Hearing Research xxx (2018) 1–7

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

Head shadow enhancement with low-frequency beamforming improves sound localization and speech perception for simulated bimodal listeners

Benjamin Dieudonné*, Tom Francart

KU Leuven — University of Leuven, Department of Neurosciences, Experimental Oto-rhino-laryngology, Herestraat 49 bus 721, 3000, Leuven, Belgium

A R T I C L E I N F O

Article history:
Received 16 January 2018
Received in revised form
27 February 2018
Accepted 6 March 2018
Available online xxx

Keywords:
Head shadow enhancement
Enhancement of interaural level differences
Sound localization
Directional hearing
Speech in noise
Speech intelligibility

A B S T R A C T

Many hearing-impaired listeners struggle to localize sounds due to poor availability of binaural cues. Listeners with a cochlear implant and a contralateral hearing aid — so-called bimodal listeners — are amongst the worst performers, as both interaural time and level differences are poorly transmitted. We present a new method to enhance head shadow in the low frequencies. Head shadow enhancement is achieved with a fixed beamformer with contralateral attenuation in each ear. The method results in interaural level differences which vary monotonically with angle. It also improves low-frequency signal-to-noise ratios in conditions with spatially separated speech and noise. We validated the method in two experiments with acoustic simulations of bimodal listening. In the localization experiment, performance improved from 50.5° to 26.8° root-mean-square error compared with standard omni-directional microphones. In the speech-in-noise experiment, speech was presented from the frontal direction. Speech reception thresholds improved by 15.7 dB SNR when the noise was presented from the cochlear implant side, improved by 7.6 dB SNR when the noise was presented from the hearing aid side, and was not affected when noise was presented from all directions. Apart from bimodal listeners, the method might also be promising for bilateral cochlear implant or hearing aid users. Its low computational complexity makes the method suitable for application in current clinical devices.

1. Introduction

Poor perception of binaural cues is a problem for many hearing-impaired listeners, leading to poor sound localization and speech understanding in noisy environments. Listeners with a cochlear implant (CI) and a hearing aid in the non-implanted ear — so-called bimodal (CI) listeners — are amongst the worst performers, as both interaural time differences (ITDs) and interaural level differences (ILDs) are poorly transmitted (Francart and McDermott, 2013). ITDs are most probably not perceived due to (1) the signal processing in clinical (CI) sound processors which neglects most temporal fine structure information, (2) tonotopic mismatch between electric and acoustic stimulation, and (3) differences between the processing delay of both devices (Francart et al., 2009b; 2011b). ILDs are also poorly perceived because (1) the head shadow is most effective at high frequencies, which are often not perceived in the non-implanted ear due to high-frequency hearing loss, (2) different dynamic range compression algorithms in both devices, and (3) different loudness growth functions for electric and acoustic stimulation (Francart et al., 2009a; 2011a). Moreover, for large angles, the natural ILD-versus-angle function becomes non-monotonic (Shaw, 1974). This means that it is physically impossible to localize sounds unambiguously for all directions with only natural ILDs.

Therefore, several authors have presented sound processing strategies to artificially enhance ILDs, resulting in improved sound localization and improved speech intelligibility in noise. Francart et al. (2009a) have shown improved sound localization in an acoustic simulation of bimodal CI listeners, by adapting the level in the hearing aid to obtain the same broadband ILD as a normal-hearing listener (Francart et al., 2009a; 2013). Lopez-Poveda et al. (2016) implemented a strategy for bilateral CI users, inspired by the contralateral medial olivocochlear reflex. Their strategy attenuated sounds in frequency regions with larger amplitude on the contralateral side, resulting in an increase in speech intelligibility

* Corresponding author.
E-mail addresses: benjamin.dieudonne@med.kuleuven.be (B. Dieudonné), tom.francart@med.kuleuven.be (T. Francart).

https://doi.org/10.1016/j.heares.2018.03.007
0378-5955/

for spatially separated speech and noise. Since both strategies are solely based on level cues that are present in the acoustic signal, they cannot solve the problem of the non-monotonic ILD-versus-angle function. Francart et al. (2011a) adapted their above-mentioned strategy by applying an artificial ILD based on the angle of incidence, to obtain a monotonic ILD-versus-angle function. They found improved sound localization for real bimodal listeners. However, this strategy relied on a priori information about the angle of incidence of the incoming sound. Brown (2014) extended the strategy by estimating the angle of incidence in different frequency regions based on ITDs, resulting in an improved speech intelligibility for bilateral CI users. Moore et al. (2016) evaluated a similar algorithm for bilateral hearing aid users, and found improved sound localization while speech perception was not significantly affected.

All above-mentioned strategies try to artificially impose an ILD based on estimations of auditory cues that are already present. Unfortunately, these estimations are either suboptimal (if based on non-monotonic ILD cues) or computationally expensive (if based on ITDs). Moreover, they can only handle multiple sound sources if these sources are temporally or spectro-temporally separated, while the spectrograms of multiple concurrent speakers most likely have some overlaps. Recently, Veugen et al. (2017) tried to improve the access to high-frequency ILDs for bimodal listeners without the need for estimations of auditory cues, by applying frequency compression in the hearing aid. However, they did not find a significant improvement in sound localization. Moreover, frequency compression might result in undesired side-effects on speech intelligibility, sound quality, envelope ITDs and interaural coherence (Simpson, 2009; Brown et al., 2016).

In this paper, we present and validate a novel method to enhance low-frequency ILDs without the need of estimations of auditory cues or distorting the incoming sound. We enhance the head shadow by supplying each ear with a fixed bilateral electronic beamformer applied in the low frequencies, attenuating sounds coming from its contralateral side (as opposed to conventional fixed (unilateral or bilateral) beamformers that attenuate sounds coming from the rear side). This results in enhanced low-frequency ILDs and resolves non-monotonic ILD-versus-angle functions. Because of its low computational complexity, our method is suitable for application in current clinical devices. As a proof-of-concept, we validate the effect of head shadow enhancement on localization and speech perception in noise for simulated bimodal listeners.

2. General methods

2.1. Head shadow enhancement

In the low frequencies (below 1500 Hz), the ear naturally has an omni-directional directivity pattern, resulting in very small ILDs (Moore, 2012, Chapter 7). We enhanced this directivity pattern with an end-fire delay-and-subtract directional microphone applied below 1500 Hz. In each ear, the beamformer attenuated sounds arriving from its contralateral side. Above 1500 Hz, we did not apply any beamforming.

To achieve contralateral attenuation in each ear, a linear microphone array in the lateral direction was realized with a data link between the left- and right-ear devices, as illustrated in Fig. 1 (a). The low-frequency gain was boosted with a first-order low-pass Butterworth filter (cutoff at 50 Hz), to compensate for the 6 dB/octave attenuation of the subtractive directional microphone (Dillon, 2001, Chapter 7).

In this set-up, the microphone spacing equals the distance between the ears, approximately 20 cm. Such a large microphone spacing yields good sensitivity of the directional microphone in low frequencies (note that a frontal directional microphone in a behind-the-ear (BTE) device is usually not active in low frequencies because of its strong high-pass characteristic (Ricketts and Henry, 2002)). On the other hand, this large spacing decreases the sensitivity at frequencies above approximately 800 Hz due to the comb filter behavior of a subtractive array (Dillon, 2001, Chapter 7): the first null in the comb filter would appear at 850 Hz when considering a microphone distance of 20 cm and a sound speed of 340 m/s, the second null at 1700 Hz, etc. This comb filtering behavior also affects the directional pattern of the beamformer. Since we only enhanced the head shadow for frequencies below 1500 Hz, the directional pattern and frequency response were not strongly affected by the comb filtering.

In Fig. 1(b) it can be seen that the method results in a cardioid-like directivity pattern for low frequencies, while the natural directivity pattern of the ear remains unchanged for frequencies above 1500 Hz. The directivity patterns are calculated as the spectral power in the respective band with a white noise signal as input to the algorithm.

2.2. Simulations of spatial hearing

Spatial hearing was simulated with head-related transfer functions (HRTFs). We measured the response of an omni-directional microphone in a BTE piece placed on the right ear of a CORTEX MK2 human-like acoustical manikin; for each angle, the left-ear HRTF was obtained by taking the HRTF from the right ear for a sound coming from the opposite side of the head (e.g., the left-ear HRTF for a sound coming from -60° equaled the right-ear HRTFs for a sound coming from $+60^\circ$). The manikin was positioned in the center of a localization arc with radius of approximately 1 m, with 13 loudspeakers (type Fostex 6301B) positioned at angles between -90° (left) and $+90^\circ$ (right) in steps of 15°. To also obtain HRTFs for sounds arriving from behind the head, we performed a second measurement in which the manikin was rotated 180°. To simulate an anechoic response, reflections were removed by truncating each HRTF after 2 ms starting from its highest peak.

2.3. Simulation of bimodal cochlear implant hearing

We simulated bimodal CI hearing according to the methods of Francart et al. (2009a).

CI listening was simulated in the left ear with a noise band vocoder to mimic the behavior of a CI processor: the input signal was sent through a filter bank; within each channel, the envelope was detected with half-wave rectification followed by a 50 Hz low-pass filter; this envelope was used to modulate a noise band of which the spectrum corresponded to the respective filter; the outputs of all channels were summed to obtain a single acoustic signal. In the localization experiment (Experiment 1), the vocoder contained 8 channels, logarithmically spaced between 125 Hz and 8000 Hz. In the speech perception experiment (Experiment 2), we lowered the number of channels to 5 (also logarithmically spaced between 125 Hz and 8000 Hz) to obtain worse speech perception, i.e., to better correspond with real CI listening. The number of channels did not have an influence on the head shadow enhancement algorithm, as both vocoders had the same effect on the long-term spectrum of any input signal.

Severe hearing loss was simulated in the right ear with a sixth order low-pass Butterworth filter with a cutoff frequency of 500 Hz, such that the response rolled off at -36 dB per octave. This corresponds with a ski-slope audiogram of a typical bimodal CI listener.

In this simulation with a vocoder in one ear and a low-pass filter in the other ear, little to no ITD cues could be used to localize