Sound localization latency in normal hearing and simulated unilateral hearing loss

Martin Eklöf a, b, *, Filip Aspa a, b, Erik Berninger a, c

a Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden
b Department of ENT, Section of Hearing Implants, Karolinska University Hospital, Stockholm, Sweden
c Department of Audiology and Neurotology, Karolinska University Hospital, Stockholm, Sweden

A R T I C L E   I N F O

Article history:
Received 29 November 2019
Received in revised form 14 May 2020
Accepted 28 May 2020
Available online 11 June 2020

Keywords:
Auditory saccade
Reaction time
Gaze
Unilateral hearing loss
Saccadic latency

A B S T R A C T

Directing gaze towards auditory events is a natural behavior. In addition to the well-known accuracy of auditory elicited gaze responses for normal binaural listening, their latency is a measure of possible critical interest and methodological importance. The aim was to develop a clinically feasible method to assess sound localization latency (SLL), and to study SLL as a function of simulated unilateral hearing loss (SUHL) and the relationship with accuracy. Eight healthy and normal-hearing adults (18–40 years) participated in this study. Horizontal gaze responses, recorded by non-invasive corneal reflection eye-tracking, were obtained during azimuthal shifts (24 trials) of a 3-min continuous auditory stimulus. In each trial, a sigmoid function was fitted to gaze samples. Latency was estimated by the abscissa corresponding to 50% of the arctangent amplitude. SLL was defined as the mean latency across trials. SLL was measured in normal-hearing and simulated SUHL conditions (SUHL30 and SUHL43, mean threshold of 30 dB HL and 43 dB HL across 0.5, 1, 2, and 4 kHz). In the normal-hearing condition, the mean ± SD SLL was 280 ± 40 ms (n = 8) with a test-retest SD = 20 ms. A linear mixed model showed a statistically significant effect of listening condition on SLL. The SUHL30 and SUHL43 conditions revealed a mean SLL of 370 ± 49 ms and 540 ± 120 ms, respectively. Repeated measures correlation analysis showed a clear relationship between SLL and the average sound localization accuracy (R² = 0.94). The rapid and reliable method to obtain SLL may be an important clinical tool for evaluation of binaural processing. Future studies in clinical cohorts are needed to assess whether SLL may reveal information about binaural processing abilities beyond that afforded by sound localization accuracy.

© 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The present study was motivated by the clinical need for rapid and language-independent tests of auditory processing of binaural suprathreshold stimuli for all ages. Towards developing such a test, localization latencies in normal-hearing and healthy adults were measured with an objective, rapid, and non-invasive eye-tracking technique. A mathematical function was used to model auditory elicited gaze patterns following an azimuthal sound shift.

Directing gaze towards a sound is a natural response with high survival and communication value (Grothe et al., 2010; Lachs et al., 2001). To localize sounds in the horizontal plane, e.g. an interesting conversation or an event outside the visual field, differences in time (interaural time difference; ITD), and level (interaural level difference; ILD), of the sound reaching the two ears are processed in the central auditory system (Middlebrooks and Green, 1991). Binaural processing is thus the basis for horizontal sound localization, which shows high acuity in humans in the frontal horizontal plane (one degree resolution for targets around midline (Mills, 1958)). While previous laboratory-oriented studies have revealed typical latencies for auditory evoked saccades and their relation to accuracy (e.g. Frens and Van Opstal, 1995), the clinical availability of techniques for measuring auditory evoked saccades is limited.

Potentially, latencies could reveal diagnostic information beyond accuracy (see for example Pichora-Fuller et al. (2016) for a discussion about latency and listening effort). It is also possible that adaption to a condition occurs in one of the measures but not the other. For example, individuals with unilateral hearing loss show an adaptation to their condition, sometimes resulting in high localization accuracy (Kumpik and King, 2019). Latencies for these
2.2. Study design

Sound localization latency was estimated in an objective way by

Methods for the determination of the saccadic response latency towards auditory targets include electro-oculography with skin electrodes (e.g. Goldring et al., 1996; Zambarbieri, 2002) or insertion of scleral coils (e.g. Fresn and Van Opstal, 1995; Populin, 2008; Populin et al., 2002; Yao and Peck, 1997), often restraining the head of the subject (Frens and Van Opstal, 1995; Zahn et al., 1978). For the head-unrestrained condition, the head often starts to turn simultaneously or soon after for contribution to saccadic amplitude and realignment of eye and head (Goldring et al., 1996; Goosens and Opstal, 1997). With the introduction of corneal reflection eye-tracking technology, eye-movements towards spatially distributed events became available without medical procedures. For auditory events, this was initially done with heads restrained (Ten Brink et al., 2014). Modern corneal reflection eye-tracking however, has a potential for clinical use given not only its non-invasiveness but also freedom of head-movements (Asp et al., 2016). While freedom of head-movements may dynamically change ITD and ILD cues, thereby affecting experimental control, it allows the study of sound localization behavior in infants and young children for whom the assessment of auditory development is important in cases of early diagnosed hearing impairment.

Plugging one human ear to "monauralize" a listener is known to decrease sound localization accuracy for various response methods (head-fixed laser pointer: Agterberg et al., 2012; touch screen: Irving and Moore, 2011; electromagnetic head orientation: Slattery and Middlebrooks, 1994) including eye gaze (Asp et al., 2018). To our knowledge, it is unknown how sound localization latency measured by gaze response latency is affected. It is known from experiments with visual targets that the latency increases with task difficulty (Warren et al., 2013; Wheelless et al., 1967). For auditory targets, we are only aware of a single study examining effects of task difficulty on saccadic response latencies in humans, demonstrating some 30 ms increase in the latency when visual distractors were present (Ten Brink et al., 2014).

Two levels of simulated unilateral hearing loss (SUHL) were induced in this study, to investigate the acute effect on localization latency. While an acute peripheral SUHL is not the same as longstanding, unilateral hearing loss during development deprives binaural hearing (Kaplan et al., 2016). For example, neurons sprout to the hearing side at MSO following experimentally induced unilateral hearing loss at birth, as shown in rats (Feng and Rogowski, 1980). This study revealed a large and statistically significant prolongation of sound localization latency (SLL) for azimuthal shifts of a broad-band sound after a SUHL in normal healthy adults and a distinct repeated measures correlation with sound localization accuracy across SUHL conditions.
fitting a mathematical sigmoid function (arctangent) to eye gaze responses evoked by instantaneous shifts of a broad-band sound in the frontal horizontal plane (Fig. 1). Listening conditions included normal binaural conditions (test and retest; retest only used for reliability analysis) and two levels of simulated unilateral hearing loss. The order of listening conditions was randomized. A linear mixed model was used to study the acute effect of simulated UHL on SLL in a within-subject repeated measures design.

This study was part of another study and the data presented here are obtained from the same measurements and subjects (Asp et al., 2018). It was approved by the regional ethical review board in Stockholm, Sweden (2013/104-31). Written informed consent was obtained and the research complied to the ethical principles stated in the declaration of Helsinki.

2.3. Sound localization test

2.3.1. Setup

The setup consisted of an auditory-visual stimulus system and an eye-tracking system (Smart Eye Pro; Smart Eye AB, Gothenburg, Sweden), described in detail elsewhere (Asp et al., 2016). The auditory-visual stimulus system comprised a personal computer (Dell Latitude E5520; Dell Inc., TX) with Windows 7, equipped with a multichannel external soundcard (AudioFire 12; Echo Audio Corporation, CA) and two external multichannel video mixers (VP-108; Kramer Electronics, Israel). The outputs from the soundcard were connected to 12 loudspeakers positioned in 10° increments in a 110-degree arc (±55° azimuth) with 1.2-m radius, and the outputs from the video mixers were connected to 7" TFT-displays mounted below each loudspeaker (loudspeaker/display-pairs; LD-pairs), feeding them with a video graphics array signal from the computer. Subjects were seated 1.2 m from the LD-pairs, which were vertically adjusted to align the center of the loudspeakers to ear-level individually. Subjects were free to move their head. Gaze alone was registered which combine head and eye-movements. The head angle was not necessarily aligned with eye-ball direction and not necessarily directed towards the current display at the start of a trial.

Each LD-pair was defined as an Area of Interest (AOI) in a three-dimensional coordinate system in the eye-tracking system (Asp et al., 2016). Objective detection of eye gaze intersection with the AOs was performed at 20 Hz, resulting in 12 possible gaze direction samples (−55 to +55° in 10° increments). The discreet visual target space motivated the use of gaze intersected AOI rather than raw gaze angle. These samples were transmitted from the eye-tracking system through a low latency network connection (User Datagram Protocol) to the audio-visual presentation system together with a time stamp. Synchronization between the eye-tracker and the audio-visual presentation systems was performed offline by adjusting the minimum difference between the time stamps from eye-tracker and the audio-visual presentation system to be 1 ms. The constant system delay of 20 ms was compensated for (the sum of the buffer size of the sound card (5.8 ms) and the delay of the eye-tracking system (14 ms)).

For each sound localization test, the following variables were stored for subsequent offline analysis (see 10.6084/m9.figshare.7856909): presented AOI (−55 to +55° azimuth), gaze intersected AOI (−55 to +55° azimuth), time stamps in the eye-tracking system and audio-visual presentation systems, respectively before off-line synchronization, and the delay estimated by the eye-tracking system.

Fig. 1. An example recording showing the 4.1-s analysis window for an azimuthal sound shift from +25 to −25° (dashed line). Samples of gaze intersected AOI are shown (open circles) together with its corresponding arctangent fit (solid line). The latency (T) in each trial was defined as the abscissa corresponding to 50% of the arctangent amplitude.
2.4.1. Optimization algorithm and constraints

Accordingly, the abscissa corresponding to 50% of the amplitude of trial (analysis window: 4.1 s/82 samples). The sound was ramped off with a 50 ms raised cosine before change in azimuth and immediately ramped on again after the shift without any other alterations to the sound.

2.4.2. Exclusion of trials

Trials with \( T = 0 \) ms were excluded due to alleged failure in the optimization. Trials with \( T > 1.6 \) s, i.e. longer than the sound-only period, were also excluded.

2.5. Normal and simulated UHL listening conditions

The subjects participated in one normal-hearing (NH) and two simulated UHL conditions. Unilateral attenuation was achieved by an earplug in the right ear (EAR Classic foam ear plug, 3M, Minneapolis, USA) and a combination of an ear plug and a circum-aural hearing protector (Bilsom 847 NST II) placed over the ear plug (Asp et al., 2018). The left ear plug and hearing protector were removed during sound localization tests. The simulated UHL was estimated by hearing thresholds with bilateral attenuation according to standard (ISO 4869-1, 2018). Hearing thresholds were measured by frequency-modulated tones in sound field in an audiometric test room using a fixed-frequency Békésy technique (Asp et al., 2018; Berninger et al., 2014). The mean ± SD simulated UHL (the average threshold of 0.5, 1, 2, and 4 kHz in dB HL) obtained with earplugs was 30 ± 5.9 dB HL, hence denoted SUHL30. The combination of earplugs and hearing protectors yielded a mean ± SD UHL of 43 ± 4.7 dB HL, hence denoted SUHL83, see supplementary material in Asp et al. (2018) for individual hearing thresholds.

2.6. Linear mixed modelling and statistical analyses

The statistical software used was R version 3.4.2 (R Foundation for Statistical Computing, Austria). The T-values in the normal-hearing condition was modeled by the effects of separation between loudspeakers (10°–110°), azimuth after shift (−55°–55°), and time elapsed during test (trial number 1–24). Due to the presence of repeated measures, we used linear mixed modeling (package lme4 version 1.1.21) with subject as random intercept. The accuracy was quantified as the absolute difference between presented AOI and the subjects’ responses. The relationships between localization accuracy (defined as the absolute difference between presented AOI and the responded gaze intersected AOI in degrees), latency (T-values), and SUHL thresholds were investigated by means of repeated measures correlation (package rmcorr version 0.3.0). Since SUHL was assumed to be constant throughout a test (as measured before the test) we aggregated the latency (i.e. average latency), and accuracy per test. The accuracy was quantified as an Error Index (EI), which is a normalized mean absolute error where zero is perfect localization and 1.0 corresponds to random performance, see Asp et al. (2018) for a formal definition. William’s test (package psych version 1.8.12) was used to study the difference between two dependent correlations with one parameter in common.

Differences in the number of included trials and the goodness of the arc tangent fit between the conditions were quantified by means of repeated measures ANOVA or Kruskal-Wallis where appropriate.

Visual inspection of gaze intersected AOI samples motivated the fitting of an arc tangent function as a model of eye movement during an azimuthal sound shift (Fig. 1). The analysis window was 4.1 s (82 samples), starting 2.5 s (50 samples) before each azimuthal sound shift which allowed enough samples for a subsequent optimization of the fit (see below). The endpoint of the analysis window coincided with the reindroduction of the visual stimulus, i.e. 1.6 s (32 samples) after the azimuthal sound shift. The period for the transition of eye gaze from one sound source to another coincides with the period for distinctly increased motor neuron firing rate of the eye muscles (cf. Sparks, 2002, Fig. 1). Accordingly, the abscissa corresponding to 50% of the amplitude of the fitted arc tangent function was defined as the latency \( T \) in each trial (Fig. 1), i.e. halfway through the localization response. SLL for each subject was defined as the mean of the T-values for each test (≤24 trials per subject and condition).

The following formula was used to fit the samples in each trial:

\[
a(t) = a_1 + (a_2 - a_1) \cdot \left( \left( \frac{\pi}{2} + \tan^{-1}(c^*(t - T)) \right) / \pi \right) \tag{1}
\]

where \( a_1 \) and \( a_2 \) (\( c^*(t) \)) are continuous variables (−55° ≤ \( a_1,2 ≤ +55° \)) corresponding to the gaze intersected AOI before and after the azimuthal sound shift. The slope parameter \( c^*(t) \) is a measure of the speed and eccentricity of the trace (0 ≤ c ≤ 130), \( t(t) \) is the time and \( T(t) \) is the latency for each trial (\( T(t) > 0 \)). Non-linear constrained optimization minimized the squared error between Equation (1) and the samples and returned optimal \( a_1, a_2, c^* \) and \( T \) for each trial (analysis window: 4.1 s/82 samples). The fitting of the function was done with open source package SciPy (version 0.16.1, http://www.scipy.org/), Python programming language (version 2.7 available at https://www.python.org/), Python Software Foundation) and the optimization was performed with the bound constraints version of the limited memory Broyden-Fletcher-Goldfarb-Shanno algorithm (Byrd et al., 1995).

An ongoing auditory-visual stimulus (a colorful cartoon movie with an accompanying melody) was presented at 63 dB(A), as measured at the position of the subject’s head. The auditory stimulus was filtered to achieve a long-term frequency spectrum corresponding to the unmodulated noise in the Hagerman sentence recognition test (equal energy in 1/3-octave bands (Hagerman, 1982)), long-term spectrum and an excerpt of the stimulus spectrogram is found in supplements Figure S1. The sound was ramped off with a 50 ms raised cosine before change in azimuth and immediately ramped on again after the shift without any other alterations to the sound.

### 2.3.2. Stimulus

The subject watched the 3-min cartoon starting at the loudspeaker/display-pair at −5° azimuth (i.e. just to the left of midline). During these 3 min 24 azimuthal sound shift 0704s were elicited (trials) with gaps in the visual part of the stimulus, see Asp et al. (2016) for details. The mean inter-trial gap was 7 s. In each trial, the sound shifted instantaneously from the loudspeaker in the current LD-pair (the target in the previous trial) to another randomly assigned loudspeaker (target). The visual stimulus stopped 170 ms before the azimuthal sound shift and was reintroduced 16 s after the azimuthal sound shift on the target loudspeaker/display-pair. The subjects were not asked to perform the task quickly but instructed to follow the auditory-visual stimulus and that the visual stimulus would temporarily pause as the auditory stimulus changed azimuth.

### 2.3.3. Test procedure

The subject watched the 3-min cartoon starting at the loudspeaker/display-pair at −5° azimuth (i.e. just to the left of midline). During these 3 min 24 azimuthal sound shift 0704s were elicited (trials) with gaps in the visual part of the stimulus, see Asp et al. (2016) for details. The mean inter-trial gap was 7 s. In each trial, the sound shifted instantaneously from the loudspeaker in the current LD-pair (the target in the previous trial) to another randomly assigned loudspeaker (target). The visual stimulus stopped 170 ms before the azimuthal sound shift and was reintroduced 16 s after the azimuthal sound shift on the target loudspeaker/display-pair. The subjects were not asked to perform the task quickly but instructed to follow the auditory-visual stimulus and that the visual stimulus would temporarily pause as the auditory stimulus changed azimuth.
3. Results

3.1. Sound localization latency in the normal-hearing condition

The mean ± SD SLL was 280 ± 40 ms (n = 8) (Table 1). No learning effect was observed (test – retest mean ± SD = −10 ± 20 ms). The 95% confidence interval of a single SLL was ±26 ms (SDtest/retest = 1.96). The intrasubject variability in T-values (average SD = 91 ms, n = 8 subjects) was more than twice as large as the SD in SLL.

3.1.1. Fitting of arctangent function

Gaze intersected AOI samples and corresponding arctangent functions are depicted for trials 1, 12, and 24 for each subject in Fig. 2. Nearly all trials (23 of 24 on average) could be modeled by an arctangent function (Table 1). Modeled arctangent functions for all trials (24 trials x 8 subjects = 192) are shown in Supplement Figures S2-S3 available online.

3.1.2. The effect of separation between loudspeakers, azimuth after shift, and time elapsed on T-values

The T-values were not affected by separation between loudspeakers (t(180) = 1.3, p = 0.2), azimuth after shift (t(177) = 0.8, p = 0.4), or time elapsed (t(177) = 0.19, p = 0.8), as demonstrated by a linear mixed model.

3.2. Effect of simulated UHL on sound localization latency

The mean ± SD SLL (n = 8) was 280 ± 40 [95% C.I. = 260; 290], 370 ± 49 [95% C.I. = 350; 390], and 540 ± 120 [95% C.I. = 500; 580] ms in the NH, SUHL30, and SUHL43 conditions, respectively (Table 1). The coefficient of variance in SLL for the NH, SUHL30 and SUHL43 condition was 0.14, 0.13 and 0.22, respectively. A linear mixed model with random intercept for subject showed a statistically significant effect of listening condition on the T-values (χ²(2) = 183, p < 0.0001). No subject showed a decrease in SLL as a function of simulated UHL.

3.2.1. Relationship between sound localization latency and hearing threshold in the experimental ear

SLL across subjects and conditions for each frequency (Fig. 3) revealed a prolongation of SLL as a function of increasing hearing threshold. This was confirmed by adding the hearing thresholds as fixed effects (χ²(6) = 26.4, p < 0.0001), which demonstrated a significant effect of the hearing threshold at 1 kHz on T-values (p < 0.05, Table 2). When the non-significant effects (Table 2) were removed from the model, the threshold at 1 kHz predicted that T-values increased by 7.4 ms/db (t(563) = 14.6, p < 0.0001) and R²GLM(M) = 0.27 (p < 0.0001).

3.2.2. Relationship between sound localization latency and accuracy

There was no correlation between localization accuracy and T-values within a test (repeated measures R² = 0.002, 0.0004, and
0.003 with p's > 0.5 for all three conditions), suggesting no accuracy/latency tradeoff.

However, when taking all conditions into consideration a statistically significant relationship between T-values and sound localization accuracy across listening conditions existed (repeated measures $R^2 = 0.13$, $p < 0.0001$), indicating prolonged T-values and decreasing localization accuracy as the SUHL thresholds increased. The relationship between sound localization latency and accuracy was further studied for the overall latency and accuracy measures (SLL versus EI, Fig. 4). The repeated measures $R^2$ for SLL versus EI was 0.94 ($p < 0.0001$). This value was statistically significantly different from the repeated measures $R^2$ for Error Index versus threshold at 1 kHz (0.77, $p < 0.0001$), and for SLL versus threshold at 1 kHz (0.82, $p < 0.0001$) (William’s test, $t(24) = 2.5$, $p = 0.02$ for both comparisons).

### 3.2.3. Fitting of arctangent function

Fig. 5 shows trials 1, 12, and 24 for each subject in SUHL30 and SUHL43, respectively.

The mean number of trials included in the analysis was 23 of 24 in the SUHL conditions, i.e., similar to the NH condition (Table 1).

Visual inspection of all trials indicated that most exclusions were related to eye-blinks during the analysis window, or unstable eye-tracking, e.g. gaze was shifting between AOs faster than what is physically possible for saccades, see for example supplementary Figure S3, subject (column) 4, trial 21. No significant difference in number of included trials existed between the 3 listening conditions (Repeated measures ANOVA $F(2,14) = 0.099$, $p = 0.9$). The variability (IQR) in T increased slightly in comparison to the NH condition (SUHL30 range: 88–270 ms and SUHL43 range: 130–330 ms). Arctangent functions for all trials (24 trials × 8 subjects × 2 conditions = 384) are shown in Supplementary Figures S4-S7 available online.

### Table 2

Linear mixed model of T-values with hearing thresholds at different frequencies as fixed effects.

| Parameter   | Regression coefficient | Standard error | p-value |
|-------------|------------------------|----------------|---------|
| Intercept   | 317.8                  | 35.8           | <0.0001 |
| 0.5 kHz     | −5.7                   | 5.1            | 0.27    |
| 1 kHz       | 13.0                   | 5.2            | 0.02    |
| 2 kHz       | −2.9                   | 3.1            | 0.38    |
| 3 kHz       | −7.1                   | 4.9            | 0.15    |
| 4 kHz       | 5.2                    | 5.3            | 0.33    |
| 6 kHz       | 2.8                    | 2.3            | 0.21    |
| SUHL30      | −31.9                  | 115.4          | 0.78    |
| SUHL43      | 70.1                   | 143.8          | 0.63    |

Fig. 3. Sound localization latency (SLL) as a function of hearing threshold (dB HL) in the experimental ear for frequencies between 0.5 and 6 kHz. The normal-hearing condition is depicted by green circles, SUHL30 by yellow circles, and SUHL43 dB HL by red circles. Dashed lines connect individual SLL across listening conditions. Vertical error bars show standard error of the mean. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
all conditions: 280 ms, SUHL30, and SUHL43, respectively. There was a statistically significant difference in RMSE between the three conditions [Kruskal Wallis test] was 2.8\(\pm\)0.72; cf.Table 1. Out of the 24 SLL measurements, 7 means were the same in the original analysis as in the subset. The maximum deviation between the original analysis and the subset in SLL was 9\% (subject 4 in SUHL43 and subject 8 in NH condition). In summary, additional exclusion criteria resulted in similar SLL across subjects and condition.

4. Discussion

This study demonstrated the feasibility of a rapid and objective eye-tracking method to assess SLL in the frontal horizontal plane in normal hearing subjects. A sigmoid function was used as a mathematical model to reflect auditory elicited gaze patterns following azimuthal sound shifts. The mathematical approach allowed the fitting of eye gaze data with low errors. The median RMSE ranged from 2.8° in normal binaural conditions to 4.0° in a simulated mild-to-moderate UHL condition. SLL increased as a function of simulated UHL, from 280 ms in the NH condition to 540 ms in the SUHL43 condition.

In studies with subjects’ heads restrained, sound localization latencies ranged between 182 and 300 ms (Frens and Van Opstal, 1995; Shaflig et al., 1998; Zahn et al., 1978). Two studies that allowed unrestrained heads of the subjects reported latencies between approximately 120 ms (n = 2, several hundred trials, Goldring et al., 1996) and 320 ms (n = 5, Populin et al., 2002). In comparison, the present study revealed a mean latency of 280 ms, that is, within the range of reported latencies obtained with heads unrestrained. The lack of effect of separation between loudspeakers on T-values is also consistent with previous studies where no effect of eccentricity was found above 15° (head restrained) (Frens and Van Opstal, 1995; Zambbarieri, 2002). The reaction time of the neck muscles is rarely faster than the eye saccade (Goldring et al., 1996; Goossens and Opstal, 1997) why latencies are expected to be similar with the head unrestrained. The current setup might, however, have allowed for longer processing of auditory cues due to freedom of head movements and a continuous stimulus.

It is known from the visual domain that the otherwise common accuracy/latency tradeoff, where increased time of motion planning increases endpoint accuracy, does not apply to saccadic behavior to visual target sequences (Wu et al., 2010). This is consistent with our findings considering the performance in each trial within a listening condition, that is, no accuracy/latency tradeoff occurred. However, as SUHL became larger there was indeed an increase in latency as well as in accuracy (cf. Fig. 4). The effect of the SUHL on accuracy in the present study was not as drastic as in previous studies (for a review see Kumpik and King, 2019). The main reason for this difference is probably related to the measurement paradigm. Unlike previous studies where subjects were instructed to face 0° azimuth between presentations of sounds of short duration, the subjects in the present study were allowed to move their heads freely throughout the test while localizing a continuous sound that changed azimuth at random time intervals. The 1.6 s sound-only period likely allowed subjects to attend to the stimulus at more than one head angle with respect to the presented azimuth before the final eye gaze position. As such, monaural head-shadow cues probably were more salient compared to cues for a brief stimulus and could have increased accuracy. Furthermore, the visibility of the twelve possible targets might have increased accuracy and decreased latency compared to other setups. However, the use of gaze intersected AOI rather than raw gaze might have decreased the RMSE but should have limited effect on the SLL measure.

The statistically significant relationship between mean latency and accuracy across conditions suggests that increasing the difficulty of an auditory target task (here achieved by unilateral ear plugging) results in increased uncertainty in the integration of
available localization cues. This uncertainty leads to a decrease in accuracy and a corresponding prolongation of the processing in the non-auditory centres involved in the determination of gaze movement. For example, perfect accuracy (EI = 0) corresponded to a latency of 270 ms, whereas poor accuracy (e.g. EI = 0.8) corresponded to a latency of 670 ms. The repeated measures correlation, which accounted for 94% of the variance in SLL, predicted that the mean EI in the SUHL43 dB condition (EI = 0.54, cf. Asp et al., 2018) corresponded to a SLL of 540 ms, which means twice the processing time compared to the normal condition (280 ms; delay in the motor pathway = 20 ms, Fuchs et al., 1985; Zambardi, 2002). In support of this finding, a corresponding statistically significant relationship has been found with the effect of decreased target luminance on visually evoked saccadic latency and accuracy in the frontal horizontal plane (e.g. Warren et al., 2013; Wheeless et al., 1967). Also, head saccades show a substantial increase in latency when one ear received simulated cochlear implant stimulation (Asp et al., 2019). The strong relation between latency and accuracy as a function of increasing task difficulty (i.e. degree of SUHL) remains to be studied in clinical cohorts. It is for example demonstrated that some listeners with permanent UHL of various degrees may achieve relatively high localization accuracy (Firszt et al., 2017; Johansson et al., 2019; Slattery and Middlebrooks, 1994) and daily training with a unilateral ear plug in normal-hearing subjects results in

Fig. 5. The 1st, 12th, and 24th trial of the sound localization test for each subject in the SUHL30 condition (A) and SUHL43 dB condition (B). The dashed lines illustrate the azimuthal sound shifts, the solid lines show the arctangent fits, and the open circles depict gaze intersected AOI samples.
increased accuracy by a reweighting of available localization cues (e.g., spectral cues, loudness cues and binaural cues) (Kumpik et al., 2010). Accuracy and latency of sound localization responses might thus be affected differently than in the acute SUHL conditions in the present study. Hence, the clinical value of SLL remains to be determined, but the importance of a combination of latency and accuracy measures has been pointed out previously (Fitts, 1966; Wickelgren, 1977).

One possible source of increased latency in the SUHL-conditions is the processing of the asymmetric binaural input at the level of the brainstem. There is a large disparity between maturation of central auditory pathways at the cortical level (Sharma et al., 2002) and maturation of sound localization abilities (Asp et al., 2016; 2011; Kumpik and King, 2019; van Deun et al., 2009). Sound localization accuracy matures during the first few years of life (Asp et al., 2016) in a quite similar way as the maturation of the auditory brainstem response (ABR) wave I–V latency (Eggermont and Salamy, 1988), whereas latencies of cortical auditory-evoked potentials decreases into adulthood. In addition, it has been shown that the ABR I–V latency is closely related to sound localization accuracy in children with unilateral hearing loss (Johansson et al., 2019). Thus, the auditory brainstem seems to have an important role for sound localization ability (Crothe et al., 2010) but the development of SLL from early childhood to adolescence has not yet been investigated.

From a physiological point of view, it is interesting to observe the similarity in latency, dynamic behavior, and overlap of anatomical structures between the acoustic middle ear reflex and SLL. At reflex threshold, the latency of the acoustic middle ear reflex for a 500 Hz pure tone is ≥240 ms, decreasing relatively rapidly to ≥120 ms 10 dB above threshold, followed by a slower decrease to ≥100 ms 25 dB above threshold (absissa corresponding to 50% amplitude) (see Borg, 1982, Fig. 1). These dynamic characteristics correspond to latency changes of 12 ms/dB for the first 10 dB above threshold, followed by a decrease of 1.3 ms/dB between 10 and 25 dB above threshold. In the present study, the increase in SLL as a function of simulated UHL threshold was ≥7.4 ms/dB at 1 kHz, that is, in the range of latency changes of the acoustic middle ear reflex in response to varying stimulus level. The similarity in latencies and dynamic characteristics might be due to the overlap of the anatomical structures for the stapedius reflex activity and horizontal sound localization. The neural organization of the acoustic reflex response is 1) auditory afferent, 2) cells in ventral cochlear nucleus with axons in trapezoid body, 3) interneurons in MSO leading to 4) stapedius motoneurons in the facial motor nucleus (Borg, 1973). Correspondingly, for the ITD: 1) auditory afferent, 2) cochlear nucleus (spherical bushy cells), 3) and MSO (Laumen et al., 2016) and for ILD: 1) auditory afferent, 2) cochlear nucleus (spherical bushy cells), 3) LSO, and 4) an inhibitory contralateral path via globular bushy cells and medial nucleus of the trapezoid body (Laumen et al., 2016). The relation between the ABR I–V latency and sound localization accuracy (Johansson et al., 2019), and the similarity in latencies and dynamic behavior between the acoustic middle ear reflex and sound localization responses, indicates that the brainstem has an important role in the processing of auditory spatial information.

While the present study did not quantify the ITD imposed by the simulated UHL, the ear plug used in the present study introduces an ITD (Kumpik et al., 2010). ITDs differing from zero are shown to increase latencies of electrophysiologically recorded postauricular muscle reflexes (Doubell et al., 2018) as well as brainstem responses (Laumen et al., 2016). Furthermore, Polyakov and Pratt (2003) has shown that ABR latencies are sensitive to spatial cues. It seems likely, therefore, that the SLL was affected by ITD besides the large ILD introduced in the simulated UHL conditions, rather than an overall decrease in audibility of the stimulus. Gabriel et al. (2010) showed that a decrease in presentation level from 75 to 55 dB(A) did not affect localization latency, further indicating that overall audibility does not change localization latency.

The corneal reflection eye-tracking technique for the assessment of SLL presented in this study may, though currently costly, be clinically feasible due to the low testing time (~3 min), high test-retest reliability, and objective evaluation of auditory evoked saccadic responses. Horizontal sound localization accuracy has been assessed from 6 months of age with the same technique as in the present study (Asp et al., 2016). Hence, SLL may possibly be determined in infants and young children (Eklof, Asp, & Berninger, presentation at ARO, Association for Research in Otolaryngology, Mid-Winter Meeting, Baltimore, MD, 2019). It remains to be confirmed in future studies in clinical cohorts whether the SLL measure may add information in the assessment of sound localization accuracy, and thereby enhance evaluation of binaural hearing in young ages, possibly at the level of the brainstem.

5. Conclusions

Sound localization latency was successfully determined by fitting an arctangent function to eye-gaze patterns evoked by azimuthal sound shifts in the frontal horizontal plane. The latency increased significantly following simulated UHL. A corresponding increase in sound localization accuracy occurred, resulting in a close relationship between SLL and sound localization accuracy across listening conditions.

Funding

This work was supported by Karolinska University Hospital and the regional agreement on medical training and clinical research (ALF) between Stockholm County Council, Karolinska Institutet, and STAF (Swedish Technical Audiological Society), and Tysta Skolan foundation. The study was presented as a poster during the 40th and 41st Midwinter Research Meetings, Association for Research in Otolaryngology, Baltimore, MD, 2017 and San Diego, CA, 2018.

Declaration of competing interests

The authors do not have any conflicting interests to declare.

Acknowledgments

We wish to thank Åke Olofsson for programming the MatLab interface for localization and Anne-Marie Jakobsson for performing the measurements.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heares.2020.108011.

References

Agerberg, M.J.H., Snik, A.F.M., Hol, M.K.S., Van Wanrooj, M.M., Van Opstal, A.J., 2012. Contribution of monaural and binaural cues to sound localization in listeners with acquired unilateral conductive hearing loss: improved directional hearing with a bone-conduction device. Hear. Res. 285, 9–18. https://doi.org/10.1016/j.heares.2012.02.012.

Asp, F., Eskilsson, G., Berninger, E., 2011. Horizontal sound localization in children with bilateral cochlear implants: effects of auditory experience and age at implantation. Otol. Neurotol. 32 (4), 558–564. https://doi.org/10.1097/MAO.0b013e318218fbd.

Asp, F., Jakobsson, A.-M., Berninger, E., 2018. The effect of simulated unilateral hearing loss on horizontal sound localization accuracy and recognition of
speech in spatially separate competing speech. Hear. Res. 357, 54–63. https://doi.org/10.1016/j.hearsres.2017.11.008.

Asp, F., Olofsson, A., Backman, V., 2016. Corneal-reflection eye-tracking technique for the assessment of horizontal sound localization accuracy from 6 months of age. Ear Hear. 37, e104–e118. https://doi.org/10.1097/AUD.0000000000000235.

Austi, S.A., Backus, B., Agterberg, M.J.H., van Opstal, A.J., van Wanrooij, M.M., 2019. Sound localization in real-time voiced cochlear-implant simulations with normalized head-movements. Trends Hear. 23 https://doi.org/10.1177/2312165919843733.

Borg, E., 1982. Time course of acoustic intra-aural muscle reflex in non-anesthetized rabbits. Acta Otolaryngol. (Stockh.), 93, 161–167. https://doi.org/10.3109/01945998209229557.

Berner, E., Olofsson, Å., Lejon, A., 2014. Click-evoked auditory brainstem responses using time domain cross-correlations between interleaved responses. Ear Hear. 35 (3), 318–329. https://doi.org/10.1016/j.earsur.2014.09.019.

Borg, E., 1973. On the neuronal organization of the acoustic middle ear reflex. Hear. Res. 372, 17–33. https://doi.org/10.1016/0194-5998(02)82910-1.

Doubell, T.P., Alsetrawi, A., Bastawrous, D.A.S., Bastawrous, M.A.S., Daibes, A., Jadalla, A., Schnupp, J.W.H., 2018. The effect of interaural timing on the posterior auricular muscle reflex in normal adult volunteers. PLoS One 13, e0194965. https://doi.org/10.1371/journal.pone.0194965.

Eggermont, J.J., Salamy, A., 1984. Maturational time course for the ABR in preterm and full term infants. Hear. Res. 33, 35–47. https://doi.org/10.1016/0378-5959(85)90119-6.

Feng, A.S., Rogowski, B.A., 1980. Effects of monaural and binaural occlusion on the morphology of neurons in the medial superior olivary nucleus of the rat. Brain Res. 189, 530–534. https://doi.org/10.1016/0006-8993(80)90112-2.

Fitzr, J.B., Reeder, R.M., Holden, L.K., 2017. Unilateral hearing loss: understanding speech recognition and localization variability-implications for cochlear implant candidacy. Ear Hear. 38, 159–173. https://doi.org/10.1097/AUD.0000000000000380.

Fitts, P.M., 1966. Cognitive aspects of information processing: III. Set for speed/accuracy trade-offs in human operators space in the primary superior colliculus. PLoS One 9, e85017. https://doi.org/10.1371/journal.pone.0085017.

Lee, J., Groh, J.M., 2012. Auditory signals evolve from hybrid- to eye-centered coordinates in the primary superior colliculus. J. Neurophysiol. 108, 227–242. https://doi.org/10.1152/jn.00227.2011.

Leigh, R.J., Kennard, C., 2004. Using saccades as a research tool in the clinical neurosciences. Brain 127, 460–477. https://doi.org/10.1093/brain/awh035.

Middlebrooks, J.C., Green, D.M., 1991. Sound localization by human listeners. Ann. Rev. Psychol. 42, 135–150. https://doi.org/10.1146/annurev.ps.42.020191.001515.

Mills, A.W., 1958. On the minimum audible angle. J. Acoust. Soc. Am. 30, 237–246. https://doi.org/10.1121/1.1917043.

O’Neil, J.N., Connelly, C.J., Limb, C.J., Ryugo, D.K., 2011. Synaptic morphology and the influence of auditory experience. Hear. Res. 279, 118–130. https://doi.org/10.1016/j.heares.2011.01.019.

Pichon, M.J., Kranz, S.E., Eckert, M.A., Edwards, B., Hornsby, B.W.Y., Hulse, L.E., Lunde, K., Lutkenmayer, M., Mackersie, C.L., Naylor, G., Phillips, N.A., Richter, M., Rudner, M., Sommers, M.S., Tremblay, K.L., Wingfield, A., 2016. Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). Ear Hear. 37, 55–75. https://doi.org/10.1097/AUD.0000000000000312.

Plamondon, R., Alimi, A.M., 1997. Speed/accuracy trade-offs in target-directed movements. Behav. Brain Sci. 20, 279–303. https://doi.org/10.1017/S0140525X97001441.

Polyakov, A., Pratt, H., 2003. Electrotopographic correlates of direction and elevation cues for sound localization in the human brain. Int. J. Audiol. 42, 140–151. https://doi.org/10.1080/1499202031000079291.

Populin, L.C., 2008. Human sound localization: measurements in untrained, head-unrestrained subjects using gaze as a pointer. Exp. Brain Res. 190, 11–30. https://doi.org/10.1007/s00221-008-1445-2.

Populin, L.C., Tolchin, D.J., Weinstein, J.M., 2002. Human gaze shifts to acoustic and visual targets. Ann. N. Y. Acad. Sci. 958, 468–473. https://doi.org/10.1111/j.1749-6632.2002.tb02876.x.

Shafrq, R., Stuart, G.W., Sandbach, J., Maruff, P., Currie, J., 1998. The gap effect and express saccades in the auditory modality. Exp. Brain Res. 118, 221–229. https://doi.org/10.1007/s002210000263.

Sharma, A., Mironov, S.P., Spahr, A.J., 2002. A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. Ear Hear. 23, 522. https://doi.org/10.1097/00003958-200205000-00003.

Slattery, W.H., Middlebrooks, J.C., 1994. Monaural sound localization: acute versus chronic unilateral hearing loss. Hear. Res. 75, 38–46. https://doi.org/10.1016/0378-5955(94)90012-9.

Sparks, D.L., 2002. The brainstem control of saccadic eye movements. Nat. Rev. Neurosci. 3, 952–964. https://doi.org/10.1038/nrn8686.

Ten Brink, A.F., Nijboer, T.C.W., Van der Stoop, N., Van der Stugel, S., 2014. The influence of vertically and horizontally aligned visual distractors on aurally guided saccadic eye movements. Exp. Brain Res. 232, 1357–1366. https://doi.org/10.1007/s00221-014-3854-6.

van Deun, L., van Wieringen, A., Van den Bogaert, T., Scherf, F., Offeciers, F.E., Van de Werf, J., 2009. Sound localization, sound lateralization, and binaural masking level differences in young children with normal hearing. Ear Hear. 30, 178–190. https://doi.org/10.1097/AUD.0b013e318194256b.

Wahrmann, D.E., Thurtell, M.J., Carroll, J.N., Wall, M., 2013. Perioperative evaluation of saccade latency, saccadic accuracy, and visual threshold for peripheral visual stimuli in young compared with older adults. Invest. Ophthal. Vis. Sci. 54, 5778. https://doi.org/10.1167/iovs.13-12039.

Wheelless, L.L., Cohen, G.H., Boynton, R.M., 1967. Luminance as a parameter of the eye-movement control system. J. Opt. Soc. Am. 57, 394. https://doi.org/10.1117/12.3000394.

Wickelgren, W.A., 1977. Speed-accuracy tradeoff and information processing dynamics. Acta Psychol. (Amst.) 41, 85–87. https://doi.org/10.1016/0001-6918(77)90017-7.

Wu, C-C., Kwon, O.-S., Kowler, E., 2010. Fitts’ Law and speed/accuracy trade-offs during sequences of saccades: implications for strategies of saccade planning. Vis. Res. 50, 2142–2157. https://doi.org/10.1016/j.visres.2010.08.008.

Zahn, J.R., Abel, L., Del’Osso, L.F., 1978. Auditory onco-car response characteristics. Sensory Process. 2, 32–37. https://doi.org/10.1007/BF00229557.

Zamboncini, D., 2002. The latency of saccades toward auditory targets in humans. Prog. Brain Res. 140, 51–59. https://doi.org/10.1016/S0079-6123(02)40041-6.