Journal of Deaf Studies and Deaf Education, 2020, 10–21

doi: 10.1093/deafed/enz036
Advance Access Publication Date: 29 October 2019
Empirical Manuscript

EMPIRICAL MANUSCRIPT

Attention Dynamics During Emotion Recognition by Deaf and Hearing Individuals

Izabela Krejtz[1,*,†], Krzysztof Krejtz[1], Katarzyna Wisiecka[2], Marta Abramczyk[3], Michał Olszanowski[1] and Andrew T. Duchowski[4]

[1]SWPS University of Social Sciences and Humanities, Chodakowska 19/31, Warsaw, Poland [2]Polish–Japanese Academy of Information Technology [3]Polish Association of the Deaf and [4]Clemson University

*Correspondence should be sent to Izabela Krejtz, SWPS University of Social Sciences and Humanities, Warsaw, Poland (e-mail: ikrejtz@swps.edu.pl)

Abstract

The enhancement hypothesis suggests that deaf individuals are more vigilant to visual emotional cues than hearing individuals. The present eye-tracking study examined ambient–focal visual attention when encoding affect from dynamically changing emotional facial expressions. Deaf ($n = 17$) and hearing ($n = 17$) individuals watched emotional facial expressions that in 10-s animations morphed from a neutral expression to one of happiness, sadness, or anger. The task was to recognize emotion as quickly as possible. Deaf participants tended to be faster than hearing participants in affect recognition, but the groups did not differ in accuracy. In general, happy faces were more accurately and more quickly recognized than faces expressing anger or sadness. Both groups demonstrated longer average fixation duration when recognizing happiness in comparison to anger and sadness. Deaf individuals directed their first fixations less often to the mouth region than the hearing group. During the last stages of emotion recognition, deaf participants exhibited more focal viewing of happy faces than negative faces. This pattern was not observed among hearing individuals. The analysis of visual gaze dynamics, switching between ambient and focal attention, was useful in studying the depth of cognitive processing of emotional information among deaf and hearing individuals.

Decoding facial expressions is a valuable source of information about another's behavior and intentions toward us (Haxby, Hoffman, & Gobbini, 2000). Typically, people effectively differentiate between emotions, with only a few fixations on the face needed to recognize expressions (Schurgin, Nelson, Iida, Ohira, Chiao, & Franconeri, 2014). Both visual and auditory signals of emotion are typically used to decode emotions (e.g., Hopyan-Misakyan, Gordon, Dennis, & Papsin, 2009); however, deaf people may effectively decode emotions with limited access to auditory signals (Jones, Gutierrez, & Ludlow, 2018; Letourneau & Mitchell, 2011; Sidera, Amadó, & Martínez, 2017).

Deaf individuals may have developed greater visual sensitivity to facial expressions, and as a result they may use different strategies of visual inspection of faces during emotion recognition. This is known as the enhancement hypothesis (Sidera et al., 2017), which also holds that deaf people rely on visual cues as a basis for interpreting another person's state of mind, and they may recognize facial emotions more effectively than hearing individuals. For example, deaf British Sign Language users outperformed hearing signers and non-signers in the faces memory and recognition task (Arnold & Murray, 1998). In line with this result McCullough and Emmorey (1997) indicate that deaf American Sign Language users have higher ability to detect subtle differences in facial features.

In contrast to the enhancement hypothesis related to sensory deprivation (Merabet & Pascual-Leone, 2010), the deficit

†Izabela Krejtz, http://orcid.org/0000-0002-9827-8371
Received December 10, 2018; revisions received July 11, 2019; accepted August 1, 2019

hypothesis (Sidera et al., 2017) holds that, due to the lack of auditory cues and opportunities to hear dialogue regarding emotion, deaf people may have difficulties in facial emotion recognition (Dyck, Farrugia, Shochet, & Holmes-Brown, 2004).

Supporting evidence for the enhancement hypothesis also comes from a growing number of studies suggesting that sensory deprivation is related to sensory cross-modal neuroplastic changes in the brain (e.g., Armstrong, Neville, Hillyard, & Mitchell, 2002; Arnold & Murray, 1998; McCullough & Emmorey, 1997; Merabet & Pascual-Leone, 2010; Pavani & Bottari, 2012). These studies show that brain regions typically linked with the lost modality receive some stimulation from the remaining sensory modalities. This cross-modal recruitment suggests compensatory mechanisms in deaf individuals.

The enhancement of visual attention is not general but rather it refers to enhancement of selected aspects of visual perception, especially peripheral vision and bottom-up attention triggered by exogenous cues (Pavani & Bottari, 2012). Several studies indicate that deaf people are characterized by enhanced selective spatial attention (Dye, Hauser, & Bavelier, 2009; Pavani & Bottari, 2012) and a wider perceptual span (Stevens & Neville, 2006). For example, Armstrong et al. (2002) showed that deaf individuals are more reactive to motion in their peripheral field of vision compared to hearing participants, whereas there were no group differences in reactions to color stimuli in the periphery.

The present study aims to compare recognition strategies underlying visual decoding of dynamic facial expression of emotions between deaf and hearing people. A significant contribution of the presented study is the use of eye-tracking indicators of visual attention to interpret attentional strategies during visual processing of facial stimuli. The study hypotheses are stated after a review of relevant literature.

Background

Recognition of Emotional Facial Expressions Among Deaf Individuals

Empirical evidence concerning visual attention among deaf people is mixed. On the one hand several studies support the deficit hypothesis indicating poorer recognition of facial emotions among deaf and hard-of-hearing (DHH) individuals compared to controls (Dyck et al., 2004; Most & Michaelis, 2012; Wang, Su, Fang, & Zhou, 2011). For example, Most and Michaelis (2012) showed that accuracy of emotion perception was higher among children with typical hearing compared to children with hearing loss in three conditions: with auditory, visual, and auditory–visual cues. Differences between hard-of-hearing and hearing children in facial expression recognition were also demonstrated in adolescence and early adulthood, but to a lower degree (Dyck et al., 2004).

On the other hand, other studies provide evidence of similar results in hearing and deaf or hard-of-hearing (D/HH) children's capacity to recognize facial emotions (Hao & Su, 2014; Hosie, Gray, Russell, Scott, & Hunter, 1998; Sidera et al., 2017; Ziv, Most, & Cohen, 2013). Hosie et al. (1998) showed comparable levels of performance for deaf and hearing children in labeling static facial expressions of emotion.

The observed variability in findings suggest either deficit or enhancement in emotion recognition may stem from methodological nuances, for example, age of participants, level of hearing loss, and experimental task. First, in line with the assumption that delay in emotion recognition is linked with language development (Dyck et al., 2004; Sidera et al., 2017), studies with adults did not show behavioral deficit in emotion recognition among deaf adults such as decreased speed or lower accuracy of emotion recognition compared to hearing adults. For example, Watanabe, Matsuda, Nishioka, and Namatame (2011) did not observe differences in emotion recognition accuracy between deaf and hearing adults. However, using eye tracking, the authors indicated between-group differences in visual decoding of emotions. Deaf participants allocated their attention for a longer duration and more frequently to the eyes than hearing participants who looked at the nose during emotion recognition.

Second, a crucial factor for emotion recognition is the level of hearing loss. Exemplar studies examined children or adults with different rate of hearing status: deaf individuals (e.g., Sidera et al., 2017); children with moderate, severe, and profound hearing loss (e.g., Most & Michaelis, 2012); and children with cochlear implants or hearing aids (e.g., Wang et al., 2011). For example, Most and Michaelis (2012) observed the lowest performance in emotion recognition among children with profound hearing loss. Groups with moderate and moderate-to-severe hearing loss did not significantly differ in emotion-perception ability from children with typical hearing.

Third, differences in stimulus presentation and instructions may influence where and how attention is allocated during visual tasks (Boot, Ensar, & Kramer, 2009). For example, Wang et al. (2011) who supported the deficit hypothesis used a visual search task in which children searched for the exact expression among five different simultaneously displayed facial expressions. Whereas Hosie et al. (1998), who adopted the enhancement hypothesis, asked children to recognize sequentially presented facial expressions.

Fourth, Anastasi and Rhodes (2005) indicated that emotional facial expressions of a person about the same age as the participant were recognized with a higher accuracy than when there was a mismatch in age between participants and people in the presented photographs. For example, Dyck et al. (2004) who supported the deficit hypothesis among hard-of-hearing children displayed facial expressions of adult males and females in the emotion recognition task. On the other hand, Ziv et al. (2013) who supported the enhancement hypothesis used a set of photographs of boys and girls in the same age group as the participants.

Fifth, some studies suggest that the differences in emotion recognition might depend on the intensity of facial expressions. For example, deaf children, aged from 7 to 13, with a cochlear implant, were better than hearing children at recognizing emotions from low-intensity photographs of angry and scared faces (Hopyan-Misakyan et al., 2009). Similarly, when children were asked to name emotional expressions of photographed fearful and disgusted faces, deaf children were more accurate than hearing children (Hosie et al., 1998).

Sixth, support for the deficit or enhancement hypotheses may depend on the type of examined emotions. For example, Sidera et al. (2017) showed a delay in the ability of deaf children to recognize frightened, surprised, and disgusted expression. However, deaf children recognized happy, sad, and angry expressions with a similar accuracy to hearing children. Wang et al. (2011) found that for both deaf and hearing children happiness was easier to recognize than anger and fear. This finding is consistent with so-called "happiness superiority effect," a phenomenon frequently observed in research on emotion recognition among hearing individuals, that is described as the tendency to prioritize positivity (Anderson, Christoff, Panitz, De Rosa, & Gabrieli, 2003; Craig, Becker, & Lipp, 2014).

Finally, mixed results for emotion recognition abilities may be at least partially determined by the character of stimuli used in the recognition task. It should be noted that almost all of the aforementioned research used static pictures. However, temporal patterns of facial expressions are useful cues for distinguishing emotions (Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004). For example, Becker et al. (2012) conducted a series of studies with hearing adults using dynamical presentation of faces which were transforming from neutral to angry or happy expressions. They found that “becoming-happy” faces were recognized faster and with higher accuracy. Contrary, the recognition of “becoming angry” faces required more time which suggests that cognitive processing during threat detection occurs at later stages of emotion recognition.

Dynamic Emotion Recognition

Deaf people use their facial expressions to emphasize lexical indications of certain emotions and linguistic information which, in spoken languages, are typically conveyed by voice (e.g., Corina, Bellugi, & Reilly, 1999). Deaf individuals might have developed unique encoding mechanisms to maximize their ability to absorb facial expressions for both affective and linguistic input during conversation (Letourneau & Mitchell, 2011). Letourneau and Mitchell (2011) observed in an eye-tracking study that deaf individuals tend to pay an extended attention to facial expressions compared to hearing individuals. Specifically, deaf adults (congenital or early deafened signers) devoted more fixations (but not first fixations) to the mouth area of still images than hearing adults did. The authors explained this effect suggesting that facial expressions convey the majority of affective information obtained from others in the absence of auditory cues (e.g., Sidera et al., 2017). This interpretation is in line with Emmorey’s and McCullough’s findings (2009) who demonstrated that although deaf adults keep their fixations on the face of their interlocutor, they are able to simultaneously perceive manual signs and facial expressions.

Facial expressions of basic emotions (happy, sad, fear/surprise, disgust/anger) are produced with constriction of different facial muscles (Jack, Garrod, & Schyns, 2014), depending on the emotional expression (Smith, Cottrell, Gosselin, & Schyns, 2005) which evolve in specific sequences over time. For example, a happy facial expression is recognized mainly by a smile which is produced by a constriction of the musculus risorius muscle. Temporal patterns of facial expressions are useful cues for distinguishing emotions (Sato et al., 2004; Smith et al., 2005), which evolve in specific sequences over time. Therefore, successful emotion categorization requires dynamical scanning of facial regions responsible for emotion production, for example, the mouth for happiness recognition (Delis et al., 2016). A review based on research with hearing adults shows that there is a dynamic advantage in emotion recognition of facial expressions, as dynamic information improves identification of more subtle facial changes (Krumhuber, Kappas, & Manstead, 2013).

Jones et al. (2018) did not observe any differences in emotion recognition from dynamic pictures between hearing and hard-of-hearing children. Their performance was similar even when movement was displayed at low levels of intensity. However, when static pictures were presented, the hard-of-hearing children performed worse than their hearing peers. This suggests a compensatory role of motion in emotion recognition among deaf children. In sum, research suggests that the use of dynamical stimuli, even with a minimal amount of movement, improves performance of deaf participants in recognizing emotion.

Eye Tracking and Visual Gaze Dynamics

Eye tracking has become an important method to study visual attention and the correspondence between attention and behavior (Duchowski, 2007). Recording of eye movements provides a clear indication of overt gaze orientation during emotion recognition. When monitoring gaze, we assume that direction of gaze corresponds to visual attention (the eye-mind hypothesis; Just & Carpenter, 1984).

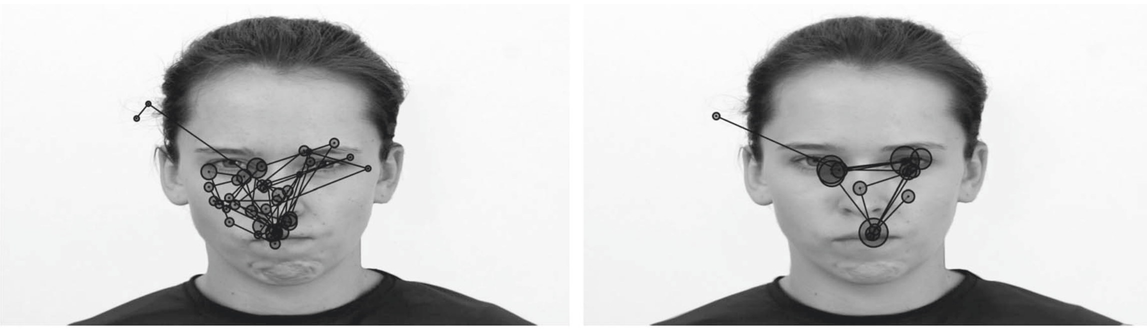
The process of viewing any stimuli is an interplay between fixations and saccades. Fixations can be understood as relative stops between two consecutive saccades, which are rapid eye movements. Typical duration of fixations is between 150 and 300 ms; however, shorter and longer fixations have been observed (Rayner, 1998). Saccade amplitude ranges between 20 and 200 ms (Fischer & Ramsperger, 1984). We acquire new information during fixations.

In terms of eye movement parameters, short fixations followed by long saccades are characteristic of ambient attention (Figure 1a), while longer fixations followed by shorter saccades are indicative of focal attention (Unema, Pannasch, Joos, & Velichkovsky, 2005) (Figure 1b). Focal attention is often treated as an indicator of deep information processing, whereas ambient attention indicates stimuli exploration (Krejtz, Duchowski, Krejtz, Szarkowska, & Kopacz, 2016; Strauch, Huckauf, Krejtz, & Duchowski, 2018). Ambient information processing is typical for visual exploration (Krejtz et al., 2016; Strauch et al., 2018), indicating a relatively fast scanning of low-level characteristics of visual stimuli (Posner, 1980).

The interplay between focal and ambient modes of visual information processing changes dynamically. At early stages of scene perception, shorter fixations and longer saccades appear to explore the scene. Once a target has been identified, longer fixations are followed by shorter saccades suggesting a change to a focal mode of processing (Irwin & Zelinsky, 2002; Velichkovsky, Joos, Helmert, & Pannasch, 2005).

At early stages of viewing time, people tend to exhibit more ambient processing while searching for distinctive aspects of the stimulus. During this time, attention is likely guided by stimuli characteristics. In later stages gaze characteristics become more focal. During this time attention tends to be internally motivated (Krejtz, Szarkowska, Krejtz, Walczak, & Duchowski, 2012; Velichkovsky et al., 2005). For example, Krejtz et al. (2018) measured ambient and focal attention during recognition of facial emotional expressions among socially anxious and non-anxious hearing participants. Their results indicated that recognition of negative emotions was less accurate, required more time, and their processing was more ambient. Anxious individuals exhibited more ambient processing of facial expressions than non-anxious individuals. Greater ambient processing was interpreted as an indicator of emotional and cognitive withdrawal during exploration of facial expression (Krejtz et al., 2018).

A mathematical expression of ambient–focal attention, relating fixation duration and saccadic amplitude immediately following a fixation has been proposed for the analysis of static and dynamic viewing (Krejtz et al., 2016). The $\mathcal{K}$ coefficient (based on the relationship between fixation duration and amplitude of subsequent saccade) was used as the indicator of gaze dynamics shifting between ambient–focal attention (Krejtz et al., 2016; Krejtz, Çöltekin, Duchowski, & Niedzielska, 2017). The $\mathcal{K}$ coefficient combines fixations and saccades into a single expression capturing the dynamic interplay of ambient and focal attention. The $\mathcal{K}$ coefficient is computed for each participant and



(a) Ambient attention (b) Focal attention

Figure 1 Exemplary scanpaths of recorded eye movements while viewing angry facial expression.

experimental trial as

$$\mathcal{K}_{\rangle} = \frac{d_i - \mu_d}{\sigma_d} - \frac{a_{i+1} - \mu_a}{\sigma_a}, \mathcal{K} = \frac{1}{n}\sum_n \mathcal{K}_i$$

where d_i and a_{i+1} are fixation duration and consecutive saccade amplitude; μ_d and μ_a are mean fixation duration and saccade amplitude, respectively; and $\sigma_d$and $\sigma_a$are the fixation duration and saccade amplitude standard deviations (SDs), respectively, computed over all n fixations and hence $n\mathcal{K}_{\rangle}$ coefficients. The $\mathcal{K}$ coefficient is expressed in SD units. For example, $\mathcal{K} = -1$SD indicates that, in the current moment of viewing, standardized fixation duration is 1 SD smaller than the standardized saccade amplitude. Consequently, $\mathcal{K} = +1$SD indicates that fixation duration is 1 SD greater than the standardized saccade amplitude. Thus, $\mathcal{K} < 0$ is interpreted as manifest of ambient attention, and $\mathcal{K} > 0$ is interpreted as the focal mode of information processing. Coefficient $\mathcal{K}$ is unbounded (has no lower or upper limit), but its distribution does not deviate from the normal distribution (Krejtz et al., 2016).

The $\mathcal{K}$ ambient/focal coefficient fosters statistical comparison between individuals and between groups. It has been validated in a parallel and serial search study using abstract stimuli (Krejtz et al., 2016) and also in a dynamic emotion recognition study among socially anxious individuals (Krejtz et al., 2018). Previous research has demonstrated the usefulness of $\mathcal{K}$ in capturing attentional dynamics of visual information processing during different cognitive and emotional tasks, such as map reading (Krejtz et al., 2017), multimedia learning (Krejtz et al., 2012), decision making (Duchowski et al., 2019), or emotion recognition (Krejtz et al., 2018).

The Present Study

The present study investigates group differences among deaf and hearing individuals in dynamics of ambient–focal attention during recognition of dynamical facial expressions. The first hypothesis predicts that during the early stages of facial expression recognition, all participants would exhibit more ambient attention compared to later stages of emotion recognition. In later stages of emotion recognition (before a decision is made), it is hypothesized that both groups will switch to more focal attention, with effects alterable by emotional expression and group.

The second prediction expects a preference for recognition of happy faces supporting the happiness superiority effect (Anderson et al., 2003; Craig et al., 2014), among both deaf and hearing participants. It is hypothesized that happy faces will be recognized with greater accuracy and with eye movement indices indicating deeper processing of happy faces than negative facial expressions (longer fixations and more focal attention to happy faces than negative ones). This prediction is based on previous findings observed in a group of socially anxious people (Krejtz et al., 2018). Krejtz et al. showed that recognition of negative emotional expressions was less accurate and less efficient compared to recognition of positive faces, as visual scanning during recognition of negative faces was more ambient. This effect was the most pronounced among socially anxious participants when looking at facial expressions of anger. We may assume that more ambient attention while looking at negative emotional facial expressions suggests less attentive processing whereas more focal attention to positive facial expressions suggests more attentive processing. Greater focal attention to happy rather than negative facial expressions supports the happiness superiority effect.

Our third hypothesis predicts that deaf and hearing individuals will differ in visual scanning of faces for signs of emotional expressions. Due to some inconsistencies in the findings of behavioral studies comparing emotion recognition among deaf and hearing people, suggesting similarities (e.g., Hao & Su, 2014), or supporting either enhancement (e.g., Arnold & Murray, 1998) or deficit hypotheses (e.g., Gray, Hosie, Russell, & Ormel, 2001), we examined group differences on an exploratory basis. We believe that eye tracking will help to clarify the nature of potential differences by explaining the complexity of the emotion recognition processes.

Finally, successful emotion recognition requires dynamical scanning of facial regions responsible for emotion production what can expedite transitioning into focal attention. Each stimulus is explored in the ambient mode at the beginning regardless of individual differences. A wider perceptual span can shorten the process of exploration (Stevens & Neville, 2006). As deaf individuals are more reactive to motion in the peripheral field (Armstrong et al., 2002), we predict that they are faster in finding crucial signals leading to a faster switching into focal attention comparing to hearing individuals. As a consequence, they will distribute their attention differently than hearing participants over relevant regions of interests (e.g., eye and mouth; Letourneau & Mitchell, 2011).

In summary, our study offers several advantages over previous investigations of emotion recognition among deaf participants. First, we use eye tracking to capture visual attention during emotion recognition. Registering paths of eye movements allows for a thorough examination of basic attentional processes that cannot be recorded by behavioral studies. We are aware of only two studies that incorporated eye tracking to study emotion recognition among adult deaf individuals (Letourneau & Mitchell, 2011; Watanabe et al., 2011); however, they did not analyze the process of gaze dynamics obtained during display of video stimuli. Second, we apply a measure of eye movement dynamics, namely $\mathcal{K}$ coefficient, allowing examination of switching between ambient and focal attention. Finally, earlier studies of emotion perception among DHH used static stimuli (e.g., Sidera et al., 2017). We believe that dynamical facial expressions (short video clips) are more ecologically valid reflecting the natural condition of emotion recognition.

Method

Participants

Thirty-four white Caucasian individuals (11 females), aged between 17 and 28 years old, ($M = 20.53$; $SD = 2.05$) participated in the study. There were 17 deaf participants recruited from the Institute for the Deaf in Warsaw. Participants were matched in age and gender with 17 hearing participants who did not declare any hearing loss, recruited from the SWPS University of Social Sciences and Humanities in Warsaw. The deaf and control groups were similar in age ($M_{deaf} = 20.18$; $SD = 1.63$; $M_{hearing} = 20.88$; $SD = 2.39$; $t[28.20] = 1.00$; $p = .324$) and gender (number of females: $n_{deaf} = 5$; $n_{hearing} = 6$; $\chi^2(1) = .13$; $p = .714$). All participants had normal or corrected to normal vision.

Among the deaf participants, six declared profound hearing loss (90 dB), nine declared severe hearing loss (70–90 dB), and two declared moderate hearing loss (60 dB). Ten participants were deaf at birth, and 7 lost their hearing at the age of 6 or earlier. The etiology of hearing loss was not collected. Thirteen participants had both hearing parents, 2 had both deaf parents, and 2 had one deaf parent. Seven deaf participants declared that sign language was their primary language; 6 participants declared that spoken language was their primary language; and 4 deaf participants declared that they did not have a language preference. On a 10-point Likert-type scale (anchored: 1 = none; 10 = proficient), deaf participants rated proficiency in their sign language ($M = 7.94$; $SD = 1.64$) and in their spoken language ($M = 6.53$; $SD = 1.77$). Twelve participants did not use any hearing aid at the time of the study.

Participation in the study was voluntary, and there were no incentives for participation. All participants signed a written consent form, as well as parents of participants under 18 years old. At the end of the experiment all participants were individually debriefed. A research assistant for the deaf group was fluent in Polish sign language. The study was approved by the review board for research involving human participants of the first author's institution (protocol number: 23/2016; date: June 14, 2016).

Experimental Task and Stimuli

The task started with a 5-point eye-tracker calibration. The average calibration error was recorded at $.47^{\circ}$ ($SD = .19$). Deaf and hearing participants did not significantly differ in calibration error, $t(24.59) = 1.85$, $p = .077$.

The experimental procedure followed the protocol described by Krejtz et al. (2018). Participants were randomly presented with 18 videos starting with a still image of a neutral face that within 10 s changed gradually to full intensity of a happy, sad, or angry expression. The videos were created with the use of morphing, that is, a computer technique that generates smooth transitions between images. For the purposes of this study we used pictures of three male and three female models taken from the Warsaw Set of Emotional Facial Expression Pictures (Olszanowski et al., 2015). The images were selected on the basis of the Facial Action Coding System (Ekman, Friesen, & Hager, 2002) and similarity of facial action unit scores (AU scores) provided for each picture. Specifically, we relied on AU12 (lip corner puller), AU15 (lip corner depressor), and AU4 (brow lowerer).

FantaMorph 5.0 software was used to synthesize dynamic emotional expressions, resulting in stimuli movies with 29 fps. Each stimulus was created from two images of the same face showing neutral and full emotional expression. The morphing effect was achieved by using over 100 facial landmarks. Each trial presented a morphed video at a constant rate across all trials and all participants.

Participants were instructed to press the space bar as soon as they decoded the emotion. This means that each video was ended by the participant once he/she recognized the morphed emotion indicative of individual emotion recognition response time. After each video the recognition question "Which emotion was presented?" appeared on the screen. Participants answered using a single choice scale (happiness, sadness, or anger); see Figure 2a. After their choice, the next video was presented. In the procedure there was no fixation cross prior to the onset of the morph clips. However, we applied a standard procedure in eye-tracking data preparation, which excludes the first fixation made after the change of stimulus.

The experimental procedure was prepared with SMI Experiment Center software. Eye movements were recorded with an SMI RED eye tracker at a sampling rate of 120 Hz. Participants were seated in front of a 22-inch monitor and $1{,}680 \times 1{,}050$ resolutionat a distance of ~57 cm from the screen; see Figure 2b.

Experimental Design

The present study is an eye-tracking experiment following a 2 (Group) × 3 (Emotion) mixed design. The first fixed factor (Group) described two independent groups: deaf and hearing participants. The second fixed factor (Emotion) was a within-subjects experimental manipulation of the facial emotional expressions presented to the participants. This factor consisted of three emotional expressions: happiness, sadness, and anger.

Results

Raw data were processed with SMI's BeGaze software. SMI's BeGaze dispersion-based algorithm was used for detection of fixations and saccades. Fixations of duration 80–1,200 ms were analyzed together with saccades of amplitude $< 10^{\circ}$ (Velichkovsky et al., 2005).

Eye movement data from each stimulus were categorized into two dynamic areas of interest (AOIs) drawn around the eyes and mouth; see Figure 2c. These two AOIs were selected bearing in mind the role of muscles around the eyes and mouth used for facial emotion expression (Bombari et al., 2013; Ekman, 1993; Vassallo, Cooper, & Douglas, 2009; Waters, 1987). The AOIs were used in the statistical analyses as a within-subjects fixed factor.

To analyze the dynamics of eye movements during scanning of dynamical facial expressions, the viewing duration of each stimulus was divided in three time segments (early, middle, and late) for each participant. Since each trial duration differed depending on participants' decision times, this division reflects the relative early, middle, and late stages of visual processing for

(a) Experimental procedure (b) Experimental setup (c) Example AOI annotations

Figure 2 Procedure schematic, AOI annotations on stimulus, and apparatus.

each participant and trial. This variable was a within-subjects fixed factor in statistical analyses.

The dependent variables used for statistical analyses were created by averaging eye-tracking data for six stimuli faces within each of the three emotional expression categories. The same averaging was done for behavioral indicators, namely emotion recognition accuracy and speed (response time). The main indicators of attention allocation were fixation count, average fixation duration, and first fixation allocation to the two distinct AOIs (the eyes and mouth). Finally, the $\mathcal{K}$ coefficient (based on the relationship between fixation duration and amplitude of subsequent saccade) was used as the indicator of dynamical visual attention shifting between ambient–focal attention (Krejtz et al., 2016, 2017).

All statistical analyses were performed in R (R Core Team, 2017). Analysis of variance (ANOVA) was used for hypotheses testing, followed by post-hoc pairwise comparisons with multivariate t correction. Whenever the homogeneity of variances assumption was not met, we report results with Greenhouse–Geisser correction. Each ANOVA result is provided with significance level and effect size expressed by the generalized eta-squared (η^2) coefficient.

Emotion Recognition: Prioritizing Happiness

To test the differences in emotion recognition accuracy and response time between the groups and different emotions, two mixed design 2 × 3 ANOVAs were conducted with the group as a between-subjects factor and facial emotional expression as a within-subject factor.

Accuracy In general, the recognition of emotional expression from morphing faces seemed to be fairly easy. Overall accuracy was close to 90% ($M = .89$; $SD = .09$). ANOVA for accuracy of emotion recognition showed that participants from both groups were similarly accurate, F (1, 38) = 1.48; $p = .232$; $\eta^2 = .02$ (deaf group: $M = .87$, $SD = .12$; hearing group: $M = .91$, $SD = .06$). The analysis revealed a significant main effect of emotional expression, F (1.58, 59.91) = 20.58, $p < .001$, $\eta^2 = .22$. Post-hoc analysis showed that accuracy for happy faces ($M = .99$; $SD = .05$) was significantly higher ($p < .001$) than for sad ($M = .85$; $SD = .16$) or angry ($M = .83$; $SD = .23$) facial expressions. The difference in accuracy between angry and sad faces was not significant. Interaction between emotional expression and group was not significant.

Response time Participants made their decisions about which emotion is expressed on the morphing face in ~5 s, on average ($M = 4797.47$ ms; $SD = 1596.94$). ANOVA showed that recognition time differed significantly between emotions, F (1.74, 65.98) = 79.66, $p < .001$, $\eta 2 = .23$. Decisions were significantly ($p < .001$) faster when examining happy faces ($M = 3533.52$ ms; $SD = 2036.02$) than sad ($M = 5643.81$ ms; $SD = 1984.18$) or angry faces ($M = 5215.07$ ms; $SD = 2278.08$). The difference in recognition time between sad and angry faces was statistically significant ($p = .031$).

The main effect of group did not reach significance; however, a tendency favoring deaf participants was observed, F (1, 38) = 2.85, $p = .09$, $\eta^2 = .06$. Deaf participants were marginally faster in their decisions ($M = 4313.12$ ms; $SD = 1426.94$) compared to the hearing group ($M = 5155.46$ ms; $SD = 165.50$). The observed tendency should be interpreted with caution, as the effect of group explains only 6% of the decision time variance. Interaction between the group and emotion was not statistically significant.

Visual Attention Allocation

Emotion recognition requires a certain amount of attention allocated to the eyes and mouth as the most important regions of facial emotional expression (Smith et al., 2005). In order to test the distribution of visual allocation to the morphing faces we considered three indicators: average fixation duration, fixation count on the eyes and mouth, as well as the first fixation location. The first fixation location may indicate which region of the face, the eyes, or the mouth was the most salient for emotion recognition.

Two independent ANOVAs of mixed design 2 (Group) × 2 (AOI) × 3 (Emotion) were conducted for average fixation duration and fixation count. Due to the nominal type of the location of the first fixation we used cross tables and independence χ^2 tests.

First fixation allocation In order to examine the differences in location of the first fixation (to the eyes or mouth) during the recognition of happiness, sadness, or anger by deaf and hearing participants, a three-way cross table was constructed (Table 1).

The chi-squared test for independence reached statistical significance, χ^2 (7) = 24.22, $p = .001$.

To interpret this significant effect, the three-way table was decomposed into 3 two-way cross tables of AOI and Group (separate for each emotion). The independence chi-squared test for happy facial expressions was only marginally significant, $\chi^2(1) = 2.925$, $p = .087$. Proportion tests showed that deaf participants had fewer first fixations allocated to mouth region than hearing, and there were no group differences in allocation of the

Table 1 Proportion tests of differences between deaf and hearing individuals in allocation of first fixation

Emotion	AOI	Deaf	Hearing	χ^2
Happiness				
	Mouth	.32	.68	14.83***
	Eyes	.49	.51	<1
Sadness				
	Mouth	.33	.67	25.59***
	Eyes	.50	.50	<1
Anger				
	Mouth	.31	.69	32.60***
	Eyes	.50	.50	<1

Note. ***$p < .001$; χ^2—one-sample proportion test.

first fixation to eyes; see Table 1 for detailed proportion values and proportion tests results.

The chi-squared test was significant for sad, $\chi^2(1) = 5.10$, $p = .024$, and angry facial expressions, $\chi^2(1) = 6.18$, $p = .013$. Proportion tests between deaf and hearing participants showed that while looking at sad and angry faces, deaf participants allocated their first fixations significantly less often to the mouth AOI than hearing participants. At the same time, the proportion of first fixations allocated to the eyes for both negative emotional facial expressions did not differ between deaf and hearing individuals. The results suggest group differences in visual strategy of different emotion recognition. The mouth region captured significantly fewer first fixations among the deaf participants than hearing participants.

Fixation count ANOVA of fixation count revealed a significant main effect of Emotion, F (1.95, 58.35) = 48.42, $p < .001$, $\eta^2 = .16$. The following post-hoc analysis showed that happy faces gathered significantly ($p < .001$) fewer fixations before the recognition of emotion ($M = 3.70$; SD = 5.07) than the sad ($M = 5.96$; SD = 5.18) or angry faces ($M = 5.73$; SD = 5.73). The difference in fixation count on sad and angry faces was not statistically significant. A statistically significant main effect of AOI, F (1, 30) = 60.29, $p < .001$, $\eta^2 = .32$, showed that the eyes attracted almost twice as many fixations ($M = 6.55$; SD = 6.06) as the mouth ($M = 3.51$; SD = 3.64).

There was a significant interaction effect of AOI and emotion, F (1.86, 55.86) = 48.71, $p < .001$, $\eta^2 = .11$; see Figure 3a. Post-hoc comparisons showed that there was a significantly ($p < .001$) larger number of fixations on the eyes than on the mouth for sad and angry faces. For happy faces this difference only reached statistical tendency, $p = .09$.

Average fixation duration The present analysis examined average fixation duration as a dependent variable. A statistically significant main effect of AOI, F (1, 30) = 10.79, $p = .003$, $\eta^2 = .07$, showed that average fixation duration was longer when looking at the mouth ($M = 378.70$ ms; SD = 243.25) than at the eyes ($M = 319.16$ ms; SD = 173.14).

The main effect of the emotional expression was marginally significant, F (1.83, 54.97) = 3.07, $p = .054$, $\eta^2 = .01$. Post-hoc tests show that average fixation duration was longer ($p = .042$) on happy faces ($M = 355.60$ ms; SD = 202.69) than on sad faces ($M = 339.38$ ms; SD = 170.57). The average fixation duration when recognizing angry faces ($M = 344.41$ ms; SD = 177.45) was neither statistically different from happy nor from sad faces.

The interaction effect between emotion expressed by the face and AOI reached significance, F (1.93, 58.02) = 15.16, $p < .001$, $\eta^2 = .07$; see Figure 3b. Pairwise comparisons showed that, when recognizing happy faces, participants made significantly ($p < .001$) longer fixations on the mouth ($M = 424.37$ ms; SD = 206.43) than on the eyes ($M = 295.38$ ms; SD = 123.09). During recognition of both sadness and anger the difference in fixation duration between the eyes and mouth was not significant (Figure 3b). Neither the main effect of group nor interaction including group reached statistical significance.

Visual Gaze Dynamics

Recognition of emotion from a dynamically changing face needs a certain pattern of attentional focus: from ambient scanning to focal visual information processing. The later stage of stimuli processing is hypothetically associated with deeper visual information processing, which facilitates and precedes decision. In order to test the hypothesis a 3-way $2 \times 3 \times 3$ mixed design ANOVA was performed with Group as the between-subjects factor and Emotion and Trial Time (divided into three equal periods) as within-subjects factors. We used the $\mathcal{K}$ coefficient as the dependent variable to test changes in ambient–focal attention dynamics.

In line with expectations, analysis showed a main effect of trial time on ambient–focal attention, F (1.94, 60.14) = 4.33, $p = .017$, $\eta^2 = .20$. Post-hoc tests showed that in the first period participants' attention was ambient ($M = -.05$; SD = 1.73) then shifting to focal attention in the second time period ($M = .09$; SD = 1.66), $p < .001$. The difference in $\mathcal{K}$ between the second and third time periods did not reach statistical significance, $p = .19$. The shift from ambient to focal attention observed during emotion recognition of dynamical facial expression is consistent with the literature (e.g., Krejtz et al., 2018; Unema et al., 2005; Velichkovsky et al., 2005).

The general pattern of visual gaze dynamics was significantly moderated by emotion, F (3.55, 110.00) = 3.97, $p = .007$, $\eta^2 = .02$. Post-hoc comparisons showed that only for happy faces the pattern of shifting from ambient to focal attention between the first ($M = -.06$; SD = 1.49) and the second trial time period ($M = .18$; SD = 1.53) was statistically significant, $p = .009$. For happy faces, during the last (third) time period, attention remained focal ($M = .16$; SD = 1.48), and it was significantly ($p < .01$) more focal than for sad ($M = -.09$; SD = 1.33) and angry faces ($M = -.11$; SD = 1.41). For sad and angry facial expressions, the shift between ambient and focal attention was not significant (although the pattern of means is consistent with that for happy faces).

The above interaction was significantly moderated by the group, F (3.55, 110.00) = 3.07, $p = .019$, $\eta^2 = .02$; see Figure 4. We decomposed the 3-way interaction 2 (Group) × 3 (Trial Time) × 3 (Emotion) into two separate ANOVAs for each group, designed as 3 (Trial Time) × 3 (Emotion). The ANOVA results for deaf participants showed significant interaction of emotion and time, F (3.24, 48.61) = 5.67, $p = .002$, $\eta^2 = .08$; see Figure 4a.

Post-hoc comparisons showed that deaf participants were significantly more focal during the last time period when looking at happy faces ($M = .28$; SD = 1.57) compared to both sad ($M = -.22$; SD = 1.41; $p = .02$) and angry faces ($M = -.21$; SD = 1.59; $p < .05$); see Figure 4a.

Post-hoc comparisons also revealed that, when looking at happy faces, deaf participants were significantly ($p = .004$) more focal during the last time period ($M = .28$; SD = 1.57) compared to

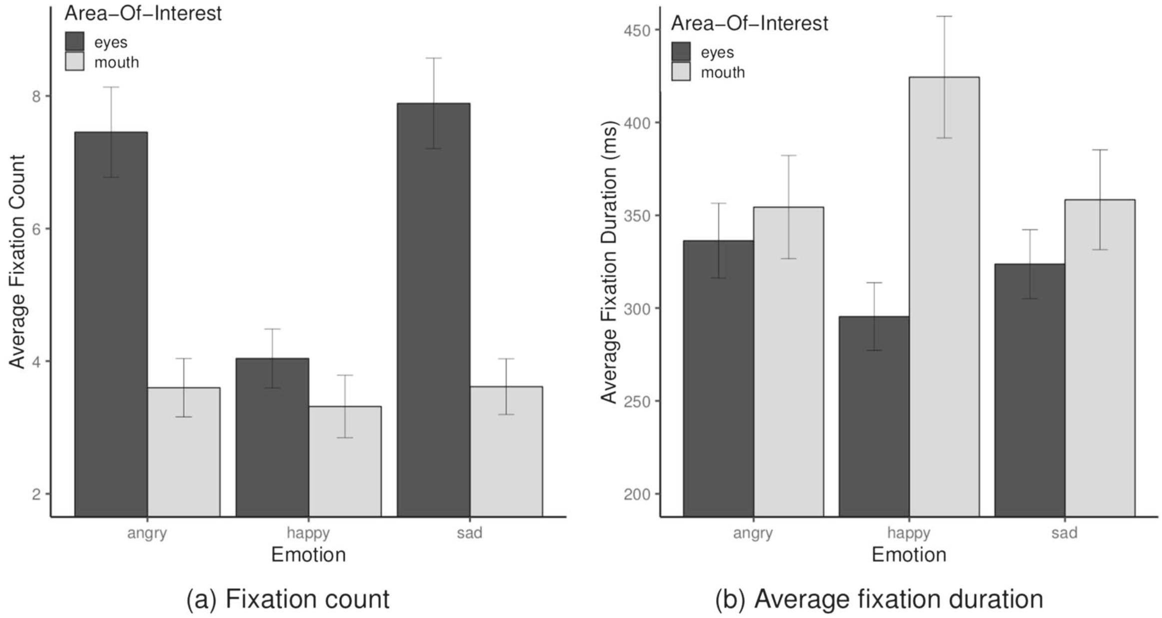


Figure 3 Fixation count and average fixation duration on the eyes and mouth during recognition of happy, sad, and angry dynamical facial expressions (error bars represent ±1 confidence intervals).

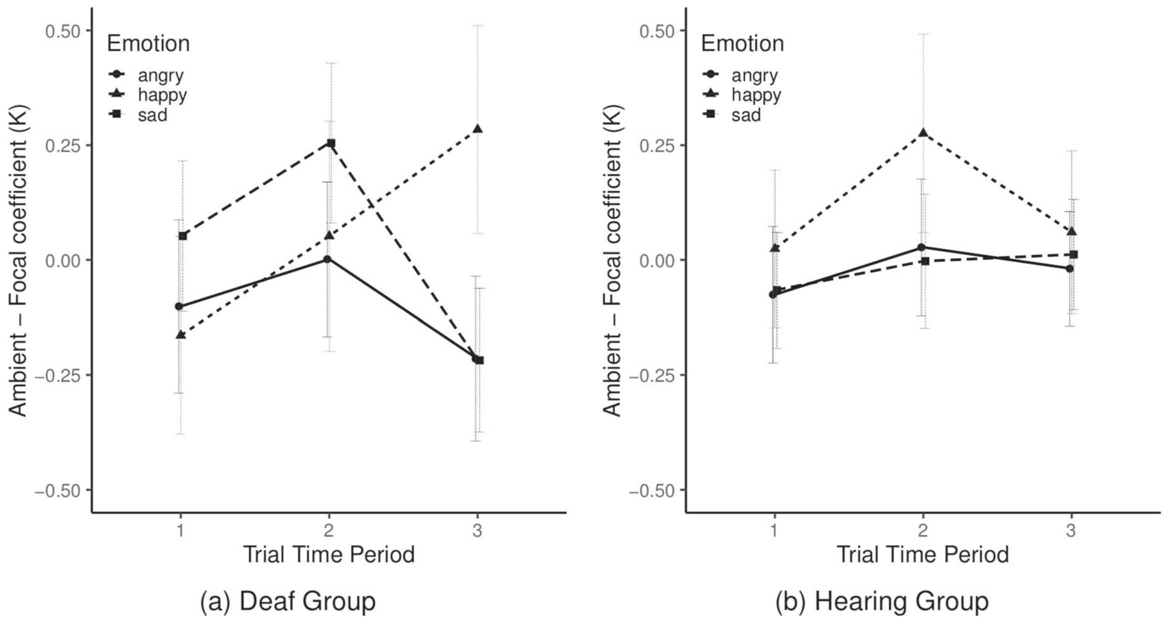


Figure 4 Dynamics of ambient–focal attention moderated by emotional facial expression and group (error bars represent ±1 confidence intervals).

the first time period ($M = -.16$; SD = 1.60). When looking at sad faces, deaf participants were significantly ($p = .01$) more ambient during the last time period ($M = -.22$; SD = 1.41) compared to the middle time period ($M = .25$; SD = 1.45). There were no significant differences when looking at angry faces across different time periods ($p > .1$).

The interaction between Emotion and Trial Time was not significant for the hearing group, F (3.01, 48.23) < 1. There was a significant main effect of Emotion, F (1.84, 29.52) = 5.97, $p = .008$, $\eta^2 = .02$. In general, hearing participants were more focal when looking at happy ($M = .11$; SD = 1.64) than sad ($M = -.02$; SD = 1.52) and angry faces ($M = -.03$, SD = 1.61); see Figure 4b.

Discussion

The main aim of the present study was to examine characteristics of visual attention of deaf and hearing individuals while they were decoding emotions from dynamically morphing facial expressions. The discussion starts with a summary of the key findings and contribution to the current literature.

Facial Expression Recognition

The present eye-tracking experiment demonstrated that happiness in dynamical facial expressions is recognized with significantly greater accuracy and in a shorter amount of time compared to both negative emotions (anger and sadness). This result is in line with a well-established finding in the literature known as the "happiness superiority effect" among typical and anxious individuals (Krejtz et al., 2018; Leyman, De Raedt, Vaeyens, & Philippaerts, 2011) and corroborates previous studies showing recognition advantage of happy faces (Kret, Stekelenburg, Roelofs, & De Gelder, 2013). In the present study, there were no group differences in accuracy of emotion recognition. However, there was a tendency for deaf participants to be quicker in recognition of emotional expressions in comparison to hearing individuals. Although the result is relatively weak in terms of effect size, it may be worth pursuing in future research on a larger sample.

In general, the findings corroborate previous studies suggesting small or no differences in emotion recognition between D/HH and hearing individuals (Hosie et al., 1998; Most & Aviner, 2009; Rieffe & Terwogt, 2000). Results support a compensatory role of motion in emotion recognition among deaf people (Jones et al., 2018). The use of dynamic facial expressions provided an advantage for deaf individuals in emotion recognition (Jones et al., 2018) probably due to their enhanced reactivity to motion in the peripheral visual field (Armstrong et al., 2002).

Attention Allocation

There were group differences in the allocation of the first fixation. Deaf individuals directed their first fixations less often toward the mouth region than did the control group. Both groups directed their gaze similarly to the eye region. The group difference in first fixation allocation is in line with eye-tracking studies on the perception of a sign interpreter (Emmorey, Thompson, & Colvin, 2008). Emmorey et al. (2008) showed that deaf people, who are fluent in sign language, allocate more attention to the eyes of the signing person than hearing individuals from deaf signing families who were using sign language in their daily lives, who fixed their eyes more often on the signing person's lips.

The region of eyes is crucial for recognition of negative emotions (sadness or anger) (Schurgin et al., 2014). Therefore, the first fixation directed less often towards the mouth AOI may facilitate slightly faster emotion recognition by deaf individuals. Previous research provided evidence of a wider perceptual span (Stevens & Neville, 2006) among deaf people compared to hearing individuals. The present results support these claims to some extent. Assuming a wider perceptual span, deaf participants, while fixating the eyes, may have monitored motion in the mouth region.

In general, happy faces received fewer fixations than angry and sad facial expressions, but the fixation durations tend to be longer on happy faces than on sad faces, suggesting deeper information processing of happiness. This finding is in line with the tendency to prioritize happiness. Interestingly, while recognizing happy faces, average fixation duration was longer on the mouth than on the eyes, whereas while recognizing sadness and anger there were no differences in fixation duration of the eyes or the mouth. There were no group differences in the frequency and duration of fixations between AOIs and emotion.

Letourneau and Mitchell (2011) found that deaf individuals fixated more on faces than hearing individuals, whereas Watanabe et al. (2011) reported that deaf participants allocated more attention to eyes compared to hearing individuals. In the both studies still images of facial emotional expressions were used with limited presentation time (max. 4 s). The present study focused on discerning focal and ambient processing of emotionally morphing faces with a longer period of time (max. 10 s). To understand visual attention processes of recognition of dynamical facial expressions, gaze dynamics were analyzed.

Dynamics of Ambient–Focal Attention

Our contribution to the understanding of recognition process is the analysis of visual gaze dynamics. We assumed that in order to recognize emotion, participants need to shift visual attention from relatively ambient scanning of the stimuli during which simple stimuli features are processed to a more focal attention with deeper processing of emotional expression. In order to examine this dynamical process, recognition response time was divided into three equal periods. That is, analysis is based on relative time epochs and not on absolute time, based on seconds. This type of approach was frequently used in previous literature on attention dynamics (e.g., Duchowski et al., 2019; Krejtz et al., 2018; Krejtz et al., 2016, 2017).

The observed eye movement dynamics followed the hypothesized pattern from ambient to focal despite the speed of emotion recognition. In line with predictions, there was a general pattern of switching from ambient to focal attention. This finding corroborates previous literature on visual gaze dynamics (Velichkovsky et al., 2005; Krejtz et al., 2016). Emotional expression and group moderated the pattern of attentional switching. While recognizing happiness, focal attention among deaf participants continued to increase simultaneously with emotion but not when decoding anger or sadness. At the last stage of emotion recognition, eye movements were significantly more focal on happy than on sad or angry facial expressions. We did not observe a similar pattern of gaze dynamics among hearing individuals (the interaction between Emotion and Time period was not significant). Hearing participants were in general more focal when recognizing happiness comparing to anger or sadness.

Study Limitations and Future Directions

Future studies should investigate the hypothesis of emotional and cognitive withdrawal from deep processing of negative emotions, for example, in a visual search experiment presenting emotional faces versus abstract stimuli (for review, see Frischen et al., 2008). The results may explain group differences (e.g., anxious versus non-anxious and deaf versus hearing) in switching between ambient and focal attention during search for certain emotions versus abstract stimuli. For anxious individuals we would expect to observe more ambient processing (indicating emotional withdrawal) for emotional stimuli compared to search for abstract stimuli. For deaf individuals, assuming that they "do not see better" but react faster to the stimuli in their environment (Pavani & Bottari, 2012), we would expect a faster switch from ambient to focal attention for both types of stimuli, compared to hearing individuals. Reactivity measures such as the speed of switch from ambient into focal attention may be more sensitive than accuracy reports when comparing deaf and hearing controls during searching for abstract and emotional stimuli. For example Dye et al. (2009) showed that, following early auditory deprivation, visual attention resources toward the periphery are slowly augmented to eventually result in a clear behavioral advantage by pre-adolescence on a selective visual attention task. However, Dye et al. used only abstract stimuli and did not use the distinction between ambient and focal attention.

Our second hypothesis referred to a preference for recognition of happy faces among the deaf and hearing participants. In line with this prediction, both groups were more focal while recognizing happiness, but deaf participants were more focal during the last period of happy face recognition. To the best of our knowledge this is the first study on ambient/focal attention studying gaze dynamics among deaf participants. Future studies should further examine the happiness superiority effect among deaf participants.

In the present study there were only visual cues for emotion recognition. However, it is worth investigating emotion recognition with and without auditory cues (elements of prosody: pitch of voice, intonation, stress, and rhythm) among individuals with hearing or residual hearing. For example, the study of Hopyan-Misakyan et al. (2009) indicates that children with a cochlear implant recognize emotion in faces but have limited perception of affective speech prosody. These difficulties may stem in part from difficulties with pitch processing (Scherer, 1995, 2003).

We are aware that at this time that there is not enough empirical evidence to fully understand the extent to which residual hearing plays a role in regulating attentional and emotional mechanisms. However, we believe that eye tracking sheds some additional light on the processes. Therefore, a replication of the present study with a control of degree of hearing loss as well as the use of hearing aids is a potentially suitable extension.

Finally, the present study used a relatively small sample size to capture between-group differences. The deaf group was relatively heterogeneous in terms of hearing loss, and they differed in preference for the use of sign or spoken language. These factors as well as the sample size may be responsible for the high variability in emotion recognition times across both groups leading to a marginally significant between-group difference. One may expect that the difference in decision times may reach statistical significance level when tested on a larger and more homogeneous sample.

Conclusions

Emotion recognition is crucial among deaf people for understanding of social signals and for maintenance of emotional contact within their social environment. In the present study, there were no substantial differences in emotion recognition between deaf and hearing individuals. However, the pattern of visual attention during recognition of emotion reflected between-group differences in the first fixation to the mouth region. Further, we observed a significant shift from ambient to focal attention during happiness detection only for deaf participants. Although we cannot argue for support of either the “enhancement” or the “deficit hypothesis” (Sidera et al., 2017), current results showed the usefulness of a dynamical approach to the investigation of attentional processes among both deaf and hearing individuals. We postulate that motion of facial expression may afford an advantage toward emotion recognition among deaf people (Jones et al., 2018). The advantage of motion may be related to potential differences in visual attention span, which calls for further investigation.

The main contribution of this paper is its approach to analysis of ambient and focal visual gaze dynamics. Such an approach provides information about the dynamics of the depth of cognitive processing. Deaf and hearing groups demonstrated deeper cognitive processing of happiness.

Acknowledgments

We are grateful to Zofia Kostrzewa and Piotr Baltyn for their help in collecting the data. We also thank Instytut Głuchoniemych im. Jakuba Falkowskiego in Warsaw for their help in obtaining the research sample.

References

1. Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review*, *12*(6), 1043–1047. doi:10.3758/BF0320644
2. Anderson, A. K., Christoff, K., Panitz, D., De Rosa, E., & Gabrieli, J. D. (2003). Neural correlates of the automatic processing of threat facial signals. *Journal of Neuroscience*, *23*(13), 5627–5633. doi:10.1523/JNEUROSCI.23-13-05627.2003
3. Armstrong, B. A., Neville, H. J., Hillyard, S. A., & Mitchell, T. V. (2002). Auditory deprivation affects processing of motion, but not color. *Cognitive Brain Research*, *14*(3), 422–434. doi:10.1016/S0926-6410(02)00211-2
4. Arnold, P., & Murray, C. (1998). Memory for faces and objects by deaf and hearing signers and hearing nonsigners. *Journal of Psycholinguistic Research*, *27*(4), 481–497. doi:10.1023/A:1023277220438
5. Becker, D. V., Neel, R., Srinivasan, N., Neufeld, S., Kumar, D., & Fouse, S. (2012). The vividness of happiness in dynamic facial displays of emotion. *PLoS One*, *7*(1), 1–6. doi:10.1371/journal.pone.0026551
6. Bombari, D., Schmid, P. C., Mast, M. S., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: The role of featural and configural face information. *Quarterly Journal of Experimental Psychology*, 66(12), 2426–2442. doi:10.1080/17470218.2013.789065
7. Boot, W. R., Becic, E., & Kramer, A. F. (2009). Stable individual differences in search strategy?: The effect of task demands and motivational factors on scanning strategy in visual search. *Journal of Vision*, *9*(3), 7–7. doi:10.1167/9.3.7
8. Corina, D. P., Bellugi, U., & Reilly, J. (1999). Neuropsychological studies of linguistic and affective facial expressions in deaf signers. *Language and Speech*, *42*(2–3), 307–331. doi:10.1177/00238309990420020801
9. Craig, B. M., Becker, S. I., & Lipp, O. V. (2014). Different faces in the crowd: A happiness superiority effect for schematic faces in heterogeneous backgrounds. *Emotion*, *14*(4), 794–803. doi:10.1037/a0036043
10. Delis, I., Chen, C., Jack, R. E., Garrod, O. G., Panzeri, S., & Schyns, P. G. (2016). Space-by-time manifold representation of dynamic facial expressions for emotion categorization. *Journal of Vision*, *16*(8), 14–14. doi:10.1167/16.8.14
11. Duchowski, A., Krejtz, K., Żurawska, J., & House, D. (2019). Using microsaccades to estimate task difficulty during visual search of layered surfaces. *IEEE Transactions on Visualization and Computer Graphics*, 1–1. doi:10.1109/TVCG.2019.2901881
12. Duchowski, A. T. (2007). *Eye tracking methodology. Theory and practice*. London: Springer. doi:10.1007/978-1-84628-609-4
13. Dyck, M. J., Farrugia, C., Shochet, I. M., & Holmes-Brown, M. (2004). Emotion recognition/understanding ability in hearing or vision-impaired children: Do sounds, sights, or words make the difference? *Journal of Child Psychology and Psychiatry*, 45(4), 789–800. doi:10.1111/j.1469-7610.2004.00272.x
14. Dye, M. W. G., Hauser, P. C., & Bavelier, D. (2009). Is visual selective attention in deaf individuals enhanced or deficient?

The case of the useful field of view. *PLoS One*, 4(5), e5640. doi:10.1371/journal.pone.0005640
15. Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*(4), 384. doi:10.1037/0003-066X.48.4.384
16. Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *The facial action coding system: The manual on CD-ROM & investigator's guide* (). Salt Lake City, UT: A Human Face.
17. Emmorey, K., & McCullough, S. (2009). The bimodal bilingual brain: Effects of sign language experience. *Brain and Language*, *109*(2–3), 124–132. doi:10.1016/j.bandl.2008.03.005
18. Emmorey, K., Thompson, R., & Colvin, R. (2008). Eye gaze during comprehension of American sign language by native and beginning signers. *Journal of Deaf Studies and Deaf Education*, *14*(2), 237–243. doi:10.1093/deafed/enn037
19. Fischer, B., & Ramsperger, E. (1984). Human express saccades: Extremely short reaction times of goal directed eye movements. *Experimental Brain Research*, *57*(1), 191–195. doi:10.1007/BF00231145
20. Frischen, A., Eastwood, J. D., & Smilek, D. (2008). Visual search for faces with emotional expressions. *Psychological Bulletin*, *134*(5), 662–676. doi:10.1037/0033-2909.134.5.662
21. Gray, C. D., Hosie, J. A., Russell, P. A., & Ormel, E. A. (2001). Emotional development in deaf children: Facial expressions, display rules, and theory of mind. In M. D. Clark, M. Marschark & M. Karchmer (Eds.), *Context, cognition, and deafness* (pp. 135–160). Washington, DC: Gallaudet University Press.
22. Hao, J., & Su, Y. (2014). Deaf children's use of clear visual cues in mindreading. *Research in Developmental Disabilities*, *35*(11), 2849–2857. http://doi.org/10.1016/j.ridd.2014.07.034
23. Haxby, J. F., Hoffman, E. A., & Gobbini, I. M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. doi:10.1016/S1364-6613(00)01482-0
24. Hopyan-Misakyan, T. M., Gordon, K. A., Dennis, M., & Papsin, B. C. (2009). Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. *Child Neuropsychology*, *15*(2), 136–146. doi:10.1080/09297040802403682
25. Hosie, J., Gray, C., Russell, P., Scott, C., & Hunter, N. (1998). The matching of facial expressions by deaf and hearing children and their production and comprehension of emotion labels. *Motivation and Emotion*, *22*(4), 293–313. doi:10.1023/A:1021352323157
26. Irwin, D. E., & Zelinsky, G. J. (2002). Eye movements and scene perception: Memory for things observed. *Perception & Psychophysics*, *64*(6), 882–895. doi:10.3758/BF03196793
27. Jack, R. E., Garrod, O. G., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology*, *24*(2), 187–192. doi:10.1016/j.cub.2013.11.064
28. Jones, A. C., Gutierrez, R., & Ludlow, A. K. (2018). The role of motion and intensity in deaf children's recognition of real human facial expressions of emotion. *Cognition and Emotion*, *32*(1), 102–115. doi:10.1080/02699931.2017.1289894
29. Just, M. A., & Carpenter, P. A. (1984). Using eye fixations to study reading comprehension. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 151–182). Hillsdale, NJ: Erlbaum.
30. Krejtz, I., Szarkowska, A., Krejtz, K., Walczak, A., & Duchowski, A. (2012). Audio description as an aural guide of children's visual attention: Evidence from an eye-tracking study. In *Proceedings of the symposium on eye tracking research and applications* (pp. 99–106). New York: ACM. doi:10.1145/2168556.2168572
31. Krejtz, K., Coltekin, A., Duchowski, A. T., & Niedzielska, A. (2017). Using coefficient K to distinguish ambient/focal visual attention during cartographic tasks. *Journal of Eye Movement Research*, *10*(2), 3. doi:10.16910/jemr.10.2.3
32. Krejtz, K., Duchowski, A., Krejtz, I., Szarkowska, A., & Kopacz, A. (2016). Discerning ambient/focal attention with coefficient K. *ACM Transactions on Applied Perception (TAP)*, *13*(3), 11. doi:10.1145/37402.37405
33. Krejtz, K., Wisiecka, K., Krejtz, I., Holas, P., Olszanowski, M., & Duchowski, A. T. (2018). Dynamics of emotional facial expression recognition in individuals with social anxiety. In *Proceedings of the 2018 ACM symposium on eye tracking research & applications* (p. 43). Warsaw: ACM. doi:10.1145/3204493.3204533
34. Kret, M. E., Stekelenburg, J. J., Roelofs, K., & De Gelder, B. (2013). Perception of face and body expressions using electromyography, pupillometry and gaze measures. *Frontiers in Psychology*, *4*, 28. doi:10.3389/fpsyg.2013.00028
35. Krumhuber, E. G., Kappas, A., & Manstead, A. S. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review*, *5*(1), 41–46. doi:10.1177/1754073912451349
36. Letourneau, S. M., & Mitchell, T. V. (2011). Gaze patterns during identity and emotion judgments in hearing adults and deaf users of American sign language. *Perception*, *40*(5), 563–575. doi:10.1068/p6858
37. Leyman, L., De Raedt, R., Vaeyens, R., & Philippaerts, R. M. (2011). Attention for emotional facial expressions in dysphoria: An eye-movement registration study. *Cognition and Emotion*, *25*(1), 111–120. doi:10.1080/02699931003593827
38. McCullough, S., & Emmorey, K. (1997). Face processing by deaf asl signers: Evidence for expertise in distinguishing local features. *Journal of Deaf Studies and Deaf Education*, *2*(4), 212–222. doi:10.1093/oxfordjournals.deafed.a014327
39. Merabet, L. B., & Pascual-Leone, A. (2010). Neural reorganization following sensory loss: The opportunity of change. *Nature Reviews Neuroscience*, *11*(1), 44. doi:10.1038/nrn2758
40. Most, T., & Aviner, C. (2009). Auditory, visual, and auditory–visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. *The Journal of Deaf Studies and Deaf Education*, 14(4), 449–464. doi:10.1093/deafed/enp007
41. Most, T., & Michaelis, H. (2012). Auditory, visual, and auditory–visual perceptions of emotions by young children with hearing loss versus children with normal hearing. *Journal of Speech, Language, and Hearing Research*, 55(4), 1148–1162. doi:10.1044/1092-4388(2011/11-0060)
42. Olszanowski, M., Pochwatko, G., Kukliński, K., Ścibor-Rylski, M., Lewinski, P., & Ohme, R. K. (2015). Warsaw set of emotional facial expression pictures: A validation study of facial display photographs. *Frontiers in Psychology*, 5, 1516. doi:10.3389/fpsyg.2014.01516
43. Pavani, F., & Bottari, D. (2012). Visual abilities in individuals with profound deafness: A critical review. In M. M. Murray, & M. T. Wallace (Eds.), *The neural bases of multisensory processes*. Boca Raton, FL: CRC Press/Taylor & Francis. doi:10.1201/b11092
44. Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. doi:10.1080/00335558008248231
45. R Core Team. (2017). *R: a language and environment for statistical computing [computer software manual]*. Vienna, Austria. Retrieved from https://www.R-project.org/

46. Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372. doi:10.1037/0033-2909.124.3.372
47. Rieffe, C., & Terwogt, M. M. (2000). Deaf children's understanding of emotions: Desires take precedence. *Journal of Child Psychology and Psychiatry*, *41*(5), 601–608. doi:10.1111/1469-7610.00647
48. Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: An fmri study. *Cognitive Brain Research*, *20*(1), 81–91. doi:10.1016/j.cogbrainres.2004.01.008
49. Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, *9*, 235–248. doi:10.1016/S0892-1997(05)80231-0
50. Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, *40*, 227–256. doi:10.1016/S0167-6393(02)00084-5
51. Schurgin, M. W., Nelson, J., Iida, S., Ohira, H., Chiao, J., & Franconeri, S. (2014). Eye movements during emotion recognition in faces. *Journal of Vision*, *14*(13), 1–16. doi:10.1167/14.13.14
52. Sidera, F., Amadó, A., & Martínez, L. (2017). Influences on facial emotion recognition in deaf children. *Journal of Deaf Studies and Deaf Education*, 164–177. doi:10.1093/deafed/enw072
53. Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, *16*(3), 184–189. doi:10.1111/j.0956-7976.2005.00801.x
54. Stevens, C., & Neville, H. (2006). Neuroplasticity as a double-edged sword: Deaf enhancements and dyslexic deficits in motion processing. *Journal of Cognitive Neuroscience*, *18*(5), 701–714. doi:10.1162/jocn.2006.18.5.701
55. Strauch, C., Huckauf, A., Krejtz, K., & Duchowski, A. T. (2018). Towards a selection mechanism integrating focal fixations, pupil size, and microsaccade dynamics [peer-reviewed conference paper]. In P. Kiefer, I. Giannopoulos, F. Göbel, M. Raubal, & A. T. Duchowski (Eds.), *Eye tracking for spatial research, proceedings of the 3rd international workshop* (pp. 9–15). Zurich: ETH Zurich. doi:10.3929/ethz-b-000222294.
56. Unema, P. J. A., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, *12*(3), 473–494. doi:10.1080/13506280444000409
57. Vassallo, S., Cooper, S. L., & Douglas, J. M. (2009). Visual scanning in the recognition of facial affect: Is there an observer sex difference? *Journal of Vision*, *9*(3), 11. doi:10.1167/9.3.11
58. Velichkovsky, B. M., Joos, M., Helmert, J. R., & Pannasch, S. (2005). Two visual systems and their eye movements: Evidence from static and dynamic scene perception. In CogSci 2005: *Proceedings of the XXVII conference of the cognitive science society* (pp. 2283–2288). July 21–23 Stresa, Italy. doi:10.1.1.486.9238
59. Wang, Y., Su, Y., Fang, P., & Zhou, Q. (2011). Facial expression recognition: Can preschoolers with cochlear implants and hearing aids catch it? *Research in Developmental Disabilities*, *32*(6), 2583–2588. doi:10.1016/j.ridd.2011.06.019
60. Watanabe, K., Matsuda, T., Nishioka, T., & Namatame, M. (2011). Eye gaze during observation of static faces in deaf people. *PLoS One*, *6*(2), e16919. doi:10.1371/journal.pone.0016919
61. Waters, K. (1987). A muscle model for animation three-dimensional facial expression. *SIGGRAPH Computer Graphics*, *21*(4), 17–24. doi:10.1145/37401.37405
62. Ziv, M., Most, T., & Cohen, S. (2013). Understanding of emotions and false beliefs among hearing children versus deaf children. *Journal of Deaf Studies and Deaf Education*, *18*(2), 161–174. doi:10.1093/deafed/ens073