand and its sensitivity to noise-induced disruption on the other hand leads to the prediction that permanent exposure to irrelevant sounds of even moderate intensity may cause enduring deficits in these developmental domains. In fact, field studies have shown that indoor noise in classrooms affects children’s language and reading acquisition (Klatte, Hellbrück, Seidel, & Leistner, 2009; Maxwell & Evans, 2000; Shield & Dockrell, 2008). Thus, experimental studies on the ISE in children are important in order to construct acoustic environments suitable for learning.

From a theoretical viewpoint, studying the developmental changes of the ISE may help to better understand its nature and to decide on the adequacy of ISE models postulated so far. In the hitherto only study comparing the
ISE in children and adults, Elliott (2002) reported a severe increase in the magnitude of the ISE on serial recall performance of visually presented digits with decreasing age. Eight-year-old second-graders were more affected by irrelevant speech and tones than older children, who in turn were more affected than adults. The age effect was most pronounced with changing state speech, which evoked a dramatic impairment in the young children. According to Elliott, the age effect demonstrates the importance of attentional control in the ISE. Young children are less able to ignore irrelevant sounds and thus are more susceptible to noise-induced disruption. This poses some problems to current models of the ISE, which assume that irrelevant sounds have automatic access to short-term memory without explicitly specifying a role of attention. In the following section, we will introduce major theories on the ISE and discuss their predictions with respect to developmental change.

THE PHONOLOGICAL LOOP MODEL

The phonological loop is a subcomponent of working memory specialised for the retention of verbal material (Baddeley, 2003b). It consists of two components, a passive store holding verbal information in a phonological code, and an active rehearsal process, refreshing the contents of the phonological store in order to prevent trace decay. The rehearsal process is also necessary for transforming visually presented items into a phonological code prior to entry into the phonological store. In contrast, heard (spoken) items gain direct access to the phonological store, without the mediation of the rehearsal process. It is assumed that background speech or speech-like sounds have obligatory access to the phonological store, where they interfere with the representation of the memory list. The mechanisms underlying this disruption are not clearly specified. Originally, it was assumed that irrelevant sounds partially mask memory traces within the phonological store (Salamé & Baddeley, 1982, 1989). This view was later abandoned, as it is not in line with major characteristics of the ISE, such as its evocation by nonspeech sounds (Elliott, 2002; Jones & Macken, 1993), the impact of the sounds’ changing state, and the finding that the ISE with speech is unaffected by the phonological similarity between the irrelevant speech and the memory items (Jones & Macken, 1995; Larsen, Baddeley, & Andrade, 2000). A more recent version of the phonological loop interpretation assumes that irrelevant sounds draw on domain-specific resources which build up a representation of the order in the to-be-remembered list (Page & Norris, 2003). However, what still remains in the modified version is the basic assumption that visual items have to be translated into a phonological
code in order to become susceptible to irrelevant sounds. This leads to the prediction that young children should be less affected than adults, as they make less use of phonological coding when memorising items presented visually (Cowan & Kail, 1996; Henry, 1991; Palmer, 2000). This is clearly not in line with Elliott’s findings that children were affected by irrelevant sounds even more than adults.

THE OBJECT-ORIENTED EPISODIC RECORD MODEL

In their object-oriented episodic record (O-OER) model, Jones and his colleagues (Jones, 1999; Jones, Beaman, & Macken, 1996) argue that the ISE results from interference between different sets of order cues. According to this view, a sound consisting of different auditory elements, such as tones or syllables, is automatically represented in short-term memory as a sequence of objects joined by linkages. These linkages weaken the associations between adjacent items in the memory list, which are generated through deliberate rehearsal. In particular, it is assumed that changing state sounds affect the efficiency of the rehearsal process because they give rise to conflicting order information in a common representational space. In contrast, steady-state sounds, i.e., sounds consisting of only one repeated element, do not contain order information and thus do not interfere with serial order retention. The changing state effect, i.e., stronger disruption with changing state when compared to steady-state sounds, has been confirmed in numerous studies (Jones, Madden, & Miles, 1992; for a review, see Hughes & Jones, 2001).

It is difficult to derive predictions concerning developmental change from the O-OER model, as the model does not address this point. Elliott (2002) argues that, since rehearsal is a precondition of the ISE in the O-OER model, young children should be less affected than older children and adults, as they make less use of rehearsal (e.g., Palmer, 2000).

The phonological loop and O-OER models both agree that irrelevant sounds have direct and automatic access to short-term memory, where they interfere with the retention of the memory list. These direct-access models do not explicitly specify a role of attention in the ISE and thus do not predict that the ISE is mediated by the attentional abilities of the participants.

THE FEATURE MODEL

According to the feature model (Nairne, 1990; Neath, 2000), short-term memory traces consist of two types of features: modality-dependent features representing raw, physical aspects of the stimuli, and modality-independent
features representing abstract aspects independent of modality, such as identification and categorisation. The detrimental effect of irrelevant speech on short-term memory is modelled by adding noise to the modality-independent features. In particular, it is assumed that modality-independent features of the list items are replaced by modality-independent features of the irrelevant speech (“feature adoption”). This mechanism is clearly confined to speech as irrelevant sound, as nonspeech sounds do not share modality-independent features with the memory list items. In addition to feature adoption, irrelevant sounds are assumed to cause attentional distraction. It is argued that changing state sounds evoke stronger disruption than steady-state sounds, since the former divert attention away from the task, whereas the latter are easy to ignore (see Cowan, 1995, for a related view). The effects of nonspeech sounds are solely attributed to the attentional burden caused by the necessity to ignore the irrelevant sounds (Neath, 2000, p. 420). Since children are less able to focus attention on task-relevant information and less able to resist interference from irrelevant stimuli (Dempster, 1993; Gumenyuk, Korzyukov, Alho, Escera, & Naatanen, 2004), a stronger effect of changing state sounds in children is expected from the feature model. Consequently, the increase in susceptibility to sound-induced impairments with decreasing age observed in Elliott’s study is in line with the feature model, as it includes a specified role of attention.

RESEARCH INTEREST

In view of its theoretical and practical significance, Elliott’s (2002) finding that the ISE is stronger in children than adults clearly merits further exploration. This is particularly true in view of some discrepancies between the results concerning the ISE in children reported by Elliott, and by Klatte et al. (2007). In Elliott’s study, serial recall of visually presented digits in second-graders was dramatically impaired by irrelevant speech. Klatte et al. used a similar task, but found only a moderate impairment (13%) in second-graders, comparable to the effects found in adults in Elliott’s and other studies (e.g., Ellermeier & Zimmer, 1997). Methodological differences concerning sound presentation between the studies might have contributed to the different outcomes. For example, Klatte et al. used an unknown foreign language as irrelevant speech, whereas Elliott used words from the native language of the participants. Furthermore, the level of the irrelevant sounds was higher in Elliott’s study as compared to Klatte et al. Thus, in Elliott’s study, the sounds were more difficult to ignore. This might have led to the disproportionate increase in the disruption evoked in children, as children are less able than adults to focus attention on the task in the
presence of irrelevant sounds. However, the Klatte et al. study did not include adults, and thus does not allow any conclusions to be made concerning developmental change. The current study aimed to compare the ISE in children and adults using sound presentation methods comparable to those of Klatte et al.

In the present study, serial recall performance for common nouns presented pictorially was measured during irrelevant speech, classroom noise without speech, and silence. Speech as irrelevant sound is the standard sound condition to explore the ISE. In contrast to Elliott’s (2002) study, an unknown foreign language was used as irrelevant speech in order to avoid distraction due to semantic content. Classroom noise without speech was included for two reasons. First, children’s cognitive performance is most often accompanied by this kind of noise. Second, Elliott attributed the increase of the ISE with decreasing age to age-related differences in attentional control. Since classroom noise contains different acoustic events occurring in an unpredictable order, it is especially difficult to ignore. Thus, if attention plays a crucial role in ISE, children should be more impaired by classroom noise than adults. Recent studies have reported significant impairments of serial recall performance due to office noise in adults (Bell & Buchner, 2007; Schlittmeier & Hellbrück, 2009). However, the sounds used in these studies also contained speech, and thus do not allow predictions about the specific effects of nonspeech classroom noise.

**EXPERIMENT 1**

In earlier studies on ISE in children, visually presented digits were used as memory list items. In contrast, the present study investigated the ISE on serial recall of common nouns presented pictorially. This task is more suitable for young children who just entered school and might have still difficulties dealing with written words or digits. Furthermore, this task allows the usage of phonological coding to be tested by varying the word length. This is important because young children might rely on visual strategies when memorising visually presented verbal items, whereas adults presumably use phonological coding in these situations (Palmer, 2000). Such strategic differences might mediate the effects of irrelevant sounds, and thus should be controlled. Serial recall performance is lower for words of long spoken duration as compared to words of short spoken duration (Baddeley, Thomson, & Buchanan, 1975). The specific origin of the word length effect is still a matter of debate. Specifically, it is unclear whether or not it reflects subvocal rehearsal (e.g., Campoy, 2008). However, there is consensus that a word-length effect in serial recall of visual items reflects phonological
coding, provided that output time is kept constant for short and long words (Baddeley, Chincotta, Stafford, & Turk, 2002).

Method

Participants

Participants were native German speakers. There were 47 (15 male) student volunteers or employees from the University of Oldenburg, aged between 20 and 40 years (median = 25 years), who either received course credit or payment for participation. Furthermore, 53 first-grade children (26 male), aged between 6 years, 4 months and 7 years, 9 months (median = 7 years, 0 months) from five elementary schools in Oldenburg (a major city in the north of Germany) took part in the study. Each child received 5 Euro for participation for the class treasury. The children and their parents and teachers gave their agreement to this study. All participants had normal or corrected to normal vision and normal hearing.

Materials

Memory materials. Two sets of digital black and white line drawings were assembled (Traeger, 2005), one representing phonologically dissimilar words of long spoken duration, and one representing phonologically dissimilar words of short spoken duration. Set size was six items for children and ten items for adults. The set of short items consisted of pictures representing the monosyllabic words Bett (bed), Hund (dog), Topf (pot), Haus (house), Ball (ball), and Schiff (ship). These were used for children and adults. For the adults the additional items Stuhl (chair), Pilz (mushroom), Herz (heart), and Kamm (comb) were used. The set of long items consisted of pictures representing the multisyllabic words Schmetterling (butterfly), Luftballon (balloon), Telefon (telephone), Zahnbus (toothbrush), Banane (banana), and Pullover (pullover). These were used for children and adults. For the adults the additional items Wasserhahn (water-tap), Trompete (trumpet), Schraubenzieher (screwdriver), and Elefant (elephant) were used.

Irrelevant sounds. Serial recall performance was measured during silence and two different sound conditions: irrelevant speech, and classroom noise without speech. The irrelevant speech consisted of a Danish newspaper article read by a professional female Danish speaker. The record contained no reverberation and no remarkable changes in loudness and intonation. The classroom noise without speech contained typical classroom sounds such as moving chairs, scraping feet, coughing, leafing through papers, rattling with writing utensils, and opening and closing school bags. The record was produced in a sound-attenuated laboratory room equipped with
school furniture with the help of 12 children and adults using an artificial head system ("Cortex MK2").

Design and procedure

The task required immediate serial recall of common nouns presented pictorially. The task was carefully explained to the participants and practiced with examples. In the following, the number of items used for adults is presented first, followed by the number used for children presented in brackets. In each trial, a pseudorandomised combination of seven out of ten (four out of six) pictures from one of the two sets was presented.

As the word-length effect does not unambiguously reflect phonological coding when output time differs between lists of short and long words (Cowan et al., 1992), the recall procedure equated output time for pictures representing short and long spoken words in the following way. For adults, each trial was represented on the answer sheet as a random arrangement of the seven pictures used in that particular trial. Serial recall was carried out by writing the numbers 1 to 7 next to the appropriate drawings on the answer sheet. For children, each trial was represented on the answer sheet as a structured arrangement of the four pictures presented in that particular trial. Two pictures were printed in the upper field and two in the lower field of the arrangement, in randomised order. Additionally, between the upper and lower field, four boxes with the digits 1, 2, 3, and 4 in ascending order were displayed. Serial recall was carried out by drawing lines from the digits to the appropriate pictures (a line from digit 1 to the first picture, from digit 2 to the second picture, etc.). For illustration see Figure 1. This procedure was chosen for the children because pilot studies revealed that writing down the numbers from 1 to 4 took them a very long time, which may increase output interference dramatically. All participants, children and adults, were given a maximum of 20 s for the recall phase.

Item conditions (short vs. long items) were varied randomly within a session, with the restriction that no condition was repeated more than twice in succession. In each sound condition, 12 (8) trials were completed, half of them with long and half with short items.

The experiment was performed in a special laboratory room with 16 individual working places in the Hearing Research Centre at the University of Oldenburg. Paperboards were placed between adjacent seats in order to ensure that the participants could not see their neighbours’ answer sheets. The room is equipped with an acoustic system which allows presentation of ambient sounds by means of twelve loudspeakers, four at each of the side walls, and two at each of the front and back walls.

The presentation of the pictures and sounds was controlled by a notebook using standard presentation software. The graphics were presented
on a screen in front of the room via a projector which was located outside the room.

The two irrelevant sound conditions (irrelevant speech and classroom noise) were varied between subjects. Each participant performed the memory task in silence and in one of the two irrelevant sound conditions in a blocked design; 27 children and 23 adults performed the task in silence and with classroom noise, 26 children and 24 adults performed the task in silence and with irrelevant speech. The order of blocks was counterbalanced across participants.

Before a block with irrelevant sound was started, this sound was introduced to the participants for 30 s. Sound level, as measured at a central seat position, was 54 dB(A) ($L_{eq}$, 20 s) in each of the two sound conditions, and ranged from 53 to 56 dB(A) between seat locations. Instruction was given to ignore the sound as task irrelevant and to focus attention on the memory task. An experimental trial started with an acoustical signal sounding like a gong, presented for 1100 ms. After a 900 ms slot, the presentation of the irrelevant sound started, followed by the exposure of the first picture after 1000 ms. The pictures were presented one after another for 1000 ms, respectively, with an interstimulus interval of 1000 ms between
them. 1000 ms after presentation of the final list item, a slide showing a copy of the participants’ paper-and-pencil answer sheet was presented as a cue to start recall. The presentation of the irrelevant sound endured until the end of the recall phase. The silent control block followed the same procedure, except that no irrelevant sound was presented.

When children performed the experiment, the conductor was assisted by two students who took care that the children used the correct answer sheet and entered their answer in the correct line. One block took about 6–7 minutes to perform. The blocks were separated by a break. All children managed to remain silent while performing the task.

Results and discussion

Performance in the serial recall task was determined by summarising the number of items correctly recalled at their serial position. Mean percentage correct scores in the silent control and the irrelevant sound conditions with respect to age and sound group are displayed in Table 1.

The first analysis was performed on the data from the silent control condition in order to verify that the age groups did not differ with respect to overall error rate and in the usage of phonological coding. A two-factorial analysis of variance (ANOVA) with age (children vs. adults) as group factor and word length as within-subject factor revealed a significant main effect of word length, $F(1, 98) = 8.08$, $MSE = 163.64$, $p < .01$, partial $\eta^2 = .08$, with mean percentage correct $M = 75.9$ ($SD = 18.62$) for short words and $M = 70.64$ ($SD = 17.52$) for long words, but no effect of age and no interaction (both $F$s < 1). This indicates that overall task difficulty was comparable between the age groups, that participants used phonological coding in order to perform the memory task, and that this usage did not differ between children and adults.

Separate two-factorial analyses of variance (ANOVAs) were performed for each sound group (i.e., the group which performed the task with irrelevant

<table>
<thead>
<tr>
<th>Sound group</th>
<th>Classroom noise (CN)</th>
<th>Irrelevant speech (IS)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Sound</td>
</tr>
<tr>
<td>Children</td>
<td>76.50 (17.06)</td>
<td>68.40 (18.04)</td>
</tr>
<tr>
<td>Adults</td>
<td>70.08 (14.29)</td>
<td>70.76 (15.69)</td>
</tr>
</tbody>
</table>
speech and the group which performed the task with classroom noise) with age as a group factor and sound condition (silent control vs. irrelevant sound) as a within-subject factor. With respect to irrelevant speech, the analysis revealed a main effect of sound condition, $F(1, 48) = 10.29$, $MSE = 141.6$, $p < .01$, partial $\eta^2 = .18$, but no effect of age, and no interaction (both $Fs < 1$). Thus, we can verify that irrelevant speech impaired memory performance and that the ISE is present in children when speech is used as irrelevant sound. This replicates prior findings (Elliott, 2002; Klatte et al., 2007). However, the present finding that the magnitude of the irrelevant speech effect did not differ between 7-year-old children and adults is in contrast with Elliott’s (2002) findings. She reported a considerably larger ISE with changing state speech in 8-year-old children compared with adults. In her study, adults achieved 78% correct in the silent control condition and 67% with speech. This is comparable to the results in the present study (Table 1). In contrast, in Elliott's study children’s performance dramatically dropped from 72% correct in the control condition, which is comparable to the present finding, to 34% correct in the irrelevant speech condition, which contrasts to the 66% reached by the children in the present study (Table 1).

The issue of developmental change in the susceptibility to irrelevant speech is crucial with respect to the models discussed in the introduction. According to the feature model, changing state speech as used in the current study, affects serial recall through both the speech sounds’ specific interference with short-term memory representations and the capture of attention resulting from the changes inherent in the speech stream. In view of the development of attentional control in childhood, one might expect that the latter of these mechanisms is more pronounced in children. Thus, from the feature model, children should be more impaired by changing state speech than adults. Elliott’s result supports this notion. In contrast, our finding does not confirm the view that changing state speech captures attention. Our result is more in line with the phonological loop and the O-OER model, which both assume that irrelevant speech gains direct access into short-term memory, without the mediation of attention.

With respect to classroom noise, the analysis yielded no significant main effects of age group ($F < 1$) and sound condition, and no interaction. As stated in the introduction, prior studies have revealed reliable effects of nonspeech sounds in adults with sequences of tones varying in frequency, and with instrumental music. This suggests that the crucial sound characteristics for ISE evocation are inherent in tones and music, but not in the classroom noise used in this study. In the framework of the O-OER model, changing state tone sequences as well as instrumental music are automatically represented as streams of objects in short-term memory, which involve order information and thus give rise to interference with retention of serial order. Classroom noise, on the contrary, might not be perceived as a
unitary stream of auditory tokens, but as a loose aggregation of unrelated events without information about serial order.

The feature model attributes the effects of nonspeech sounds solely to attentional distraction. Since classroom noise contains different acoustic events occurring in unpredictable order, it is especially difficult to ignore. Therefore, according to this model, we may expect an ISE with classroom noise in children and adults, larger for the former than for the latter. This was not found in the present ANOVA. However, visual inspection of Table 1 suggests that classroom noise might impair performance in children. A pairwise comparison between children’s performance in silence versus classroom noise in fact revealed a significant ISE, $t(26) = 2.27, p < .05$, whereas in adults no such effect was evident. But still, the feature model cannot explain the absence of the ISE in adults. This is especially unexpected, since classroom noise is presumably even harder to ignore than the regular tone sequences which evoked an ISE in Elliott’s and other studies (Divin et al., 2001; Jones & Macken, 1993). Thus, the current results do not confirm the assumption of the feature model that effects of nonspeech sounds result exclusively from a binding of attention. We will return to this point in the General Discussion.

The most important result of Experiment 1 is that the dramatic increase in the disruption due to irrelevant speech in young children reported by Elliott (2002) was not replicated. Note, however, that there are considerable differences in methodology between Elliott’s and the present study, which may account for the contradictory results. The potential impact of the differences in sound presentation methods will be considered in the general discussion. A further difference between the studies is that of list-length specification. Elliott measured individual span prior to the irrelevant sound experiment, and in the latter, she administered lists according to the individual’s span. In the current study, in contrast, fixed list lengths were used in both age groups. Although the analysis confirmed that overall task difficulty was comparable between the age groups, one might argue that adjusting list length to individual span might be more appropriate to ensure equal task difficulty across participants.

Therefore, we decided to replicate Experiment 1, but this time we equated task difficulty for the two groups by adjusting sequence lengths to participants’ individual span. Furthermore, we aimed to make the tasks as comparable as possible between the age groups. In Experiment 1, the recall procedure differed between the age groups. This was motivated by results from pilot studies which indicated that writing the digits 1 to 4 is effortful and takes much time for many first graders. In Experiment 2, second- and third-grade children served as participants, who were able to use the same recall procedure as the adults. Finally, in Experiment 2, sound conditions (silence, irrelevant speech, and classroom noise) were varied within subjects, in order to increase the probability of detecting potential age differences in the magnitude of the ISE.
EXPERIMENT 2

Method

Participants

Participants were native German speakers. There were 24 (11 male) student volunteers or University employees from the University of Kaiserslautern, aged between 20 and 35 years (median = 22 years), who either received course credit or payment for participation. Furthermore, 21 second-to third-grade children (11 male), aged between 7 years, 5 months and 9 years, 6 months (median = 8 years, 5 months) took part in this study. Twelve of them were recruited from a childcare centre in Kaiserslautern (a major city in the west of Germany), and the others from different schools nearby. Children received a 5 Euro voucher, sweets, and giveaways for their participation. The children, their parents and educators as well as the City of Kaiserslautern gave their agreement to this study. All participants had normal or corrected to normal vision and normal hearing. After first completion of data collection, we found four participants, two children and two adults, who gave 100% correct responses in two or more conditions, which may reflect a ceiling effect. This may have resulted from an underestimation of the individual memory span, a problem also mentioned by Buchner, Bell, Rothermund, and Wentura (2008, p. 623). The data of these four subjects were replaced by those of two newly recruited children and adults, respectively. The final data set was collected from the participants described earlier in this section.

Materials

As memory material the same pictures representing monosyllabic words described in Experiment 1 were used. Word length was not varied. The sounds used for the irrelevant sound conditions were identical to those used in Experiment 1.

Design and procedure

Prior to the experiment proper, the participant’s memory span for the pictorial items was assessed. For the memory span test, memory lists were created that varied in length from three to ten items. Three lists of each list length were presented in ascending order. Presentation and recall procedures in the memory span task were identical to those in the serial recall task in Experiment 1, except that the children recalled the order of the items in the same way as the adults did. Each trial was represented on the answer sheets as a random arrangement of the three to ten pictures presented in that particular trial. The order of the items was recalled by numbering the
pictures on the answer sheet. Span was defined as the highest list length for which the individual participant recalled at least one of three sequences correctly (cf. Buchner et al., 2008, and Elliott, 2002).

For the experiment proper, the same task as in Experiment 1 was performed, with the following exceptions: For both age groups, memory lists were drawn from a set of 10 items representing monosyllabic words, and the list length corresponded to the participant’s individual memory span. Furthermore, a self-paced procedure was chosen, i.e., adults started the following trial individually by mouse click after finishing the preceding one; for children, the trials were started by the experimenter who was sitting next to them.

The participants performed three blocks of 10 trials each. One block of trials was performed in silence, one with irrelevant speech, and one with classroom noise. In both age groups, order of sound conditions was counterbalanced across individuals, such that each sound condition occurred equally often in the first, second, and third experimental block. Prior to the first block, the irrelevant sounds were introduced to the participant for 30 s. The sounds were delivered via two loudspeakers (Genelec 8030 APM) located on the left and right side of the computer screen. Sound levels, measured at the position of the participants’ head, were 55 dB(A) (L eq, 20 s) in both conditions. Irrelevant sounds started 1 s before onset of the first list item and endured during recall. Instruction was given to ignore the sound as task irrelevant and to focus attention on the memory task.

All participants were tested individually in a single session lasting about 45 minutes. The adults and nine children were tested in a laboratory room at the University of Kaiserslautern. Twelve children were tested in a quiet office room in a childcare centre in Kaiserslautern.

Results and discussion

The mean span was $M = 4.81\ (SD = 0.75)$ for children, and $M = 7.12\ (SD = 1.12)$ for adults. Span ranged from 4 to 6 items for the children, and from 5 to 9 items for the adults. Mean percentage correct scores for serial recall performance with respect to sound condition and age are depicted in Figure 2.

A two-factorial ANOVA with age as a between-subjects factor and sound condition (silence, classroom noise, irrelevant speech) as a within-subjects factor yielded a main effect of sound condition, $F(2, 86) = 27.86, MSE = 88.05, p < .001$, partial $\eta^2 = .39$, but no effect of age and no interaction ($F < 1$ in both cases). Mean percentage correct scores were $M = 79.8\ (SD = 16.0)$ in the silent control condition, $M = 77.51\ (SD = 14.83)$ with classroom noise, and $M = 66.02\ (SD = 16.78)$ with irrelevant speech. Bonferroni-corrected post hoc tests confirmed lower performance with irrelevant speech as
compared with silence \( (p < .001) \) and classroom noise \( (p < .001) \), whereas the latter two did not differ.

The results of Experiment 2 are clear-cut. Irrelevant speech impaired performance, whereas classroom noise had no effect. The magnitude of the disruption evoked by irrelevant speech did not differ with age. This clearly replicates the finding of an equal effect of irrelevant speech in children and adults in Experiment 1, and thus suggests that this finding is independent from the method of list-length specification. As in Experiment 1, the disruption evoked by irrelevant speech in children was much weaker than that reported by Elliott (2002). Classroom noise had no effect on performance in any group, even when compared separately.

**GENERAL DISCUSSION**

The present study examined the effects of irrelevant sounds on serial recall performance in children and adults. The only study addressing the comparison of the ISE in children and adults conducted to date (Elliott, 2002), revealed a dramatic increase in the detrimental impact of irrelevant sounds with decreasing age. This suggests that the ISE is present in children
and is an issue of developmental change. This is a crucial finding with respect to theoretical questions and practical implications, which clearly merits further research.

In the current study, the effects of narrative speech and classroom noise without speech on serial recall of common nouns presented pictorially were analysed in children and adults. The main results can be summarised as follows. In Experiment 1, serial recall performance for visual-pictorial items was impaired by irrelevant speech in first-grade children and adults. The magnitude of the irrelevant speech effect did not differ with age. Classroom noise had no effect and did not interact with age in the overall analysis. Separate analyses in each group, however, proved that children performed better in silence as compared with classroom noise. In Experiment 2, the finding that the irrelevant speech effect is independent of age was replicated, this time using sequence lengths adjusted to individual spans. Second- to third-grade children and adults were equally affected by irrelevant speech. Classroom noise had no effect on either the children or adults.

In the following we will discuss the ISE for speech and classroom noise separately, starting with the latter. We do so, because, as we will outline, we believe that two different mechanisms are initiated by these sounds.

Comparing the effect of classroom noise across Experiments 1 and 2, we may conclude that classroom noise impairs performance in first-graders, whereas older children and adults are unaffected. This result is best accounted for by assuming that classroom noise attracts attention. Young children are more susceptible to distraction when compared to older children and adults, and are thus more prone to noise-induced impairments. This explanation is in line with the feature model, which attributes the effects of nonspeech sounds to a capture of attention. However, the feature model attributes the effects of all nonspeech sounds to attention capture. With this assumption, it is difficult to explain why the classroom noise used in the present study did not evoke an ISE in adults, whereas regular tone sequences evoke a robust disruption (Divin et al., 2001; Elliott, 2002; Jones & Macken, 1993). Since classroom noise contains different acoustic events in unpredictable order, it is presumably even more capable of diverting attention away from the task than tone sequences. Thus, in view of the null effect of classroom noise in adults, it seems unlikely that the effect of tone patterns results from a capture of attention.

The O-OER model assumes that the ISE depends on the sound’s changing state characteristic. Following this model, classroom noise has no effect because it does not constitute a coherent auditory stream but is perceived as a succession of unconnected events. Therefore, classroom noise does not set up a representation of an ordered sequence in short-term memory, and thus does not interfere with serial order retention.
A similar explanation may be derived from the enhanced version of the phonological loop model (Page & Norris, 2003), which also attributes the ISE to interference with serial order retention due to conflicting order information automatically set up by ordered auditory material.

However, both direct-access models do not specify a role of attention in the ISE and thus cannot explain the current finding that young children, but not adults were affected by classroom noise. Furthermore, they exhibit difficulties in accounting for the effects of office noise (Bell & Buchner, 2007; Schlittmeier & Hellbrück, 2009) and nonspeech environmental sounds (Buchner et al., 2008) on serial recall performance. The sounds used in these studies were comparable to the classroom noise used here, in that they consisted of a montage of acoustical events from different sources. Thus, if we assume that classroom noise does not impair serial recall performance because it does not contain order information, the same should hold for office noise.

We may conclude that none of the common ISE models can account for the variety of effects found for nonspeech sounds. We should consider, however, that the term “nonspeech sound” refers to a heterogeneous category of any sound except speech. In our view, it is not promising to attribute the effects of all kinds of nonspeech sounds to one single mechanism. Rather, depending on its characteristics, a sound may evoke domain-specific interference (i.e., a changing state effect) or a capture of attention, or a mixture of both. Coherent auditory streams consisting of changing elements, such as sequences of different tones played by the same instrument, evoke a changing state effect. In contrast, sounds which consist of diverse auditory elements from different sources, such as classroom noise, office noise or mixtures of other environmental sounds, impair performance via attention capture. In these sounds, each element signifies the occurrence of a new event, which might be in some way important for the system.

The most important result is that the increase in the magnitude of the irrelevant speech effect in children compared with adults reported by Elliott (2002) was not found in the current study. In fact, the children’s increase in error rate due to irrelevant speech was by no means as strong as that reported by Elliott. In the current study, children’s performance dropped by 9% in Experiment 1 and 14% in Experiment 2 in the presence of irrelevant speech, whereas in Elliott’s study a 38% drop was found in 8-year-old children. Note that the magnitude of the disruption found in adults was approximately the same in both studies. Elliott found a decrement of adult performance by 11%, which is comparable with the findings of the current study.

The revealed age independency of the irrelevant speech effect clearly contradicts the feature model’s assumption that the ISE with changing state speech reflects an attention-capture process. This age independency was also
found in prior studies comparing the irrelevant speech effect in old and young adults (Beaman, 2005b; Belleville, Rouleau, van der Linden, & Collette, 2003; Rouleau & Bellville, 1996; see Bell & Buchner, 2007, for a review). According to the inhibitory deficit theory, old adults exhibit difficulties in preventing task-irrelevant information from entry into working memory. In line with this, it has been shown that old adults are less able than young adults to ignore task-irrelevant sounds when processing auditory or visual information (Lustig, Hacker, & Tonev, 2001). Thus, if attentional control mediates the irrelevant speech effect, one should not only find stronger impairments in children as compared with adults, but also stronger impairments in old adults as compared with young adults.

In addition to these developmental studies, there is further evidence against the attention-capture account of the disruption evoked by changing state speech. First, if attention capture is crucial, there should be habituation, which, however, did not occur in a number of studies (Ellermeier & Zimmer, 1997; Hellbrück, Kuwano, & Namba, 1996; Jones, Macken, & Mosdell, 1997; Tremblay & Jones, 1998). Second, according to the attention capture account, the disruption should increase with the number of different elements in the irrelevant speech sequence. Contrary to this prediction, it was shown that two alternating syllables evoke a reliable disruption which does not increase with adding further syllables (Hughes & Jones, 2005; Tremblay & Jones, 1998). Third, the finding that the irrelevant speech effect is independent from sound intensity in a wide dynamic range (Colle, 1980; Ellermeier & Hellbrück, 1998) is clearly not in line with an attention capture account. Finally, the attention capture account predicts that changing state speech should interfere with any attention-demanding task, i.e., the disruption should be domain-general. However, there is evidence that tasks involving serial order retention are particularly vulnerable to disruption evoked by changing state speech (Beaman & Jones, 1997; Jones & Macken, 1993).

So far we may conclude that no evidence for a role of attention in the ISE has been found, at least when typical ISE stimuli, e.g., narrative speech, syllables, or tones, were used. It should be mentioned, however, that some studies, which varied meaning-related aspects of the irrelevant speech stimuli, in fact reported evidence for an influence of attention. In particular, low frequency words were found to evoke a stronger disruption when compared to high-frequency words (Buchner & Erdfelder, 2005), and words of negative valence were identified to evoke a stronger disruption when compared to neutral words (Buchner, Mehl, Rothermund, & Wentura, 2006). Note, however, that in these studies, low frequency words and neutral words also evoked a reliable disruption, and that a number of studies including the present one revealed a reliable ISE with irrelevant speech that was meaningless to the participants. Altogether, we may thus conclude that,
under certain conditions, meaning-related aspects of the irrelevant speech may in fact capture attention, as in an extreme is the case when the name of participant is presented as irrelevant speech stimulus (Moray, 1959), but that the effect of semantics is not the primary source of disruption. It may rather reflect an independent additive contribution to the impairment evoked by irrelevant speech.

Results from other studies also support our view that specific interference and attention capture independently contribute to ISE (Hughes, Vachon, & Jones 2005, 2007). Hughes, Vachon, and Jones (2007), for instance, reported independent contributions of a speech sound’s changing state characteristic (repetitions of a single syllable vs. sequences of different syllables) and deviations in voice to impairments in serial recall performance; the former was concluded to reflect specific interference with serial order retention, the latter was interpreted to reflect attention capture. What does this mean with respect to the developmental view? As children are less able than adults to ignore sounds and to focus attention on the task, the impact of attentional factors is expected to be stronger in children. This could explain the contradictory results concerning age-related differences in Elliott’s (2002) and the present study. As already outlined, the sounds in Elliott’s study were harder to ignore as compared to those in the present study. We may assume that this has a rather small impact in adults, but a large one in children, leading to a disproportionate increase in the disruption for children. In more detail, in the present study, the level of the irrelevant speech was about 55 dB(A). In Elliott’s study, speech level was 72 dB(A). Furthermore, in the present study an unknown foreign language was used as irrelevant speech, whereas Elliott used words from the native language of the participants. Although the irrelevant speech effect has been shown to be independent of intensity (Colle, 1980; Ellermeier & Hellbrück, 1998) and meaningfulness (Buchner, Irmen, & Erdfelder, 1996; Klatte et al., 1995) in adults, this may not hold for children. If, as Elliott argues, the ISE in children stems from their inability to focus attention on the task in the presence of irrelevant sounds, children should be more impaired by meaningful speech and loud sounds. Moreover, in Elliott’s study sounds were localised within the head, a hearing sensation resulting from headphone presentation, which contrasts to the loudspeaker presentation in the present study. Furthermore, the participants in the current study were exposed to just one (Experiment 1) or two (Experiment 2) irrelevant sound conditions which were varied in a blocked design. In Elliott’s study, sound conditions were varied randomised, i.e., each participant was exposed to five auditory conditions varying quasirandomly from trial to trial. Again, this increases the difficulty in ignoring the sounds, and thus might lead to a disproportionate increase in the disruption in children. Indirect evidence for this assumption is derived from developmental studies on signal detection in noise. It has been shown
that age-related differences in the ability to detect a tone in noise are severely increased when the signal is embedded in an acoustic background that changes randomly on each presentation (Allen & Wightman, 1995; Oh, Wightman, & Lutfi, 2001). This effect has been attributed to age differences in the ability to ignore irrelevant sounds under conditions of uncertainty (Oh et al., 2001, p. 2893).

To summarise, the present results show that under conditions which minimise the impact of attentional factors, such as moderate sound intensity and blocked presentation of the background sounds, children and adults are equally affected by irrelevant speech. The present findings support our view that the effects of irrelevant sounds on serial recall performance result from two separate mechanisms. On the one hand, irrelevant speech and nonspeech sounds with a changing state characteristic have automatic access to short-term memory, where they interfere with the representation of the to-be-remembered list. This mechanism operates adult-like in elementary school children. On the other hand, irrelevant sounds may distract attention. The role of attention in the ISE depends on characteristics of the sounds, and on the attentional abilities of the participants. In line with this, in the present study classroom noise impaired performance in first-graders, whereas older children and adults were unaffected.

In future studies, care should be taken to avoid a superimposition of these mechanisms of impairment (i.e., domain-specific interference and attention capture) when comparing the ISE in children and adults. Experiments should be designed in a way to minimise distraction, or include a control task that demands attention but does not involve short-term memory, as has been done in studies with adults (e.g., Beaman & Jones, 1997; Hughes et al., 2007).

Results on ISE also have practical implications. As stated at the beginning of this paper, short-term memory plays an important role in children’s everyday cognition. In view of the finding that background speech of low to moderate intensity impairs children’s short-term memory, it is reasonable to assume that irrelevant speech also impairs other tasks in which short-term memory is involved, such as reading and spelling in reading beginners, listening comprehension, and learning the phonological structure of new words. In fact, effects of irrelevant speech on children’s listening comprehension and phonological processing have recently been reported (Klatte et al., 2007). Furthermore, field studies have shown that chronic exposure to irrelevant sounds, i.e., indoor noise in classrooms, has enduring effects on children’s language, prereading skills, and academic attainment (Klatte et al., 2009; Maxwell & Evans, 2000; Shield & Dockrell, 2008). In order to prevent such outcomes, optimal interior acoustics in the classrooms as well as pedagogical interventions are necessary. Caregivers, teachers, and parents should take care to adjust the acoustical conditions to the learning activities
of the children. When children are occupied with tasks involving short-term memory, background sounds such as speech, singing, and lively instrumental music should be avoided. It should be kept in mind that simply reducing the loudness level of the background sounds does not eliminate the disruptive effects.

REFERENCES


