Title: The role of spatial selective attention in the processing of affective prosodies in congenitally blind adults: An ERP study

Short Title: Processing of affective prosodies in congenitally blind adults

Authors and authors' affiliations: Pavlos Topalidis[1], Artyom Zinchenko[1], Julia C. Gädeke[2] & Julia Föcker[2,3]

[1] Department of Psychology and Educational Sciences, Ludwig Maximilian University, Munich Germany

[2] Biological Psychology and Neuropsychology, University of Hamburg, Germany

[3] University of Lincoln, School of Social Sciences, United Kingdom

***Corresponding author:**

Julia Föcker
School of Psychology
College of Social Sciences
University of Lincoln
United Kingdom
E-mail: JFöcker@lincoln.ac.uk

Keywords: congenitally blind, sensory deprivation, plasticity, auditory, attention, emotion, event-related potentials, human voice

Highlights:

- Congenitally blind individuals detected auditory syllables more efficiently compared to sighted controls
- Blind individuals revealed an enhanced auditory N100 amplitude compared to sighted controls
- The auditory N100 was higher in amplitude when the spatial location was attended compared to unattended
- This was true for all emotions in the blind, but only for the fearful and threatening emotions in sighted controls
- Differences between blind and sighted individuals were mainly observed at posterior electrodes

Title: The role of spatial selective attention in the processing of affective prosodies in congenitally blind adults: An ERP study

Short Title: Processing of affective prosodies in congenitally blind adults

Authors and authors' affiliations: Pavlos Topalidis[1], Artyom Zinchenko[1], Julia C. Gädeke[2] & Julia Föcker[2,3]

[1] Department of Psychology and Educational Sciences, Ludwig Maximilian University, Munich Germany

[2] Biological Psychology and Neuropsychology, University of Hamburg, Germany

[3] University of Lincoln, School of Social Sciences, United Kingdom

***Corresponding author:**

Julia Föcker
School of Psychology
College of Social Sciences
University of Lincoln
United Kingdom
E-mail: JFöcker@lincoln.ac.uk

Keywords: congenitally blind, sensory deprivation, plasticity, auditory, attention, emotion, event-related potentials, human voices

Abstract

The question whether spatial selective attention is necessary in order to process vocal affective prosody has been controversially discussed in sighted individuals: whereas some studies argue that attention is required in order to process emotions, other studies conclude that vocal prosody can be processed even outside the focus of spatial selective attention. Here, we asked whether spatial selective attention is necessary for the processing of affective prosodies after visual deprivation from birth. For this purpose, pseudowords were presented at the left or right loudspeaker and spoken in happy, neutral, fearful or threatening prosodies. Congenitally blind individuals (N = 8) and sighted controls (N=13) had to attend to one of the loudspeakers and detect rare pseudowords presented at the attended loudspeaker during EEG recording. Emotional prosody of the syllables was task-irrelevant. Blind individuals outperformed sighted controls by being more efficient in detecting deviant syllables at the attended loudspeaker. Higher auditory N1 amplitude was observed in blind individuals compared to sighted controls. Additionally, sighted controls showed enhanced attention-related ERP amplitudes in response to fearful and threatening voices during the time range of the N1. By contrast, blind individuals revealed enhanced ERP amplitudes in attended relative to unattended locations irrespective of the affective valence in all time windows (110-350 ms). These effects were mainly observed at posterior electrodes. The results provide evidence for “emotion-general“ auditory spatial selective attention effects in congenitally blindness and provide further indirect support for the idea of reorganization of the voice processing brain system following visual deprivation from birth.

Words: 250/250

1. Introduction

Human voices and vocalizations play an essential role in social interactions and communication as they allow us to not only process speech, but also to draw conclusions about other people's affective state, age, gender and even a person's body size (Lavan et al., 2019; Pisanski et al., 2017; Schweinberger et al., 2014; Skuk & Schweinberger, 2013; Zinchenko et al., 2015, 2017). Processing of human voices becomes particularly important in blind individuals as vocal features can be identified even from long distances. Some characteristics of human voice processing have been extensively studied in blind individuals, such as auditory perceptual skills (Arnaud et al., 2018; Jiang et al., 2014; Röder et al., 1999b); auditory memory (Amedi et al., 2003; Bull et al., 1983; Rokem & Ahissar, 2009), person identification (Fairhall, et al. 2017; Föcker et al., 2012, 2015; Hölig et al., 2014a, 2014b), language (Röder et al., 2003; Schild & Friedrich, 2018), auditory localization and spatial selective attention (Amadeo et al., 2019; Doucet et al., 2005; Muchnik et al., 1991; Röder et al., 1999a). Surprisingly, the nature of human affective voice processing undergoing neural plastic reorganization after visual deprivation – and more importantly – the processing of emotional features– are rather unknown so far (Fairhall et al. 2017; Klinge et al., 2010a).

One of the methods to study attention- and emotion-related processes is electroencephalography (EEG), which is known for its high temporal resolution. In sighted individuals, it was shown that emotions can modulate auditory event-related potentials (ERPs) as early as 100 and 200 ms after stimulus onset (N1; P2) but also during later processing stages such as between 260-350 ms (see also Gädeke et al., 2013; Liu et al., 2012; Pinheiro et al., 2013). For instance, ERP responses to emotional vocalizations differed from ERPs to neutral vocalizations at around 120 ms (Jessen & Kotz, 2011) and 150 ms poststimulus (Sauter & Eimer, 2010). Additionally, Pinheiro et al., (2013) observed an enhanced negativity to neutral compared to angry spoken words in the time range of the auditory N1 that has been in-terpreted as emotional evaluation of incoming sensory information (Kokinous et al., 2015). This implies that the processing of the affective quality of the signal happens very early during sensory processing, possibly due to its high relevance for survival and social interactions. To sum up, while electrophysiological correlates of emotional prosody processing are relatively well studied in sighted individuals, these processes are less understood in blind individuals.

Interestingly, a growing number of studies have shown improved auditory localization skills in blind individuals using behavioral, electrophysiological and brain imaging studies (Collignon et al., 2006; Doucet et al., 2005; Muchnik et al., 1991; Röder et al., 2007, 1999). In some of those studies, blind and sighted participants attended to a sound source in space and detected rare target stimuli while ignoring more frequent auditory standards and other (task-irrelevant) rare deviant stimuli presented at the same or other loudspeakers (e.g., Röder et al., 1999a). As a result, the authors found that blind relative to sighted participants could localize spatial positions of targets significantly further away in the periphery (Röder et al., 1999a). In line with these findings, Röder et al. (1999a) reported that blind relative to sighted participants showed a more pronounced ERP negativity (N1) in response to more peripheral sources of audio stimuli (see also Amadeo et al., 2019; Föcker et al., 2012; Röder et al., 2007 for an enhanced auditory N1 in blind individuals).

In a previous study, Röder et al. (2007) asked 8 congenitally blind individuals and 12 sighted controls to attend either to the left or right loudspeaker at which auditory stimuli were presented and to concentrate either on a long or short time interval which separated the two auditory stimuli (S1 and S2) from each other. The authors examined the length of audio refractory period across the two groups. Refractory periods are defined as time periods during which the cell is not able to generate further action potentials. Interestingly, congenitally blind individuals showed a more pronounced ERP negativity for the second auditory stimulus (S2), suggesting shorter auditory refractory periods in the blind compared to the sighted controls (Röder et al., 2007). This implies that blind participants had an advantage in the processing of auditory stimuli. Correspondingly, another study has shown that the auditory N1 recovered faster in the blind than in the sighted controls when the interstimulus interval between the two auditory stimuli was varied (Röder et al., 1999b). As the neural generators for the auditory N1 are thought to originate in the primary and secondary auditory cortices (Näätänen & Picton, 1987), an enhanced excitability of the auditory cortex might contribute to enhanced perceptual skills in the blind.

To summarize, there is consistent evidence that blind individuals show generally more efficient processing of auditory information (Fine & Park, 2018; but see Collignon et al., 2009, and Singh et al., 2018 for a further discussion). However, there is a lack of research on whether spatial selective attention is necessary in order to process affective prosodies after visual deprivation from birth. This question is of interest, as blind individuals rely much more on vocal cues and could potentially be more efficient in detecting emotional features, even outside the focus of spatial attention. By contrast, in the sighted population there is convincing evidence that emotions can be processed within and even outside of the focus of spatial selective attention (Grandjean et al., 2005; Holmes et al., 2003; Mothes-Lasch et al., 2011; Pessoa et al., 2002; Pessoa & Ungerleider, 2004; Vuilleumier & Schwartz, 2001). In one pioneering study, Grandjean and coauthors (2005) examined whether processing of emotional prosody depends on selective attention to the voice. Participants listened to auditory utterances pronounced with either threatening or neutral tone of voice in a dichotic listening task. Specifically, participants were asked to attend either to the left or right ear and identify the gender of a speaker at the target-ear and ignore the voices presented in the unattended ear. Results showed that activations in response to threatening utterances in the middle part of the right superior temporal sulcus occurred irrespective of the attended location, indicating that the brain could still detect emotional prosody from voices presented at the non-attended location (Grandjean et al., 2005). By contrast, other studies have challenged the "automaticity" hypothesis of emotional processing. For instance, Pessoa and colleagues (2002) indicated that especially under high load conditions, attention is necessary in order to process emotional features. Evidence for the hypothesis that spatial attention modulates the degree of emotional voice processing as a function of emotional valence was observed in an EEG experiment: Auditory pseudowords (neutral, happy, threatening, and fearful) have been presented at two different loudspeakers and participants were asked to detect rare deviant syllables (e.g. "giki", "fefi") at the attended location while ignoring all standard pseudowords presented at the same location and all deviants and standards at the non-attended location (Gädeke et al., 2013). Emotional valence of the pseudowords was task-irrelevant. As a result, the authors found more pronounced negativity in response to attended versus unattended voices specifically in the time range of the auditory N1 and especially for fearful voices. This implies that processing of emotional information modulates early but not later stages of information processing. Importantly, these authors also showed emotion-specific brain activations at both attended and unattended locations in this early time-window, suggesting that emotions can be processed even outside the focus of selective spatial attention in sighted individuals.

In order to investigate whether spatial selective attention is necessary to process emotional prosody in blind individuals, we used the well-established paradigm outlined above (Gädeke et al., 2013) and applied it to congenitally blind individuals. In more detail, we used an auditory spatial attention paradigm in which participants were asked to detect rare bisyllabic pseudowords (e.g. "fefi"; "giki", "nane") at the attended loudspeaker and ignore the same infrequently presented syllables at the unattended loudspeaker as well as more frequently presented pseudowords (e.g. “baba”, “dede”, “fafa”) at both loudspeakers (see Figure 1).

INSERT FIGURE 1 HERE

Pseudoword were presented in four different emotions (neutral, happy, fearful and threatening). The emotion of the pseudowords was task-irrelevant. In sighted controls it has been shown that attention modulates the processing of emotional prosody at early perceptual processing stages by showing a more pronounced negativity to attended fearful voices compared to unattended fearful voices. We asked (1) whether congenitally blind individuals would show the same attentional capture by negative stimuli (e.g. fearful) as sighted controls in the time range of the auditory N1 and whether there would be a main effect of *Emotion* and *Attention* in later time windows (>150ms) similar to sighted controls. We were also interested (2) whether congenitally blind individuals outperform sighted controls in distinguishing targets (“giki”, “fefi”, “nane”) from more frequently presented standard voices (e.g. “baba”, “dede”, “fafa”) at the attended location and if so, (3) at which processing stages would congenitally blind individuals differ from sighted controls (early versus late). This question was motivated by previous work that compared different temporal processing stages between congenitally blind individuals and sighted controls (Föcker et al., 2012, 2015; Röder et al., 1999). It was found that congenitally blind relative to sighted controls show different patterns and topographical distributions of auditory event-related potentials, e.g., N1, N2b, mismatch negativity (MMN), Auditory-evoked Contralateral Occipital Positivity (ACOP) recorded at posterior electrodes (e.g., Pz), which was linked to cortical reorganization of the auditory system in the blind (Alho et al., 1993; Amadeo et al., 2019; Föcker et al., 2012; Hötting et al., 2004; Kujala et al., 1992; Röder et al., 1999a; see also: Leclerc, Saint-Amour, Lavoie, Lassonde, & Lepore, 2000). Therefore, (4) we aimed to investigate at which electrodes (central versus posterior) differences in emotional processing are mostly pronounced in congenitally blind and sighted controls.

We hypothesized (1) that if emotions are processed outside the focus of spatial selective attention, we should observe emotion-related ERP modulations independently of the focus of spatial selective attention. That is, ERPs for fearful, threatening and happy voices should show different patterns of activity (i.e., amplitudes) relative to neutral voices within both attended and unattended conditions. However, if attention is required to process emotional valence, we expected different modulations of ERPs with regards to emotional valence for spatially attended and unattended stimuli similar to sighted controls in the time range of the auditory N1 to fearful human voices (Gädeke et al., 2013). In more detail, Gädeke and colleagues (2013) showed that sighted individuals revealed a more pronounced N1 negativity in response to attended relative to unattended fearful human voices, which might be an index of an enhanced suppression of spatially irrelevant human fearful voices and an enhanced capture of attention to fearful voices presented at the attended location. We did not expect any interaction between attention and emotion at later processing stages (see Gädeke et al., 2013). Regarding question (2), we hypothesized that blind individuals would be more efficient in processing human voices at the attended speaker compared to sighted controls (Klinge et al., 2010a).

Similarly to previous studies (Röder et al., 1999a, 1999b; 2007) we expected more enhanced auditory N1 amplitudes in the congenitally blind individuals compared to sighted controls (3). Finally, (4) we expected to find group-specific differences between congenitally blind and sighted controls at more posterior electrode sites as observed in previous research (Amadeo et al., 2019; Föcker et al., 2012; Röder et al., 1999a,b).

2. Results

In the following, results are presented including N = 8 congenitally blind individuals and N = 13 sighted controls. We first describe the behavioral results followed by the event-related potential (ERP) results.

2.1. Behavioral results

For the behavioral results we report the ANOVA including the factors *Emotion* (happy, neutral, fearful, threatening) and the between subject factor *Group* (congenitally blind individuals versus sighted controls) on d'prime and Inverse Efficiency scores (IE scores). IE scores combine both reaction times and correct responses (Townsend & Ashby, 1987; Spence et al., 2001) and have been used as we aimed to follow the same procedure as reported in Gädeke et al., 2013. Percent correct (PC), mean reaction times (RT), d'prime as well as IE scores are reported in Table 1.

INSERT TABLE 1 HERE

2.1.1. D-prime scores

As expected, blind individuals outperformed sighted controls in distinguishing targets from more frequently presented standards at the attended location (main effect of *Group*: $F(1,19) = 19.557$, $P < .001$, $\eta^2 = .507$, blind individuals: mean d' = 2.7, SE = .113; sighted controls: mean d' = 2.1, SE = .088, see Figure 2 C). Moreover, all participants could better detect targets at the attended location when spoken in a neutral prosody compared to happy, threatening or fearful emotions (main effect of *Emotion*: $F(3,57) = 22.625$, $P < .001$, $\eta^2 =$

.544, neutral = 3.089, SE = .154; threatening = 2.125, SE = .108; happy = 2.135, SE = .057; fearful = 2.512, SE = .095; all *Ps* < .001, see Figure 2 A). Furthermore, d-prime for fearful human voices were higher compared to threatening and happy voices (*P* < .007). The interaction between *Emotion* and *Group* was not significant (*F*(3,57) = .85, *P* = .445).

2.1.2. Inverse Efficiency (IE) scores

Participants responded more efficiently in the neutral condition compared to the threatening vocal prosody (main effect of *Emotion: F*(3,57) = 5.898, *P* = .008; η^2 = .237; mean neutral = 1405 ms, SE=141; mean happy: 1639 ms, SE = 92; mean threatening: 1997 ms, SE = 206, mean fearful: 1459 ms, SE = 93, *P* = .001; see Figure 2B). Moreover, blind individuals responded more efficiently to target voices compared to sighted controls (main effect of *Group: F*(1,19) = 8.093, *P* = .010, η^2 = .299; sighted controls*:* mean: 1922 ms, SE=128; blind individuals: mean*:* 1328 ms, SE=164, see Figure 2D). The interaction between *Emotion* and *Group* was not significant (*F*(3,57) = .165, *P* = .826).

INSERT FIGURE 2 HERE

2.2. ERP results

For the ERP analysis, we used a 2 (*Group*: congenitally blind, sighted control) * 4 (*Emotion:* neutral, happy, fearful, threatening) * 2 (*Attention:* attended, unattended) repeated measures analysis of variance. We first run an ANOVA at the central electrode M4 given that the auditory vertex potential is maximal in amplitude at this site. Based on our hypotheses that differences between congenitally blind and sighted controls would be mainly observed at the posterior electrode M7, we run an additional

ANOVA at the posterior electrode M7 (corresponding to Pz). We finalized this result section by reporting the results of the four-way interaction between the factors *Attention, Emotion, Electrode* and *Group*.

Figure 3 summarizes the ERPs recorded to human voices presented at the attended versus unattended loudspeaker averaged across all participants (Figure 3 A, D), the ERPs averaged separately for each emotional condition (neutral, happy, threatening, fearful) across all participants (Figure 3 B, E), and the ERPs averaged separately in all congenitally blind and all sighted controls (Figure 3 C, F) at the central electrode M4 and the posterior electrode M7 (corresponding to electrode Pz of the 10-20 system). Figure 4 shows the difference waves (attended minus unattended) as a function of emotional valence (neutral, happy, threatening and fearful) and the topographical distribution of the attention effect (E,F) for the three time windows separately for congenitally blind (B, D) and sighted controls (A, C) at electrode M4 and M7). Figure 5 illustrates the mean amplitudes for spatial attended (red dashed line) and unattended locations (black solid line) in the time range of the N1 plotted as a function of emotions (fearful, happy, neutral and threatening) separately for congenitally blind (A,B) and sighted controls (C,D) at electrodes M4 and M7.

To foresee the results, we observed significant main effects of *Emotion* and *Attention* in the ERP amplitudes of congenitally blind individuals across all time windows. By contrast, in sighted controls, the interaction between *Attention* and *Emotion* with mean ERP amplitudes as dependent measurement was significant in the first time window, but not in the second or third time windows.

INSERT FIGURE 3 HERE

INSERT FIGURE 4 HERE

Time Window: 110-150 ms

Electrode M4

The ANOVA including the factors *Emotion*, *Attention* and *Group* on mean ERP amplitudes revealed a main effect of *Attention* and a main effect of *Emotion* (main effect of *Attention*: $F(1,19) = 21.855$, $P < .001$, $\eta^2 = .535$; main effect of *Emotion*: $F(3,57) = 20.648$, $P < .001$, $\eta^2 = .521$). ERPs were more negative to attended compared to unattended human voices (mean attended: -3.202 µV, SE = .425; mean unattended: -2.598 µV, SE = .380, see Figure 3 A). Moreover, ERPs to neutral prosodies revealed a more pronounced negativity compared to all other emotions (mean neutral: - 3.578 µV, SE = .408, mean happy: -2.151µV, SE = .357; mean threatening: -2.892µV, SE = .473, mean fearful: -2.98µV, SE = .407, see Figure 3 B). Additionally, ERPs recorded to happy voices revealed a less pronounced negativity compared to all other voices (all *P*s < .012). The main effect of *Group* was significant ($F(1,19) = 5.013$, $P = .037$, $\eta 2 = .209$, see Figure 3 C). The auditory N1 amplitude was more negative in congenitally blind individuals compared to sighted controls (congenitally blind: mean: -3.791 µV, SE = .626; sighted controls: mean: -2.009 µV, SE = .491).

The interaction between *Emotion*Attention*Group* was not significant ($F(3,57) = 1.153$, $P = .331$).

Correlation between behavioral performance and auditory N1

Electrode M4

In order to investigate, whether higher auditory N1 amplitudes were associated with improved performance in congenitally blind individuals but not in the sighted controls, we calculated the correlations between mean amplitudes of the auditory N1 and the behavioral performance (IE scores and d'). The correlations between the auditory N1 and IE scores and the auditory N1 and d' were not significant within each group (*Blind individuals:* IE scores: r

= -.195, *P* = .643; *Blind individuals d prime*: r = .230, *P* = .584; *Sighted Controls:* IE scores: r = -.009, *P* = .977; *Sighted Controls:* d prime: r = .258. *P* = .396).

Time Window: 190-260 ms

Electrode M4

The main effect of *Attention* was significant *(Attention: F*(1,19) = 38.447, *P* < .001*;* η^2= .669*)*. ERPs to the attended condition were more negative compared to the unattended condition *(mean attended:* 2.891µV, SE = .700*; mean unattended:* 4.834µV, SE = .693, see Figure 3 A*)*. Additionally, the main effect of *Emotion* was significant *(F*(3,57) = 31.933, *P* < *.001;* η^2= .627). ERPs revealed a more pronounced positivity to the threatening voices compared to all other prosodies (mean threatening: 4.88µV, SE = .717; mean neutral: 3.58µV, SE = .659; mean fearful: 2.922µV, SE = .687; mean happy: 4.067µV, SE = .698, *P* < .001, see Figure 3B). Additionally, ERPs to fearful voices were more negative compared to all other voices (all *Ps* < .006). The interaction between *Emotion*, *Attention* and *Group* was not significant (*F*(3,57) = .461, *P* = .684).

Time Window: 260-350 ms

M4

The ANOVA including the factors *Attention, Emotion* and *Group* on mean ERP amplitudes revealed a main effect of *Attention (Attention: F(1,19) = 67.967, P < .001;* η^2= .782*)*. ERPs of the attended condition were more negative compared to ERPs of the unattended condition *(mean attended:* -.144µV, *SE* = .651; *mean unattended:* 2.210 µV, *SE* = .629, see Figure 3A*)*. Additionally, the main effect of *Emotion* was significant *(F*(3,57) = 23.257, *P < .001;* η^2= .550). ERPs revealed a more pronounced positivity to the happy voices compared to all other prosodies (mean happy: 2.115µV, SE = .612; mean neutral: .688µV, SE = .619;

mean fearful: .280µV, SE = .660; mean threatening: 1.049µV, SE = .667, $P < .001$, see Figure 3B). Additionally, ERPs to threatening voices revealed a more pronounced positivity compared to fearful voices ($P = .021$). The interaction between *Emotion, Attention* and *Group* was not significant ($F(3,57) = .585$, $P = .594$).

Time Window: 110-150 ms

Electrode M7 (Pz)

The overall ANOVA including the factors *Attention, Emotion* and *Group on mean ERP amplitudes* revealed a main effect of *Attention* ($F(1,19) = 31.248$, $P < .001$, $\eta^2 = .\ 0.622$) with a more pronounced negativity in the attended compared to unattended condition (mean attended: -2.596 µV, SE = .233; mean unattended: -2.089 µV, SE = .203; see Figure 3 D). Additionally, the main effect of *Emotion* was significant ($F(3,57) = 11.758$, $P < .001$, $\eta^2 = .382$) with a more pronounced negativity in the neutral condition compared to the threatening and happy condition (mean neutral: -2.841 µV, SE = .273; mean happy: -1.945 µV, SE= .185; mean threatening: -2.286 µV, SE= .249; mean fearful: -2.299 µV, SE= .217; all *P*s <.006; see Figure 3 E). Moreover, ERPs recorded to happy voices revealed a less pronounced negativity compared to fearful and neutral voices (all *P*s < .006). The main effect of *Group* was not significant ($F(1,19) = 2.789$, $P = .111$, $\eta^2 = .\ 128$).

Importantly, the interaction between the factors *Emotion, Attention* and *Group* was significant ($F(3,57) = 2.975$, $P = .048$, $\eta^2 = .135$).

In the congenitally blind individuals, we observed no significant *Emotion* by *Attention* interaction ($F(3,21) = .895$, $P = .424$, $\eta^2 = .113$). However, this interaction was significant in sighted controls ($F(3,36) = 4.066$, $P = .018$, $\eta^2 = .253$). Subordinate ANOVAs confirmed that the effect of *Emotion* was significant at both the attended and unattended location but the higher F value for the attended condition ($F(3,36) = 10.109$, $P < .001$, $\eta^2 = .457$) than the unattended condition (*F*(3,36) = 5.956, *P* = .004, η^2 = .332) suggests a stronger *Emotion* effect at the attended location (see Figure 5). Post hoc t-tests showed a difference between ERPs of the attended and unattended condition only for the fearful and threatening voices (fearful condition attended versus unattended: *t*(12) = 3.708, *P* = .003, mean attended = -2.51 mV, SE: .22, mean unattended = -1.67mV, SE: 1.07; threatening condition attended versus unattended: *t*(12) = 2.875, *P* = .014, mean attended = -2.10 mV, SE: .31, unattended = -1.42 mV, SE: .25; all other *P*'s > .36, see Figure 4 C,D).

INSERT FIGURE 5 HERE

Correlation between behavioral performance and auditory N1

Electrode M7

Similarly to M4, we correlated mean N1 amplitudes and the behavioral performance (IE scores and d'). Neither of the effects reached significance (*Blind individuals:* IE scores: r = -.199, *P* = .637; *Blind individuals d'*: r = -.147, *P* = .729; *Sighted Controls:* IE scores: r = .04, *P* = .896; *Sighted Controls:* d': r = .055. *P* = .857).

Time Window: 190-260 ms

M7

The overall ANOVA including the factors *Attention*, *Emotion*, and *Group* on mean ERP amplitudes revealed a main effect of *Attention* (*F*(1,19) = 26.936, *P* < .001, η2= .586) with a more pronounced negativity to ERPs in the attended compared to unattended condition (mean attended: 2.89 µV, SE = .700; mean unattended: 4.83 µV, SE = .69; see Figure 3 D). Additionally, the main effect of *Emotion* was significant (*F*(3,57) = 14.823, *P* < .001, η^2=

.438) with a more pronounced positivity to threatening voices compared to all other emotions (mean neutral: .788 µV, SE= .344; mean happy: mean = .947 µV, SE: .365, mean threatening: = 1.384 µV, SE = .395, mean fearful: mean = .317 µV, SE= .362, *P* < .039, see Figure 3 E). Moreover, ERPs to fearful voices were more negative compared to all other emotions (all *P*s < .019).

We also observed an interaction between the factors *Attention*Group* (*F*(1,19) = 7.717, *P* = .012, η2 = .289) showing stronger differences between the attended and unattended condition in blind individuals compared to the sighted controls (blind individuals: mean attended: -.887 µV, SE = .469, mean unattended: 1.235 µV, SE = .178; *t*(7) = -4.009, *P* = .005; sighted controls: mean attended: 1.222 µV, SE = .597; mean unattended: 1.864 µV, SE = .489, *t*(12) = -2.418, *P* = .032; see also more posterior shift of the attention effect in the topographies in congenitally blind individuals, Figure 4 E,F). The interaction between *Attention*, *Emotion* and *Group* and the main effect of *Group* were not significant (*Attention*Emotion*Group: F*(3,57) = .334, *P* = .794; main effect of *Group*: *F*(1,19) = 3.763, *P* = .067).

Time Window: 260-350 ms

M7

Similar to the first and the second time windows, a main effect of *Attention* and a main effect of *Emotion* were observed (main effect of *Attention*: *F*(1,19) = 6.841, *P* = .017, η^2= .265; main effect of *Emotion*: *F*(3,57) = 15.718, *P* < .001, η^2= .453, see Figures 3 D, E). ERPs to human voices presented at the attended location were more negative compared to human voices presented at the unattended location (mean attended: .162 µV, SE = .463, mean unattended: 1.149 µV, SE = .293, *P* < .001). ERPs to happy human voices revealed a more pronounced positivity compared to all other emotions (mean neutral: .444 µV SE = .347;

mean happy: 1.276 µV SE = .312, mean threatening: .68, SE = .341 µV, mean fearful: .223 µV, SE = .405, *P* < .004).

The interaction between the factors *Attention* and *Group* was significant (*F*(1,19) = 6.36, *P* = .021, η^2= .251). Separate ANOVAs run in each *Group* revealed a significant main effect of *Attention* in congenitally blind, but not in sighted controls (congenitally blind: *F*(1,7) = 13.596, *P* = .008, η^2= .66; blind individuals: mean attended: -.287 µV, SE: .547; mean unattended: 1.65 µV, SE: 465; sighted controls: *F*(1,12) = .005 *P* = .944, sighted controls mean attended: .612 µV, SE: 640; mean unattended: .647 µV, SE: 359).

Finally, the interaction between *Emotion*, *Attention* and *Group* was not significant (interaction between *Emotion*, *Attention* and *Group*: *F*(3,57) = .878, *P* = .443).

Note also that the critical 4-way interaction of *Attention* (attended versus unattended), *Emotion* (neutral, happy, threatening, fearful), *Electrode* (M4, M7) and *Group* (congenitally blind versus sighted controls) on mean ERP amplitudes was significant in the first time window only (*F*(3,57) = 6.258, *P* = .003, η^2= .248, for all other time windows: *P* > .7). Confirming the similarity across the two electrodes, the 4-way interaction was not significant for the second and third time windows (see Table 2).

INSERT TABLE 2 HERE

63
64
65

3. Discussion

The goal of the present study was to understand whether spatial selective attention is necessary for processing of affective prosodies after visual deprivation from birth. Therefore, we aimed at identifying the time course and underlying processing stages that differ in congenitally blind adults compared to sighted controls and that potentially provide enhanced auditory emotional processing capacities. Moreover, we tried to understand if, similar to sighted controls, congenitally blind individuals suppress irrelevant fearful voices and attend to relevant fearful human voices at the attended location during early processing stages (auditory N1; see Gädeke et al., 2013). This effect was demonstrated by a more pronounced negativity to attended relative to unattended fearful human voices in sighted controls (see Gädeke et al., 2013). Finally, we analyzed whether the group differences in orienting spatial selective attention to different emotional voices are distributed at posterior electrodes (Amadeo et al., 2019; Föcker et al., 2012; Hötting et al., 2004; Röder et al., 1999a; see also: Leclerc et al., 2000).

For this purpose, an auditory oddball paradigm was run in which participants had to detect rare deviant syllables at the attended location and ignore deviant syllables at the unattended location as well as all standard syllables at both locations. We observed that congenitally blind individuals were more efficient compared to sighted controls in detecting deviant syllables at the attended spatial location. Those group effects cannot be due to gender or age differences as both groups did not differ in this respect. This result pattern contributes to a large range of studies reporting superior auditory skills in the blind, such as pitch discrimination and auditory spectral cues (Doucet et al., 2005; Gougoux et al., 2004; Wan et al., 2010), human echolocation (Schenkman & Nilsson, 2010), auditory language processing (Röder et al., 2003; Schild & Friedrich, 2018), auditory memory (Amedi et al., 2003; Rokem & Ahissar, 2009), auditory spatial selective attention (Hugdahl et al., 2004; Kujala et al., 1995; 1997; Lessard et al., 1998; Röder et al., 1999a,b) and processing of auditory vocal prosody

(Klinge et al., 2020). Particularly, those results confirm findings of enhanced auditory spatial selective attention in blind individuals (demonstrated in higher d primes) and point to the fact that blind individuals might not be distracted by the emotional valence of the voices when attending to a specific spatial location (Hugdahl et al., 2004; Kujala et al., 1995; 1997; Lessard et al., 1998; Röder et al., 1999a,b).

Early perceptual processing

Consistently with previous studies in this area, we found an enhanced N1 amplitude in blind individuals compared to sighted controls (Amadeo et al., 2019; Doucet et al., 2005; Muchnik et al., 1991; Röder et al., 1999a). This group difference in the N1 mirrors facilitated behavioral performance in congenitally blind individuals. It might be speculated, that an improved representation of auditory perceptual features (as measured via N1) contributes to more efficient task processing at the attended location in blind participants. Other studies have argued that there is a more efficient perceptual encoding in the blind as reflected in shorter N1 latencies and shorter recovery periods of auditory ERPs (Elbert et al., 2020; Röder et al., 1996).

On the other hand, four out of 13 blind participants were excluded in the current experiment from data analysis because they were not able to perceptually discriminate the two female vocal identities, which has been set as a test of basic hearing abilities and was used criteria to be included in data analysis (see Gädeke et al., 2013; Bull et al., 1983; see Föcker et al., 2012; Hölig et al., 2014 a, 2014 b; for a better voice identification performance in congenitally blind compared to sighted controls). Some studies did report impaired performance on auditory tasks in blind individuals (Cappagli & Gori 2016; Finocchietti et al. 2015; Gori et al. 2014; Menard et al. 2015; Voss, 2016), while the others found no difference from sighted participants (Collignon et al. 2011, 2013; Voss & Zatorre 2012). The heterogeneity of results reported in auditory tasks in blind individuals may be due to different task requirements (e.g. voice identification versus detecting target syllables in the current experiment; see also King, 2014; de Borst & de Gelder, 2019, p.2860) or different training protocols (see Föcker et al., 2012).

In the time range of the auditory N1, we observed a main effect of *Attention* in blind individuals: similar to sighted controls, congenitally blind individuals showed a difference between attended and unattended fearful voices. However, this effect was not specifically tight to negative human voices such as fear or threat as in sighted controls (see Gädeke et al., 2013 for a further discussion for sighted controls). Thus, while attention effects for most of the emotional voices were observed relatively late in sighted individuals (> 150 ms), the main effect of spatial attention to emotional stimuli was already established in blind individuals and quite similar across all emotions including happy, fearful, neutral and threatening. Spatial selective attention might act as a mechanism that allows processing of emotions at the attended location and suppressing irrelevant information at the unattended location. It might be argued that congenitally blind individuals have an “improved and more efficient” spatial filter system in order to process and distinguish relevant from irrelevant information irrespective of the type of emotion.

We argue that the emotional valence of auditory stimuli might be partially extracted automatically (in the absence of at least spatial attention) in the congenitally blind. This is shown by the main effect of *Emotion* in congenitally blind individuals, which suggests that emotions are processed in the attended and the unattended channel in a similar way. This corresponds to findings reported by Klinge et al. (2010a) in congenitally blind individuals: In this study, congenitally blind participants and sighted controls had to discriminate either the emotional prosody (happy, threatening, neutral, fearful: emotion discrimination task) or the first vowel of each stimulus (a, e, i, o: vowel discrimination task) while functional brain activity was recorded (Klinge et al., 2010a). As a result, blind individuals showed higher proficiency in discriminating voice prosodies, they were faster in emotion discrimination compared to sighted controls and showed higher activation in occipital cortex to *all* emotional vocal stimuli (Klinge et al., 2010a). This group of participants also showed higher amygdala activation in response to threatening and fearful compared to neutral voices. Moreover, amygdala activation was observed *irrespective* of the underlying task (emotion versus vowel discrimination task), indicating that this activation is not related to explicit emotion detection, but is rather automatically driven by the emotional valence of the stimulus.

It has to be noticed that quite long inter-stimulus intervals (ISIs) were applied in the current experiment and it is well known that N1 attention effects are elicited if short ISIs are employed (see also Gädeke et al., 2013 for a similar discussion). Of course, we cannot exlude the fact that participants might have additional resources left over to attend to the other (i.e., task-irrelevant) loudspeaker. However, we argue that this is unlikely. Spatial selective attention effects were already established in the first time window, especially in the congenitally blind individuals to all emotions, which suggest specific enhancement of the processing of vocal stimuli by spatial attention even when long ISIs are applied. Nevertheless, future studies could examine this idea more explicitly by additonally taxing participants' attentional resources and testing whether participants would still be able to show emotion-specific processing at the unattended spatial locations.

Interestingly, ERPs were modulated by emotional valence in both sighted controls and congenitally blind individuals in a similar way in the time range of the auditory N1, suggesting that emotions itself are similarly processed in both groups. ERPs showed a more pronounced negativity to neutral voices compared to threatening, happy or fearful human voices for both groups. This is in line with previous studies (Liu et al., 2012; Pinheiro et al., 2013) and might suggest that salient acoustic cues direct the emotional evaluation. Interestingly, enhanced N1 amplitudes to neutral voices might reflect improved voice detection of neutral stimuli another indication that enhanced amplitudes mirror better task performance.

Later processing stages

In the second time window (190-260 ms), we observed that both, congenitally blind and sighted controls showed a main effect of *Attention* and a main effect of *Emotion* (without any interactions). However, the difference in ERPs to the attended versus unattended condition was much stronger in congenitally blind individuals compared to sighted controls. Interestingly, unlike congenitally blind individuals, sighted controls did not show any attention effect in the time window 260-350 ms at the posterior electrode M7, suggesting a more sustained attention effect over time in the blind compared to sighted controls especially at posterior electrodes. This is also shown by the more posterior topographical distribution of the attention effect in congenitally blind compared to sighted controls which might point to a reorganization of the voice processing system in congenitally blind individuals. These more posterior topographies of auditory evoked potentials have been also shown in other studies (Amadeo et al., 2019: Auditory-evoked Contralateral Occipital Positivity (ACOP); Föcker et al., 2012; Hötting et al., 2004; Leclerc et al., 2000; Röder et al., 1999a). Therefore, we argue that attention was not necessary to process emotional valence of the voices at these later time windows in both sighted and congenitally blind individuals. However, spatial selective attention – even at this late processing stage – is much more enhanced in blind individuals compared to sighted controls.

Neural reorganization of the emotional voice processing system

It has been suggested that an intramodal reorganization in blind individuals might contribute to enhanced performance in several auditory perceptual tasks (Röder et al., 2007). For instance, brain imaging studies reported cortical reorganization of the auditory cortex as a neural mechanism to understand the shorter auditory N1 latencies (Elbert et al., 2002; Stevens & Weaver, 2009). Besides changes within unisensory brain areas (also called intramodal plasticity in auditory brain structure, Röder & Neville, 2003), other studies observed neural plastic changes in multisensory regions (De Volder et al., 1997; Röder et al., 1999a), including the functional connections between auditory and visual brain areas (Bavelier & Neville, 2002; Klinge et al., 2010b) and additional recruitment of visual cortices during auditory processing (crossmodal plasticity, Merabet & Pascual-Leone, 2010; Fairhall et al., 2017) which has been suggested to facilitate performance of the blind including voice processing (Gougoux et al., 2009). For instance, Gougoux and coauthors (2009) have shown higher voice specific activation in the left superior temporal sulcus (STS) in congenitally blind individuals compared to sighted controls (Gougoux et al., 2009). This increased recruitment of the STS was correlated with their performance in a voice discrimination task (Gougoux et al., 2009). Thus, it might be speculated that visual deprivation from birth leads to a reorganization of the multisensory zone in the STS.

Several other brain imaging studies have shown a crossmodal reorganization in human voice processing tasks, such as a higher activation in the right fusiform gyrus in congenitally blind and even late blind individuals when asked to indicate the age of a voice (see Hölig et al., 2014a, 2014b). This activation has been even observed when onset of blindness starts later in life suggesting that neural reorganization can also be observed in the more mature human brain (Hölig et al., 2014b). Klinge et al. (2010a) observed an enhanced performance of the congenitally blind in auditory discrimination tasks that was paralleled by occipital cortex activation, which was absent in the sighted controls. Even though further studies are needed to understand the exact location of neural plastic reorganization in the current task, we assume that also the recruitment of visual brain areas is involved in the current voice discrimination task in congenitally blind individuals. This assumption is based on the fact, that attention effects, even at later processing stages are observed at more posterior electrodes which is usually atypical for auditory ERPs.

3.1. Conclusion

These results provide evidence for enhanced auditory spatial selective attention irrespective of the emotional valence in the absence of vision from birth and point to a reorganization of the auditory voice processing system following congenital blindness.

4. Experimental Procedure

4.1. Participants

Thirteen congenitally blind individuals participated in the experiment. This sample size was based on a highly relevant previous work in this area (e.g., Gädeke et al., 2013 who included 13 sighted controls in the same paradigm; see also Röder et al., 2007). Five participants had to be excluded from data analysis due to the following reasons: (1) four participants had to be excluded due to very low performance in discriminating human voices (d prime < .04) see also Gädeke et al., 2013 for a similar approach), (2) one participant had too many artifacts in the EEG data recordings (less than 40 % of trials remaining). The final sample consisted of eight congenitally blind individuals (mean age: 26 years, age range: 23-29 years, four female). Please note that comparable sample sizes of blind individuals (N = 8) have been reported in previous studies e.g., de Borst & de Gelder, (2005); Easton et al., (1998); Föcker et al., (2015); Hampson & Duffy, (1984); Matteau et al., (2010); Röder et al., (1999a; 2007); Szucs & Csepe, 2005; Vercillo, Burr, & Gori, (2016). Six participants were students at the University of Marburg, Germany, one participant was a businessman, and another participant was a service operator.

All blind participants were totally blind or did not have more than rudimentary sensitivity for brightness differences without any pattern vision. In all cases, blindness was due to peripheral deficits. More specifically, blindness was due to the following reasons: retinopathia pigmentosa (N = 3), retina degeneration (N=2), too high levels of oxygen in the incubator (N=1). For two participants, the reasons for blindness (peripheral defect) were unknown (N = 2). All participants were German native speakers and reported normal hearing and no history of neurological illness. Eight blind participants were compared with 13 sighted controls (mean age: 23 years, age range: 20—28 years, seven females; see Gädeke et al., 2013). Congenitally blind individuals and sighted controls did not differ in gender or age (gender distribution blind individuals: 4 females and 4 males; sighted controls:7 females, 6 males; $\chi 2$ = .0294, *P* = .864; mean age blind individuals: 26 years, SD = 2.43 years; mean age sighted controls: 23 years, SD = 2.61 years, *t*(19) = 1.60, *P* >.05). Sighted participants had normal or corrected to normal vision. All participants were blindfolded throughout the experiment.

All participants were recruited from the local community or towns near the city of Marburg and received monetary compensation for their participation. Written informed consent was given by each participant prior to the beginning of the experiment. This study was in accordance with the Declaration of Helsinki and approved by the Ethics committee of the medical association of Marburg.

4.2. Stimulus Material

The stimulus material, training and experimental procedure were identical to the procedure reported in Gädeke et al. (2013). The stimulus material has been rated by a separate group of 24 University students (see Gädeke et al., 2013). Nine disyllable pseudo-words spoken by two actresses in four emotional prosodies (neutral, happy, threatening and fearful) were selected for the purpose of the study (9 x 2 x 4 = 72 different stimuli). Pseudowords consisting of two different syllables were classified as deviant stimuli (such as fefi), while the remaining six vocal stimuli with the same two syllables belonged to the standard stimuli (such as fefe). Deviant syllables presented at the attended location (for instance right loudspeaker) are called targets throughout the manuscript. Mean stimulus duration for neutral human voices was 632 ms, SE = 35, for happy human voices 575 ms, SE = 57, for threatening human voices 602 ms, SE = 56 and for fearful human voices 518 ms, SE = 44. We run the Kruskal Wallis test (see Zinchenko et al., 2015 for comparable procedures) in order to compare the stimulus duration of the targets and standards. Results show that the duration between different emotional stimuli does not significantly differ from each other (targets: $\chi2 = 6$, $P > 0.1$; standards: $\chi2 = 3.66$, $P > .2$, df = 3). The characteristics of the stimulus material (duration, pitch, intensity, valence, intensity and dominance ratings) are reported in tables 3 and 4. Pitch was calculated using the Praat phonetics software package (Boersma & Weenink, 2012) developed for Phonetic or Phonological research. Praat uses an autocorrelation method for pitch analysis based on a robust algorithm for periodicity detection, that has been optimised for speech analysis, proposed by Boersma (2001). For further information see http://www.fon.hum.uva.nl/praat/manual/Sound__To_Pitch__ac____.html

INSERT TABLE 3

INSERT TABLE 4

4.3. Procedure

4.3.1. Experiment

Two loudspeakers were positioned in front of the participant at a distance of 1.4 m, one 45 degrees to the left and one 45 degrees to the right of the participant. All stimuli were presented with an equal probability and in randomized order from the left and right loudspeakers. The time intervals between the onset of the presentation of any two successive voices (i.e., stimulus onset asynchronies: SOA) varied between 1300 ms to 1700 ms (see Figure 1).

INSERT FIGURE 5 HERE

Participants' task was to attend to stimuli which were presented at one of two spatial positions (left or right speaker) spoken by one of the two female speakers. Whenever participants detected one of the deviant stimuli spoken by the attended voice and presented at the attended position (i.e., targets), participants had to lift the right or left index finger out of a light gate. After half of the trials, the response hand was switched (from left to right index finger or vice versa). Emotional prosody of the syllables was task-irrelevant. In total there were four experimental conditions (attend voice I versus attend voice II and attend left vs. attend right loudspeaker).[1] The experiment consisted of 16 blocks; each block lasted six to seven minutes. The following four experimental conditions were presented: condition 1: attend left speaker, attend voice 1; condition 2: attend right speaker, attend voice 1; condition 3: attend left speaker, attend voice 2, condition 4: attend right speaker, attend voice 2. (p.22). A block comprised 192 standard stimuli (80%) and 48 deviant stimuli (20%), 24 of which were targets (5 %). Every two blocks participants were instructed to attend to the other location (e.g., from left to right). Only spatial attention effects with regards to the different emotional prosodies were analyzed.

All participants were blindfolded throughout the experiment and a chin rest was used to restrict head movements. Moreover, participants were instructed to avoid excessive blinking dur-

---

[1] "Originally, the main experiment comprised an additional orthogonally manipulated factor (Gaedeke et al., 2013). Participants had to selectively attend to one voice only. However, the voices of two female actors were too similar and participants did not manage to distinguish between them. Even after excluding participants (N = 4) with very low performance in discriminating the voices (d' < .04), mean d' was low (d' = .67, SE = .08) (see Gaedeke et al., 2013). In the current experiment, we applied the same criteria to congenitally blind indi-viduals and sighted controls, in order to guarantee that there were no pre-existing differ-ences based on any auditory task performed (i.e. voice identification)" (see Gaedeke et al., 2013, p. 14).

ing the blocks. The EEG experiment without any breaks took approximately 1.5 hours. The whole experimental session including breaks, practice and the electrode preparation and removal, lasted between 5 and 6 hours.

4.3.2. Training

In order to familiarize participants with all voice stimuli and experimental procedure, participants had to take part in a training session, one or two days prior to the actual experiment. They were asked to discriminate the voices of the two and the experimental procedure. We did not analyze the factor voice in the current experiment. Participants who were not able to distinguish the two actors were excluded from data analysis (criterion d-prime = 0.04). Further details of the training are provided in Gädeke et al. (2013).

4.4. ERP data

The data acquisition and EEG recording was identical to Gädeke et al. (2013). For the EEG recording 61 Ag/AgCl electrodes were used, mounted equidistantly in an elastic cap (Falk Minow Services, Munich). A bipolar horizontal electrooculogram (HEOG) recording was obtained by attaching two electrodes to the outer canthi of the eyes, and the vertical EOG (VEOG) was monitored by placing an electrode under the right eye against the common reference. The right earlobe electrode was used as reference electrode during recording, but offline all channels were re-referenced to the averaged left and right earlobe references. The ground electrode was placed on a position at the middle of the forehead (below Fpz).

Participant’s skin was prepared by using Every (Meditec SRI, Negernbotel) and alcohol. Electrogel (Electrocap International, Ohio, USA) served as the electrolyte for all electrodes. Impedances were kept below 5 kΩ for scalp recordings and below 10 kΩ for EOG recordings. Signal amplification was made possible by using two SynAmps-amplifiers (NeuroScan, Inc. Sterling, USA). The sample rate was 500 Hz and the bandpass was set to 0.1 – 100 Hz.

For the ERP analysis, the EEG was averaged for time epoch -100 ms (pre-stimulus) to 1000 ms (post- stimulus), for each participant and condition. The prestimulus interval was defined as baseline. Only segments following standard stimuli were analyzed, while segments with responses to standard stimuli were discarded. Segments containing eye movements artifacts, defined as a larger difference of 120 μV between two sample points within a segment of the vertical or horizontal EOG or M_1 electrode, were not included in the analysis. Segments containing muscle activity artifacts (voltage channel differences of more than 160 μV between two adjacent sample points) as well as amplifier saturation (maximal voltage difference less than 0.5 μV over a time epoch of at least 100 ms) were eliminated prior to averaging. Participants with a rejection rate of higher than 40% of the epochs were discarded (see Gädeke et al., 2013 for a comparable data analysis approach). For ERP analysis, we used a 2 (groups: blind, control) * 4 (emotions: neutral, happy, fearful, threatening) * 2 (attention: attended, unattended) * 2 (electrodes: M4, M7) repeated measures design. The rationale behind choosing two midline electrodes M4 and M7 is that no mid-line electrode has been included in the electrode clusters. Please note that the analysis including electrode clusters is now reported in the supplement. The central electrode M4 has been chosen as auditory vertex potentials are known to be maximal in amplitude at central scalp electrodes (see also Figure 1). Moreover, M4 has been investigated in Gädeke et al., 2013 in sighted individuals. Electrode M7 has been selected as this is a more posterior electrode (corresponding to the Pz electrode of the 10-20 system). Research in blind individuals has shown that the auditory N1 recorded at posterior electrodes is modulated differently in congenitally blind individuals compared to sighted controls (Amadeo et al., 2019; Föcker et al., 2012; Hötting et al., 2004; Röder et al., 1999a; see also: Leclerc et al., 2000). Statistical analysis of mean amplitudes was performed for the following three time epochs (same time windows as for the cluster analysis): first time window (110–150 ms), second time window (190–260 ms), and third time window (260–350 ms) and are reported below. Greenhouse-Geisser –corrected p-values are reported. In order to prevent an inflation of the alpha error, the Bonferroni-correction was applied in case of violation of sphericity assumptions for behavioral and EEG data.

The bootstrapping analysis with replacement was conducted with R (R Core Team 2018; version 3.6.1).

4.5. Behavioral Data

D-prime was calculated in order to estimate the performance accuracy for discriminating the positions as a function of emotional prosody: d’ = z(p(hit)) − z(p(FA)) (Green & Swets, 1966). The hit rate was defined as the number of correct responses to deviant stimuli presented at the attended position divided by the total number of deviants presented at the attended position. The false alarm rate (FA rate) was defined as the number of incorrect responses to deviant stimuli presented at the unattended position divided by the total number of deviants at the unattended position. Mean reaction times (RT) and percent correct (PC) were also calculated for each condition and participant. In order to account for potential speed-accuracy trade-offs, the Inverse Efficiency Scores (IES) were calculated for each condition by dividing RT by PC (Townsend & Ashby, 1987; Spence et al., 2001). Trials with reaction times below 200 ms or exceeding 1700 ms were disregarded (see also Gädeke et al., 2013 for a similar procedure).

Analysis of Variance (ANOVAs) with repeated measurement factor *Emotion* (four levels: neutral, happy, threatening, and fearful) and the between subject factor *Group* (congenitally blind versus sighted controls) were run for the dependent variables d-prime (d’) and inverse efficiency scores (IEs). A main effect of Emotion was further analyzed with t-tests (two-tailed) for dependent samples.

Note also that we performed a bootstrapping analysis (with replacement) by randomly selecting 8 sighted controls and comparing them against 8 blind participants to account for sample size differences (in 1000 iterations). The results of the bootstrapping analysis test were largely comparable to our main analysis and are not reported in the main text (see supplement for a more detailed description and results of this analysis).

Acknowledgements:

We are grateful to Professor Dr. Brigitte Röder for providing us the data sets. This study was supported by a grant of the German Research Foundation of Brigitte Röder [DFG, Ro1226/4-3] and European Research Council [ERC-2009-AdG 249425-CriticalBrainChanges] to Brigitte Röder.

Author contributions

PT: Analyzing the data, writing and editing the paper. **AZ:** Analyzing the data, editing and writing the paper. **JG:** Designing the experiment**,** running the EEG experiment, editing and writing the paper, **JF:** Analyzing the data, editing and writing the paper.

Declarations of interest:

none

Figure Legends:

Figure 1 A-B. Experimental design (A). Pseudowords are presented at either the left or the right loudspeaker. Participants were asked to respond to targets (deviant syllables, example fefi) at the attended loudspeaker (Attend left or Attend right loudspeaker). In this case, the participant had to attend to the left loudspeaker and respond to deviant syllables (non-identical syllables) at the left side. All participants were blindfolded throughout the experiment. Experimental setup (B). Two loudspeakers were positioned in front of the participant at a distance of 1.4 m, one 45 degrees to the left and one 45 degrees to the right of the participant. All stimuli were presented with an equal probability and in randomized order from the left and right loudspeakers.

Figure 2. **Behavioral data (A-D). A**) **d-prime**: ***Main effect of Emotion.*** Higher d-prime scores were observed to neutral human voices compared to happy, threatening and fearful human voices. D-prime for fearful human voices were higher compared to threatening and happy voices. **B) Inverse Efficiency Scores:** Main effect of Emotion. Inverse efficiency scores were significant lower in the neutral condition compared to the threatening condition. **C)** ***Main effect of Group***. Congenitally blind individuals reached higher d-prime values compared to sighted controls. **D)** Main effect of *Group*, with lower inverse efficiency scores in congenitally blind individuals compared to sighted controls.

Figure 3. A-C. A) ERPs recorded to human voices presented at the attended (red dashed line) and unattended location (black solid line) averaged across all participants. **B)** ERPs recorded

to the different emotional prosodies (neutral = black, fearful = red, happy = dashed blue line, threatening = dotted green line) averaged across all participants**. C)** ERPs averaged separately across all participants in congenitally blind and sighted controls. ERPs are shown at the central electrode M4 and the posterior electrode M7 (corresponds to Pz according to the 10-20 system, see electrode montage). ERPs reveal a more pronounced negativity in congenitally blind individuals (dashed red line) compared to sighted controls. Selected time windows are shaded in grey.

Figure 4. Difference waves (attended minus unattended) in sighted controls (upper row, left side, A) and congenitally blind (upper row, right side, B) separately for neutral (black line), fearful (blue line), threatening (green dotted line), and happy (red dashed line) human voices at central electrode M4 (A,B) and posterior electrode M7 (C,D, see electrode montage). The ERPs of the selected time window (N1 (110-150 ms) are zoomed in a higher resolution as shown in the orange circle. Lower Row: Topographical distribution of the attention effect (attended minus unattended) across all emotions separately for each time window (110-150 ms, 190-260 ms, 260-350 ms) and separately for sighted controls (E) and congenitally blind (F). Selected time windows are shaded in grey.

Figure 5: Mean amplitudes for spatial attended (red dashed line) and unattended locations (black solid line) in the time range of the N1 plotted as a function of emotions (fear, happy, neutral and threat) separately for congenitally blind (A,B) and sighted controls (C,D) at electrodes M4 and M7 (** = P < .01, * = P < .05), bars represent standard errors of the mean. For congenitally blind individuals, the main effect of *Attention* is shown. For sighted controls, the interaction between *Emotion* and *Attention* is presented (see Results section).

References

Amadeo, M. B., Störmer, V. S., Campus, C., Gori, M., 2019. Peripheral sounds elicit stronger activity in contralateral occipital cortex in blind than sighted individuals. *Scientific reports*, 9.

Amedi, A., Raz, N., Pianka, P., Malach, R., Zohary, E., 2003. Early ‘visual’cortex activation correlates with superior verbal memory performance in the blind. *Nature neuroscience, 6*(7), 758.

Arnaud, L., Gracco, V., Ménard, L., 2018. Enhanced perception of pitch changes in speech and music in early blind adults. *Neuropsychologia*, *117*, 261-270.

Bavelier, D., Neville, H. J., 2002. Cross-modal plasticity: where and how?. *Nature Reviews Neuroscience*, *3*(6), 443.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. Glot Inter-national 5:9/10, 341-345

Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer [Computer program]. Version 6.1.09, retrieved 26 January 2020 from http://www.praat.org/

Bull, R., Rathborn, H., Clifford, B. R., 1983. The voice-recognition accuracy of blind listeners. *Perception*, *12*(2), 223-226.

Cappagli, G., & Gori, M. (2016). Auditory spatial localization: Developmental delay in children with visual impairments. Research in developmental disabilities, 53, 391-398.

Collignon, O., Renier, L., Bruyer, R., Tranduy, D., & Veraart, C. (2006). Improved selective and divided spatial attention in early blind subjects. Brain research, 1075(1), 175-182.

Collignon, O., Voss, P., Lassonde, M., & Lepore, F. (2009). Cross-modal plasticity for the spatial processing of sounds in visually deprived subjects. Experimental brain research, 192(3), 343.

Collignon, O., Dormal, G., Albouy, G., Vandewalle, G., Voss, P., Phillips, C., & Lepore, F. (2013). Impact of blindness onset on the functional organization and the connectivity of the occipital cortex. Brain, 136(9), 2769-2783.

Collignon, O., Vandewalle, G., Voss, P., Albouy, G., Charbonneau, G., Lassonde, M., & Lepore, F. (2011). Functional specialization for auditory–spatial processing in the occipital cortex of congenitally blind humans. Proceedings of the National Academy of Sciences, 108(11), 4435-4440.

de Borst, A. W., & de Gelder, B. (2019). Mental imagery follows similar cortical reorganization as perception: intra-Modal and cross-modal plasticity in congenitally blind. Cerebral Cortex, 29(7), 2859-2875.

De Volder, A. G., Bol, A., Blin, J., Robert, A., Arno, P., Grandin, C., ... Veraart, C., 1997. Brain energy metabolism in early blind subjects: neural activity in the visual cortex. *Brain research*, *750*(1-2), 235-244.

Doucet, M. E., Guillemot, J. P., Lassonde, M., Gagné, J. P., Leclerc, C., Lepore, F., 2005.

Blind subjects process auditory spectral cues more efficiently than sighted individuals. *Experimental brain research, 160*(2), 194-202.

Easton, R. D., Greene, A. J., DiZio, P., & Lackner, J. R., 1998. Auditory cues for orientation and postural control in sighted and congenitally blind people. *Experimental brain research, 118*(4), 541-550.

Elbert, T., Sterr, A., Rockstroh, B., Pantev, C., Müller, M. M., Taub, E., 2002. Expansion of the tonotopic area in the auditory cortex of the blind. *Journal of Neuroscience, 22*(22), 9941-9944.

Fairhall, S. L., Porter, K. B., Bellucci, C., Mazzetti, M., Cipolli, C., Gobbini, M. I. (2017). Plastic reorganization of neural systems for perception of others in the congenitally blind. *Neuroimage, 158,* 126-135.

Fine, I., Park, J. M., 2018. Blindness and Human Brain Plasticity. *Annual review of vision science,* 4, 337-356.

Finocchietti, S., Cappagli, G., & Gori, M. (2017). Auditory spatial recalibration in congenital blind individuals. *Frontiers in neuroscience, 11,* 76.

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.

Föcker, J., Hölig, C., Best, A., Röder, B., 2015. Neural plasticity of voice processing: Evidence from event-related potentials in late-onset blind and sighted individuals. *Restorative Neurology and Neuroscience, 33*(1), 15-30.

Gädeke, J. C., Föcker, J., Röder, B., 2013. Is the processing of affective prosody influenced by spatial attention? an ERP study. *BMC neuroscience, 14*(1), 14.

Gougoux, F., Belin, P., Voss, P., Lepore, F., Lassonde, M., Zatorre, R. J., 2009. Voice perception in blind persons: a functional magnetic resonance imaging study. *Neuropsychologia, 47*(13), 2967-2974.

Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R. J., Belin, P., 2004. Neuropsychology: pitch discrimination in the early blind. *Nature, 430*(6997), 309.

Gori, M., Sandini, G., Martinoli, C., & Burr, D. C. (2014). Impairment of auditory spatial localization in congenitally blind human subjects. *Brain, 137*(1), 288-293.

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., Vuilleumier, P., 2005. The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature neuroscience, 8*(2), 145.

Green, D. M., Swets, J. A., 1966. Signal detection theory and psychophysics (Vol. 1). New York: Wiley.

Hampson, P. J., Duffy, C., 1984. Verbal and spatial interference effects in congenitally blind and sighted subjects. *Canadian Journal of Psychology, 38*(3), 411.

Hölig, C., Föcker, J., Best, A., Röder, B., Büchel, C., 2014a. Brain systems mediating voice identity processing in blind humans. *Human brain mapping, 35*(9), 4607-4619.

Hölig, C., Föcker, J., Best, A., Röder, B., Büchel, C., 2014b. Crossmodal plasticity in the fusiform gyrus of late blind individuals during voice recognition. *Neuroimage, 103,* 374-382.

Holmes, A., Vuilleumier, P., Eimer, M., 2002. The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. *Cognitive Brain Research, 16*(2), 174-184.

Hötting, K., Rösler, F., Röder, B., 2004. Altered auditory-tactile interactions in congenitally blind humans: an event-related potential study. *Experimental brain research, 159*(3), 370-381.

Hugdahl, K., Ek, M., Takio, F., Rintee, T., Tuomainen, J., Haarala, C., & Hämäläinen, H. (2004). Blind individuals show enhanced perceptual and attentional sensitivity for identification of speech sounds. *Cognitive brain research, 19*(1), 28-32.

Jessen, S., Kotz, S. A., 2011. The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. Neuroimage, 58(2), 665-674.

Jiang, F., Stecker, G. C., Fine, I., 2014. Auditory motion processing after early blindness. *Journal of vision, 14*(13), 4-4.

King, A. J. (2014). What happens to your hearing if you are born blind? *Brain, 137*(1), 6-8

Klinge, C., Eippert, F., Röder, B., Büchel, C., 2010b. Corticocortical connections mediate primary visual cortex responses to auditory stimulation in the blind. *Journal of Neuroscience, 30*(38), 12798-12805.

Klinge, C., Röder, B., Büchel, C., 2010a. Increased amygdala activation to emotional auditory stimuli in the blind. *Brain, 133*(6), 1729-1736.

Kokinous, J., Kotz, S. A., Tavano, A., Schröger, E., 2015. The role of emotion in dynamic audiovisual integration of faces and voices. *Social Cognitive and Affective Neuroscience, 10*(5), 713-720.

Kujala, T., Alho, K., Kekoni, J., Hämäläinen, H., Reinikainen, K., Salonen, O., ... Näätänen, R., 1995. Auditory and somatosensory event-related brain potentials in early blind humans. *Experimental brain research, 104*(3), 519-526.

Kujala, T., Lehtokoski, A., Alho, K., Kekoni, J., Näätänen, R., 1997. Faster reaction times in the blind than sighted during bimodal divided attention. *Acta Psychologica, 96*(1-2), 75-82.

Lavan, N., Burton, A. M., Scott, S. K., McGettigan, C., 2019. Flexible voices: Identity perception from variable vocal signals. *Psychonomic bulletin & review, 26*(1), 90-102.

Leclerc, C., Saint-Amour, D., Lavoie, M. E., Lassonde, M., Lepore, F., 2000. Brain functional reorganization in early blind humans revealed by auditory event-related potentials. *Neuroreport, 11*(3), 545-550.

Lessard, N., Paré, M., Lepore, F., Lassonde, M., 1998. Early-blind human subjects localize sound sources better than sighted subjects. *Nature, 395*(6699), 278.

Liu, T., Pinheiro, A. P., Deng, G., Nestor, P. G., McCarley, R. W., Niznikiewicz, M. A., 2012. Electrophysiological insights into processing nonverbal emotional vocalizations.

NeuroReport,*23*(2), 108-112.

Matteau, I., Kupers, R., Ricciardi, E., Pietrini, P., Ptito, M., 2010. Beyond visual, aural and haptic movement perception: hMT+ is activated by electrotactile motion stimulation of the tongue in sighted and in congenitally blind individuals. *Brain research bulletin*, *82*(5-6), 264-270.

Ménard, L., Trudeau-Fisette, P., Côté, D., & Turgeon, C. (2015). Acoustic and articulatory correlates of speaking condition in blind and sighted speakers. In ICPhS.

Merabet, L. B., Pascual-Leone, A., 2010. Neural reorganization following sensory loss: the opportunity of change. *Nature Reviews Neuroscience*, *11*(1), 44.

Mothes-Lasch, M., Mentzel, H. J., Miltner, W. H., Straube, T., 2011. Visual attention modulates brain activation to angry voices. *Journal of Neuroscience, 31*(26), 9594-9598.

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.

Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, *24*(4), 375-425.

Pessoa, L., Ungerleider, L. G., 2004. Neuroimaging studies of attention and the processing of emotion-laden stimuli. *Progress in brain research*, *144*, 171-182.

Pessoa, L., McKenna, M., Gutierrez, E., Ungerleider, L. G., 2002. Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences, 99*(17), 11458-11463.

Pinheiro, A. P., Del Re, E., Mezin, J., Nestor, P. G., Rauber, A., McCarley, R. W., ... Niznikiewicz, M. A., 2013. Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: an electrophysiological investigation. *Psychological Medicine, 43*(3), 603-618.

Pisanski, K., Feinberg, D., Oleszkiewicz, A., Sorokowska, A., 2017. Voice cues are used in a similar way by blind and sighted adults when assessing women’s body size. *Scientific reports, 7*(1), 10329.

R Core Team 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Röder, B., Demuth, L., Streb, J., Rösler, F., 2003. Semantic and morpho-syntactic priming in auditory word recognition in congenitally blind adults. *Language and Cognitive Processes*, *18*(1), 1-20.

Röder, B., Krämer, U. M., Lange, K., 2007. Congenitally blind humans use different stimulus selection strategies in hearing: an ERP study of spatial and temporal attention. *Restorative Neurology and Neuroscience*, *25*(3-4), 311-322.

Röder, B., Rösler, F., Hennighausen, E., & Näcker, F. (1996). Event-related potentials during auditory and somatosensory discrimination in sighted and blind human subjects. Cognitive Brain Research, 4(2), 77-93.

Röder, B., Rösler, F., Neville, H. J., 1999b. Effects of interstimulus interval on auditory event-related potentials in congenitally blind and normally sighted humans. *Neuroscience letters*, *264*(1-3), 53-56.

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

Rokem, A., Ahissar, M., 2009. Interactions of cognitive and auditory abilities in congenitally blind individuals. *Neuropsychologia*, *47*(3), 843-848.

Sauter, D. A., Eimer, M., 2010. Rapid detection of emotion from human vocalizations. *Journal of cognitive neuroscience*, *22*(3), 474-481.

Schenkman, B. N., Nilsson, M. E., 2010. Human echolocation: Blind and sighted persons' ability to detect sounds recorded in the presence of a reflecting object. *Perception*, *39*(4), 483-501.

Schild, U., Friedrich, C. K., 2018. What determines the speed of speech recognition? Evidence from congenitally blind adults. *Neuropsychologia*, *112*, 116-124.

Schweinberger, S. R., Kawahara, H., Simpson, A. P., Skuk, V. G., Zäske, R., 2014. Speaker perception. Wiley Interdisciplinary Reviews: Cognitive Science, 5(1), 15-25.

Singh, A. K., Phillips, F., Merabet, L. B., Sinha, P., 2018. Why does the cortex reorganize after sensory loss?. *Trends in cognitive sciences*, *22*(7), 569-582.

Skuk, V. G., Schweinberger, S. R., 2013. Adaptation aftereffects in vocal emotion perception elicited by expressive faces and voices. *PloS one*, *8*(11), e81691.

Spence, C., Kingstone, A., Shore, D. I., Gazzaniga, M. S., 2001. Representation of visuotactile space in the split brain. *Psychological Science*, *12*(1), 90-93.

Stevens, A. A., Weaver, K. E., 2009. Functional characteristics of auditory cortex in the blind. *Behavioural brain research*, *196*(1), 134-138.

Szűcs, D., Csépe, V., 2005. The parietal distance effect appears in both the congenitally blind and matched sighted controls in an acoustic number comparison task. *Neuroscience letters*, *384*(1-2), 11-16.

Townsend, J. T., Ashby, F. G., 1978. Methods of modeling capacity in simple processing systems. *Cognitive theory*, *3*, 211-252.

Vercillo, T., Burr, D., Gori, M., 2016. Early visual deprivation severely compromises the auditory sense of space in congenitally blind children. *Developmental psychology*, *52*(6), 847.

Voss, P. (2016). Auditory spatial perception without vision. Frontiers in psychology, 7, 1960.

Voss, P., & Zatorre, R. J. (2012). Occipital cortical thickness predicts performance on pitch and musical tasks in blind individuals. *Cerebral Cortex*, *22*(11), 2455-2465.

Vuilleumier, P., Schwartz, S., 2001. Emotional facial expressions capture attention. *Neurology*, *56*(2), 153-158.

Wan, C. Y., Wood, A. G., Reutens, D. C., Wilson, S. J., 2010. Early but not late-blindness leads to enhanced auditory perception. *Neuropsychologia*, *48*(1), 344-348.

Zinchenko, A., Kanske, P., Obermeier, C., Schröger, E., Kotz, S. A., 2015. Emotion and goal-directed behavior: ERP evidence on cognitive and emotional conflict. *Social cognitive and affective neuroscience*, *10*(11), 1577-1587.

Zinchenko, A., Obermeier, C., Kanske, P., Schröger, E., Kotz, S. A., 2017. Positive emotion impedes emotional but not cognitive conflict processing.*Cognitive, Affective, & Behavioral Neuroscience*,*17*(3), 665-677.

Title: The role of spatial selective attention in the processing of affective prosodies in congenitally blind adults: An ERP study

Short Title: Processing of affective prosodies in congenitally blind adults

Authors and authors' affiliations: Pavlos Topalidis[1], Artyom Zinchenko[1], Julia C. Gädeke[2] & Julia Föcker[2,3]

[1] Department of Psychology and Educational Sciences, Ludwig Maximilian University, Munich Germany

[2] Biological Psychology and Neuropsychology, University of Hamburg, Germany

[3] University of Lincoln, School of Social Sciences, United Kingdom

***Corresponding author:**

Julia Föcker
School of Psychology
College of Social Sciences
University of Lincoln
United Kingdom
E-mail: JFöcker@lincoln.ac.uk

Keywords: congenitally blind, sensory deprivation, plasticity, auditory, attention, emotion, event-related potentials, human voices

Abstract

The question whether spatial selective attention is necessary in order to process vocal affective prosody has been controversially discussed in sighted individuals: whereas some studies argue that attention is required in order to process emotions, other studies conclude that vocal prosody can be processed even outside the focus of spatial selective attention. Here, we asked whether spatial selective attention is necessary for the processing of affective prosodies after visual deprivation from birth. For this purpose, pseudowords were presented at the left or right loudspeaker and spoken in happy, neutral, fearful or threatening prosodies. Congenitally blind individuals (N = 8) and sighted controls (N=13) had to attend to one of the loudspeakers and detect rare pseudowords presented at the attended loudspeaker during EEG recording. Emotional prosody of the syllables was task-irrelevant. Blind individuals outperformed sighted controls by being more efficient in detecting deviant syllables at the attended loudspeaker. Higher auditory N1 amplitude was observed in blind individuals compared to sighted controls. Additionally, sighted controls showed enhanced attention-related ERP amplitudes in response to fearful and threatening voices during the time range of the N1. By contrast, blind individuals revealed enhanced ERP amplitudes in attended relative to unattended locations irrespective of the affective valence in all time windows (110-350 ms). These effects were mainly observed at posterior electrodes. The results provide evidence for “emotion-general“ auditory spatial selective attention effects in congenitally blindness and provide further indirect support for the idea of reorganization of the voice processing brain system following visual deprivation from birth.

Words: 250/250

1. Introduction

Human voices and vocalizations play an essential role in social interactions and communication as they allow us to not only process speech, but also to draw conclusions about other people’s affective state, age, gender and even a person’s body size (Lavan et al., 2019; Pisanski et al., 2017; Schweinberger et al., 2014; Skuk & Schweinberger, 2013; Zinchenko et al., 2015, 2017). Processing of human voices becomes particularly important in blind individuals as vocal features can be identified even from long distances. Some characteristics of human voice processing have been extensively studied in blind individuals, such as auditory perceptual skills (Arnaud et al., 2018; Jiang et al., 2014; Röder et al., 1999b); auditory memory (Amedi et al., 2003; Bull et al., 1983; Rokem & Ahissar, 2009), person identification (Fairhall, et al. 2017; Föcker et al., 2012, 2015; Hölig et al., 2014a, 2014b), language (Röder et al., 2003; Schild & Friedrich, 2018), auditory localization and spatial selective attention (Amadeo et al., 2019; Doucet et al., 2005; Muchnik et al., 1991; Röder et al., 1999a). Surprisingly, the nature of human affective voice processing undergoing neural plastic reorganization after visual deprivation – and more importantly – the processing of emotional features– are rather unknown so far (Fairhall et al. 2017; Klinge et al., 2010a).

One of the methods to study attention- and emotion-related processes is electroencephalography (EEG), which is known for its high temporal resolution. In sighted individuals, it was shown that emotions can modulate auditory event-related potentials (ERPs) as early as 100 and 200 ms after stimulus onset (N1; P2) but also during later processing stages such as between 260-350 ms (see also Gädeke et al., 2013; Liu et al., 2012; Pinheiro et al., 2013). For instance, ERP responses to emotional vocalizations differed from ERPs to neutral vocalizations at around 120 ms (Jessen & Kotz, 2011) and 150 ms poststimulus (Sauter & Eimer, 2010). Additionally, Pinheiro et al., (2013) observed an enhanced negativity to neutral compared to angry spoken words in the time range of the auditory N1 that has been in-terpreted as emotional evaluation of incoming sensory information (Kokinous et al., 2015). This implies that the processing of the affective quality of the signal happens very early during sensory processing, possibly due to its high relevance for survival and social interactions. To sum up, while electrophysiological correlates of emotional prosody processing are relatively well studied in sighted individuals, these processes are less understood in blind individuals.

Interestingly, a growing number of studies have shown improved auditory localization skills in blind individuals using behavioral, electrophysiological and brain imaging studies (Collignon et al., 2006; Doucet et al., 2005; Muchnik et al., 1991; Röder et al., 2007, 1999). In some of those studies, blind and sighted participants attended to a sound source in space and detected rare target stimuli while ignoring more frequent auditory standards and other (task-irrelevant) rare deviant stimuli presented at the same or other loudspeakers (e.g., Röder et al., 1999a). As a result, the authors found that blind relative to sighted participants could localize spatial positions of targets significantly further away in the periphery (Röder et al., 1999a). In line with these findings, Röder et al. (1999a) reported that blind relative to sighted participants showed a more pronounced ERP negativity (N1) in response to more peripheral sources of audio stimuli (see also Amadeo et al., 2019; Föcker et al., 2012; Röder et al., 2007 for an enhanced auditory N1 in blind individuals).

In a previous study, Röder et al. (2007) asked 8 congenitally blind individuals and 12 sighted controls to attend either to the left or right loudspeaker at which auditory stimuli were presented and to concentrate either on a long or short time interval which separated the two auditory stimuli (S1 and S2) from each other. The authors examined the length of audio refractory period across the two groups. Refractory periods are defined as time periods during which the cell is not able to generate further action potentials. Interestingly, congenitally blind individuals showed a more pronounced ERP negativity for the second auditory stimulus (S2), suggesting shorter auditory refractory periods in the blind compared to the sighted controls (Röder et al., 2007). This implies that blind participants had an advantage in the processing of auditory stimuli. Correspondingly, another study has shown that the auditory N1 recovered faster in the blind than in the sighted controls when the interstimulus interval between the two auditory stimuli was varied (Röder et al., 1999b). As the neural generators for the auditory N1 are thought to originate in the primary and secondary auditory cortices (Näätänen & Picton, 1987), an enhanced excitability of the auditory cortex might contribute to enhanced perceptual skills in the blind.

To summarize, there is consistent evidence that blind individuals show generally more efficient processing of auditory information (Fine & Park, 2018; but see Collignon et al., 2009, and Singh et al., 2018 for a further discussion). However, there is a lack of research on whether spatial selective attention is necessary in order to process affective prosodies after visual deprivation from birth. This question is of interest, as blind individuals rely much more on vocal cues and could potentially be more efficient in detecting emotional features, even outside the focus of spatial attention. By contrast, in the sighted population there is convincing evidence that emotions can be processed within and even outside of the focus of spatial selective attention (Grandjean et al., 2005; Holmes et al., 2003; Mothes-Lasch et al., 2011; Pessoa et al., 2002; Pessoa & Ungerleider, 2004; Vuilleumier & Schwartz, 2001). In one pioneering study, Grandjean and coauthors (2005) examined whether processing of emotional prosody depends on selective attention to the voice. Participants listened to auditory utterances pronounced with either threatening or neutral tone of voice in a dichotic listening task. Specifically, participants were asked to attend either to the left or right ear and identify the gender of a speaker at the target-ear and ignore the voices presented in the unattended ear. Results showed that activations in response to threatening utterances in the middle part of the right superior temporal sulcus occurred irrespective of the attended location, indicating that the brain could still detect emotional prosody from voices presented at the non-attended location (Grandjean et al., 2005). By contrast, other studies have challenged the “automaticity” hypothesis of emotional processing. For instance, Pessoa and colleagues (2002) indicated that especially under high load conditions, attention is necessary in order to process emotional features. Evidence for the hypothesis that spatial attention modulates the degree of emotional voice processing as a function of emotional valence was observed in an EEG experiment: Auditory pseudowords (neutral, happy, threatening, and fearful) have been presented at two different loudspeakers and participants were asked to detect rare deviant syllables (e.g. “giki”, “fefi”) at the attended location while ignoring all standard pseudowords presented at the same location and all deviants and standards at the non-attended location (Gädeke et al., 2013). Emotional valence of the pseudowords was task-irrelevant. As a result, the authors found more pronounced negativity in response to attended versus unattended voices specifically in the time range of the auditory N1 and especially for fearful voices. This implies that processing of emotional information modulates early but not later stages of information processing. Importantly, these authors also showed emotion-specific brain activations at both attended and unattended locations in this early time-window, suggesting that emotions can be processed even outside the focus of selective spatial attention in sighted individuals.

In order to investigate whether spatial selective attention is necessary to process emotional prosody in blind individuals, we used the well-established paradigm outlined above (Gädeke et al., 2013) and applied it to congenitally blind individuals. In more detail, we used an auditory spatial attention paradigm in which participants were asked to detect rare bisyllabic pseudowords (e.g. “fefi”; “giki”, “nane”) at the attended loudspeaker and ignore the same infrequently presented syllables at the unattended loudspeaker as well as more frequently presented pseudowords (e.g. “baba”, “dede”, “fafa”) at both loudspeakers (see Figure 1).

INSERT FIGURE 1 HERE

Pseudoword were presented in four different emotions (neutral, happy, fearful and threatening). The emotion of the pseudowords was task-irrelevant. In sighted controls it has been shown that attention modulates the processing of emotional prosody at early perceptual processing stages by showing a more pronounced negativity to attended fearful voices compared to unattended fearful voices. We asked (1) whether congenitally blind individuals would show the same attentional capture by negative stimuli (e.g. fearful) as sighted controls in the time range of the auditory N1 and whether there would be a main effect of *Emotion* and *Attention* in later time windows (>150ms) similar to sighted controls. We were also interested (2) whether congenitally blind individuals outperform sighted controls in distinguishing targets (“giki”, “fefi”, “nane”) from more frequently presented standard voices (e.g. “baba”, “dede”, “fafa”) at the attended location and if so, (3) at which processing stages would congenitally blind individuals differ from sighted controls (early versus late). This question was motivated by previous work that compared different temporal processing stages between congenitally blind individuals and sighted controls (Föcker et al., 2012, 2015; Röder et al., 1999). It was found that congenitally blind relative to sighted controls show different patterns and topographical distributions of auditory event-related potentials, e.g., N1, N2b, mismatch negativity (MMN), Auditory-evoked Contralateral Occipital Positivity (ACOP) recorded at posterior electrodes (e.g., Pz), which was linked to cortical reorganization of the auditory system in the blind (Alho et al., 1993; Amadeo et al., 2019; Föcker et al., 2012; Hötting et al., 2004; Kujala et al., 1992; Röder et al., 1999a; see also: Leclerc, Saint-Amour, Lavoie, Lassonde, & Lepore, 2000). Therefore, (4) we aimed to investigate at which electrodes (central versus posterior) differences in emotional processing are mostly pronounced in congenitally blind and sighted controls.

We hypothesized (1) that if emotions are processed outside the focus of spatial selective attention, we should observe emotion-related ERP modulations independently of the focus of spatial selective attention. That is, ERPs for fearful, threatening and happy voices should show different patterns of activity (i.e., amplitudes) relative to neutral voices within both attended and unattended conditions. However, if attention is required to process emotional valence, we expected different modulations of ERPs with regards to emotional valence for spatially attended and unattended stimuli similar to sighted controls in the time range of the auditory N1 to fearful human voices (Gädeke et al., 2013). In more detail, Gädeke and colleagues (2013) showed that sighted individuals revealed a more pronounced N1 negativity in response to attended relative to unattended fearful human voices, which might be an index of an enhanced suppression of spatially irrelevant human fearful voices and an enhanced capture of attention to fearful voices presented at the attended location. We did not expect any interaction between attention and emotion at later processing stages (see Gädeke et al., 2013). Regarding question (2), we hypothesized that blind individuals would be more efficient in processing human voices at the attended speaker compared to sighted controls (Klinge et al., 2010a).

Similarly to previous studies (Röder et al., 1999a, 1999b; 2007) we expected more enhanced auditory N1 amplitudes in the congenitally blind individuals compared to sighted controls (3). Finally, (4) we expected to find group-specific differences between congenitally blind and sighted controls at more posterior electrode sites as observed in previous research (Amadeo et al., 2019; Föcker et al., 2012; Röder et al., 1999a,b).

2. Results

In the following, results are presented including N = 8 congenitally blind individuals and N = 13 sighted controls. We first describe the behavioral results followed by the event-related potential (ERP) results.

2.1. Behavioral results

For the behavioral results we report the ANOVA including the factors *Emotion* (happy, neutral, fearful, threatening) and the between subject factor *Group* (congenitally blind individuals versus sighted controls) on d'prime and Inverse Efficiency scores (IE scores). IE scores combine both reaction times and correct responses (Townsend & Ashby, 1987; Spence et al., 2001) and have been used as we aimed to follow the same procedure as reported in Gädeke et al., 2013. Percent correct (PC), mean reaction times (RT), d'prime as well as IE scores are reported in Table 1.

INSERT TABLE 1 HERE

2.1.1. D-prime scores

As expected, blind individuals outperformed sighted controls in distinguishing targets from more frequently presented standards at the attended location (main effect of *Group*: $F(1,19) = 19.557$, $P < .001$, $\eta^2 = .507$, blind individuals: mean d' = 2.7, SE = .113; sighted controls: mean d' = 2.1, SE = .088, see Figure 2 C). Moreover, all participants could better detect targets at the attended location when spoken in a neutral prosody compared to happy, threatening or fearful emotions (main effect of *Emotion*: $F(3,57) = 22.625$, $P < .001$, $\eta^2 =$

.544, neutral = 3.089, SE = .154; threatening = 2.125, SE = .108; happy = 2.135, SE = .057; fearful = 2.512, SE = .095; all *Ps* < .001, see Figure 2 A). Furthermore, d-prime for fearful human voices were higher compared to threatening and happy voices (*P* < .007). The interaction between *Emotion* and *Group* was not significant (*F*(3,57) = .85, *P* = .445).

2.1.2. Inverse Efficiency (IE) scores

Participants responded more efficiently in the neutral condition compared to the threatening vocal prosody (main effect of *Emotion: F*(3,57) = 5.898, *P* = .008; η² = .237; mean neutral = 1405 ms, SE=141; mean happy: 1639 ms, SE = 92; mean threatening: 1997 ms, SE = 206, mean fearful: 1459 ms, SE = 93, *P* = .001; see Figure 2B). Moreover, blind individuals responded more efficiently to target voices compared to sighted controls (main effect of *Group: F*(1,19) = 8.093, *P* = .010, η² = .299; sighted controls*:* mean: 1922 ms, SE=128; blind individuals: mean*:* 1328 ms, SE=164, see Figure 2D). The interaction between *Emotion* and *Group* was not significant (*F*(3,57) = .165, *P* = .826).

INSERT FIGURE 2 HERE

2.2. ERP results

For the ERP analysis, we used a 2 (*Group*: congenitally blind, sighted control) * 4 (*Emotion:* neutral, happy, fearful, threatening) * 2 (*Attention:* attended, unattended) repeated measures analysis of variance. We first run an ANOVA at the central electrode M4 given that the auditory vertex potential is maximal in amplitude at this site. Based on our hypotheses that differences between congenitally blind and sighted controls would be mainly observed at the posterior electrode M7, we run an additional

ANOVA at the posterior electrode M7 (corresponding to Pz). We finalized this result section by reporting the results of the four-way interaction between the factors *Attention, Emotion, Electrode* and *Group*.

Figure 3 summarizes the ERPs recorded to human voices presented at the attended versus unattended loudspeaker averaged across all participants (Figure 3 A, D), the ERPs averaged separately for each emotional condition (neutral, happy, threatening, fearful) across all participants (Figure 3 B, E), and the ERPs averaged separately in all congenitally blind and all sighted controls (Figure 3 C, F) at the central electrode M4 and the posterior electrode M7 (corresponding to electrode Pz of the 10-20 system). Figure 4 shows the difference waves (attended minus unattended) as a function of emotional valence (neutral, happy, threatening and fearful) and the topographical distribution of the attention effect (E,F) for the three time windows separately for congenitally blind (B, D) and sighted controls (A, C) at electrode M4 and M7). Figure 5 illustrates the mean amplitudes for spatial attended (red dashed line) and unattended locations (black solid line) in the time range of the N1 plotted as a function of emotions (fearful, happy, neutral and threatening) separately for congenitally blind (A,B) and sighted controls (C,D) at electrodes M4 and M7.

To foresee the results, we observed significant main effects of *Emotion* and *Attention* in the ERP amplitudes of congenitally blind individuals across all time windows. By contrast, in sighted controls, the interaction between *Attention* and *Emotion* with mean ERP amplitudes as dependent measurement was significant in the first time window, but not in the second or third time windows.

INSERT FIGURE 3 HERE

INSERT FIGURE 4 HERE

Time Window: 110-150 ms

Electrode M4

The ANOVA including the factors *Emotion, Attention* and *Group* on mean ERP amplitudes revealed a main effect of *Attention* and a main effect of *Emotion* (main effect of *Attention*: *F*(1,19) = 21.855, *P* < .001, η^2= .535; main effect of *Emotion*: *F*(3,57) = 20.648, *P* < .001, η^2= .521). ERPs were more negative to attended compared to unattended human voices (mean attended: -3.202 µV, SE = .425; mean unattended: -2.598 µV, SE = .380, see Figure 3 A). Moreover, ERPs to neutral prosodies revealed a more pronounced negativity compared to all other emotions (mean neutral: - 3.578 µV, SE = .408, mean happy: -2.151µV, SE = .357; mean threatening: -2.892µV, SE = .473, mean fearful: -2.98µV, SE = .407, see Figure 3 B). Additionally, ERPs recorded to happy voices revealed a less pronounced negativity compared to all other voices (all *P*s < .012). The main effect of *Group* was significant (*F*(1,19) = 5.013, *P* = .037, $\eta 2$= .209, see Figure 3 C). The auditory N1 amplitude was more negative in congenitally blind individuals compared to sighted controls (congenitally blind: mean: -3.791 µV, SE = .626; sighted controls: mean: -2.009 µV, SE = .491).

The interaction between *Emotion*Attention*Group* was not significant (*F*(3,57) = 1.153, *P* = .331).

Correlation between behavioral performance and auditory N1

Electrode M4

In order to investigate, whether higher auditory N1 amplitudes were associated with improved performance in congenitally blind individuals but not in the sighted controls, we calculated the correlations between mean amplitudes of the auditory N1 and the behavioral performance (IE scores and d'). The correlations between the auditory N1 and IE scores and the auditory N1 and d' were not significant within each group (*Blind individuals:* IE scores: r

= -.195, *P* = .643; *Blind individuals d prime*: r = .230, *P* = .584; *Sighted Controls:* IE scores: r = -.009, *P* = .977; *Sighted Controls:* d prime: r = .258. *P* = .396).

Time Window: 190-260 ms

Electrode M4

The main effect of *Attention* was significant *(Attention: F*(1,19) = 38.447, *P* < .001*;* η^2= .669*)*. ERPs to the attended condition were more negative compared to the unattended condition *(mean attended:* 2.891µV, SE = .700*; mean unattended:* 4.834µV, SE = .693, see Figure 3 A*)*. Additionally, the main effect of *Emotion* was significant *(F*(3,57) = 31.933, *P < .001;* η^2= .627). ERPs revealed a more pronounced positivity to the threatening voices compared to all other prosodies (mean threatening: 4.88µV, SE = .717; mean neutral: 3.58µV, SE = .659; mean fearful: 2.922µV, SE = .687; mean happy: 4.067µV, SE = .698, *P* < .001, see Figure 3B). Additionally, ERPs to fearful voices were more negative compared to all other voices (all *Ps* < .006). The interaction between *Emotion, Attention* and *Group* was not significant (*F*(3,57) = .461, *P* = .684).

Time Window: 260-350 ms

M4

The ANOVA including the factors *Attention, Emotion* and *Group* on mean ERP amplitudes revealed a main effect of *Attention (Attention: F(1,19) = 67.967, P < .001;* η^2= .782*)*. ERPs of the attended condition were more negative compared to ERPs of the unattended condition *(mean attended:* -.144µV, *SE* = .651; *mean unattended:* 2.210 µV, *SE* = .629, see Figure 3A*)*. Additionally, the main effect of *Emotion* was significant *(F*(3,57) = 23.257, *P < .001;* η^2= .550). ERPs revealed a more pronounced positivity to the happy voices compared to all other prosodies (mean happy: 2.115µV, SE = .612; mean neutral: .688µV, SE = .619;

mean fearful: .280µV, SE = .660; mean threatening: 1.049µV, SE = .667, $P < .001$, see Figure 3B). Additionally, ERPs to threatening voices revealed a more pronounced positivity compared to fearful voices ($P = .021$). The interaction between *Emotion, Attention* and *Group* was not significant ($F(3,57) = .585$, $P = .594$).

Time Window: 110-150 ms

Electrode M7 (Pz)

The overall ANOVA including the factors *Attention, Emotion* and *Group on mean ERP amplitudes* revealed a main effect of *Attention* ($F(1,19) = 31.248$, $P < .001$, $\eta^2 = .\ 0.622$) with a more pronounced negativity in the attended compared to unattended condition (mean attended: -2.596 µV, SE = .233; mean unattended: -2.089 µV, SE = .203; see Figure 3 D). Additionally, the main effect of *Emotion* was significant ($F(3,57) = 11.758$, $P < .001$, $\eta^2 = .382$) with a more pronounced negativity in the neutral condition compared to the threatening and happy condition (mean neutral: -2.841 µV, SE = .273; mean happy: -1.945 µV, SE= .185; mean threatening: -2.286 µV, SE= .249; mean fearful: -2.299 µV, SE= .217; all *P*s <.006; see Figure 3 E). Moreover, ERPs recorded to happy voices revealed a less pronounced negativity compared to fearful and neutral voices (all *P*s < .006). The main effect of *Group* was not significant ($F(1,19) = 2.789$, $P = .111$, $\eta^2 = .\ 128$).

Importantly, the interaction between the factors *Emotion, Attention* and *Group* was significant ($F(3,57) = 2.975$, $P = .048$, $\eta^2 = .135$).

In the congenitally blind individuals, we observed no significant *Emotion* by *Attention* interaction ($F(3,21) = .895$, $P = .424$, $\eta^2 = .113$). However, this interaction was significant in sighted controls ($F(3,36) = 4.066$, $P = .018$, $\eta^2 = .253$). Subordinate ANOVAs confirmed that the effect of *Emotion* was significant at both the attended and unattended location but the higher F value for the attended condition ($F(3,36) = 10.109$, $P < .001$, $\eta^2 = .457$) than the unattended condition ($F(3,36) = 5.956$, $P = .004$, $\eta^2 = .332$) suggests a stronger *Emotion* effect at the attended location (see Figure 5). Post hoc t-tests showed a difference between ERPs of the attended and unattended condition only for the fearful and threatening voices (fearful condition attended versus unattended: $t(12) = 3.708$, $P = .003$, mean attended = -2.51 mV, SE: .22, mean unattended = -1.67mV, SE: 1.07; threatening condition attended versus unattended: $t(12) = 2.875$, $P = .014$, mean attended = -2.10 mV, SE: .31, unattended = -1.42 mV, SE: .25; all other *P*'s > .36, see Figure 4 C,D).

INSERT FIGURE 5 HERE

Correlation between behavioral performance and auditory N1

Electrode M7

Similarly to M4, we correlated mean N1 amplitudes and the behavioral performance (IE scores and d'). Neither of the effects reached significance (*Blind individuals:* IE scores: r = -.199, *P* = .637; *Blind individuals d'*: r = -.147, *P* = .729; *Sighted Controls:* IE scores: r = .04, *P* = .896; *Sighted Controls:* d': r = .055. *P* = .857).

Time Window: 190-260 ms

M7

The overall ANOVA including the factors *Attention, Emotion,* and *Group* on mean ERP amplitudes revealed a main effect of *Attention* ($F(1,19) = 26.936$, $P < .001$, $\eta^2= .586$) with a more pronounced negativity to ERPs in the attended compared to unattended condition (mean attended: 2.89 µV, SE = .700; mean unattended: 4.83 µV, SE = .69; see Figure 3 D). Additionally, the main effect of *Emotion* was significant ($F(3,57) = 14.823$, $P < .001$, $\eta^2=$

.438) with a more pronounced positivity to threatening voices compared to all other emotions (mean neutral: .788 µV, SE= .344; mean happy: mean = .947 µV, SE: .365, mean threatening: = 1.384 µV, SE = .395, mean fearful: mean = .317 µV, SE= .362, $P < .039$, see Figure 3 E). Moreover, ERPs to fearful voices were more negative compared to all other emotions (all *P*s < .019).

We also observed an interaction between the factors *Attention*Group* ($F(1,19) = 7.717$, $P = .012$, $\eta^2 = .289$) showing stronger differences between the attended and unattended condition in blind individuals compared to the sighted controls (blind individuals: mean attended: -.887 µV, SE = .469, mean unattended: 1.235 µV, SE = .178; $t(7) = -4.009$, $P = .005$; sighted controls: mean attended: 1.222 µV, SE = .597; mean unattended: 1.864 µV, SE = .489, $t(12) = -2.418$, $P = .032$; see also more posterior shift of the attention effect in the topographies in congenitally blind individuals, Figure 4 E,F). The interaction between *Attention*, *Emotion* and *Group* and the main effect of *Group* were not significant (*Attention*Emotion*Group:* $F(3,57) = .334$, $P = .794$; main effect of *Group*: $F(1,19) = 3.763$, $P = .067$).

Time Window: 260-350 ms

M7

Similar to the first and the second time windows, a main effect of *Attention* and a main effect of *Emotion* were observed (main effect of *Attention*: $F(1,19) = 6.841$, $P = .017$, $\eta^2 = .265$; main effect of *Emotion*: $F(3,57) = 15.718$, $P < .001$, $\eta^2 = .453$, see Figures 3 D, E). ERPs to human voices presented at the attended location were more negative compared to human voices presented at the unattended location (mean attended: .162 µV, SE = .463, mean unattended: 1.149 µV, SE = .293, $P < .001$). ERPs to happy human voices revealed a more pronounced positivity compared to all other emotions (mean neutral: .444 µV SE = .347;

mean happy: 1.276 µV SE = .312, mean threatening: .68, SE = .341 µV, mean fearful: .223 µV, SE = .405, $P < .004$).

The interaction between the factors *Attention* and *Group* was significant ($F(1,19) = 6.36$, $P = .021$, $\eta^2 = .251$). Separate ANOVAs run in each *Group* revealed a significant main effect of *Attention* in congenitally blind, but not in sighted controls (congenitally blind: $F(1,7) = 13.596$, $P = .008$, $\eta^2 = .66$; blind individuals: mean attended: -.287 µV, SE: .547; mean unattended: 1.65 µV, SE: 465; sighted controls: $F(1,12) = .005$ $P = .944$, sighted controls mean attended: .612 µV, SE: 640; mean unattended: .647 µV, SE: 359).

Finally, the interaction between *Emotion, Attention* and *Group* was not significant (interaction between *Emotion, Attention* and *Group*: $F(3,57) = .878$, $P = .443$).

Note also that the critical 4-way interaction of *Attention* (attended versus unattended), *Emotion* (neutral, happy, threatening, fearful), *Electrode* (M4, M7) and *Group* (congenitally blind versus sighted controls) on mean ERP amplitudes was significant in the first time window only ($F(3,57) = 6.258$, $P = .003$, $\eta^2 = .248$, for all other time windows: $P > .7$). Confirming the similarity across the two electrodes, the 4-way interaction was not significant for the second and third time windows (see Table 2).

INSERT TABLE 2 HERE

---

3. Discussion

The goal of the present study was to understand whether spatial selective attention is necessary for processing of affective prosodies after visual deprivation from birth. Therefore, we aimed at identifying the time course and underlying processing stages that differ in congenitally blind adults compared to sighted controls and that potentially provide enhanced auditory emotional processing capacities. Moreover, we tried to understand if, similar to sighted controls, congenitally blind individuals suppress irrelevant fearful voices and attend to relevant fearful human voices at the attended location during early processing stages (auditory N1; see Gädeke et al., 2013). This effect was demonstrated by a more pronounced negativity to attended relative to unattended fearful human voices in sighted controls (see Gädeke et al., 2013). Finally, we analyzed whether the group differences in orienting spatial selective attention to different emotional voices are distributed at posterior electrodes (Amadeo et al., 2019; Föcker et al., 2012; Hötting et al., 2004; Röder et al., 1999a; see also: Leclerc et al., 2000).

For this purpose, an auditory oddball paradigm was run in which participants had to detect rare deviant syllables at the attended location and ignore deviant syllables at the unattended location as well as all standard syllables at both locations. We observed that congenitally blind individuals were more efficient compared to sighted controls in detecting deviant syllables at the attended spatial location. Those group effects cannot be due to gender or age differences as both groups did not differ in this respect. This result pattern contributes to a large range of studies reporting superior auditory skills in the blind, such as pitch discrimination and auditory spectral cues (Doucet et al., 2005; Gougoux et al., 2004; Wan et al., 2010), human echolocation (Schenkman & Nilsson, 2010), auditory language processing (Röder et al., 2003; Schild & Friedrich, 2018), auditory memory (Amedi et al., 2003; Rokem & Ahissar, 2009), auditory spatial selective attention (Hugdahl et al., 2004; Kujala et al., 1995; 1997; Lessard et al., 1998; Röder et al., 1999a,b) and processing of auditory vocal prosody

(Klinge et al., 2020). Particularly, those results confirm findings of enhanced auditory spatial selective attention in blind individuals (demonstrated in higher d primes) and point to the fact that blind individuals might not be distracted by the emotional valence of the voices when attending to a specific spatial location (Hugdahl et al., 2004; Kujala et al., 1995; 1997; Lessard et al., 1998; Röder et al., 1999a,b).

Early perceptual processing

Consistently with previous studies in this area, we found an enhanced N1 amplitude in blind individuals compared to sighted controls (Amadeo et al., 2019; Doucet et al., 2005; Muchnik et al., 1991; Röder et al., 1999a). This group difference in the N1 mirrors facilitated behavioral performance in congenitally blind individuals. It might be speculated, that an improved representation of auditory perceptual features (as measured via N1) contributes to more efficient task processing at the attended location in blind participants. Other studies have argued that there is a more efficient perceptual encoding in the blind as reflected in shorter N1 latencies and shorter recovery periods of auditory ERPs (Elbert et al., 2020; Röder et al., 1996).

On the other hand, four out of 13 blind participants were excluded in the current experiment from data analysis because they were not able to perceptually discriminate the two female vocal identities, which has been set as a test of basic hearing abilities and was used criteria to be included in data analysis (see Gädeke et al., 2013; Bull et al., 1983; see Föcker et al., 2012; Hölig et al., 2014 a, 2014 b; for a better voice identification performance in congenitally blind compared to sighted controls). Some studies did report impaired performance on auditory tasks in blind individuals (Cappagli & Gori 2016; Finocchietti et al. 2015; Gori et al. 2014; Menard et al. 2015; Voss, 2016), while the others found no difference from sighted participants (Collignon et al. 2011, 2013; Voss & Zatorre 2012). The heterogeneity of results reported in auditory tasks in blind individuals may be due to different task requirements (e.g. voice identification versus detecting target syllables in the current experiment; see also King, 2014; de Borst & de Gelder, 2019, p.2860) or different training protocols (see Föcker et al., 2012).

In the time range of the auditory N1, we observed a main effect of *Attention* in blind individuals: similar to sighted controls, congenitally blind individuals showed a difference between attended and unattended fearful voices. However, this effect was not specifically tight to negative human voices such as fear or threat as in sighted controls (see Gädeke et al., 2013 for a further discussion for sighted controls). Thus, while attention effects for most of the emotional voices were observed relatively late in sighted individuals (> 150 ms), the main effect of spatial attention to emotional stimuli was already established in blind individuals and quite similar across all emotions including happy, fearful, neutral and threatening. Spatial selective attention might act as a mechanism that allows processing of emotions at the attended location and suppressing irrelevant information at the unattended location. It might be argued that congenitally blind individuals have an “improved and more efficient” spatial filter system in order to process and distinguish relevant from irrelevant information irrespective of the type of emotion.

We argue that the emotional valence of auditory stimuli might be partially extracted automatically (in the absence of at least spatial attention) in the congenitally blind. This is shown by the main effect of *Emotion* in congenitally blind individuals, which suggests that emotions are processed in the attended and the unattended channel in a similar way. This corresponds to findings reported by Klinge et al. (2010a) in congenitally blind individuals: In this study, congenitally blind participants and sighted controls had to discriminate either the emotional prosody (happy, threatening, neutral, fearful: emotion discrimination task) or the first vowel of each stimulus (a, e, i, o: vowel discrimination task) while functional brain activity was recorded (Klinge et al., 2010a). As a result, blind individuals showed higher proficiency in discriminating voice prosodies, they were faster in emotion discrimination compared to sighted controls and showed higher activation in occipital cortex to *all* emotional vocal stimuli (Klinge et al., 2010a). This group of participants also showed higher amygdala activation in response to threatening and fearful compared to neutral voices. Moreover, amygdala activation was observed *irrespective* of the underlying task (emotion versus vowel discrimination task), indicating that this activation is not related to explicit emotion detection, but is rather automatically driven by the emotional valence of the stimulus.

It has to be noticed that quite long inter-stimulus intervals (ISIs) were applied in the current experiment and it is well known that N1 attention effects are elicited if short ISIs are employed (see also Gädeke et al., 2013 for a similar discussion). Of course, we cannot exlude the fact that participants might have additional resources left over to attend to the other (i.e., task-irrelevant) loudspeaker. However,we argue that this is unlikely. Spatial selective attention effects were already established in the first time window, especially in the congenitally blind individuals to all emotions, which suggest specific enhancement of the processing of vocal stimuli by spatial attention even when long ISIs are applied. Nevertheless, future studies could examine this idea more explicitly by additonally taxing participants' attentional resources and testing whether participants would still be able to show emotion-specific processing at the unattended spatial locations.

Interestingly, ERPs were modulated by emotional valence in both sighted controls and congenitally blind individuals in a similar way in the time range of the auditory N1, suggesting that emotions itself are similarly processed in both groups. ERPs showed a more pronounced negativity to neutral voices compared to threatening, happy or fearful human voices for both groups. This is in line with previous studies (Liu et al., 2012; Pinheiro et al., 2013) and might suggest that salient acoustic cues direct the emotional evaluation. Interestingly, enhanced N1 amplitudes to neutral voices might reflect improved voice detection of neutral stimuli another indication that enhanced amplitudes mirror better task performance.

Later processing stages

In the second time window (190-260 ms), we observed that both, congenitally blind and sighted controls showed a main effect of *Attention* and a main effect of *Emotion* (without any interactions). However, the difference in ERPs to the attended versus unattended condition was much stronger in congenitally blind individuals compared to sighted controls. Interestingly, unlike congenitally blind individuals, sighted controls did not show any attention effect in the time window 260-350 ms at the posterior electrode M7, suggesting a more sustained attention effect over time in the blind compared to sighted controls especially at posterior electrodes. This is also shown by the more posterior topographical distribution of the attention effect in congenitally blind compared to sighted controls which might point to a reorganization of the voice processing system in congenitally blind individuals. These more posterior topographies of auditory evoked potentials have been also shown in other studies (Amadeo et al., 2019: Auditory-evoked Contralateral Occipital Positivity (ACOP); Föcker et al., 2012; Hötting et al., 2004; Leclerc et al., 2000; Röder et al., 1999a). Therefore, we argue that attention was not necessary to process emotional valence of the voices at these later time windows in both sighted and congenitally blind individuals. However, spatial selective attention – even at this late processing stage – is much more enhanced in blind individuals compared to sighted controls.

Neural reorganization of the emotional voice processing system

It has been suggested that an intramodal reorganization in blind individuals might contribute to enhanced performance in several auditory perceptual tasks (Röder et al., 2007). For instance, brain imaging studies reported cortical reorganization of the auditory cortex as a neural mechanism to understand the shorter auditory N1 latencies (Elbert et al., 2002; Stevens & Weaver, 2009). Besides changes within unisensory brain areas (also called intramodal plasticity in auditory brain structure, Röder & Neville, 2003), other studies observed neural plastic changes in multisensory regions (De Volder et al., 1997; Röder et al., 1999a), including the functional connections between auditory and visual brain areas (Bavelier & Neville, 2002; Klinge et al., 2010b) and additional recruitment of visual cortices during auditory processing (crossmodal plasticity, Merabet & Pascual-Leone, 2010; Fairhall et al., 2017) which has been suggested to facilitate performance of the blind including voice processing (Gougoux et al., 2009). For instance, Gougoux and coauthors (2009) have shown higher voice specific activation in the left superior temporal sulcus (STS) in congenitally blind individuals compared to sighted controls (Gougoux et al., 2009). This increased recruitment of the STS was correlated with their performance in a voice discrimination task (Gougoux et al., 2009). Thus, it might be speculated that visual deprivation from birth leads to a reorganization of the multisensory zone in the STS.

Several other brain imaging studies have shown a crossmodal reorganization in human voice processing tasks, such as a higher activation in the right fusiform gyrus in congenitally blind and even late blind individuals when asked to indicate the age of a voice (see Hölig et al., 2014a, 2014b). This activation has been even observed when onset of blindness starts later in life suggesting that neural reorganization can also be observed in the more mature human brain (Hölig et al., 2014b). Klinge et al. (2010a) observed an enhanced performance of the congenitally blind in auditory discrimination tasks that was paralleled by occipital cortex activation, which was absent in the sighted controls. Even though further studies are needed to understand the exact location of neural plastic reorganization in the current task, we assume that also the recruitment of visual brain areas is involved in the current voice discrimination task in congenitally blind individuals. This assumption is based on the fact, that attention effects, even at later processing stages are observed at more posterior electrodes which is usually atypical for auditory ERPs.

3.1. Conclusion

These results provide evidence for enhanced auditory spatial selective attention irrespective of the emotional valence in the absence of vision from birth and point to a reorganization of the auditory voice processing system following congenital blindness.

4. Experimental Procedure

4.1. Participants

Thirteen congenitally blind individuals participated in the experiment. This sample size was based on a highly relevant previous work in this area (e.g., Gädeke et al., 2013 who included 13 sighted controls in the same paradigm; see also Röder et al., 2007). Five participants had to be excluded from data analysis due to the following reasons: (1) four participants had to be excluded due to very low performance in discriminating human voices (d prime < .04) see also Gädeke et al., 2013 for a similar approach), (2) one participant had too many artifacts in the EEG data recordings (less than 40 % of trials remaining). The final sample consisted of eight congenitally blind individuals (mean age: 26 years, age range: 23-29 years, four female). Please note that comparable sample sizes of blind individuals (N = 8) have been reported in previous studies e.g., de Borst & de Gelder, (2005); Easton et al., (1998); Föcker et al., (2015); Hampson & Duffy, (1984); Matteau et al., (2010); Röder et al., (1999a; 2007); Szucs & Csepe, 2005; Vercillo, Burr, & Gori, (2016). Six participants were students at the University of Marburg, Germany, one participant was a businessman, and another participant was a service operator.

All blind participants were totally blind or did not have more than rudimentary sensitivity for brightness differences without any pattern vision. In all cases, blindness was due to peripheral deficits. More specifically, blindness was due to the following reasons: retinopathia pigmentosa (N = 3), retina degeneration (N=2), too high levels of oxygen in the incubator (N=1). For two participants, the reasons for blindness (peripheral defect) were unknown (N = 2). All participants were German native speakers and reported normal hearing and no history of neurological illness. Eight blind participants were compared with 13 sighted controls (mean age: 23 years, age range: 20—28 years, seven females; see Gädeke et al., 2013). Congenitally blind individuals and sighted controls did not differ in gender or age (gender distribution blind individuals: 4 females and 4 males; sighted controls:7 females, 6 males; $\chi 2 = .0294$, $P = .864$; mean age blind individuals: 26 years, SD = 2.43 years; mean age sighted controls: 23 years, SD = 2.61 years, $t(19) = 1.60$, $P >.05$). Sighted participants had normal or corrected to normal vision. All participants were blindfolded throughout the experiment.

All participants were recruited from the local community or towns near the city of Marburg and received monetary compensation for their participation. Written informed consent was given by each participant prior to the beginning of the experiment. This study was in accordance with the Declaration of Helsinki and approved by the Ethics committee of the medical association of Marburg.

4.2. Stimulus Material

The stimulus material, training and experimental procedure were identical to the procedure reported in Gädeke et al. (2013). The stimulus material has been rated by a separate group of 24 University students (see Gädeke et al., 2013). Nine disyllable pseudo-words spoken by two actresses in four emotional prosodies (neutral, happy, threatening and fearful) were selected for the purpose of the study (9 x 2 x 4 = 72 different stimuli). Pseudowords consisting of two different syllables were classified as deviant stimuli (such as fefi), while the remaining six vocal stimuli with the same two syllables belonged to the standard stimuli (such as fefe). Deviant syllables presented at the attended location (for instance right loudspeaker) are called targets throughout the manuscript. Mean stimulus duration for neutral human voices was 632 ms, SE = 35, for happy human voices 575 ms, SE = 57, for threatening human voices 602 ms, SE = 56 and for fearful human voices 518 ms, SE = 44. We run the Kruskal Wallis test (see Zinchenko et al., 2015 for comparable procedures) in order to compare the stimulus duration of the targets and standards. Results show that the duration between different emotional stimuli does not significantly differ from each other (targets: $\chi2 = 6$, $P > 0.1$; standards: $\chi2 = 3.66$, $P > .2$, df = 3). The characteristics of the stimulus material (duration, pitch, intensity, valence, intensity and dominance ratings) are reported in tables 3 and 4. Pitch was calculated using the Praat phonetics software package (Boersma & Weenink, 2012) developed for Phonetic or Phonological research. Praat uses an autocorrelation method for pitch analysis based on a robust algorithm for periodicity detection, that has been optimised for speech analysis, proposed by Boersma (2001). For further information see http://www.fon.hum.uva.nl/praat/manual/Sound__To_Pitch__ac____.html

INSERT TABLE 3

INSERT TABLE 4

4.3. Procedure

4.3.1. Experiment

Two loudspeakers were positioned in front of the participant at a distance of 1.4 m, one 45 degrees to the left and one 45 degrees to the right of the participant. All stimuli were presented with an equal probability and in randomized order from the left and right loudspeakers. The time intervals between the onset of the presentation of any two successive voices (i.e., stimulus onset asynchronies: SOA) varied between 1300 ms to 1700 ms (see Figure 1).

INSERT FIGURE 5 HERE

Participants’ task was to attend to stimuli which were presented at one of two spatial positions (left or right speaker) spoken by one of the two female speakers. Whenever participants detected one of the deviant stimuli spoken by the attended voice and presented at the attended position (i.e., targets), participants had to lift the right or left index finger out of a light gate. After half of the trials, the response hand was switched (from left to right index finger or vice versa). Emotional prosody of the syllables was task-irrelevant. In total there were four experimental conditions (attend voice I versus attend voice II and attend left vs. attend right loudspeaker).[1] The experiment consisted of 16 blocks; each block lasted six to seven minutes. The following four experimental conditions were presented: condition 1: attend left speaker, attend voice 1; condition 2: attend right speaker, attend voice 1; condition 3: attend left speaker, attend voice 2, condition 4: attend right speaker, attend voice 2. (p.22). A block comprised 192 standard stimuli (80%) and 48 deviant stimuli (20%), 24 of which were targets (5 %). Every two blocks participants were instructed to attend to the other location (e.g., from left to right). Only spatial attention effects with regards to the different emotional prosodies were analyzed.

All participants were blindfolded throughout the experiment and a chin rest was used to restrict head movements. Moreover, participants were instructed to avoid excessive blinking dur-

---

[1] “Originally, the main experiment comprised an additional orthogonally manipulated factor (Gaedeke et al., 2013). Participants had to selectively attend to one voice only. However, the voices of two female actors were too similar and participants did not manage to distinguish between them. Even after excluding participants (N = 4) with very low performance in discriminating the voices (d’ < .04), mean d’ was low (d’ = .67, SE = .08) (see Gaedeke et al., 2013). In the current experiment, we applied the same criteria to congenitally blind indi-viduals and sighted controls, in order to guarantee that there were no pre-existing differ-ences based on any auditory task performed (i.e. voice identification)” (see Gaedeke et al., 2013, p. 14).

ing the blocks. The EEG experiment without any breaks took approximately 1.5 hours. The whole experimental session including breaks, practice and the electrode preparation and removal, lasted between 5 and 6 hours.

4.3.2. Training

In order to familiarize participants with all voice stimuli and experimental procedure, participants had to take part in a training session, one or two days prior to the actual experiment. They were asked to discriminate the voices of the two and the experimental procedure. We did not analyze the factor voice in the current experiment. Participants who were not able to distinguish the two actors were excluded from data analysis (criterion d-prime = 0.04). Further details of the training are provided in Gädeke et al. (2013).

4.4. ERP data

The data acquisition and EEG recording was identical to Gädeke et al. (2013). For the EEG recording 61 Ag/AgCl electrodes were used, mounted equidistantly in an elastic cap (Falk Minow Services, Munich). A bipolar horizontal electrooculogram (HEOG) recording was obtained by attaching two electrodes to the outer canthi of the eyes, and the vertical EOG (VEOG) was monitored by placing an electrode under the right eye against the common reference. The right earlobe electrode was used as reference electrode during recording, but offline all channels were re-referenced to the averaged left and right earlobe references. The ground electrode was placed on a position at the middle of the forehead (below Fpz).

Participant’s skin was prepared by using Every (Meditec SRI, Negernbotel) and alcohol. Electrogel (Electrocap International, Ohio, USA) served as the electrolyte for all electrodes. Impedances were kept below 5 kΩ for scalp recordings and below 10 kΩ for EOG recordings. Signal amplification was made possible by using two SynAmps-amplifiers (NeuroScan, Inc. Sterling, USA). The sample rate was 500 Hz and the bandpass was set to 0.1 – 100 Hz.

For the ERP analysis, the EEG was averaged for time epoch -100 ms (pre-stimulus) to 1000 ms (post- stimulus), for each participant and condition. The prestimulus interval was defined as baseline. Only segments following standard stimuli were analyzed, while segments with responses to standard stimuli were discarded. Segments containing eye movements artifacts, defined as a larger difference of 120 µV between two sample points within a segment of the vertical or horizontal EOG or M_1 electrode, were not included in the analysis. Segments containing muscle activity artifacts (voltage channel differences of more than 160 µV between two adjacent sample points) as well as amplifier saturation (maximal voltage difference less than 0.5 µV over a time epoch of at least 100 ms) were eliminated prior to averaging. Participants with a rejection rate of higher than 40% of the epochs were discarded (see Gädeke et al., 2013 for a comparable data analysis approach). For ERP analysis, we used a 2 (groups: blind, control) * 4 (emotions: neutral, happy, fearful, threatening) * 2 (attention: attended, unattended) * 2 (electrodes: M4, M7) repeated measures design. The rationale behind choosing two midline electrodes M4 and M7 is that no mid-line electrode has been included in the electrode clusters. Please note that the analysis including electrode clusters is now reported in the supplement. The central electrode M4 has been chosen as auditory vertex potentials are known to be maximal in amplitude at central scalp electrodes (see also Figure 1). Moreover, M4 has been investigated in Gädeke et al., 2013 in sighted individuals. Electrode M7 has been selected as this is a more posterior electrode (corresponding to the Pz electrode of the 10-20 system). Research in blind individuals has shown that the auditory N1 recorded at posterior electrodes is modulated differently in congenitally blind individuals compared to sighted controls (Amadeo et al., 2019; Föcker et al., 2012; Hötting et al., 2004; Röder et al., 1999a; see also: Leclerc et al., 2000). Statistical analysis of mean amplitudes was performed for the following three time epochs (same time windows as for the cluster analysis): first time window (110–150 ms), second time window (190–260 ms), and third time window (260–350 ms) and are reported below. Greenhouse-Geisser –corrected p-values are reported. In order to prevent an inflation of the alpha error, the Bonferroni-correction was applied in case of violation of sphericity assumptions for behavioral and EEG data.

The bootstrapping analysis with replacement was conducted with R (R Core Team 2018; version 3.6.1).

4.5. Behavioral Data

D-prime was calculated in order to estimate the performance accuracy for discriminating the positions as a function of emotional prosody: d’ = z(p(hit)) − z(p(FA)) (Green & Swets, 1966). The hit rate was defined as the number of correct responses to deviant stimuli presented at the attended position divided by the total number of deviants presented at the attended position. The false alarm rate (FA rate) was defined as the number of incorrect responses to deviant stimuli presented at the unattended position divided by the total number of deviants at the unattended position. Mean reaction times (RT) and percent correct (PC) were also calculated for each condition and participant. In order to account for potential speed-accuracy trade-offs, the Inverse Efficiency Scores (IES) were calculated for each condition by dividing RT by PC (Townsend & Ashby, 1987; Spence et al., 2001). Trials with reaction times below 200 ms or exceeding 1700 ms were disregarded (see also Gädeke et al., 2013 for a similar procedure).

Analysis of Variance (ANOVAs) with repeated measurement factor *Emotion* (four levels: neutral, happy, threatening, and fearful) and the between subject factor *Group* (congenitally blind versus sighted controls) were run for the dependent variables d-prime (d’) and inverse efficiency scores (IEs). A main effect of Emotion was further analyzed with t-tests (two-tailed) for dependent samples.

Note also that we performed a bootstrapping analysis (with replacement) by randomly selecting 8 sighted controls and comparing them against 8 blind participants to account for sample size differences (in 1000 iterations). The results of the bootstrapping analysis test were largely comparable to our main analysis and are not reported in the main text (see supplement for a more detailed description and results of this analysis).

Acknowledgements:

We are grateful to Professor Dr. Brigitte Röder for providing us the data sets. This study was supported by a grant of the German Research Foundation of Brigitte Röder [DFG, Ro1226/4-3] and European Research Council [ERC-2009-AdG 249425-CriticalBrainChanges] to Brigitte Röder.

Author contributions

PT: Analyzing the data, writing and editing the paper. **AZ:** Analyzing the data, editing and writing the paper. **JG:** Designing the experiment**,** running the EEG experiment, editing and writing the paper, **JF:** Analyzing the data, editing and writing the paper.

Declarations of interest:

none

Figure Legends:

Figure 1 A-B. Experimental design (A). Pseudowords are presented at either the left or the right loudspeaker. Participants were asked to respond to targets (deviant syllables, example fefi) at the attended loudspeaker (Attend left or Attend right loudspeaker). In this case, the participant had to attend to the left loudspeaker and respond to deviant syllables (non-identical syllables) at the left side. All participants were blindfolded throughout the experiment. Experimental setup (B). Two loudspeakers were positioned in front of the participant at a distance of 1.4 m, one 45 degrees to the left and one 45 degrees to the right of the participant. All stimuli were presented with an equal probability and in randomized order from the left and right loudspeakers.

Figure 2. **Behavioral data (A-D). A**) **d-prime**: ***Main effect of Emotion.*** Higher d-prime scores were observed to neutral human voices compared to happy, threatening and fearful human voices. D-prime for fearful human voices were higher compared to threatening and happy voices. **B) Inverse Efficiency Scores:** Main effect of Emotion. Inverse efficiency scores were significant lower in the neutral condition compared to the threatening condition. **C)** ***Main effect of Group***. Congenitally blind individuals reached higher d-prime values compared to sighted controls. **D)** Main effect of *Group*, with lower inverse efficiency scores in congenitally blind individuals compared to sighted controls.

Figure 3. A-C. A) ERPs recorded to human voices presented at the attended (red dashed line) and unattended location (black solid line) averaged across all participants. **B)** ERPs recorded to the different emotional prosodies (neutral = black, fearful = red, happy = dashed blue line, threatening = dotted green line) averaged across all participants**. C)** ERPs averaged separately across all participants in congenitally blind and sighted controls. ERPs are shown at the central electrode M4 and the posterior electrode M7 (corresponds to Pz according to the 10-20

system, see electrode montage). ERPs reveal a more pronounced negativity in congenitally blind individuals (dashed red line) compared to sighted controls. Selected time windows are shaded in grey.

Figure 4. Difference waves (attended minus unattended) in sighted controls (upper row, left side, A) and congenitally blind (upper row, right side, B) separately for neutral (black line), fearful (blue line), threatening (green dotted line), and happy (red dashed line) human voices at central electrode M4 (A,B) and posterior electrode M7 (C,D, see electrode montage). The ERPs of the selected time window (N1 (110-150 ms) are zoomed in a higher resolution as shown in the orange circle. Lower Row: Topographical distribution of the attention effect (attended minus unattended) across all emotions separately for each time window (110-150 ms, 190-260 ms, 260-350 ms) and separately for sighted controls (E) and congenitally blind (F). Selected time windows are shaded in grey.

Figure 5: Mean amplitudes for spatial attended (red dashed line) and unattended locations (black solid line) in the time range of the N1 plotted as a function of emotions (fear, happy, neutral and threat) separately for congenitally blind (A,B) and sighted controls (C,D) at electrodes M4 and M7 (** = P < .01, * = P < .05), bars represent standard errors of the mean. For congenitally blind individuals, the main effect of *Attention* is shown. For sighted controls, the interaction between *Emotion* and *Attention* is presented (see Results section).

References

Amadeo, M. B., Störmer, V. S., Campus, C., Gori, M., 2019. Peripheral sounds elicit stronger activity in contralateral occipital cortex in blind than sighted individuals. *Scientific reports*, 9.

Amedi, A., Raz, N., Pianka, P., Malach, R., Zohary, E., 2003. Early ‘visual’cortex activation correlates with superior verbal memory performance in the blind. *Nature neuroscience, 6*(7), 758.

Arnaud, L., Gracco, V., Ménard, L., 2018. Enhanced perception of pitch changes in speech and music in early blind adults. *Neuropsychologia*, *117*, 261-270.

Bavelier, D., Neville, H. J., 2002. Cross-modal plasticity: where and how?. *Nature Reviews Neuroscience*, *3*(6), 443.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. Glot Inter-national 5:9/10, 341-345

Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer [Computer program]. Version 6.1.09, retrieved 26 January 2020 from http://www.praat.org/

Bull, R., Rathborn, H., Clifford, B. R., 1983. The voice-recognition accuracy of blind listeners. *Perception*, *12*(2), 223-226.

Cappagli, G., & Gori, M. (2016). Auditory spatial localization: Developmental delay in children with visual impairments. Research in developmental disabilities, 53, 391-398.

Collignon, O., Renier, L., Bruyer, R., Tranduy, D., & Veraart, C. (2006). Improved selective and divided spatial attention in early blind subjects. Brain research, 1075(1), 175-182.

Collignon, O., Voss, P., Lassonde, M., & Lepore, F. (2009). Cross-modal plasticity for the spatial processing of sounds in visually deprived subjects. Experimental brain research, 192(3), 343.

Collignon, O., Dormal, G., Albouy, G., Vandewalle, G., Voss, P., Phillips, C., & Lepore, F. (2013). Impact of blindness onset on the functional organization and the connectivity of the occipital cortex. Brain, 136(9), 2769-2783.

Collignon, O., Vandewalle, G., Voss, P., Albouy, G., Charbonneau, G., Lassonde, M., & Lepore, F. (2011). Functional specialization for auditory–spatial processing in the occipital cortex of congenitally blind humans. Proceedings of the National Academy of Sciences, 108(11), 4435-4440.

de Borst, A. W., & de Gelder, B. (2019). Mental imagery follows similar cortical reorganization as perception: intra-Modal and cross-modal plasticity in congenitally blind. Cerebral Cortex, 29(7), 2859-2875.

De Volder, A. G., Bol, A., Blin, J., Robert, A., Arno, P., Grandin, C., ... Veraart, C., 1997. Brain energy metabolism in early blind subjects: neural activity in the visual cortex. *Brain research*, *750*(1-2), 235-244.

Doucet, M. E., Guillemot, J. P., Lassonde, M., Gagné, J. P., Leclerc, C., Lepore, F., 2005.

Blind subjects process auditory spectral cues more efficiently than sighted individuals. *Experimental brain research, 160*(2), 194-202.

Easton, R. D., Greene, A. J., DiZio, P., & Lackner, J. R., 1998. Auditory cues for orientation and postural control in sighted and congenitally blind people. *Experimental brain research, 118*(4), 541-550.

Elbert, T., Sterr, A., Rockstroh, B., Pantev, C., Müller, M. M., Taub, E., 2002. Expansion of the tonotopic area in the auditory cortex of the blind. *Journal of Neuroscience, 22*(22), 9941-9944.

Fairhall, S. L., Porter, K. B., Bellucci, C., Mazzetti, M., Cipolli, C., Gobbini, M. I. (2017). Plastic reorganization of neural systems for perception of others in the congenitally blind. *Neuroimage, 158,* 126-135.

Fine, I., Park, J. M., 2018. Blindness and Human Brain Plasticity. *Annual review of vision science,* 4, 337-356.

Finocchietti, S., Cappagli, G., & Gori, M. (2017). Auditory spatial recalibration in congenital blind individuals. *Frontiers in neuroscience, 11,* 76.

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.

Föcker, J., Hölig, C., Best, A., Röder, B., 2015. Neural plasticity of voice processing: Evidence from event-related potentials in late-onset blind and sighted individuals. *Restorative Neurology and Neuroscience, 33*(1), 15-30.

Gädeke, J. C., Föcker, J., Röder, B., 2013. Is the processing of affective prosody influenced by spatial attention? an ERP study. *BMC neuroscience, 14*(1), 14.

Gougoux, F., Belin, P., Voss, P., Lepore, F., Lassonde, M., Zatorre, R. J., 2009. Voice perception in blind persons: a functional magnetic resonance imaging study. *Neuropsychologia*, *47*(13), 2967-2974.

Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R. J., Belin, P., 2004. Neuropsychology: pitch discrimination in the early blind. *Nature*, *430*(6997), 309.

Gori, M., Sandini, G., Martinoli, C., & Burr, D. C. (2014). Impairment of auditory spatial localization in congenitally blind human subjects. *Brain, 137*(1), 288-293.

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., Vuilleumier, P., 2005. The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature neuroscience, 8*(2), 145.

Green, D. M., Swets, J. A., 1966. Signal detection theory and psychophysics (Vol. 1). New York: Wiley.

Hampson, P. J., Duffy, C., 1984. Verbal and spatial interference effects in congenitally blind and sighted subjects. *Canadian Journal of Psychology*, *38*(3), 411.

Hölig, C., Föcker, J., Best, A., Röder, B., Büchel, C., 2014a. Brain systems mediating voice identity processing in blind humans. *Human brain mapping, 35*(9), 4607-4619.

Hölig, C., Föcker, J., Best, A., Röder, B., Büchel, C., 2014b. Crossmodal plasticity in the fusiform gyrus of late blind individuals during voice recognition. *Neuroimage*, *103*, 374-382.

Holmes, A., Vuilleumier, P., Eimer, M., 2002. The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. *Cognitive Brain Research*, *16*(2), 174-184.

Hötting, K., Rösler, F., Röder, B., 2004. Altered auditory-tactile interactions in congenitally blind humans: an event-related potential study. *Experimental brain research*, *159*(3), 370-381.

Hugdahl, K., Ek, M., Takio, F., Rintee, T., Tuomainen, J., Haarala, C., & Hämäläinen, H. (2004). Blind individuals show enhanced perceptual and attentional sensitivity for identification of speech sounds. *Cognitive brain research*, *19*(1), 28-32.

Jessen, S., Kotz, S. A., 2011. The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. Neuroimage, 58(2), 665-674.

Jiang, F., Stecker, G. C., Fine, I., 2014. Auditory motion processing after early blindness. *Journal of vision*, *14*(13), 4-4.

King, A. J. (2014). What happens to your hearing if you are born blind? *Brain*, *137*(1), 6-8

Klinge, C., Eippert, F., Röder, B., Büchel, C., 2010b. Corticocortical connections mediate primary visual cortex responses to auditory stimulation in the blind. *Journal of Neuroscience*, *30*(38), 12798-12805.

Klinge, C., Röder, B., Büchel, C., 2010a. Increased amygdala activation to emotional auditory stimuli in the blind. *Brain*, *133*(6), 1729-1736.

Kokinous, J., Kotz, S. A., Tavano, A., Schröger, E., 2015. The role of emotion in dynamic audiovisual integration of faces and voices. *Social Cognitive and Affective Neuroscience*, *10*(5), 713-720.

Kujala, T., Alho, K., Kekoni, J., Hämäläinen, H., Reinikainen, K., Salonen, O., ... Näätänen, R., 1995. Auditory and somatosensory event-related brain potentials in early blind humans. *Experimental brain research*, *104*(3), 519-526.

Kujala, T., Lehtokoski, A., Alho, K., Kekoni, J., Näätänen, R., 1997. Faster reaction times in the blind than sighted during bimodal divided attention. *Acta Psychologica*, *96*(1-2), 75-82.

Lavan, N., Burton, A. M., Scott, S. K., McGettigan, C., 2019. Flexible voices: Identity perception from variable vocal signals. *Psychonomic bulletin & review*, *26*(1), 90-102.

Leclerc, C., Saint-Amour, D., Lavoie, M. E., Lassonde, M., Lepore, F., 2000. Brain functional reorganization in early blind humans revealed by auditory event-related potentials. *Neuroreport*, *11*(3), 545-550.

Lessard, N., Paré, M., Lepore, F., Lassonde, M., 1998. Early-blind human subjects localize sound sources better than sighted subjects. *Nature*, *395*(6699), 278.

Liu, T., Pinheiro, A. P., Deng, G., Nestor, P. G., McCarley, R. W., Niznikiewicz, M. A., 2012. Electrophysiological insights into processing nonverbal emotional vocalizations.

NeuroReport,23(2), 108-112.

Matteau, I., Kupers, R., Ricciardi, E., Pietrini, P., Ptito, M., 2010. Beyond visual, aural and haptic movement perception: hMT+ is activated by electrotactile motion stimulation of the tongue in sighted and in congenitally blind individuals. *Brain research bulletin*, *82*(5-6), 264-270.

Ménard, L., Trudeau-Fisette, P., Côté, D., & Turgeon, C. (2015). Acoustic and articulatory correlates of speaking condition in blind and sighted speakers. In ICPhS.

Merabet, L. B., Pascual-Leone, A., 2010. Neural reorganization following sensory loss: the opportunity of change. *Nature Reviews Neuroscience*, *11*(1), 44.

Mothes-Lasch, M., Mentzel, H. J., Miltner, W. H., Straube, T., 2011. Visual attention modulates brain activation to angry voices. *Journal of Neuroscience, 31*(26), 9594-9598.

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.

Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, *24*(4), 375-425.

Pessoa, L., Ungerleider, L. G., 2004. Neuroimaging studies of attention and the processing of emotion-laden stimuli. *Progress in brain research*, *144*, 171-182.

Pessoa, L., McKenna, M., Gutierrez, E., Ungerleider, L. G., 2002. Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences, 99*(17), 11458-11463.

Pinheiro, A. P., Del Re, E., Mezin, J., Nestor, P. G., Rauber, A., McCarley, R. W., ... Niznikiewicz, M. A., 2013. Sensory-based and higher-order operations contribute to abnormal emotional prosody processing in schizophrenia: an electrophysiological investigation. *Psychological Medicine, 43*(3), 603-618.

Pisanski, K., Feinberg, D., Oleszkiewicz, A., Sorokowska, A., 2017. Voice cues are used in a similar way by blind and sighted adults when assessing women's body size. *Scientific reports, 7*(1), 10329.

R Core Team 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Röder, B., Demuth, L., Streb, J., Rösler, F., 2003. Semantic and morpho-syntactic priming in auditory word recognition in congenitally blind adults. *Language and Cognitive Processes*, *18*(1), 1-20.

Röder, B., Krämer, U. M., Lange, K., 2007. Congenitally blind humans use different stimulus selection strategies in hearing: an ERP study of spatial and temporal attention. *Restorative Neurology and Neuroscience*, *25*(3-4), 311-322.

Röder, B., Rösler, F., Hennighausen, E., & Näcker, F. (1996). Event-related potentials during auditory and somatosensory discrimination in sighted and blind human subjects. Cognitive Brain Research, 4(2), 77-93.

63
64
65

Röder, B., Rösler, F., Neville, H. J., 1999b. Effects of interstimulus interval on auditory event-related potentials in congenitally blind and normally sighted humans. *Neuroscience letters*, *264*(1-3), 53-56.

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

Rokem, A., Ahissar, M., 2009. Interactions of cognitive and auditory abilities in congenitally blind individuals. *Neuropsychologia*, *47*(3), 843-848.

Sauter, D. A., Eimer, M., 2010. Rapid detection of emotion from human vocalizations. *Journal of cognitive neuroscience*, *22*(3), 474-481.

Schenkman, B. N., Nilsson, M. E., 2010. Human echolocation: Blind and sighted persons' ability to detect sounds recorded in the presence of a reflecting object. *Perception*, *39*(4), 483-501.

Schild, U., Friedrich, C. K., 2018. What determines the speed of speech recognition? Evidence from congenitally blind adults. *Neuropsychologia*, *112*, 116-124.

Schweinberger, S. R., Kawahara, H., Simpson, A. P., Skuk, V. G., Zäske, R., 2014. Speaker perception. Wiley Interdisciplinary Reviews: Cognitive Science, 5(1), 15-25.

Singh, A. K., Phillips, F., Merabet, L. B., Sinha, P., 2018. Why does the cortex reorganize after sensory loss?. *Trends in cognitive sciences*, *22*(7), 569-582.

Skuk, V. G., Schweinberger, S. R., 2013. Adaptation aftereffects in vocal emotion perception elicited by expressive faces and voices. *PloS one*, *8*(11), e81691.

Spence, C., Kingstone, A., Shore, D. I., Gazzaniga, M. S., 2001. Representation of visuotactile space in the split brain. *Psychological Science*, *12*(1), 90-93.

Stevens, A. A., Weaver, K. E., 2009. Functional characteristics of auditory cortex in the blind. *Behavioural brain research*, *196*(1), 134-138.

Szűcs, D., Csépe, V., 2005. The parietal distance effect appears in both the congenitally blind and matched sighted controls in an acoustic number comparison task. *Neuroscience letters, 384*(1-2), 11-16.

Townsend, J. T., Ashby, F. G., 1978. Methods of modeling capacity in simple processing systems. *Cognitive theory, 3*, 211-252.

Vercillo, T., Burr, D., Gori, M., 2016. Early visual deprivation severely compromises the auditory sense of space in congenitally blind children. *Developmental psychology, 52*(6), 847.

Voss, P. (2016). Auditory spatial perception without vision. Frontiers in psychology, 7, 1960.

Voss, P., & Zatorre, R. J. (2012). Occipital cortical thickness predicts performance on pitch and musical tasks in blind individuals. *Cerebral Cortex, 22*(11), 2455-2465.

Vuilleumier, P., Schwartz, S., 2001. Emotional facial expressions capture attention. *Neurology, 56*(2), 153-158.

Wan, C. Y., Wood, A. G., Reutens, D. C., Wilson, S. J., 2010. Early but not late-blindness leads to enhanced auditory perception. *Neuropsychologia*, *48*(1), 344-348.

Zinchenko, A., Kanske, P., Obermeier, C., Schröger, E., Kotz, S. A., 2015. Emotion and goal-directed behavior: ERP evidence on cognitive and emotional conflict. *Social cognitive and affective neuroscience*, *10*(11), 1577-1587.

Zinchenko, A., Obermeier, C., Kanske, P., Schröger, E., Kotz, S. A., 2017. Positive emotion impedes emotional but not cognitive conflict processing.*Cognitive, Affective, & Behavioral Neuroscience*,*17*(3), 665-677.

Figure1

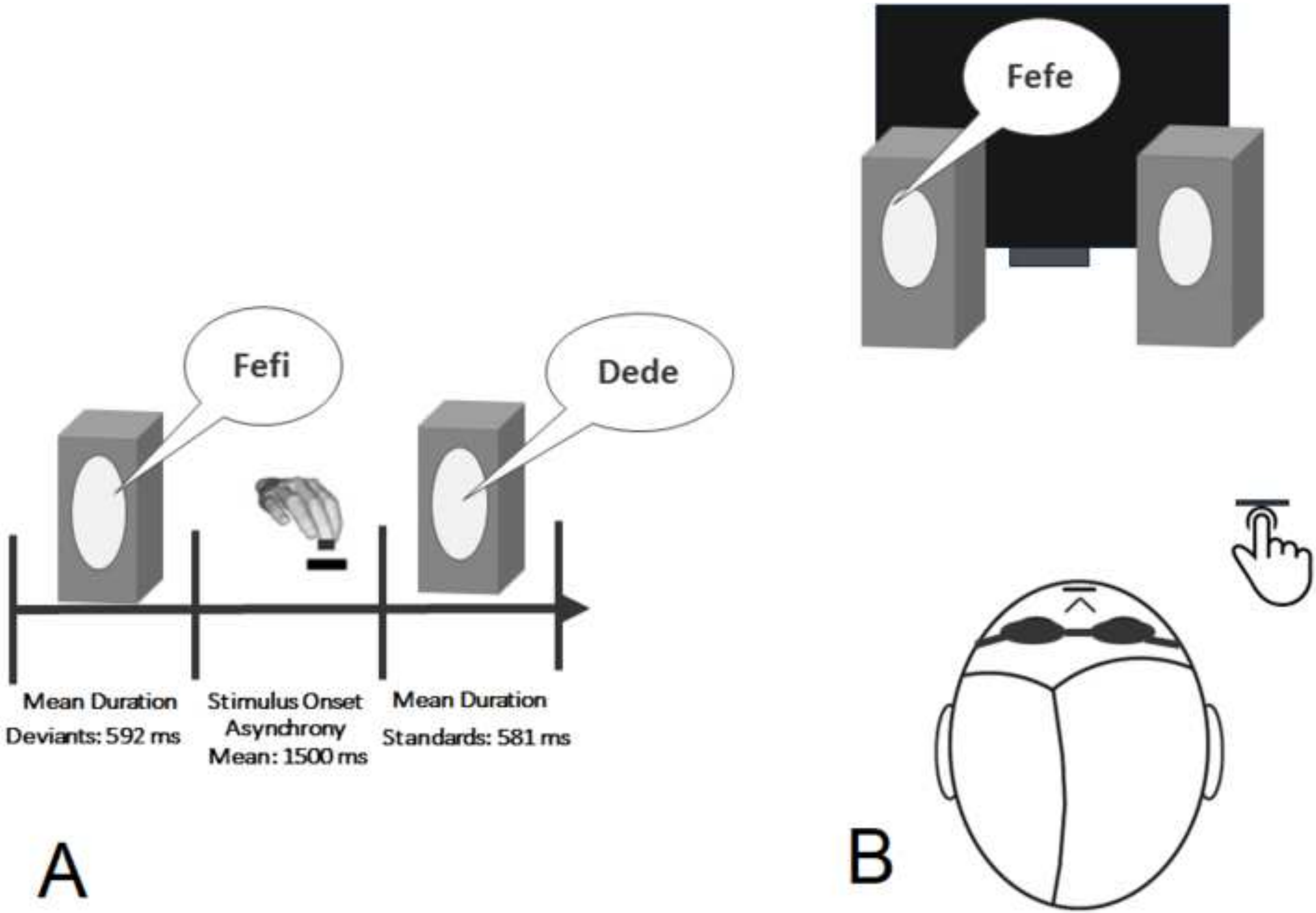
Fefi
Dede
Mean Duration
Deviants: 592 ms
Stimulus Onset
Asynchrony
Mean: 1500 ms
Mean Duration
Standards: 581 ms
A
Fefe
B

Figure2

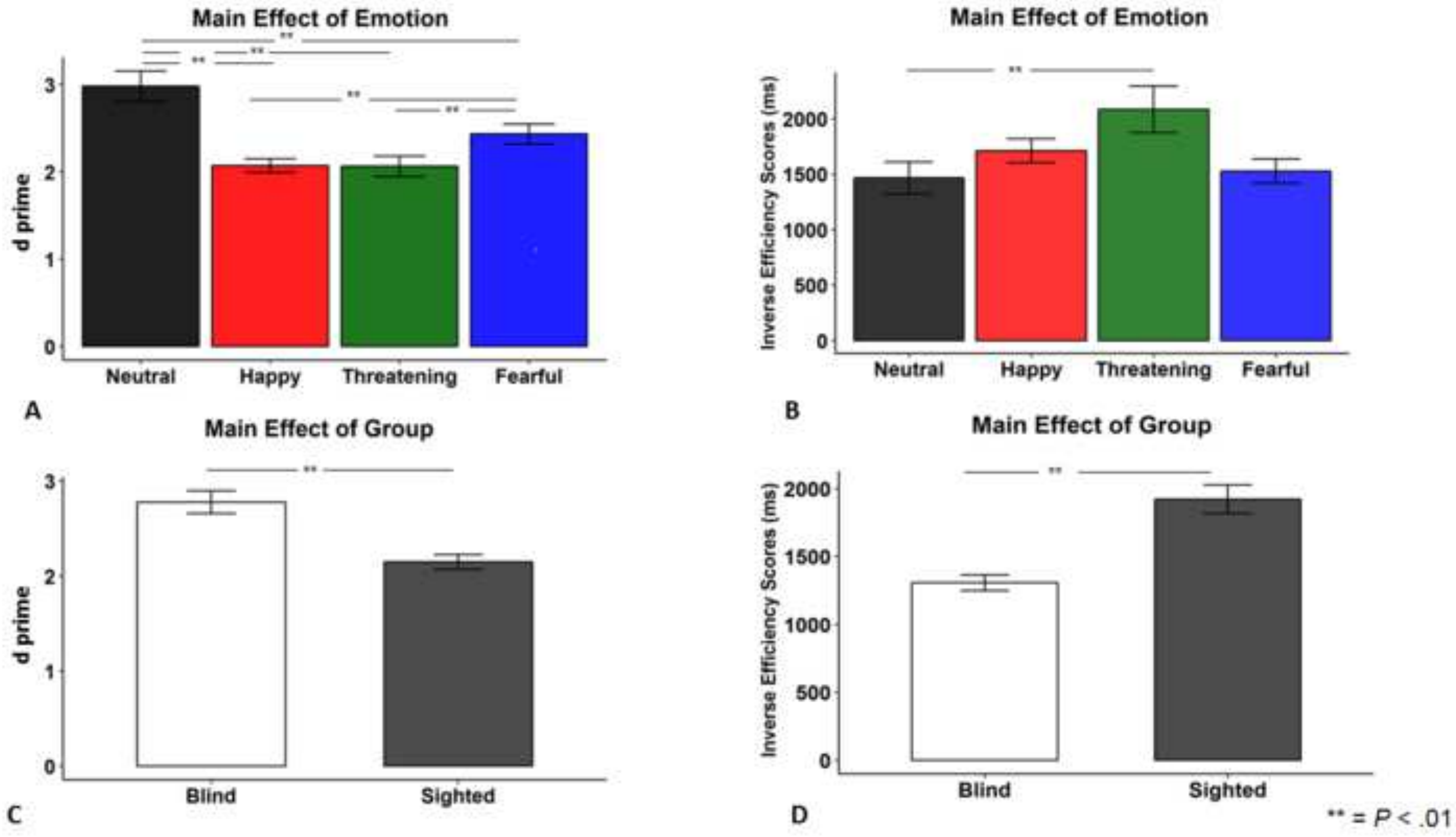
Main Effect of Emotion
**
**
**
**
**
d prime
3
2
1
0
Neutral
Happy
Threatening
Fearful
A
Main Effect of Emotion
**
Inverse Efficiency Scores (ms)
2000
1500
1000
500
0
Neutral
Happy
Threatening
Fearful
B
Main Effect of Group
**
d prime
3
2
1
0
Blind
Sighted
C
Main Effect of Group
**
Inverse Efficiency Scores (ms)
2000
1500
1000
500
0
Blind
Sighted
D
** = P < .01

Figure3

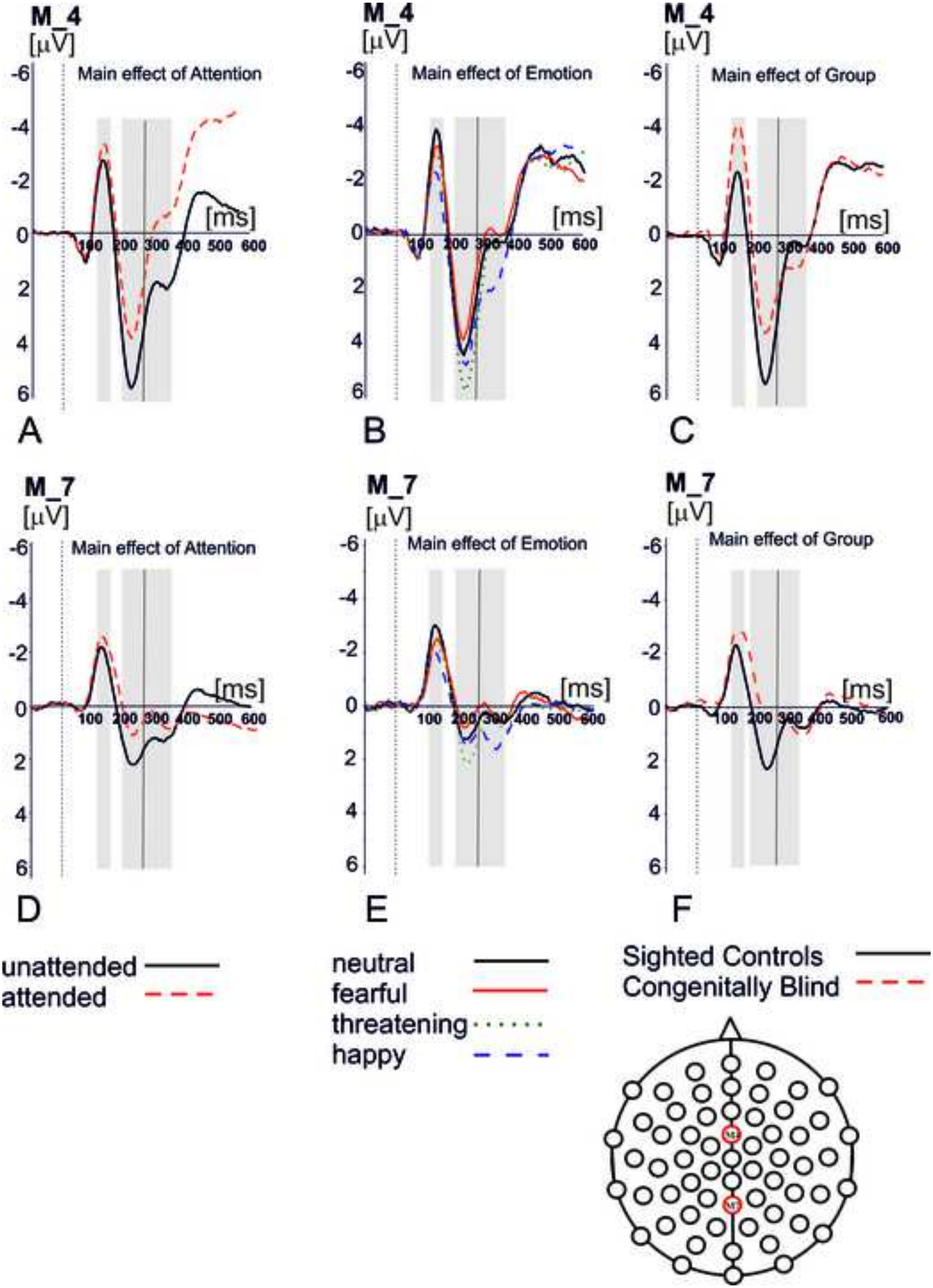
M_4
[μV]
Main effect of Attention
Main effect of Emotion
Main effect of Group
-6
-4
-2
0
2
4
6
[ms]
100 200 300 400 500 600
A
B
C
M_7
[μV]
Main effect of Attention
Main effect of Emotion
Main effect of Group
D
E
F
unattended
attended
neutral
fearful
threatening
happy
Sighted Controls
Congenitally Blind
M4
M7

Figure4

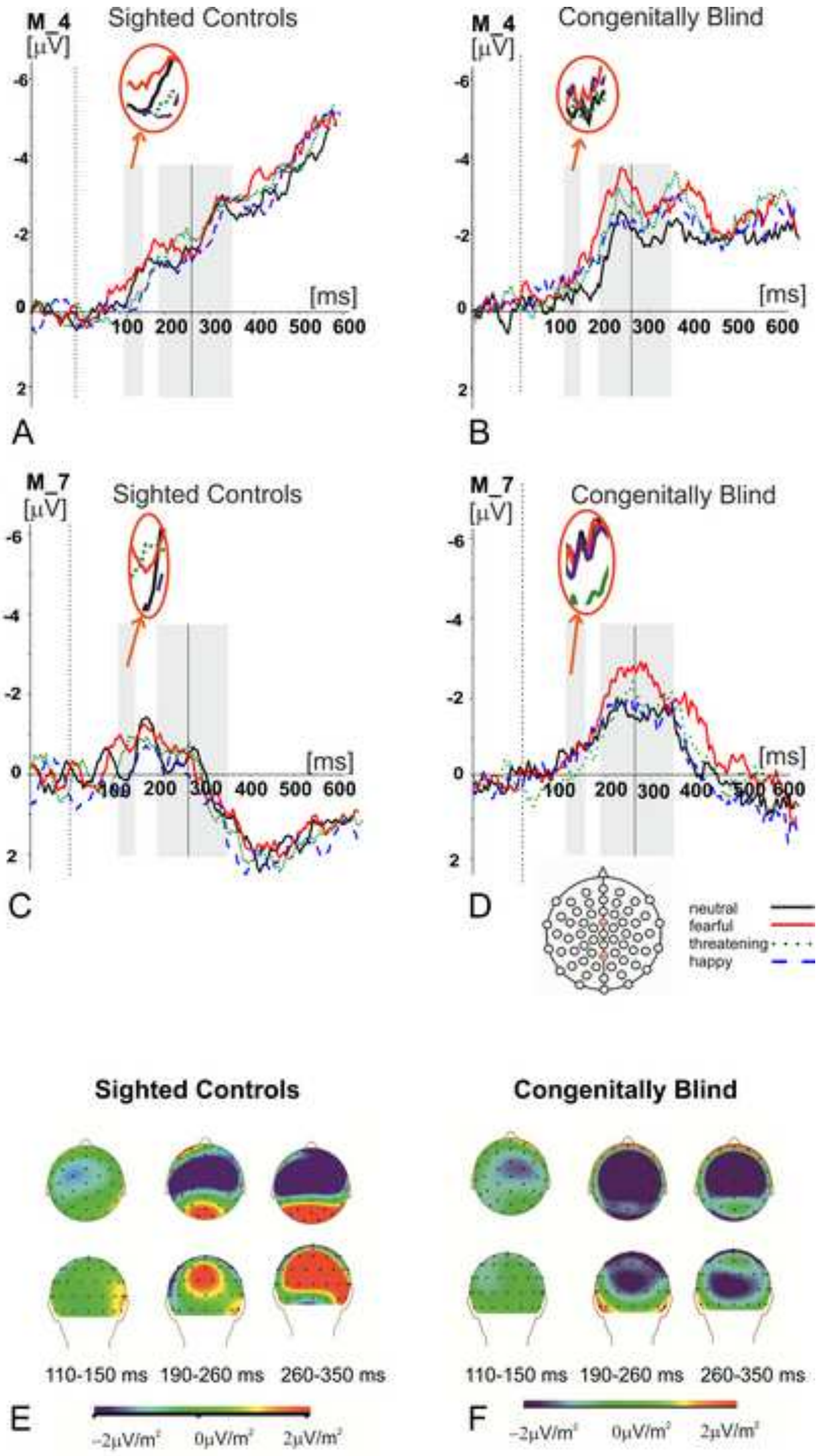
M_4
[μV]
Sighted Controls
-6
-4
-2
0
2
[ms]
100 200 300 400 500 600
A
M_4
[μV]
Congenitally Blind
-6
-4
-2
0
2
[ms]
100 200 300 400 500 600
B
M_7
[μV]
Sighted Controls
-6
-4
-2
0
2
[ms]
100 200 300 400 500 600
C
M_7
[μV]
Congenitally Blind
-6
-4
-2
0
2
[ms]
100 200 300 400 500 600
D
neutral
fearful
threatening
happy
Sighted Controls
110-150 ms
190-260 ms
260-350 ms
E
-2μV/m²
0μV/m²
2μV/m²
Congenitally Blind
110-150 ms
190-260 ms
260-350 ms
F
-2μV/m²
0μV/m²
2μV/m²

Figure5

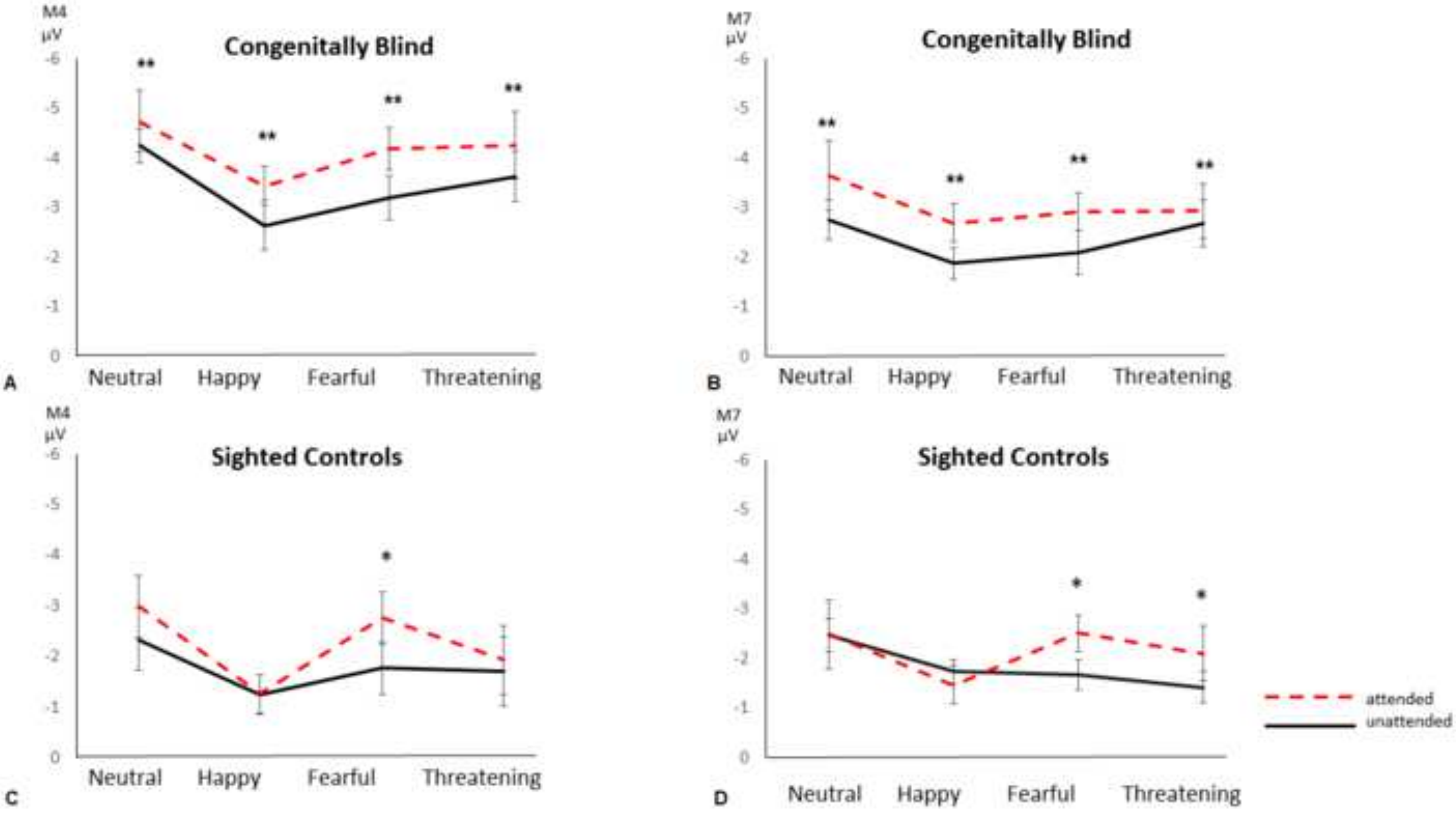
M4
µV
Congenitally Blind
-6
-5
-4
-3
-2
-1
0
**
**
**
**
Neutral
Happy
Fearful
Threatening
A
M7
µV
Congenitally Blind
-6
-5
-4
-3
-2
-1
0
**
**
**
**
Neutral
Happy
Fearful
Threatening
B
M4
µV
Sighted Controls
-6
-5
-4
-3
-2
-1
0
*
Neutral
Happy
Fearful
Threatening
C
M7
µV
Sighted Controls
-6
-5
-4
-3
-2
-1
0
*
*
Neutral
Happy
Fearful
Threatening
D
attended
unattended

Table1

Table 1. Mean reaction times (ms) and mean accuracy (%) for each emotional prosody (neutral, happy, threatening, fearful) and Group (Blind, Sighted) with standard errors of the mean.

	Emotions								**Group**			
	Neutral		**Happy**		**Threatening**		**Fearful**		**Blind**		**Sighted**	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Percent Correct (%)	79	4	56	2	58	4	65	3	75	1	57	3
Reaction Times (ms)	1059	20	910	24	1064	19	927	18	956	26	1010	24
D-prime	3.089	.15	2.135	.057	2.125	.108	2.512	.095	2.7	.113	2.1	.88
IE scores	1405	141	1639	92	1997	206	1459	93	1328	164	1922	12

Table2

Table 2. Auditory ERPs

Factors				**Time Epoch**					
	110-150 ms			**190-260 ms**			**260-350 ms**		
	F	**p**	**η^2**	**F**	**p**	**η^2**	**F**	**p**	**η^2**
(A) Across Participants									
Overall ANOVA									
Group	5.20	0.034	0.215	2.51	0.130	0.117	0.08	0.771	0.005
Emotion	17.94	< .001	0.486	25.76	< .001	0.576	22.55	< .001	0.543
Emotion ✻ Group	1.42	0.251	0.070	1.40	0.255	0.069	0.34	0.778	0.018
Attention	29.40	< .001	0.608	37.97	< .001	0.667	31.6	< .001	0.625
Attention ✻ Group	2.22	0.152	0.105	4.74	0.042	0.200	2.35	0.141	0.110
Electrode	2.85	0.107	0.131	33.54	< .001	0.638	0.66	0.424	0.034
Electrode ✻ Group	2.62	0.122	0.121	0.06	0.798	0.004	0.20	0.654	0.011
Emotion ✻ Attention	1.73	0.184	0.084	2.47	0.078	0.115	0.90	0.429	0.045
Emotion ✻ Attention ✻ Group	1.47	0.239	0.072	0.49	0.680	0.024	0.80	0.474	0.041
Emotion ✻ Electrode	9.21	< .001	0.327	22.42	< .001	0.541	9.90	< .001	0.343
Emotion ✻ Electrode ✻ Group	0.09	0.933	0.005	3.83	0.019	0.168	1.83	0.161	0.088
Attention ✻ Electrode	1.20	0.287	0.059	6.66	0.018	0.260	19.73	< .001	0.510
Attention ✻ Electrode ✻ Group	0.59	0.449	0.031	1.97	0.176	0.094	10.36	0.005	0.353
Emotion ✻ Attention ✻ Electrode	0.27	0.784	0.014	1.11	0.348	0.055	0.70	0.532	0.036
Emotion ✻ Attention ✻ Electrode ✻ Group	6.25	0.003	0.248	0.05	0.976	0.003	0.42	0.706	0.022
Electrode M4									
Group	5.01	0.037	0.209	1.46	0.242	0.071	0.144	0.709	.008
Emotion	20.64	<.001	0.521	31.93	<.001	0.627	23.25	< .001	0.550
Emotion ✻ Group	1.15	0.331	0.057	2.34	0.103	0.110	0.770	0.510	0.039
Attention	21.85	<.001	0.535	38.44	<.001	0.669	67.96	< .001	0.782
Attention ✻ Group	0.836	0.372	0.042	1.92	0.181	0.092	0.01	0.892	0.001
Emotion ✻ Attention	1.837	0.166	0.088	2.41	0.86	0.113	0.70	0.527	0.036
Emotion ✻ Attention ✻ Group	1.15	0.331	0.057	0.461	0.684	0.024	0.58	0.594	0.030
Electrode M7									
Group	2.22	0.120	0.105	3.76	0.067	0.165	0.006	0.939	0.000
Emotion	11.76	<.001	0.382	14.82	< .001	0.438	15.718	< .001	0.453
Emotion ✻ Group	1.45	0.246	0.071	0.52	0.640	0.027	0.11	0.934	0.006
Attention	31.25	<.001	0.622	26.93	< .001	0.586	6.841	0.017	0.265
Attention ✻ Group	4.27	0.053	0.183	7.71	0.012	0.289	6.36	0.021	0.251
Emotion ✻ Attention	1.89	0.152	0.090	0.82	0.484	0.41	0.970	0.403	0.049
Emotion ✻ Attention ✻ Group	2.97	0.048	0.135	0.33	0.794	0.017	0.87	0.443	0.044
(B) Blind Participants									
Overall ANOVA									
Emotion	4.71	0.028	0.403	7.347	0.005	0.559	5.24	0.012	0.429
Attention	18.12	0.004	0.721	20.34	0.003	0.744	27.62	0.001	0.798
Electrode	14.68	0.006	0.677	11.36	0.012	0.619	0.67	0.438	0.088
Emotion ✻ Attention	0.49	0.611	0.066	1.95	0.181	0.242	0.857	0.452	0.109
Emotion ✻ Electrode	4.64	0.026	0.399	7.25	0.004	0.509	1.4	0.278	0.167
Attention ✻ Electrode	0.04	0.841	0.006	0.54	0.484	0.072	0.74	0.416	0.097
Emotion ✻ Attention ✻ Electrode	2.34	0.134	0.252	0.620	0.538	0.081	0.797	0.464	0.102
Electrode M4									
Emotion	5.62	0.015	0.446	8.85	0.002	0.558	5.21	0.009	0.427
Attention	10.84	0.013	0.608	20.39	0.003	0.744	36.72	< .001	0.840
Emotion ✻ Attention	0.50	0.611	0.067	1.92	0.194	0.216	0.61	0.575	0.081
Electrode M7									
Emotion	3.44	0.064	0.330	4.20	0.034	0.375	4.25	0.034	0.378
Attention	28.79	0.001	0.804	16.07	0.005	0.697	13.60	0.008	0.660
Emotion ✻ Attention	0.89	0.424	0.113	1.55	0.242	0.181	1.15	0.349	0.141

(C) Sighted Participants									
Overall ANOVA									
Emotion	20.65	< .001	0.633	24.33	< .001	0.670	24.20	< .001	0.669
Attention	10.57	0.007	0.468	13.37	0.003	0.528	9.565	0.009	0.444
Electrode	0.002	0.962	0.000	25.34	< .001	0.679	0.083	0.777	0.007
Emotion ✻ Attention	3.02	0 .040	0.221	0.828	0.468	0.065	0.812	0.480	0.063
Emotion ✻ Electrode	5.78	0.008	0.325	23.83	< .001	0.665	13.34	< .001	0.547
Attention ✻ Electrode	2.25	0.159	0.158	10.58	0.007	0.469	34.64	< .001	0.743
Emotion ✻ Attention ✻ Electrode	4.92	0.019	0.291	0.619	0.567	0.049	0.14	0.886	0.012
Electrode M4									
Emotion	21.27	< .001	0.639	32.508	< .001	0.730	24.998	< .001	0.676
Attention	10.26	0.008	0.461	16.552	0.002	0.580	39.684	< .001	0.768
Emotion ✻ Attention	3.12	0.050	0.206	0.949	0.414	0.073	0.586	0.585	0.047
Electrode M7									
Emotion	12.84	< .001	0.517	13.43	< .001	0.528	15.04	< .001	0.556
Attention	7.39	0.001	0.804	5.84	0.032	0.328	0.005	0.944	0.000
Emotion ✻ Attention	4.07	0.018	0.253	0.598	0.601	0.047	0.747	0.499	0.059

ANOVA results **(a)** across participants including the factors **Attention** (attended vs unattended), **Emotion** (happy, neutral, threatening, ferarful), **Electrode** (M4 vs. M7), and **Group** (Blind vs Sighted) as well as separate analysis of M4 and M7 electrodes;
(b) Analysis for Blind; **(c)** Analysis for Sighted. The results are depicted separately for all three time epochs.

Table 3: Item statistics: Mean (M) and Standard error of the mean (SE) of duration, pitch, intensity, valence ratings, dominance ratings and arousal ratings of standard stimuli in the different emotional prosodies merged across the voices of the two actors.

Emotional prosody	**Duration (ms)**		**Pitch (Hz)**		**Intensity * (dB)**	
	M	SE	M	SE	M	SE
neutral	632	35	176	3.78	62.31	0.01
happy	575	57	271	17.45	62.25	0.001
threatening	602	56	244	8.78	62.15	0.19
fearful	518	44	252	11.24	62.25	0.02
	Valence rating (1–7)		**Dominance rating (1–7)**		**Arousal rating (1–7)**	
	M	SE	M	SE	M	SE
neutral	4.77	0.27	4.29	0.13	4.59	0.43
happy	5.37	0.08	4.56	0.12	4.84	0.32
threatening	1.92	0.13	6.30	0.07	4.81	0.38
fearful	2.76	0.14	2.23	0.14	4.63	0.21

Table4

Table 4: Item statistics: Mean (M) and Standard error of the mean (SE) of duration, pitch, intensity, valence ratings, dominance ratings and arousal ratings of deviant stimuli in the different emotional prosodies merged across the voices of the two actors.

Emotional Prosody	**Duration (ms)**		**Pitch (Hz)**		**Intensity * (dB)**	
	M	SE	M	SE	M	SE
neutral	731	60	186	2.69	62.21	0.01
happy	490	42	363	14.57	62.36	0.01
threatening	721	96	297	10.14	62.30	0.03
fearful	426	44	312	14	62.5	0.06
Emotional Prosody	**Valence rating (1–7)**		**Dominance rating (1–7)**		**Arousal rating (1–7)**	
	M	SE	M	SE	M	SE
neutral	4.83	0.51	4.50	0.15	3.98	0.50
happy	5.17	0.09	4.43	0.17	5.06	0.43
threatening	1.74	0.15	6.51	0.10	5.53	0.10
fearful	2.60	0.14	1.97	0.12	4.56	0.18

***Credit Author Statement**

Author contributions

PT: Analyzing the data, writing and editing the paper. **AZ:** Analyzing the data, editing and writing the paper. **JG:** Designing the experiment**,** running the EEG experiment, editing and writing the paper, **JF:** Analyzing the data, editing and writing the paper.