The Disruptive Effects of Changing-State Sound and Emotional Prosody on Verbal Short-Term Memory in Blind, Visually Impaired, and Sighted Listeners

Kattner, Florian; Fischer, Marieke; Caling, Alliza Lejano; Cremona, Sarah; Ihle, Andreas; Hodgson, Timothy; Föcker, Julia

How to cite

KATTNER, Florian et al. The Disruptive Effects of Changing-State Sound and Emotional Prosody on Verbal Short-Term Memory in Blind, Visually Impaired, and Sighted Listeners. In: Journal of cognitive psychology, 2023, vol. 36, n° 1, p. 28–41. doi: 10.1080/20445911.2023.2186771

This publication URL: https://archive-ouverte.unige.ch/unige:179368

Publication DOI: 10.1080/20445911.2023.2186771

The Disruptive Effects of Changing-State Sound and Emotional Prosody on Verbal Short-Term Memory in Blind, Visually Impaired, and Sighted Listeners

Florian Kattner[1,2*⁑], Marieke Fischer[2], Alliza Lejano Caling[3], Sarah Cremona[3], Andreas Ihle[4,5], Timothy Hodgson[3], & Julia Föcker[3*]

[1]*Health and Medical University*, Potsdam, Germany

[2]*Technical University of Darmstadt*, Darmstadt, Germany

[3]*University of Lincoln*, College of Social Science, Lincoln, United Kingdom

[4]*University of Geneva*, Geneva, Switzerland

[5] *Swiss National Centre of Competence in Research LIVES – Overcoming vulnerability: Life course perspectives*, Lausanne and Geneva, Switzerland

[*]The first and last author contributed equally to this work.

Author Note

⁑*Corresponding author's address:*

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

Abstract

Previous findings suggest that blind listeners are less susceptible to auditory distraction in a verbal serial recall task, compared to sighted individuals. However, it is unclear whether this is due to more selective encoding / filtering of auditory information (e.g., efficient perceptual streaming) or to enhanced attentional control. To test these alternatives, the interference induced by changing-state sound was compared with the disruptive effect of emotional speech prosody (happy, angry, fearful intonations) in blind (n = 17), visually impaired (n = 23), and sighted (n = 89) individuals. While a reduced changing-state effect suggests more efficient perceptual filtering, a reduced emotional prosody effect indicates enhanced attentional control. Blind participants were able to recall more items compared to sighted and visually impaired individuals. Moreover, in sighted and visually impaired participants, the changing-state effect was found to be enhanced with angry prosody, whereas blind individuals were less susceptible to changing-state sound regardless of the prosody. The results also suggest group differences in prosody processing, as both visually impaired and blind participants were better able to ignore fearful speech, compared to other prosodies. The findings suggest that profound visual sensory deprivation leads not only to a capacity increment in verbal short-term memory, but enables listeners to shield memory against auditory distracters, suggesting more efficient perceptual streaming.

Keywords: auditory distraction; changing-state effect; emotional prosody; blindness; visual impairment

The Disruptive Effects of Changing-State Sound and Emotional Prosody on Verbal Short-Term Memory in Blind, Visually Impaired, and Sighted Listeners

It has long been known that performance in a cognitive task is susceptible to the presence of task-irrelevant and unattended background sound. Most research on auditory distraction has focused on a task that required participants to memorize the serial order of visually presented items (e.g., digits) for either immediate or delayed recall. Such serial recall from short-term memory suffers particularly when speech is presented as a distracter sound (e.g., Colle & Welsh, 1976; Ellermeier & Zimmer, 1997; Salamé & Baddeley, 1982), but other types of sound such as random tone sequences or music were also found to be disruptive (e.g., Jones & Macken, 1993; Salamé & Baddeley, 1989). It has been discussed to what extend these irrelevant sound effects can be explained with (a) specific interference with a particular process required for the focal task (e.g., serial rehearsal) or (b) attentional capture due to either the meaning of the irrelevant sound (e.g., self-relevant or emotional information) or a violation of expectations regarding the auditory background (e.g., and auditory deviant). More specifically, the disruptive effect of changing-state sound (i.e., when acoustic changes occur between successive sounds in a sequence) is assumed to be due to the pre-attentive process of organizing auditory information (streaming; Bregman, 1990), which interferes with a specific deliberate process such as serial rehearsal of the to-be-remembered items (interference-by-process; Hughes & Marsh, 2017; Jones & Tremblay, 2000). In contrast, an attentional capture account (e.g., Bell et al., 2012; Cowan, 1995; Marsh et al., 2014; Röer et al., 2011) assumes that irrelevant sound diverts attentional resources from the focal task either as a result of violation of the predictive model (e.g., a deviant sound in an otherwise predictable sequence) or due to the properties (e.g., semantic aspects) of a particular sound itself (predictabiliy-based vs. stimulus-specific attentional capture; see Eimer et al., 1996; Hughes & Marsh, 2020; Kattner et al., 2022).

According to the duplex-mechanism account, both mechanisms may be involved in different situations, depending on the properties of the sound, the requirements of the focal task, and the cognitive abilities of the individual listener (Hughes, 2014; Hughes et al., 2005, 2007, 2013). For instance, it has been shown that the disruptive effect of changing-state sequences of tones or syllables is restricted to tasks that require serial-order processing (Beaman & Jones, 1997; Hughes & Marsh, 2020; Jones & Macken, 1993), while the disruptive effect of auditory deviants (e.g. an unexpected change of the acoustic profile – voice, tempo, prosody – or semantic category) may be less dependent on serial-order processing (Hughes et al., 2007; Kattner & Ellermeier, 2018; Vachon et al., 2017, 2020). In addition, it has been reported that the deviation effect (attentional capture) may be related to individual working memory capacity (measured with the operation span task), whereas the changing-state effect is not (Hughes et al., 2013; Sörqvist et al., 2010, 2012). However, other studies indicated that both types of auditory distraction may be unrelated to working memory capacity (Körner et al., 2017), or that not all types of auditory deviants may be equally controllable (i.e., different change detection mechanisms; Sörqvist et al., 2013).

Generally, the changing-state effect is assumed to be a more automatic type of auditory distraction (e.g., due to interference between perceptual streaming and seriation processes) and less susceptible to top-down control. However, there is some indication that the disruptive effect of irrelevant speech in serial recall, which is typically explained with the changing-state nature of speech, may be related to individual differences in auditory selective attention. For instance, it has been found that task-irrelevant, free-running speech (in a foreign language) does not disrupt serial recall in blind individuals – with highly trained auditory attentional skills (Kattner & Ellermeier, 2014). Moreover, there is some indication that a five-session training of auditory selective attention using a dichotic listening task – requiring continuous filtering of irrelevant auditory information – reduces the magnitude of the irrelevant speech effect on serial recall in sighted participants by about a third (compared to a "placebo training" with an auditory duration disrimination task; Kattner & Ellermeier, 2020). These findings indicate that more selective auditory filtering may help to reduce the specific interference produced by changing-state sounds such as speech (i.e., these individuals may be able to specifically encode only the relevant information in short-term memory, while efficiently filtering the irrelevant acoustical changes). Alternatively, it could be assumed that the auditory-attentional training may have strengthened central attention, which can be used to reduce the degree of distraction that is due to attentional capture, but not the more peripheral and task-specific interference-by-process. That is, blind listeners and individuals trained in auditory selective attention may be capable of using top-down attentional control in order to reduce or eliminate the diversion of attentional resources from the focal task (or attention can be redirected back to the focal task more quickly) by irrelevant sounds that are either meaningful to the individual (e.g., due to emotional prosody) or that violate a predictive model based on previous stimulation. We note that both types of attentional capture may be unlikely to account for the absence of an irrelevant speech effect in blind listeners (Kattner & Ellermeier, 2014), because in that study the same excerpt of Finnish speech (unintelligible to the participants) was presented 48 times, thus generating a sound which does not have any meaning to the participants and which is highly predictable. Nevertheless, it is still necessary to demonstrate that the enhanced auditory processing skills of blind listeners help to reduce the task-specific interference produced by task-irrelevant changing-state sound.

Several studies have shown improved auditory perception and memory in blind individuals. It has been found that (early and congenitally) blind individuals outperform sighted individuals on various auditory perceptual tasks such as pitch discrimination (Gougoux et al., 2004), absolute pitch identification (Hamilton et al., 2004), and auditory temporal-order judgments (Stevens & Weaver, 2005), speech-in-noise discrimination (Muchnik et al., 1991), and sound localization (Gougoux et al., 2005; Lessard et al., 1998; Röder et al., 1999), even though there are typically no differences in audiometric thresholds (e.g., Collignon et al., 2006). In addition, blind individuals have been reported to have a larger digit span (e.g., blind children; Hull & Mason, 1995) and better memory for non-verbal auditory information such as voices or environmental sounds (Bull et al., 1983; Röder & Rösler, 2003). In another study, congenitally blind listeners demonstrated better recognition memory for words that were presented in a previous incidental encoding phase (the task during encoding was to judge whether the word was a semantically appropriate or inappropriate sentence ending). Importantly, the presentation of words that would be recognized later was accompanied by larger (late positive) ERP amplitudes in blind individuals, suggesting more efficient encoding of auditory verbal materials (Röder et al., 2001). Consistent with these findings, faster learning rates for name-voice associations were observed in blind individuals (together with enhanced ERP amplitudes for target voices), suggesting superiority at early voice processing stages (Föcker et al., 2012). To assess whether the memory benefits in blind individuals are due to more efficient (auditory) speech perception, verbal short-term memory was tested for pseudo-words embedded in noise, and the speech-to-noise ratio was manipulated to control for differences in speech recognition thresholds (Rokem & Ahissar, 2009). While the short-term memory for pseudo-words was better in blind compared to sighted participants, this advantage was eliminated completely when the stimuli were equivalent in terms speech recognition thresholds (80% identification of pseudo-words). Taken together, the results of these studies suggest that the short-term memory advantage in blind individuals results from more efficient auditory stimulus encoding.

In addition to blind individuals' performance benefits in auditory perception and memory, there is also some indication for differences in the processing of auditory emotional information. For instance, blind participants were found to respond faster than sighted participants to voices in a prosody discrimination task (accompanied by activation in the occipital cortex), and the faster detection of emotional prosody seems to be related to a stronger activation in the amygdala by fearful and angry voices (compared to neutral voices) in congenitally blind participants compared to sighted controls (Klinge et al., 2010). Furthermore, there is evidence for a more efficient detection of deviant syllables (e.g., "fefi" rather than "fefe") at an attended location and enhanced attention-related ERP amplitudes (auditory N1) regardless of the task-irrelevant emotional prosody (neutral, happy, fearful, threatening voices) in blind participants, whereas sighted participants showed enhanced ERP amplitudes only in response to fearful and threatening voices (Topalidis et al., 2020). According to the authors, these findings suggest a reorganization of the voice-processing system following congenital visual deprivation leading to emotion-independent effects of (spatial) auditory selective attention.

The purpose of the present study is to investigate potential differences between blind, visually impaired, and sighted participants with regard the processing of task-irrelevant emotional prosody during a serial recall task (i.e., nonsense syllables spoken with either neutral, happy, fearful, or angry intonation). It has been found previously that intensely fluctuating emotional prosody in irrelevant speech (i.e., "angry" articulations of words or sentences, which were associated with stronger psychoacoustical fluctuation strength) increases the distraction of performance in a serial recall task – but not in a non-serial short-term memory task – in sighted participants (Kattner & Ellermeier, 2018). In line with previous findings, it could be expected that emotional prosodies are identified more readily in blind individuals (Klinge et al., 2010), enabling them to reduce the distracting effect on serial recall. Further, if the experience of visual deprivation leads to an emotion-independent auditory attention system (Topalidis et al., 2020), then it could be predicted that the disruptive effect of angry prosody should be reduced in blind individuals. As it is unclear what degree of visual deprivation is necessary to lead to an efficient reorganization of the voice-processing system (see Topalidis et al., 2020), blind individuals were compared also to visually impaired participants who reported to perceive visual contrast (in addition to light/dark perception).

To further clarify whether the absence of an irrelevant speech effect in blind listeners (Kattner & Ellermeier, 2014) is related to more efficient auditory perceptual organization processes (as opposed to strengthened attentional control), the disruptive effect of task-irrelevant changing-state syllables (relative to steady-state syllables) on serial recall was compared between sighted, visually impaired and blind participants. Specifically, if blindness is related to more efficient encoding of auditory-verbal information (e.g., Rokem & Ahissar, 2009), then it could be expected that (verbal) changing-state sound may not be encoded in a way that interferes with the process of serial rehearsal of the to-be-remembered items.

Method

Participants

Two separate simulation-based power analyses were conducted to estimate the required sample sizes to observe (1) a modulation of irrelevant sound effects (quiet vs. steady-state vs. changing-state) by visual status (sighted, visually-impaired, and blind participants), and (2) a difference in prosody effects (neutral vs. emotional intonations) between the three groups (and possible interactions between prosody and changing-state effects)[1].

~~Based on the previously reported effect sizes for the steady-state effect (Cohen's d_z = 0.24-0.34) and changing-state effect (d_z = 0.35-0.61) in sighted participants (compare Bell et al., 2019, Exp. 1; Hughes et al., 2013, Exp. 1; Kattner & Bryce, 2022, Exp. 3) and assuming the absence of irrelevant speech effects in blind listeners (Kattner & Ellermeier, 2014), data were sampled from nine different normal distributions as defined by a 3 (group) × 3 (sound) mixed factors design (sound as a within-subjects factor). Specifically, for sighted group, effect sizes of d_z = 0.25 and d_z = 0.375 were assumed for steady-state and changing-state effects, respectively (μ_{quiet} = .74, μ_{steady} = .68, $\mu_{changing}$ = .59, σ = .12). For the blind group, we assumed no steady-state or changing-state effects (d_z = 0), similar average accuracy, but greater variance compared to the sighted group (μ = .67, σ = .14). For the visually impaired individuals, our best guess was to assume slightly reduced steady-state (d_z = 0.10) and changing-state effects (d_z = 0.25) compared to the sighted, as well as lower average accuracy, and higher variance compared (μ_{quiet} = .648, μ_{steady} = .62, $\mu_{changing}$ = .55, σ = .14). Due to the restricted availability of blind and visually impaired participants (together with a more complicated recruitment and data collection),~~

[1] We note that group differences in (a) auditory distraction (steady-state and changing-state effects) and (b) modulations of the changing-state effect by emotional prosody were tested with two separate analyses because the quiet control condition, which is required to test for a steady-state effect, cannot combined with the prosody factor, which is expected to modulate the changing-state effect.

unequal sample sizes were simulated. Data sampling was conducted with the total sample size varying between 30 and 300, but with four times more participants in the sighted group (20 to 200) compared to the blind and visually impaired groups (5 to 50 each). For each sample size, the simulation was repeated with 1000 iterations, and the proportion of significant effects ($\alpha < .05$) was used as an estimate of the statistical power to obtain the main effects of sound and group, and their interaction. The first power simulations revealed that a total sample size of 120 participants, consisting of 80 sighted, 20 blind, and 20 visually impaired participants will be sufficient to demonstrate the expected main effects of sound and group as well as the crucial interaction between group and sound with a statistical power of 80% or more (see Fig. S1A in the Supplemental Materials for an illustration of the power as a function of different unequal sample sizes). A second power simulation was conducted to estimate the minimum sample size to demonstrate differences in prosody effects between sighted and visually-impaired/blind individuals in a 3 (group) x 2 (state) x 4 (prosody) mixed-factors design with state and prosody as repeated-measures factors. Based on the differences in auditory distraction as a function of speech prosody observed in a previous study (irrelevant speech effects of $d_z = 0.45$ for angry vs. $d_z = 0.35$ for neutral prosody; Kattner & Ellermeier, 2018, Exp. 1), we expected increased changing-state effect in sighted individuals with angry prosody ($d_z = 0.45$ for steady-state vs. changing-state) compared to neutral prosody ($d_z = 0.375$), whereas happy and fearful intonations are expected to produce intermediate effect sizes of ($d_z = 0.40$). Assuming more efficient prosody processing due to blindness (Klinge et al., 2010; Topalidis et al., 2020), we simulated no changing-state effect regardless of the prosody in blind participants, but less distraction (higher memory accuracy during both steady- and changing-state speech) by utterances with fearful or angry prosody ($d_z = 0.25$ compared to neutral prosody; assuming greater variance, $\sigma = .14$,

~~compared to the sighted sample, σ = .12) – that is we expected a group × prosody interaction. For visually impaired individuals, we simulated reduced changing-state effect across all prosodies (d_z = 0.25, as in the first simulation), but also a general prosody effect on serial recall accuracy (*d* = 0.25 compared to neutral speech, σ = .14). A simulated data collection with 1000 iterations~~The second power simulation revealed that the~~a~~ sample size ~~of 80 sighted, 20 blind, and 20 visually impaired participants~~ is also sufficient to demonstrate main effects of group and state as well as the group × state and group × prosody interactions (α < .05) with a statistical power of 80% or more (see Fig. S1B in the Supplemental Materials for the relationship between power and sample size).

A total of one hundred twenty-nine participants were recruited in Germany and the United Kingdom to take part in this online experiment. To recruit visually impaired and blind participants, organizations such as the Hessian Association of the Blind and Visually Impaired ("Blinden- und Sehbehindertenbund Hessen e.V.") and the South Lincolnshire Blind Society were contacted. The sighted group comprised *n* = 89 participants (22 from Germany, 67 from the UK, 64 women, 23 men, 2 other), the visually-impaired group comprised *n* = 23 participants (14 from Germany, 9 from the UK, 17 women, 5 men, 1 unknown gender), and the blind group comprised 17 participants (14 from Germany, 3 from the UK, 10 female, 7 male). The blind individuals reported to have no perception of visual contrast, and the onset of blindness was either at birth (*n* = 11 congenitally blind), early (*n* = 2 blind since their first year of life), or late (*n* = 4 blindness onset after the age of 12 years). The most frequent causes of blindness were glaucoma, genetic defects, retinoblastoma, and retinitis pigmentosa. The visually impaired participants reported to have either full or partial perception of visual contrast. The onset of visual impairment ranged from birth (*n* = 16) to the age of 32 years. Causes of visual impairment were retinopathy of prematurity, optic nerve atrophy, retinitis pigmentosa, and hypoplasia. Additional information on the reported causes, the onsets of blindness / visual impairment (as well as their age), and the visual abilities can be found in Table S1 in the supplemental materials. Ages ranged between 21 and 72 years, with a mean age of 26.5 years (SD = 11.28) in the sighted group, 44.4 years (SD = 13.0) in the visually impaired group, and 42.8 (SD = 13.1) in the blind group. These group differences in age were significant, $F(2, 126) = 29.18$, $p < .001$, $\eta^2_G = .317$, with both blind and visually impaired participants being older than sighted participants ($p < .001$), whereas there was no significant age difference between blind and visually impaired participants ($p = .688$). Across all groups, 86.8% of participants were right-handed, 10.9% were left-handed, and 2.3% were ambidextrous.

The study was approved by the Ethics Committee of Technical University of Darmstadt (EK 19-2020, discussed and approved in the meeting of May 14, 2020) and the University of Lincoln's Ethics Committee (ethics approval code: 2020-3415). All experimental procedures adhered to the Declaration of Helsinki (2013). All participants provided written informed consent and they were able to participate in a raffle with the chance to win one of two £50 vouchers.

Apparatus

The experiment was designed as an online study and programmed in PsyToolkit (Stoet, 2010, 2017). The study could be run either on a desktop or laptop computer, but not on mobile phones or tablets, using any web browser except Safari. Most participants (75.2%) used Google Chrome to run the task, followed by Microsoft Edge/Internet Explorer (19.3%), Firefox (3.1%), and Opera (1%).

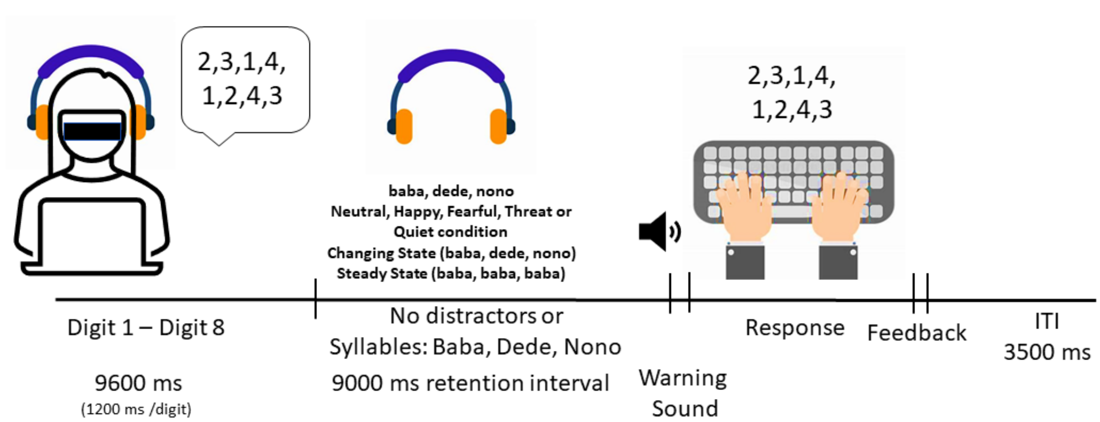
Participants were required to use headphones for the study. To make sure that participants were wearing headphones (not loudspeakers), the experiment started with a headphone-screening test (Woods et al., 2017). Prior to this test, continuous pink noise was presented and participants were asked to adjust the volume of their headphones to a comfortable level. In the headphone-screening test three 200-Hz tones (1000 ms each, separated by 500-ms intervals) were presented successively on each trial, and participants had to indicate which of the three tones was “softer” than the other two by hitting the respective number key (“1”, “2”, or “3”). A brief rising-pitch or falling-pitch whistling sound (296 and 307 ms) was presented as feedback after each response, indicating that the response was correct or incorrect, respectively. The level of one tone was 6 dB lower than the level of the two other tones, but one of the two tones with the higher level had the phase reversed between the left and right channels. This phase reversal reduces the sound pressure level in air (due to acoustical interference), thus allowing detection of the low-intensity tone only when headphones are used (if the stereo signal is presented in anti-phase, this should reduce the level that reaches the ear when using loudspeakers, thus making it difficult to detect the softer tone; see Woods et al., 2017). The headphone test was passed only if five or six responses were correct within a block of six trials. If fewer responses were correct, the test continued with the next block until either a block was passed or five blocks were completed, whichever came first. If the participant did not pass the headphone test within five blocks, a spoken message informed the participant that the study could not be continued due to an insufficient audio equipment. Participants were allowed to restart the experiment at any time.

Stimuli

The irrelevant speech sequences were created from recordings of two-syllable pseudo-words (“baba”, “nono”, “dede”) taken from a previous study (Gädeke et al., 2013). These pseudo-words were spoken by an actress with either neutral, happy, angry, or fearful emotional intonation (prosody). For steady-state sequences, the same randomly drawn pseudo-word was repeated nine times, whilst changing-state sequences consisted of nine pseudo-words drawn randomly with replacement (from the same prosody category). The pseudo-words were presented at a rate of one item per second, thus creating 9000-ms sequences. The actual duration of the spoken pseudo-words ranged between 400 and 700 ms (*neutral*: 516 ms, 557 ms, 651 ms; *happy*: 402 ms, 702 ms, 587 ms; *fearful*: 442 ms, 556 ms, 485 ms; *angry*: 399 ms, 556 ms, 485 ms for "baba", "dede", and "nono", respectively).

Figure 1.

Procedure of the auditory serial recall task. Eight digits were randomly drawn with replacement from 1-4 and presented via headphones. Blindfolded participants were asked to memorize their serial order for a retention interval of 9000-ms during which different types of irrelevant sound were presented via headphones until the digits were to be entered manually with the four number keys of the keyboard. Auditory feedback was presented before the next trial started.



Design and Procedure

A 3 (visual status: sighted, visually impaired, blind) x 2 (state of sound: changing-state, steady-state) x 4 (emotional prosody: neutral, happy, angry, fearful) mixed-factors design was implemented with state of sound and emotional prosody as within-subjects factors. Each combination of state of sound and emotional prosody was presented on six trials each. Together with six additional control trials without irrelevant sound (silence), this resulted in a total of 54 trials, which were presented in fully random order.

All participants were asked to wear headphones throughout the entire experiment. The main task started after the informed consent, some demographical questions, and the headphone screening test (see above). Instructions for the serial recall task were presented as html text, which could be read by the screen reader (for the blind listeners). The procedure of the serial recall task is illustrated in Figure 1. At the beginning of the task, an audio message informed all participants to put on the blindfold now, and to start the task by pressing the space bar. On each trial, a sequence of eight pre-recorded spoken digits was drawn randomly from the digits 1-4 (with replacement) and presented binaurally at a rate of 1200 ms per digit. The digits were spoken by either a German or English native speakers, depending on the sample (Technical University of Darmstadt or University of Lincoln). After the eighth digit, a retention interval of 9000 ms followed containing the irrelevant sound (or silence), before participants were asked to recall the series of digits in the presented order[2]. A 350 Hz tone was presented to prompt participants to enter the digits using the number keys (1-4) on their keyboard. After the eight digits had been entered, auditory feedback was presented in the form of a sequence of 0-8 rising whistling sounds (296 ms each), with the number of whistles corresponding to the number of correct digits. The next trial started automatically after an inter-trial interval of 3500 ms. Participants could take five short breaks during the task (after 8, 17, 26, 35, and 44 trials). An audio message informed the participants about the breaks and that they could continue at any

[2] To avoid partial acoustical masking, irrelevant sound was not presented during the acoustical presentation of to-be-remembered items. Studies have shown that irrelevant sound during the retention interval is as disruptive as during the encoding phase (e.g., Macken, 1999; Röer et al., 2014).

time by pressing the space bar. After 54 trials of serial recall, participants were debriefed and thanked for their participation.

Results

Serial recall accuracy was determined as the proportion of digits that were recalled in the correct serial position. In addition to the classical frequentist analyses of variance (ANOVA), analogue Bayesian ANOVAs were conducted in R (using the aov_ez() and lmBF() functions from the {afex} and {BayesFactors} packages, respectively), including random slopes for all repeated measures factors (as recommended by Oberauer, 2022; van den Bergh et al., 2022) in order to derive inclusion Bayes factors (BF_{Incl}) to estimate the contribution of each main effect and interaction term (i.e., the likelihood of the data given that the term is included in the model). For the individual contrasts, simple Bayes factors BF_{10} are reported indicating the likelihood of the alternative hypothesis (a difference between the two conditions) relative to the null hypothesis (no difference).

Figure 2 illustrates the different effects of auditory distraction separately for the sighted, visually impaired, and blind participants. A 3 (sound: quiet, steady-state, changing-state) × 3 (group: sighted, visually impaired, blind) mixed-factors ANOVA with sound as a repeated-measures factor and group as a between-subjects factor (Huynh-Feldt corrections of the degrees of freedom were used to compensate for violations of the sphericity assumption in the main effect of sound and the group × sound interaction, Mauchly's $W = .805$, $p < .001$, $\varepsilon = 0.847$) revealed a significant main effect of group, $F(2, 126) = 5.33$, $p = .006$, $\eta^2_G = .069$, $BF_{Incl} = 5.57$, with better recall performance in blind ($M = .75$, $SD = .15$) compared to visually impaired ($M = .63$, $SD = .16$) and sighted individuals ($M = .63$, $SD = .13$). Pairwise post-hoc comparisons (using "Holm" corrections; Holm, 1979) revealed that blind listeners' accuracy in the serial recall task was significantly higher compared to sighted, $d = 0.57$, $t(126) = 3.22$, $p = .005$, $BF_{10} = 25.64$, and visually impaired individuals, $d = 0.47$, $t(126) = 2.64$, $p = .019$, $BF_{10} = 2.63$, whereas there was no significant difference between sighted and visually impaired participants, $t(126) = -0.03$, $p = .975$, $BF_{10} = 0.24$. In addition, the ANOVA revealed a significant (and highly likely) main effect of sound, $F(1.69, 213.46) = 9.04$, $p < .001$, $\eta^2_G = .008$, $BF_{Incl} = 45200$, with better overall performance in the quiet ($M = .66$, $SD = .16$) and steady-state ($M = .65$, $SD = .16$) conditions compared to changing-state ($M = .62$, $SD = .15$). Here, Holm-corrected post-hoc comparisons revealed a significant irrelevant speech effect (quiet vs. steady-state and changing-state combined), $d_z = 0.24$, $t(126) = 2.66$, $p = .018$, $BF_{10} = 6.68$, and a significant changing-state effect (steady-state vs. changing-state), $d_z = 0.33$, $t(126) = 3.76$, $p < .001$, $BF_{10} = 15690.05$, but there was no significant (and Bayesian evidence against a) steady-state effect (quiet vs. steady-state), $d_z = 0.11$, $t(126) = 1.24$, $p = .217$, $BF_{10} = 0.17$. There was no interaction between sound and group, $F(3.39, 213.46) = 0.62$, $p = .622$, $\eta^2_G = .001$, $BF_{Incl} = 0.251$, indicating that steady-state and changing-state effects did not differ as a function of the visual deprivation status when not accounting for the prosody of the irrelevant utterances (the Bayesian ANOVA revealed that the most likely model with two main effects for group and sound was about 16.74 times more likely than a model that also contained the interaction term)[3].

---

[3] We note that the inclusion of age as a covariate in the analysis did not reveal a significant main effect of age, $F(10, 116) = 0.79$, $p = .640$, $\eta^2_G = 0.057$, nor did age interact with the type of irrelevant sound, $F(16.78, 194.65) = 0.72$, $p = .782$, $\eta^2_G = 0.007$, while there were still

Figure 2.

Serial recall accuracy (proportion of digits recalled in the correct serial position) after retention in quiet or during irrelevant speech (changing-state or steady-state syllables) in sighted, visually impaired, and blind listeners. Error bars depict +/- 1 standard error of the mean.

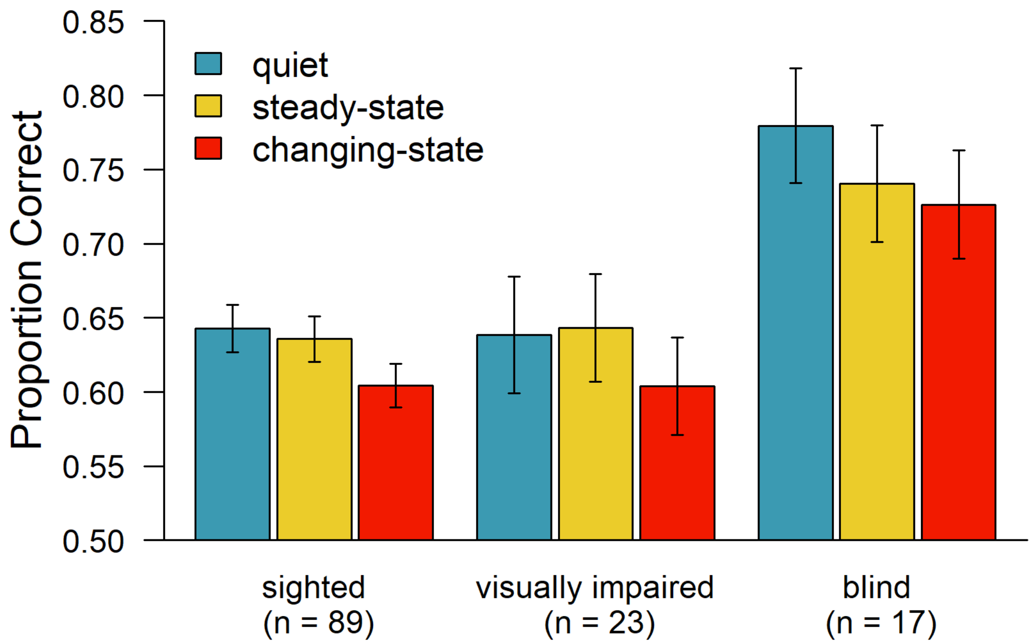


Figure 3 illustrates the changing-state effects in all three groups separately for the different emotional prosodies of irrelevant speech. The contribution of emotional prosody in irrelevant speech was analyzed with a separate 3 (group) × 2 (state: steady, changing) × 4 (prosody: neutral, fearful, happy, angry) mixed-factors ANOVA with state and prosody as repeated-measures factors (minimal Huynh-Feldt corrections ~~of~~ were applied to the degrees of freedom to compensate for the small and non-significant violations of the sphericity assumption

---

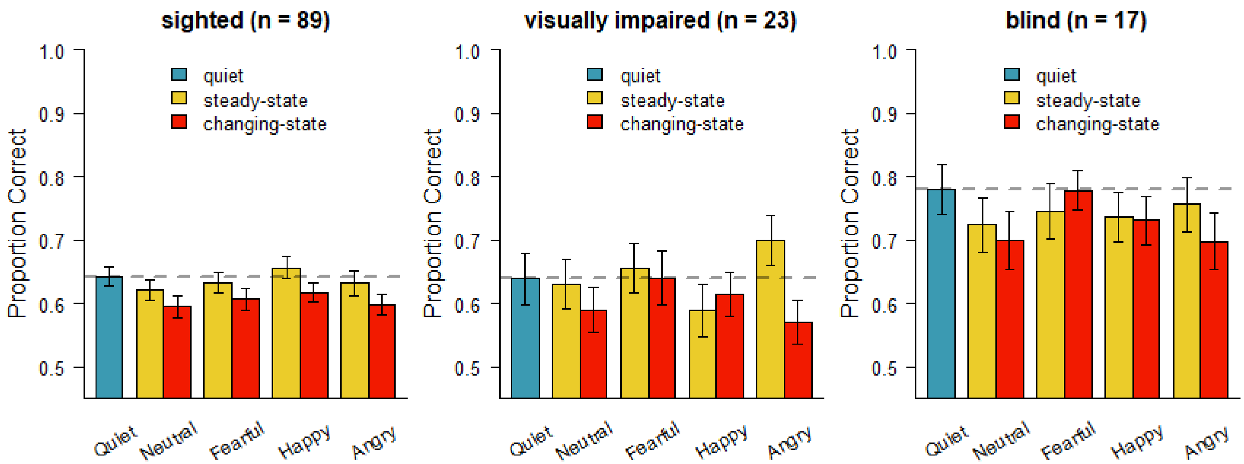
main effects of group, $F(2,116) = 3.21$, $p = .044$, $\eta^2_G = 0.047$, and sound, $F(1.68, 194.65) = 4.84$, $p = .013$, $\eta^2_G = 0.05$.

in the state × prosody and the three-way interaction, Mauchly's $W = .956$, $p = .342$, $\varepsilon = 0.996$; note that the quiet condition is not included in this analysis because it cannot be combined factorially with prosody) and group as a between-subjects factor. As expected, the analysis revealed again a significant main effect of group (and anecdotal Bayesian evidence), $F(2, 126) = 4.52$, $p = .013$, $\eta^2_G = .051$, $BF_{Incl} = 2.25$, and a main effect of state (i.e., an overall changing-state effect with extreme Bayesian evidence), $F(1, 126) = 14.10$, $p < .001$, $\eta^2_G = .005$, $BF_{Incl} = 8170$, but there was no interaction between group and state, $F(2, 126) = 0.72$, $p = .489$, $\eta^2_G < .001$, $BF_{Incl} = 0.59$. In addition, the analysis revealed a significant main effect of prosody (though with inconclusive Bayesian evidence), $F(3, 378) = 3.71$, $p = .012$, $\eta^2_G = .003$, $BF_{Incl} = 1.00$, with a significant increase in recall accuracy during fearful emotional intonations compared to neutral prosody, $t(126) = 3.52$, $d_z = 0.31$, $p = .004$, $BF_{10} = 5.07$ (Holm-adjusted for six pairwise comparisons). No other prosody contrast was significant, $p > .37$, but there was moderate Bayesian evidence for a difference between neutral and happy prosody, $BF_{10} = 3.18$. Importantly, the ANOVA also revealed a significant interaction between prosody and group (anecdotal Bayesian evidence), $F(6, 378) = 2.22$, $p = .041$, $\eta^2_G = .004$, $BF_{Incl} = 1.28$, indicating that the effect of emotional prosody differed as a function of the degree of visual deprivation. More specifically, a Holm-corrected contrast analysis of prosody effects in sighted individuals (each corrected for three tests) revealed a significant difference between neutral and happy prosody, $d_z = 0.27$, $t(126) = 3.03$, $p = .008$, $BF_{10} = 8.79$, but not between neutral and fearful, $t(126) = 1.33$, $p = .373$, $BF_{10} = 0.27$, or angry prosody, $t(126) = 0.77$, $p = .464$, $BF_{10} = 0.15$. In contrast, blind participants were significantly better during fearful speech than during neutral speech, $d_z = 0.22$, $t(126) = 2.44$, $p = .048$, $BF_{10} = 3.49$, but there was no difference between neutral and happy, $t(126) = 1.00$, $p = .639$, $BF_{10} = 0.43$, or between neutral and angry speech,

$t(126) = 0.72$, $p = .639$, $BF_{10} = 0.31$. Visually impaired participants in turn showed no significant differences between neutral and fearful, $t(126) = 2.17$, $p = .095$, $BF_{10} = 1.30$, neutral and happy, $t(126) = -0.39$, $p = .695$, $BF_{10} = 0.23$, or neutral and angry prosody, $t(126) = 1.44$, $p = .306$, $BF_{10} = 0.65$. The ANOVA further revealed a significant interaction between state and prosody, $F(2.99, 376.66) = 5.72$, $p < .001$, $\eta^2_G = .005$, $BF_{Incl} = 0.741$, with the changing-state effect – collapsed across all three groups – being significant with angry prosody, $d_z = 0.48$, $t(126) = 5.38$, $p < .001$, $BF_{10} = 4415.09$, but not with neutral, $d_z = 0.20$, $t(126) = 2.22$, $p = .085$, $BF_{10} = 2.87$ (note the anecdotal Bayesian evidence) fearful, $t(126) = 0.19$, $p > .999$, $BF_{10} = 0.25$, and happy prosody, $t(126) = 0.50$, $p > .999$, $BF_{10} = 1.06$. Interestingly, this interaction was further qualified by a significant three-way interaction between group, state, and prosody, $F(5.98, 376.66) = 3.07$, $p = .006$, $\eta^2_G = .005$, $BF_{Incl} = 0.79$, indicating that sighted participants showed a clear changing-state effect with angry prosody, $d_z = 0.24$, $t(126) = 2.65$, $p = .027$, $BF_{10} = 4.77$, and happy prosody, $d_z = 0.28$, $t(126) = 3.16$, $p = .008$, $BF_{10} = 14.36$, whereas the effect did not reach statistical significance with neutral, $d_z = 0.18$, $t(126) = 2.01$, $p = .093$, $BF_{10} = 0.77$, and fearful prosody, $d_z = 0.16$, $t(126) = 1.84$, $p = .093$, $BF_{10} = 0.51$ (all p-values Holm-corrected for four tests). Similarly, visually impaired participants demonstrated a significant changing-state effect only with angry prosody, $d_z = 0.45$, $t(126) = 5.04$, $p < .001$, $BF_{10} = 101.72$, but not with neutral, $d_z = 0.14$, $t(126) = 1.58$, $p = .347$, $BF_{10} = 0.50$, fearful, $d_z = 0.04$, $t(126) = 0.50$, $p = .621$, $BF_{10} = 0.24$, and happy prosody, $d_z = -0.09$, $t(126) = -1.06$, $p = .59$, $BF_{10} = 0.31$. However, in blind participants the changing state effect was neither significant with angry prosody, $d_z = 0.18$, $t(126) = 1.98$, $p = .20$, $BF_{10} = 1.03$, nor with any other prosody, $t(126) < 1.00$, $p > .997$, $0.25 < BF_{10} < 0.51$, suggesting that these individuals were less susceptible to disruption by changing-state sound (compare Figure 3).

Figure 3.

Serial recall accuracy (proportion of digits recalled in the correct serial position) in sighted, visually impaired, and blind listeners when different types of irrelevant sound were presented. Performance in quiet is considered a control condition as indicated by the horizontal dashed line. Error bars depict +/- 1 standard error of the mean.



Discussion

Short-Term Memory Capacity

The present study has shown that blind individuals have a remarkably better short-term memory for verbal serial information compared to sighted and visually impaired individuals (average recall accuracy of 77.9% vs. 64.3% and 63.9% for a sequence of eight digits in quiet). This finding suggests that the experience of profound visual deprivation may lead to a re-organization of the auditory-verbal processing system enabling the individual to encode and retain information more efficiently compared to individuals with no or less visual deprivation. The blind listeners' enhanced short-term memory capacity is thus compatible with previous findings of improved associative learning and recognition memory for auditory information (e.g., pseudowords, voices, environmental sounds), which has been explained in terms of more efficient perceptual encoding (Bull et al., 1983; Föcker et al., 2012; Röder & Rösler, 2003; Rokem & Ahissar, 2009). The blind listeners' short-term memory advantage in the present study is particularly impressive given that the blind participants were significantly older than the sighted participants. While age was found to have no statistical effect on serial recall, the blind participants' older age would have been expected to reduce (rather than enhance) performance in the serial recall task (e.g., Dobbs & Rule, 1989; Maylor et al., 1999; Salthouse, 1996), and thus have worked against their short-term memory advantage. Hence, the observed findings suggests that the short-term memory improvement due to an extended experience of visual deprivation may even have compensated possible age-related declines in short-term memory capacity. It is also interesting to note that the short-term memory advantage was observed only in blind, but not in visually impaired participants of equal age (i.e., individuals who were able to perceive visual contrast). This indicates that the auditory improvement and compensatory changes affecting cognitive functions (e.g., Röder et al., 2001; Röder & Rösler, 2003) may be induced only by profound visual deprivation.

In it important to note that due to the quasi-experimental study design, it is not safe to draw a clear causal conclusion with regard to the direct effect of a lack of visual experience or blindness on memory capacity. It is thus not entirely clear whether and which other variables that differ between blind, visually impaired, and sighted individuals (e.g., prior experience with short-term memory tasks, motivational differences, or age – which has been discussed above) may account for the observed differences in serial recall accuracy. In addition to cognitive benefits and a possible reorganization of verbal short-term memory in blind individuals, it is certainly possible that motivational factors may also have contributed to the blind participants' performance in the present study. In particular, since the study was conducted online, all participants were offered technical support (by phone), and this was used predominantly blind participants. While this additional direct contact to the experimenter was required to get the tasks running properly, it may also have enhanced the participants' commitment with the experiment and motivated them to show their best performance (it is possible that they may also have felt monitored by the experimenter).

Auditory Distraction of Serial Recall

The present study further investigated the disruptive effect of different types of irrelevant speech (steady-state and changing-state utterances) on serial recall in sighted, visually impaired, and blind individuals. Collapsed across the three groups, changing-state speech produced more disruption than steady-state speech, whereas there was less evidence for distraction by steady-state speech compared to silence (in contrast to Bell et al., 2019). However, we note that due to the relatively small sample sizes, there is more variability (noise) in the visually impaired and blind participants' data, and the statistical power to demonstrate effects of auditory distractors (due to the state or the prosody) is certainly reduced compared to the sighted sample. Nevertheless, there seems to be a systematic discrepancy in the pattern of significance as well as in the Bayesian evidence for changing-state effects as a function of both the status of visual deprivation and the emotional prosody. In general, both blind and visually impaired individuals demonstrated less disruption of serial recall by certain types of auditory distracters, indicating that they may be capable of shielding focal verbal or cognitive processing (i.e., retaining information in short-term memory) against the interference produced by irrelevant auditory information more efficiently.

Although task-irrelevant speech per se was found to disrupt serial recall also in blind participants (in contrast to Kattner & Ellermeier, 2014, who did not find a difference in recall accuracy between speech and white noise), there seems to be essentially no changing-state effect in blind individuals. That is, while changing-state speech (e.g., “baba-dede-nono”) produced more interference than steady-state speech (e.g., “baba-baba-baba”) in sighted and visually impaired, it did not in blind participants. Nevertheless, blind listeners also seem to be susceptible to irrelevant sound as they were about equally disrupted by steady-state and changing-state speech, compared to silence (see Figure 3). Hence, assuming that white noise can be considered a steady-state sound condition, the finding that blind individuals demonstrated no changing-state effect in the present study (changing-state speech vs. steady-state speech) appears to be consistent with the earlier finding of no irrelevant speech effect (changing-state speech vs. white noise) in blind individuals (Kattner & Ellermeier, 2014)

This pattern suggests that blind individuals may use more efficient or less automatic auditory streaming processes (Bregman, 1990), enabling them to avoid the specific interference between an irrelevant order of auditory objects (as contained in changing-state sound) with the to-be-remembered sequence of digits. More specifically, it is possible that the blind individuals’ more selective auditory encoding abilities enable them to filter task-irrelevant acoustical changes prior to the formation of an auditory stream, thus avoiding interference with serial-order processing. At the same time, it was found that both steady-state and changing-state speech disrupted serial recall compared to performance in quiet, suggesting that the mere presence of irrelevant auditory information demands some attentional resources and/or increases the perceptual load, which disrupts cognitive processing in blind listeners (e.g., due to attentional capture). The fact that there was essentially no difference between different types of speech and noise (white noise, steady-state, changing-state; see Kattner & Ellermeier, 2014) indicates that the temporal pattern of auditory information may be less important to the blind individuals' auditory processing system. It is also interesting to see that the pattern of auditory distraction (i.e., the changing-state effect) in visually impaired individuals – though not significant – resembles the pattern in sighted individuals more than the pattern in blind individuals. This indicates that extended visual deprivation may be necessary to induce a reorganization of the auditory processing system that enables more selective stream segregation of irrelevant sound (see Topalidis et al., 2020, for similar findings with regard to the reorganization of the voice-processing system in blind individuals).

Interestingly, although the Bayesian evidence is inconclusive in this regard, the degree of auditory distraction, and in particular the disruptive effect of changing-state speech, also seems to depend on the emotional prosody of the irrelevant utterances. In line with previous findings, an enhanced changing-state effect was observed in sighted participants when the utterances were spoken with angry or happy prosody, compared to neutral (and fearful) prosody (Kattner & Ellermeier, 2018). Similarly, visually impaired participants also demonstrated a large and reliable changing-state effect only with angry speech prosody, but not with other emotional (less threatening) or neutral prosodies. In contrast, blind demonstrated no changing-state effect regardless of the emotional prosody of the irrelevant speech. Hence, while we replicated the enhanced changing-state disruption with angry prosody in sighted participants, the modulation of the changing-state effect by emotional prosody seems to decrease with increasing visual deprivation. The observations in sighted and visually impaired individuals appear to be inconsistent with the results of a previous study, which showed stronger distraction by angry speech compared to neutral speech in sighted participants (Kattner & Ellermeier, 2018). This prosody-related amplification of the changing-state effect has been explained with the greater fluctuation strength in angry utterances (a psychoacoustical measure of amplitude and frequency modulations), facilitating the formation of discrete auditory objects during auditory scene analysis, and thus producing more interference with serial-order processing. It is possible that stronger fluctuations in amplitude and frequency may be required for visually impaired individuals to elicit a changing-state effect (i.e., the formation of irrelevant order cues during the perceptual streaming process may be prevented with milder fluctuations). However, in that former study, meaningful speech (either lists of words or full sentences) was presented as irrelevant sound, whereas artificial sequences of spoken syllables were presented in the present study. It is possible that emotionally intonated syllables may differ from emotional prosody in sentential speech in terms of the degree of acoustical intensity / frequency changes, and the perceived intensity of emotional prosody may differ between the studies. Therefore, it is still not entirely clear whether the enhanced distraction by angry speech (compared to neutral speech) was due to more interference-by-process (e.g., a stronger changing-state effect as a result of the more pronounced acoustical fluctuations in amplitude and spectrum) or due to stronger attentional capture by angry prosody information. However, the fact that the changing-state effect was absent and not modulated in blind participants suggests that profound visual deprivation may enable individuals to filter changing-state information regardless of the potentially attention-capturing emotional prosody information contained in irrelevant utterances.

Processing of Emotional Prosody

In addition to these prosody-dependent differences in the changing-state effect, the present data also suggest that emotional prosody cues per se are processed differently in sighted, blind and visually impaired individuals. Interestingly, the presence of fearful prosody in irrelevant speech improved serial recall in both visually impaired and blind participants, but not in sighted individuals, when compared with neutral prosody. On the other hand, happy emotional prosody was found to slightly improve serial recall accuracy in sighted individuals (note that the changing-state effect was also enhanced), but not in visually impaired and blind participants. It is important to note that the Bayesian evidence for a difference in prosody effects as a function of the visual status is only anecdotal and the interpretation of these differences in speculative and should be taken with caution. However, there was some indication of cognitive enhancement due to emotional prosody, and this might be consistent with other recent findings of emotional facilitation in the visual and auditory domain (Bocanegra & Zeelenberg, 2009; Stewart et al., 2022; Sussman et al., 2013). Under certain circumstances, cognitive performance may thus benefit from the presence of emotional prosody (in particular fearful voices), and individual differences in terms of auditory training or visual deprivation may account for these effects. The present findings are also in line with previous observations suggesting that both visually impaired and blind participants process emotional prosody information in voices differently than sighted individuals (Klinge et al., 2010; Topalidis et al., 2020). Specifically, blind and visually impaired individuals seem to encode fearful voices in a way that produced less disruption of serial recall compared to neutral or other prosodies, whereas sighted individuals' short-term memory did not benefit from the presence of fearful prosody.

Limitations

One limitation of the current study is the relatively small sample size of blind and visually impaired participants (although it is above average compared to other studies investigating blind or visually impaired individuals, e.g., Föcker et al., 2012; Gougoux et al., 2004; Rokem & Ahissar, 2009), and that the blind sample did not contain enough individuals with a late onset of blindness or visual impairment to further investigate the impact of the age of onset of blindness on short-term memory capacity, auditory distraction, and prosody effects. It could be argued that late blind individuals may be distracted more by emotional background syllables than congenitally and early blind individuals due to their past experience with visual information (and the use of the visual system) preventing the formation of a more efficient auditory-attentional processing system at early ages. Therefore, future experiments need to focus on larger sample sizes in the specific groups in order to understand how the onset of blindness might account for differences in auditory distractibility.

Due to the online nature of the study, not all variables can be controlled as strictly as in a lab environment (e.g., the sound pressure levels, unintended background noise, whether a blindfold was worn). Importantly, recent studies have documented that many of the benchmark effects of auditory distraction (and in particular the changing-state effect) can be studied quite reliably in an online setting producing similar effect sizes as in a well-controlled laboratory (Elliott et al., 2022; Kattner & Bryce, 2022). In addition, there are also studies with older adults producing robust results in an online setting, which has been recommended as an advantageous alternative to lab-based studies (Haas et al., 2022). The present study thus complements these recent demonstrations showing that online studies can be used even with special populations such as blind and visually impaired individuals.

Conclusion

The results of the present study not only demonstrate that blind listeners' verbal short-term memory has a higher capacity and may be less distracted by irrelevant sound, compared to sighted and visually impaired listeners. More specifically, a reliable changing-state effect was observed only in sighted participants, but not in blind individuals. Interestingly, blind participants demonstrated a general irrelevant speech effect, with impaired performance during speech (regardless of the state) compared to silence, suggesting that they may be susceptible to the portion of distraction that is due to attentional capture, but not to the specific interference produced by changing-state sound. To that effect, the present findings are consistent with a previous study showing no disruption by changing-state speech, compared to continuous noise, in blind listeners (Kattner & Ellermeier, 2014). Furthermore, the changing-state effect was found to depend also on the emotional prosody of the irrelevant utterances. In line with another previous study (Kattner & Ellermeier, 2018), the changing-state speech in sighted participants was more pronounced when the irrelevant utterances were spoken with an angry or happy prosody, compared to neutral or fearful prosody. However, while visually impaired individuals showed a similar modulation of the changing state effect by emotional prosody (i.e., more disruption with angry prosody), there was no changing-state effect regardless of the emotional prosody in blind individuals. In addition, the study indicates general differences in the processing of emotional speech prosody, as both visually impaired and blind participants demonstrated better overall memory performance during fearful irrelevant speech (compared to neutral or other emotional prosodies). This may suggest that visually deprived listeners – in contrast to sighted participants – may use prosody cues in a way to categorize approaching (happy, angry) vs. non-approaching (fearful) motivational tendencies at an earlier processing state.

Data Availability Statement

The stimulus materials, experimental code (PsyToolkit), demographic information (e.g., the cause of blindness or visual impairment) as well as the individual data and analysis scripts can be retrieved from an Open Science Framework (OSF) repository at the following link: https://osf.io/savke/?view_only=38a50de259c74b21b0c81fbb7c6bef8c

References

Beaman, C. P., & Jones, D. M. (1997). Role of serial order in the irrelevant speech effect: Tests of the changing-state hypothesis. *Journal of Experimental Psychology: Learning Memory and Cognition*, *23*(2), 459–471. https://doi.org/10.1037/0278-7393.23.2.459

Bell, R., Röer, J. P., Dentale, S., & Buchner, A. (2012). Habituation of the irrelevant sound effect: Evidence for an attentional theory of short-term memory disruption. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(6), 1542–1557. https://doi.org/10.1037/a0028459

Bell, R., Röer, J. P., Lang, A. G., & Buchner, A. (2019). Distraction by steady-state sounds: Evidence for a graded attentional model of auditory distraction. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(4), 500–512. https://doi.org/10.1037/xhp0000623

Bocanegra, B. R., & Zeelenberg, R. (2009). Dissociating Emotion-Induced Blindness and Hypervision. *Emotion*, *9*(6), 865–873. https://doi.org/10.1037/a0017749

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

Bull, R., Rathborn, H., & Clifford, B. R. (1983). The voice-recognition accuracy of blind listeners. *Perception*, *12*(2), 223–226. https://doi.org/10.1068/p120223

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior*, *15*(1), 17–31. https://doi.org/10.1016/S0022-5371(76)90003-7

Collignon, O., Renier, L., Bruyer, R., Tranduy, D., & Veraart, C. (2006). Improved selective and divided spatial attention in early blind subjects. *Brain Research*, *1075*, 175–182.

https://doi.org/10.1016/j.brainres.2005.12.079

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

Dobbs, A. R., & Rule, B. G. (1989). Adult age differences in working memory. *Psychology and Aging*, *4*(4), 500–503. https://doi.org/10.1037/0882-7974.4.4.500

Eich, E. (1995). Mood as a Mediator of Place Dependent Memory. *Journal of Experimental Psychology: General*, *124*(3), 293–308. https://doi.org/10.1037/0096-3445.124.3.293

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

Ellermeier, W., & Zimmer, K. (1997). Individual differences in susceptibility to the "irrelevant speech" effect. *Journal of the Acoustical Society of America*, *102*, 2191–2199. https://doi.org/10.1121/1.419596

Elliott, E. M., Bell, R., Gorin, S., Robinson, N., & Marsh, J. E. (2022). Auditory distraction can be studied online! A direct comparison between in-person and online experimentation. *Journal of Cognitive Psychology*. https://doi.org/https://doi.org/10.1080/20445911.2021.2021924

Föcker, J., Best, A., Hölig, C., & Röder, B. (2012). The superiority in voice processing of the blind arises from neural plasticity at sensory processing stages. *Neuropsychologia*, *50*(8), 2056–2067. https://doi.org/10.1016/j.neuropsychologia.2012.05.006

Gädeke, J. C., Föcker, J., & Röder, B. (2013). Is the processing of affective prosody influenced by spatial attention? An ERP study. *BMC Neuroscience*, *14*(14), 1–15. https://doi.org/10.1186/1471-2202-14-14

Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R. J., & Belin, P. (2004). Pitch discrimination in the early blind. *Nature 2004 430:6997*, *430*(6997), 309–309. https://doi.org/10.1038/430309a

Gougoux, F., Zatorre, R. J., Lassonde, M., Voss, P., & Lepore, F. (2005). A functional neuroimaging study of sound localization: Visual cortex activity predicts performance in early-blind individuals. *PLoS Biology*, *3*(2), 0324–0333. https://doi.org/10.1371/journal.pbio.0030027

Haas, M., Scheibe, S., El Khawli, E., Künzi, M., Ihle, A., Ballhausen, N., Framorando, D., Kliegel, M., & Zuber, S. (2022). Online assessment of cognitive functioning across the adult lifespan using the eCOGTEL: a reliable alternative to laboratory testing. *European Journal of Ageing*, *19*(3), 609–619. https://doi.org/10.1007/s10433-021-00667-x

Hamilton, R. H., Pascual-Leone, A., & Schlaug, G. (2004). Absolute pitch in blind musicians. *NeuroReport*, *15*(5), 803–806. https://doi.org/10.1097/00001756-200404090-00012

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. https://www.jstor.org/stable/4615733

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

Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Vachon, F., & Jones, D. M. (2013). Cognitive control of auditory distraction: impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology. Human Perception and Performance*, *39*(2), 539–553. https://doi.org/10.1037/a0029064

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

Evidence from perceptual-motor interference in verbal serial recall. *Journal of Experimental Psychology: Learning Memory and Cognition*, *43*(4), 537–551. https://doi.org/10.1037/xlm0000325

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

Hughes, R. W., Vachon, F., & Jones, D. M. (2005). Auditory attentional capture during serial recall: Violations at encoding of an algorithm-based neural model? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(4), 736–749. https://doi.org/10.1037/0278-7393.31.4.736

Hughes, R. W., Vachon, F., & Jones, D. M. (2007). Disruption of short-term memory by changing and deviant sounds: Support for a duplex-mechanism account of auditory distraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(6), 1050–1061. https://doi.org/10.1037/0278-7393.33.6.1050

Hull, T., & Mason, H. (1995). Performance of blind children on digit-span tests. *Journal of Visual Impairment and Blindness*, *89*(2 I), 166–169. https://doi.org/10.1177/0145482x9508900213

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 369–381. https://doi.org/10.1037/0278-7393.19.2.369

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

https://doi.org/10.3758/BF03214370

Kattner, F., & Bryce, D. (2022). Attentional control and metacognitive monitoring of the effects of different types of task-irrelevant sound on serial recall. *Journal of Experimental Psychology: Human Perception and Performance*, *48*(2), 139–158. https://doi.org/10.1037/xhp0000982

Kattner, F., & Ellermeier, W. (2014). Irrelevant speech does not interfere with serial recall in early blind listeners. *Quarterly Journal of Experimental Psychology*, *67*(11), 2207–2217. https://doi.org/10.1080/17470218.2014.910537

Kattner, F., & Ellermeier, W. (2018). Emotional prosody of task-irrelevant speech interferes with the retention of serial order. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(8), 1303–1312. https://doi.org/10.1037/xhp0000537

Kattner, F., & Ellermeier, W. (2020). Distraction at the cocktail party: Attenuation of the irrelevant speech effect after a training of auditory selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(1), 10–20. https://doi.org/10.1037/xhp0000695

Kattner, F., Richardson, B. H., & Marsh, J. E. (2022). The Benefit of Foreknowledge in Auditory Distraction Depends on the Intelligibility of pre-exposed Speech. *Auditory Perception & Cognition*, 1–18. https://doi.org/10.1080/25742442.2022.2089525

Klinge, C., Röder, B., & Büchel, C. (2010). Increased amygdala activation to emotional auditory stimuli in the blind. *Brain*, *133*(6), 1729–1736. https://doi.org/10.1093/BRAIN/AWQ102

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

Lessard, N., Paré, M., Lepore, F., & Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature*, *395*(6699), 278–280. https://doi.org/10.1038/26228

Marsh, J. E., Röer, J. P., Bell, R., & Buchner, A. (2014). Predictability and distraction: Does the neural model represent postcategorical features? *PsyCh Journal*, *3*(1), 58–71. https://doi.org/10.1002/pchj.50

Maylor, E. A., Vousden, J. I., & Brown, G. D. A. (1999). Adult age differences in short-term memory for serial order: Data and a model. *Psychology and Aging*, *14*(4), 572–594. https://doi.org/10.1037/0882-7974.14.4.572

Muchnik, C., Efrati, M., Nemeth, E., Malin, M., & Hildesheimer, M. (1991). Central Auditory Skills In Blind And Sighted Subjects. *Scandinavian Audiology*, *20*(1), 19–23. https://doi.org/10.3109/01050399109070785

Oberauer, K. (2022). The Importance of Random Slopes in Mixed Models for Bayesian Hypothesis Testing. *Psychological Science*, *33*(4), 648–665. https://doi.org/10.1177/09567976211046884

Röder, B., & Rösler, F. (2003). Memory for environmental sounds in sighted, congenitally blind and late blind adults: Evidence for cross-modal compensation. *International Journal of Psychophysiology*, *50*(1–2), 27–39. https://doi.org/10.1016/S0167-8760(03)00122-3

Röder, B., Rösler, F., & Neville, H. J. (2001). Auditory memory in congenitally blind adults: A behavioral-electrophysiological investigation. *Cognitive Brain Research*, *11*, 289–303. https://doi.org/10.1016/S0926-6410(01)00002-7

Röder, B., Teder-Sälejärvi, W., Sterr, A., Rösler, F., Hillyard, S. A., & Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature*, *400*(6740), 162–166.

https://doi.org/10.1038/22106

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

Rokem, A., & Ahissar, M. (2009). Interactions of cognitive and auditory abilities in congenitally blind individuals. *Neuropsychologia*, *47*(3), 843–848. https://doi.org/10.1016/j.neuropsychologia.2008.12.017

Salamé, P., & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, *21*, 150–164. https://doi.org/10.1016/S0022-5371(82)90521-7

Salamé, P., & Baddeley, A. D. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology Section A*, *41A*(1), 107–122. https://doi.org/10.1080/14640748908402355

Salthouse, T. A. (1996). The Processing-Speed Theory of Adult Age Differences in Cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Sörqvist, P., Halin, N., & Hygge, S. (2010). Individual differences in susceptibility to the effects of speech on reading comprehension. *Applied Cognitive Psychology*, *24*, 67–76. https://doi.org/10.1002/acp.1543

Sörqvist, P., Marsh, J. E., & Nöstl, A. (2013). High working memory capacity does not always attenuate distraction: Bayesian evidence in support of the null hypothesis. *Psychonomic Bulletin and Review*, *20*(5), 897–904. https://doi.org/10.3758/s13423-013-0419-y

Sörqvist, P., Nöstl, A., & Halin, N. (2012). Working memory capacity modulates habituation rate: Evidence from a cross-modal auditory distraction paradigm. *Psychonomic Bulletin and*

Review, *19*, 245–250. https://doi.org/10.3758/s13423-011-0203-9

Stevens, A. A., & Weaver, K. (2005). Auditory perceptual consolidation in early-onset blindness. *Neuropsychologia*, *43*(13), 1901–1910. https://doi.org/10.1016/j.neuropsychologia.2005.03.007

Stewart, E. K., Chen, V. V., Butler, B. E., & Mitchell, D. G. V. (2022). Emotional Distraction and Facilitation Across Sense and Time. *Emotion*. https://doi.org/10.1037/emo0001138

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

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

Sussman, T. J., Heller, W., Miller, G. A., & Mohanty, A. (2013). Emotional Distractors Can Enhance Attention. *Psychological Science*, *24*(11), 2322–2328. https://doi.org/10.1177/0956797613492774

Topalidis, P., Zinchenko, A., Gädeke, J. C., & Föcker, J. (2020). The role of spatial selective attention in the processing of affective prosodies in congenitally blind adults: An ERP study. *Brain Research*, *1739*, 146819. https://doi.org/10.1016/j.brainres.2020.146819

Vachon, F., Labonté, K., & Marsh, J. E. (2017). Attentional capture by deviant sounds: A noncontingent form of auditory distraction? *Journal of Experimental Psychology: Learning Memory and Cognition*, *43*(4), 622–634. https://doi.org/10.1037/xlm0000330

Vachon, F., Marsh, J. E., & Labonté, K. (2020). The Automaticity of Semantic Processing Revisited: Auditory Distraction by a Categorical Deviation. *Journal of Experimental*

Psychology: General, *149*(7), 1360–1397. https://doi.org/10.1037/xge0000714

van den Bergh, D., Wagenmakers, E., & Aust, F. (2022). *Bayesian Repeated-Measures ANOVA: An Updated Methodology Implemented in JASP*. https://psyarxiv.com/fb8zn/

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