Enhancement of interaural level differences improves sound localization in bimodal hearing

Tom Francart,^{a)} Anneke Lenssen, and Jan Wouters

ExpORL,

Dept. Neurosciences,

Katholieke Universiteit Leuven,

O & N 2,

Herestraat 49 bus 721,

B-3000 Leuven,

Belgium

(Dated: August 25, 2011)

Abstract

Users of a cochlear implant together with a contralateral hearing aid - so-called bimodal listeners - have difficulties with localizing sound sources. This is mainly due to the distortion of interaural time and level difference cues (ITD and ILD), and limited ITD sensitivity. An algorithm is presented that enhances ILD cues. Horizontal plane sound-source localization performance of 6 bimodal listeners was evaluated in (1) a real sound field with their clinical devices, (2) in a virtual sound field, under direct computer control and (3) in a virtual sound field with ILD enhancement. The results in the real sound field did not differ significantly from the results in the virtual field, and ILD enhancement improved localization performance by 4-10 degrees absolute error, relative to a mean absolute error of 28 degrees in the condition without ILD enhancement.

PACS numbers: 4364Me, 4366Qp, 4366Pn, 4366Ts

I. INTRODUCTION

Due to the success of cochlear implants (CIs), implantation criteria are changing and an increasing number of patients have residual hearing in the ear contralateral to their CI. In many cases these patients use a hearing aid (HA) together with their CI. This is called bimodal stimulation. The additional use of a HA has been shown to be advantageous in several areas, such as music appreciation, sound quality, speech perception, and, sometimes, sound source localisation (for a recent review, see Ching *et al.* (2007)). Since localization in the horizontal plane is mainly based on binaural cues, acoustic stimulation of the non-implanted ear would be expected to be advantageous. However, despite the binaural stimulation, the advantages for localization are limited.

In normal-hearing (NH) listeners, sound-source localization in the horizontal plane is mainly governed by two binaural cues: interaural time and level differences (ITD and ILD). These cues can be supplemented by monaural spectral cues (Wightman and Kistler, 1997). ILDs are due to the head shadow effect, i.e., the attenuation when the head is situated in between the sound source and the ear. Due to the acoustic properties of the head and pinnae, the ILD is strongly dependent on frequency. Generally, the ILD increases with increasing frequency (Feddersen *et al.*, 1957). For sound sources in the far field (i.e., further away than approximately 1 m), ILDs are considered useful for localization at frequencies higher than about 1500 Hz (Moore, 2003). ITD cues are mainly functional at lower frequencies, because as soon as half of the wavelength of the sound equals the distance between the eardrums, the ITD becomes ambiguous. If both ITD and ILD cues are available, the ITD cues prevail (Wightman and Kistler, 1992).

Localization performance using only a single CI is very poor. Grantham *et al.* (2008) tested six CI listeners and found very poor performance (a localization error of at least 40 degrees) for three listeners and near-chance performance for three others, whereas NH listeners obtained an average score of 5.6°. Various studies of bimodal listening, however,

^{a)}Tom.Francart@med.kuleuven.be

have shown that localization performance does improve in some listeners, though in general only very slightly (Dunn *et al.*, 2005; Seeber *et al.*, 2004; Ching *et al.*, 2004). Different parameters of the test setup such as number of loudspeakers, response paradigm, visual cues and acoustic characteristics of the room can have large influences on localization results. As different test setups, fitting procedures and localization error measures were used in these studies, we can only compare the results qualitatively. Dunn *et al.* (2005) reported that 2 out of 12 bimodal listeners could localize sounds. Seeber *et al.* (2004) tested 11 subjects and reported that 1 subject showed very good localization performance, 2 performed above chance level, 4 could only discriminate the left and right side and 4 showed no localization ability at all. The 4 subjects with the best residual hearing performed best on the localization tasks. Ching *et al.* (2004) tested 18 adults and reported that 12 showed benefit with the addition of a contralateral HA but 6 did not.

While some bimodal listeners are sensitive to ITD cues in the laboratory (Francart *et al.*, 2009a, 2011), these cues are probably not available from their clinical devices because (1) fine timing cues are not preserved by current clinical speech processors, (2) there is no overall synchronization between the CI and HA, and (3) ITD perception for bimodal listeners depends on precise loudness balance, which is not satisfied with current CI and HA devices.

Bimodal listeners are sensitive to ILD (Francart *et al.*, 2008a), and ILD cues are preserved by the CI signal processing, but are distorted or cannot be used in current bimodal systems for several reasons: (1) automatic gain control circuits operate independently at the two sides and therefore reduce ILD cues (Keidser *et al.*, 2006), (2) ILD cues are physically largest at frequencies higher than about 1500 Hz, while the residual hearing of many bimodal listeners only extends up to about that frequency, leading to an impairment-dependent ILD that is smaller than the broadband ILD and (3) while for narrowband stimuli the binaural bimodal equal loudness function can be linear (Francart *et al.*, 2008a), generally loudness growth functions are different at the two ears. This means that a sound presented in front of a subject may be balanced at a certain intensity but the sound image may be shifted to

a side at another intensity. Localization with only ILD and spectral cues is possible (Francart, 2008; Seeber and Fastl, 2008), but performance depends on the magnitude of the ILD available in each particular stimulus, which is largely dependent on the spectral content. Not only can the magnitude of the ILD cue be small, but typically the ILD-versus-angle function is also non-monotonic for angles larger than around 60° (Shaw, 1974).

In a previous study (Francart *et al.*, 2009b), we have shown that these problems can be partly remedied by artificially enhancing ILD cues, i.e., introducing artificially large ILDs at low frequencies. The localization performance of six NH subjects was measured through simulations of bimodal hearing, using a noise band vocoder to simulate the CI side and a low-pass filter to simulate the hearing impairment at the other side. In one condition the stimuli were used without further processing and in another condition, the level at the non-CI ear was amplified/attenuated by half of the original value of the broadband ILD in dB. In this way the ILD is enlarged, but stays stimulus-spectrum-dependent and the ILD-versus-angle function stays non-monotonic.¹ Two test signals were used: a broadband white noise and a telephone alerting signal. We found that the listeners fairly quickly learned to use the artificially enlarged ILD cues. For the telephone alerting signal no significant improvement in localization performance was found but for the white noise signal performance improved significantly by 14 degrees RMS error, relative to 48 degrees RMS error in the reference condition. The lack of improvement for the telephone alerting signal was due to the fact that the ILD cues naturally present in this signal were fairly small and were therefore still not very large after application of the ILD enhancement algorithm.

In the current study, we implemented a new ILD enhancement algorithm based on the same principles, but made it function independently of the spectral content of the signal (and thus the ILD of the original stimulus) and eliminated the non-monotonicity in the ILD-versus-angle function. This was done by determining the angle of incidence of the sound and selecting the appropriate ILD from an artificial ILD-versus-angle function. The algorithm further took into account each subject's own impairment-dependent ILD..

We compared the localization performance of 6 bimodal listeners (in contrast to NH

listeners tested in the previous study) in three conditions. In a first condition (termed “SF”), the subjects’ own CI and HA were used, while in two other conditions (termed “VF-NoEnh” and “VF-Enh”) preprocessed stimuli were presented through experimental devices under direct computer control. The latter gave us control over the exact stimulation patterns. In the VF-Enh condition, ILD enhancement was applied, and in the VF-NoEnh condition it was not. In both conditions, the cues for spatial hearing were simulated with non-individualized head related transfer functions (HRTFs) so giving a virtual auditory field, a technique frequently used to study localization in NH listeners (Wightman and Kistler, 1989). To reduce testing time, only one test signal was used. The telephone alerting signal was chosen because (1) some subjects did not like the sound of a white noise signal and had problems paying attention to it, probably due to the absence of temporal modulations, (2) it allows us to assess whether the new algorithm solves the problems associated with the previous one, and (3) it is a more ecologically relevant signal than a white noise. Prior to the localisation tests, loudness was balanced between the electric and acoustic stimuli used. For the VF-Enh condition the bilateral equal loudness function was measured and used to compensate for differences in loudness growth between the modalities.

II. METHODS

A. Subjects

Six subjects were recruited from the clinical population of the University Hospital Maastricht (AZM) and the University Hospital Leuven (UZLeuven). They were volunteers and signed an informed consent form. This study was approved by the local medical ethical committees. They used a CI manufactured by Cochlear (TM) and a HA contralateral to their CI on a daily basis. S12 had a Contour Advance electrode array; the other subjects had a Contour array. S07 used an Esprit3G clinical processor, the other subjects used a Freedom processor. Their unaided pure tone audiograms are shown in figure 1. Detailed subject information is listed in table I. S12 also participated in pilot experiments for the

FIG. 1. Unaided pure tone audiograms of the non-implanted ear for the subjects that participated in the current study. Note that the vertical axis starts at 60 dBHL.

current study. For S09 the 8 most basal electrodes were disabled due to a partial CI insertion. Consequently her processor's electrode-frequency allocation was different and the most apical electrode stimulated a less apical region in the cochlea than in the other subjects.

The subjects came to the lab for at least four test sessions of 90 minutes. They all had previous experience with psychophysical experiments.

B. ILD enhancement algorithm

The aim of the ILD enhancement algorithm was to improve sound source localisation by enhancing ILDs based on the angle of incidence of the sound. It worked by (1) determining the angle of incidence, (2) estimating the subject’s impairment-dependent ILD, and (3) amplifying or attenuating the signals based on a artificial level-versus-angle function. The ILD-versus-angle function is obtained by subtracting the level-versus-angle functions of the two ears.

In a real-time processor the angle of incidence would be estimated based on the microphone signals. However, as in our current setup the angle of incidence was known, part (1) was skipped. Part (2) and (3) of the algorithm are described in the sections below.

1. *Estimating the impairment-dependent ILD*

Depending on the signal’s spectral content and the subject’s residual hearing, a (small) ILD may be present. We need to estimate it to be able to determine the extra gain to be applied to reach the target ILD. The impairment-dependent ILD was estimated as the difference between the broadband root mean square (RMS) level at the CI side and the RMS level at the acoustic side after low-pass filtering. Using individual settings of the cut-off frequency of the low-pass filter (6th-order Butterworth), this method takes into account the main parameter of the residual hearing of the subjects, namely the cut-off frequency of their high-frequency dead region (Moore *et al.*, 2010). It was estimated based on the audiogram as the frequency above which pure tone thresholds were higher than around 100 dBHL. The cut-off frequency used for each subject is given in table II. This could easily be implemented in a speech processor.

2. *Level-versus-angle function*

The objective is to provide ILD cues that are clearly perceptible and vary monotonically with angle of incidence. Therefore the difference in ILD between two angles that should be

FIG. 2. Broadband ILD present in the telephone alerting signal after filtering with HRTFs. In the conditions labelled 1000 and 2000 Hz, the left-ear signal was filtered with a low pass (LP) filter with these respective cut-off frequencies. This is a simulation of high-frequency hearing loss at the acoustically stimulated side (left).

distinguishable should be larger than the subject's just noticeable difference (JND) in ILD and the function should vary monotonically with angle.

In figure 2 (full black line), broadband average unweighted ILDs are shown as a function of angle of incidence for the telephone alerting signal filtered with the HRTFs. In this case, the ILD-versus-angle function is non-monotonic from about 60°. Therefore listeners who

only have access to ILD cues could in this case not consistently estimate the direction of stimuli at angles larger than 60° . When low-pass filtering the signal destined for the acoustically stimulated left ear (the dashed lines in the figure), thus simulating high-frequency hearing loss, the problem becomes worse because ILD cues are mainly present in the higher frequencies. In this case, after filtering the left-ear signal with a cut-off frequency of 1000 Hz (dash-dotted line), ILD cues became much smaller for stimuli from the right hand side.

Some ILD-versus-angle functions for different devices and signals are shown in figure 3. The function depends on the anatomy of the particular head and torso, the shells and microphones of the devices that were used (Cf. difference between the BTE HA / White Noise and Freedom / White Noise conditions), and the frequency content of the signal (Cf. Freedom / White Noise versus Freedom / Telephone). Most functions are monotonic up to around 60° and flatten and/or become non-monotonic for larger angles. If ITD cues are not available, as is the case with current CIs, localization of sound sources beyond 60° will be poor.

We designed an artificial level-versus-angle function, to be used irrespective of stimulus spectrum, based on the natural broadband level-versus-angle function for the Freedom device (see figure 4). The natural function was obtained by measuring the level of a white noise signal filtered by the HRTF of the Freedom device placed on an artificial head. The artificial function followed the Freedom function from 90° down to -30° and then smoothly decreased to the maximal head shadow by shape-preserving piecewise cubic interpolation. The maximal difference between 0° and -90° was set to 11 dB, leading to a maximal ILD of 12.4 dB, which is a compromise between sufficiently large ILD cues and available dynamic range. The resulting artificial ILD-versus-angle function is shown in figure 3. In a previous study (Francart *et al.*, 2008a), we found that the JND in ILD of 10 bimodal listeners varied between 1 and 3 dB with a median of 1.7 dB. With a JND of 1.7 dB, and for a linear ILD-versus-angle function, about 7 discrete steps could be discriminated for a maximal ILD of 12.4 dB, which given that we use 7 loudspeakers at each side, would be sufficient for the current localization task. However, in practice, JNDs for some subjects were higher and

FIG. 3. ILD-versus-angle functions for a white noise stimulus with a BTE HA shell and with the Freedom device, for the telephone alerting signal used in the current study with the Freedom device, and for the artificial ILD function used in the current study. All measurements were made on the same artificial head.

angle discrimination performance is known to deteriorate for larger angles (Yost and Dye, 1988).

The ILD of the stimulus was set to the desired value by amplifying or attenuating the signals at the input of the CI or HA signal processing, by an amount equal to the difference between the determined level (Cf. section II.B.1) and the target level from the level-versus-angle function.

FIG. 4. Level-versus-angle function for a white noise sent through the Freedom device, derived from the same data as shown in figure 3, and artificial levels used in the current study.

C. Apparatus

The subject was seated in a chair in the middle of an array of loudspeakers placed at a distance of 1 m from the subject. The chair was adjusted such that the cones of the loudspeakers were at ear height. A total of 13 identical speakers were positioned at 15° intervals, spanning 180° in front of the subject, from -90 to $+90^\circ$. The loudspeakers were labelled from left to right with the numbers 1 to 13.

In the sound-field condition (SF), active loudspeakers (Fostex 6301B) were used and connected to two 8-channel sound cards (RME Multiface II). The subjects' clinical speech processor and HA were used with unmodified settings.

In the virtual-field conditions (VF-NoEnh and VF-Enh), preprocessed stimuli were presented under direct computer control. Electric stimuli were presented using an L34 experimental speech processor, controlled by the NICv2 interface provided by Cochlear. Acoustic stimuli were presented using an RME Multiface II sound card, using either Etymotic ER-3A insert phones or a ReSound Viking HA transducer with audio input cable. The latter was used if sufficient levels could not be achieved with the insert phones. The frequency response of the Viking HA transducer with audio input cable was measured using a Bruel&Kjaer 4152 artificial ear and equalized at the end of the processing chain by inverse filtering. The insert phones had a flat frequency response (± 3 dB) in the relevant frequency range (500 – 3000 Hz). It was verified that the output sound pressure level increased linearly and at equal rate as the input signal level for all levels used in the current study.

Synchronization of the electric and acoustic signals was ensured by connecting the second output channel of the sound card to the trigger-in connection of the experimental speech processor. The drift between the clocks of the speech processor and the sound card was measured and compensated for. The electric signal was delayed by an additional 1.5 ms to approximately compensate for the acoustic travelling wave delay (Francart *et al.*, 2009a).

Both the balancing and localization experiments were controlled by the APEX 3 program, developed at ExpORL (Francart *et al.*, 2008b). The stimuli were created using Matlab.

D. Stimulus

The stimulus used in this study was the alerting signal of an old-fashioned telephone. It was 250 ms long, ramped on and off over 50 ms; its spectrum is shown in figure 5. The main acoustic energy is situated between 500 and 3000 Hz with a peak just below 1000 Hz. This signal was either presented unprocessed through one of 13 loudspeakers at 60 dBA, or used

FIG. 5. Spectrum of the telephone alerting signal used.

as the input signal for the processing in the virtual field conditions VF-NoEnh and VF-Enh.

In the VF-NoEnh and VF-Enh conditions, the input signal (Cf. section II.D) was filtered by the left and right HRTFs corresponding to the target angle of incidence. HRTFs were measured in the same room as where the tests took place by placing a Cochlear (TM) Freedom speech processor on an artificial head (Cortex MK2, 01dB-Metravib) and measuring its impulse response for all angles used in the current experiments. Signals were recorded from the audio-output socket of the processor, which is normally used by parents or clinicians to check the function of the microphone.

In the VF-NoEnh condition, the signal was first filtered by the left and right HRTFs corresponding to the target angle of incidence. If the Viking HA transducer was used, the signal to be presented at the acoustic side was additionally filtered with its inverse transfer function, and the signal for the CI side was processed according to the ACE strategy, as implemented in the Nucleus Matlab Toolbox. The Nucleus Matlab Toolbox accurately simulates the processing in the clinical devices without preprocessing such as AGC or noise reduction. The input level to the Nucleus Matlab Toolbox was set to a level corresponding to 60 dBA for a sensitivity setting of 12 of a Freedom processor, and slightly adjusted in the first phase of the loudness balancing procedure (see section II.E.1). The AGC in the clinical devices provides linear amplification with limiting, so should not distort level cues for sounds at a comfortable level.

In the VF-Enh condition, the same processing as in the VF-NoEnh condition was done, but ILD enhancement as described in section II.B was added. Additionally, the HA block now incorporated the equal loudness function (ELF) to correct the level at the acoustic side according to the ELF measured in the loudness balancing procedure.

Note that no compression/AGC was implemented in the acoustic signal path, in contrast to some of the clinical devices used in the SF condition.

E. Procedure

Localization experiments were performed in three different conditions: sound field (SF), virtual field without ILD enhancement (VF-NoEnh) and virtual field with ILD enhancement (VF-Enh). Before the localization experiments, loudness balancing was performed.

1. Loudness balancing

In the SF condition, the subject's clinical CI speech processor and HA were used. They were balanced in loudness by centering the sound image for the stimulus by letting the subjects adjust the loudness controls of their devices. Other than that, the fitting of the

devices was kept unmodified. During balancing the stimulus was always presented from loudspeaker 7 (in front of the subject). First the CI was turned off, and the HA was adjusted such that the stimulus was perceived at a comfortable level. Then the HA was turned off, and the CI's sensitivity control was adjusted such that the stimulus was perceived at a comfortable level, similar to the level perceived with only the HA. Finally the two devices were turned on and the subject freely adjusted them to achieve a balanced auditory sensation for a stimulus at 0° .

In the VF-NoEnh and VF-Enh conditions, preprocessed stimuli were presented through experimental devices under computer control. To take into account differences in loudness growth between the ears, the bilateral ELF was determined using the telephone alerting signal, processed as described in section II.D with HRTFs for 0° . Different stimulus intensities were obtained by varying the intensity of the signal at the input of the processing chain. First threshold and comfort levels were determined for the acoustic and electric signals separately. Then several levels were chosen for the electric signal to encompass most of the dynamic range (three to six, depending on the subject's dynamic range), and the subject was asked to balance the level of the acoustic signal with each of them. In one balancing run, the sound level at the CI side was kept the same and the subject adjusted the acoustic level using a touch screen in steps of 2 and 5 dB. The subjects were instructed to adjust the level such that the sound came from straight ahead or at least sounded centred in the head. They were encouraged to try different balances such that at least once per run the sound was perceived at the left and at least once at the right hand side. Balancing was performed at least twice² for each selected electric level and the start level of the acoustic signal was chosen such that it sounded either much softer or much louder than the electric signal, such that a potential bias due to start level was minimized. When for every selected electric level the corresponding acoustic levels were obtained, a linear fit was performed to find the ELF. As an example, the result from the balancing procedure for S12 is shown in figure 6, the slope and intercept were 0.75 and -24 dB respectively.

For subjects S02 and S09 the measured slope was rather small (0.53 and 0.50), which

FIG. 6. Loudness balancing results for S12. The circles indicate individual results from single runs of the loudness balancing procedure. The full line indicates average results per electric level and the dashed line is the result of a linear fit.

is probably related to their very limited residual hearing (see figure 1). In preliminary localization experiments with the shallow slopes, no benefit of ILD enhancement was found. This was probably due to the fact that a small slope of the ELF led to small level differences at the acoustic side. For these two subjects, we increased the slope steepness to the values shown in table II, such that the resulting ELF crossed the monaurally determined threshold and comfort levels. Most subjects could easily and efficiently do the task during the balancing

experiment. For S23 the experimenter controlled the procedure, as the subject could not do it consistently himself. For each subject a straight line was fitted through the resulting data points, with the sound pressure level in dB at the electric side on the horizontal axis and the sound pressure level in dB at the acoustic side on the vertical axis (see figure 6). The slope, intercept and R^2 shown in table II were obtained. The slope indicates the acoustic level adjustment (dB) needed to match a 1 dB change in level applied at the input of the CI processor. The intercept indicates the general balance between the acoustic and electric side, which is a function of the subjects' residual hearing thresholds.

In the VF-NoEnh condition, input levels were set such that for the stimulus from 0° the electric signal was at a comfortable level and the acoustic signal was at the corresponding level according to the ELF. For stimuli from other angles, levels were governed by the HRTFs used, i.e., the ILDs present in the stimuli.

In the VF-Enh condition, levels for the stimulus from 0° were set similarly to the VF-NoEnh condition, but for other angles the level was determined by calculating the desired level (from the level-versus-angle function) and converting it to an acoustic level using the ELF (see section II.D). That is, loudness growth at the electric side was unmodified and loudness growth at the acoustic side was changed to match the electric side, taking into account the measured ELF.

2. Localization

During the localization experiments, the subjects were instructed to look straight ahead during stimulus presentation, and after stimulus presentation look at the apparent sound-source location and say the number indicated on the loudspeaker. In the SF condition, sounds were played from the loudspeakers, and in the VF-NoEnh and VF-Enh conditions sounds were presented through the experimental speech processor and acoustic transducer.

After each presentation, feedback was given by turning on a LED on top of the correct speaker for 2 s. During a single run, a sound was presented 3 times from each speaker, in

random order. Levels were randomly roved over a range of ± 2.5 dB.

For the VF-NoEnh and VF-Enh conditions, four runs were performed to give each subject a chance to adapt to the new ILD cues. The conditions were presented in random order, but runs of one condition were not interleaved with those of other conditions, to avoid confusion of the different cues.

The main performance measure used was the absolute error. It is defined as the average absolute difference between stimulus angle and response angle in degrees, for each stimulus presented during a run ($\frac{\sum_i |R_i - S|}{N}$, with R_i the response angle in degrees for a certain trial, S the stimulus angle in degrees, and N the total number of trials). Completely random responses in our setup would give an absolute error of 64 degrees. We also used the bias, which is the difference between the average response and stimulus angle ($\frac{\sum_i R_i}{N} - S$). The absolute error includes the uncertainty of the listener while the bias only indicates the difference between the actual source location and the average location perceived by the listener. Both measures are calculated per angle of incidence and are then averaged over multiple angles

As unfamiliar HRTFs and ILD cues were introduced in the virtual field conditions, we assessed possible learning effects in the results³. Visual inspection of result tracks indicated a possible small learning effect after the first run. We therefore considered the first run for the VF-NoEnh and VF-Enh conditions to be a training run and excluded it from further analysis.

III. RESULTS

The subjects generally indicated that the virtual-field stimuli sounded very similar to the SF stimuli. Most of them found localization in the SF condition very challenging. All subjects reported a fused percept in all conditions.

The results for each subject are shown in figure 7. The plots indicate the response as a function of angle of incidence, averaged over 2 runs (SF) or 3 runs (VF-NoEnh and VF-Enh).

FIG. 7. Localization results for each subject in the different conditions. The average response as a function of angle of incidence is shown and the total length of each error bar indicates twice the standard deviation of the response.

The difference between the average response and the stimulus angle is called the bias. The error bars indicate the standard deviation of the responses at each angle, which is a measure of consistency of the answer. As an example, we will discuss the results of S12 and S13, who were respectively the best and worst performers in all conditions. In the SF condition, S12 experienced differences in perceived azimuth up to around 30° , but for larger angles the response-versus-stimulus function flattened, indicating that he could not hear differences between those angles. This flattening is closely related to the ILD-versus-angle function for the Freedom device with telephone signal, as shown in figure 3. Note that the error bars, which indicate the consistency of the answers, are relatively small (given the sometimes large bias). In the VF-NoEnh condition, the bias is reduced and the range over which mean responses vary monotonically with azimuth is increased from about 30 to 45 degrees°. In the VF-Enh condition, the flattening has nearly disappeared and the bias is much smaller for angles beyond 45 degrees. The same tendencies are present for S13, but as the error bars are much larger, the overall effect is less clear. Note that the size of the error bars is not reduced by enhancing ILD cues⁴.

In figure 8 the average absolute error is shown per subject and condition, averaged over 2 runs (SF) or 3 runs (VF-NoEnh and VF-Enh). The absolute error averaged over the six subjects was 31.5° , 28.4° and 20.6° respectively for conditions SF, VF-NoEnh and VF-Enh. The bar labelled “NH” gives the average absolute errors in SF for a total of 23 runs of the same condition with 12 NH listeners, collected in a previous study (Francart, 2008). For S09, performance was not measured in the SF condition because the stimulus was not audible through the HA. For the other subjects a repeated measures ANOVA was performed with factors of condition (SF vs. VF-NoEnh) and repetition. There was no statistically-significant effect of condition ($p = 0.99$) nor of repetition ($p = 0.85$).

Figure 9 shows the difference in absolute error per angle between the VF-NoEnh and VF-Enh conditions, averaged over all subjects. Before averaging, the angles were arranged such that the CI was on the left hand side. Positive values indicate that performance was better in the VF-Enh condition. It is clear that the main performance improvement was for angles

FIG. 8. Average absolute error for each subject and condition. The total length of each error bar indicates twice the standard deviation. The bar labelled “NH” shows the average result of 23 runs of the SF condition in NH subjects (Francart, 2008).

greater than 45° . A second repeated measures ANOVA was performed to compare the two VF conditions; the factors were condition (VF-NoEnh vs. VF-Enh) and repetition. There was no statistically-significant effect of repetition nor of the interaction between repetition and condition, but there was a significant effect of condition: performance for VF-Enh was significantly better than for VF-NoEnh ($F(1; 25) = 33.15; p < 0.01$). The improvement across individuals ranged from 4 to 10 degrees with a mean of 8 degrees.

FIG. 9. Difference in absolute error per angle between the VF-NoEnh and VF-Enh conditions, averaged over all subjects. Positive values indicate that performance was better in the VF-Enh condition. The total length of each error bar indicates twice the standard deviation.

IV. DISCUSSION

We measured the localization performance of 6 bimodal listeners in a sound field condition with their own devices (SF), and two computer-controlled conditions: a virtual field condition without ILD enhancement (VF-NoEnh) and a virtual field condition with ILD enhancement (VF-Enh). ILD enhancement was achieved by changing ILD cues based on the angle of incidence of the sound and an artificial ILD-versus-angle function. The effect

of the ILD enhancement algorithm was assessed by comparing between the results of the VF-NoEnh and VF-Enh conditions.

Performance was very subject-dependent in all conditions. This reflects a combination of effects such as differences in sensitivity to ILD, differences in ability to relate the ILD to spatial location, and the amount of residual hearing. In the SF condition, additional differences between subjects can be caused by differences in devices used, since the subjects' clinical devices were used with unmodified fitting. The position of the microphone and the pre-processing in the subjects' own HAs, particularly the automatic gain control (AGC), can have influence on the resulting spectral and ILD cues.

Comparing between the SF and VF-NoEnh conditions, overall performance did not differ significantly, which validates the simulation used in the virtual-field conditions. Differences for individual subjects could have been caused by differences between the devices used: in the VF-NoEnh condition the HRTFs of the Freedom device were used for the signals at the two ears, which may have resulted in large spectral differences compared to the subject's own HAs. These spectral differences could result in differences in spectral and ILDs cues, either of which could result in better or worse performance. The effect of differing spectral cues was probably small as the signal was spectrally not very broad (500 – 3000 Hz). Monaural spectral cues useful for localization in the horizontal frontal plane are mainly present at frequencies higher than 4 kHz (Musicant and Butler, 1984), which is also above the limit of the subjects' residual hearing. Additionally, monaural spectral cues play a negligible role in determining lateral angle (Macpherson and Middlebrooks, 2002).

While ILD enhancement decreased the absolute error at larger angles, the localization error relative to the bias (the error bars in figure 7) was not improved. This is attributed to the subjects' limited ILD discrimination abilities. While no learning effect was found beyond the first run of each condition, performance with ILD enhancement may further improve with extended exposure to the artificial ILD cues.

In the current study the impairment-dependent ILD was estimated using a low-pass filter on the acoustically stimulated side. Ideally loudness models would be used to determine the

interaural loudness difference already present in the signals. Suitable loudness models are described by Moore and Glasberg (2004) for acoustic stimulation and McKay *et al.* (2003) for electric stimulation. While these models are well established, their use would greatly increase the computational complexity, and estimating their subject-dependent parameters would be time consuming.

For localization in the horizontal plane, NH listeners make use of at least 3 different auditory cues: ITDs, ILDs and spectral cues (Musicant and Butler, 1984). Our bimodal listeners could probably not use ITD cues because (1) in previous studies we found that except for S02 and S12 they were not sensitive to ITD with controlled stimulation (Francart *et al.*, 2011, 2009a), and (2) in the speech processing strategy used most timing cues are lost, which is illustrated by Seeber and Fastl (2008), who found that bilateral CI listeners do not make use of ITD cues for localization when using the CIS strategy⁵. Additionally, we tested the localization performance of S02 in an extra condition similar to VF-NoEnh, but with the magnitude of the HRTFs set to 1 at all frequencies, and without ILD enhancement, such that only ITD cues could have remained. In this condition, S02 could not distinguish at all between the different angles. Our subjects may also have utilized spectral cues, but as mentioned above, their use was probably limited. Additionally, in laboratory experiments, monaural level cues can potentially be used (Francart, 2008). We used level roving to reduce this cue. If a subject would have attempted to use monaural level cues, one could expect a correlation between the rove and the subject’s response in each trial, such as a response more to the center or to one particular side with increasing level. We did not find any significant correlations between rove in each trial and several outcome measures such as average response and bias, so monaural level cues were probably not used or only to a very limited extent. Therefore the main cue used in the current study was probably ILD.

In our previous study with NH listeners (Francart *et al.*, 2009b), enhancement of ILD cues resulted in a significant improvement in localization error for the wide band noise stimulus, but not for the telephone stimulus. RMS errors were used to report localization performance. In the current study, absolute errors were used to avoid emphasizing large

errors. Calculation of the average absolute errors for the without and with ILD amplification conditions of the previous study yielded, respectively, 37° and 26° (noise stimulus) and 32° and 31° (telephone stimulus). These values are of comparable magnitude to the values found in the current study. The lack of significant improvement for the telephone stimulus in the previous study was probably due to the way the algorithm was implemented. In our previous study, the ILD already present in the signal was amplified, while in the current study the ILD is selected based on the angle of incidence of the stimulus, eliminating the dependence on the ILD present in the stimulus and eliminating the non-monotonicity in the ILD-versus-angle function.

Possible future improvements to the ILD enhancement algorithm include the use of loudness models to determine the interaural loudness difference due to ILDs contained in the signals, and the design of complete artificial HRTFs that are compatible with the constraints of bimodal stimulation (such as the amount of residual hearing and sensitivity to differences in spectrum), instead of only changing the broadband ILD. For implementation in a speech processor, a direction of arrival estimation algorithm is required, which is currently still a challenging problem in complex environments. A communication link between the devices at the two sides would also be required and dynamically changing ILD cues should be taken into account.

V. CONCLUSIONS

While horizontal-plane localization is possible with only ILD cues, for bimodal listeners the usefulness of ILDs is reduced by a lack of high-frequency acoustic hearing. Additionally, ILDs are ambiguous for angles larger than around 60° . An algorithm was developed that enhances ILDs based on having knowledge of the angle of arrival. The algorithm was tested in six bimodal listeners using a virtual auditory space method. ILD enhancement improved localization performance in the horizontal plane by 4-10 $^\circ$ absolute error. The improvement was mainly due to more accurate localization of sounds originating towards the side of the

listener.

Acknowledgments

We thank our test subjects for their enthusiastic participation. We thank the staff of the university hospital of Leuven and Maastricht for their kind cooperation. Funding was provided by IWT-Vlaanderen and Cochlear, project 080304. Additionally, Tom Francart was sponsored by a Post Doctoral Mandate (BOF) from the K.U.Leuven and Anneke Lenssen was sponsored by the EU-ITN project AUDIS. We are grateful to Prof. Hugh McDermott for his useful comments on an earlier version of the manuscript. We thank the associate editor, Dr. Michael Akeroyd, and two anonymous reviewers for their constructive remarks to improve the manuscript.

Endnotes

1. ILD is considered here as a purely broadband cue that can be perceived even if the frequency content differs between the ears. While in NH subjects there is evidence that the ILD is perceived per frequency band, their ILD detection performance does not decrease dramatically when an interaural frequency shift is present (Francart and Wouters, 2007).
2. If the difference between the first and second measurement seemed above average, the procedure was repeated.
3. A repeated measures ANOVA of the absolute error with factors repetition and condition was performed and indicated no significant effect of repetition ($F(3, 35) = 2.06, p = 0.12$).
4. A repeated measures ANOVA of the standard deviation of response per angle was performed with factors of condition (VF-NoEnh or VF-Enh) and stimulus angle (-90 to 90 in steps of 15 degrees). No significant effect was found of either condition ($p = 0.20$) or angle ($p = 0.32$).
5. In the current study the ACE strategy was used in contrast to the CIS strategy used by Seeber and Fastl (2008). Both strategies only potentially encode ITD information in the envelopes. Given the narrower filters in ACE’s filterbank and its maxima selection, it is unlikely that ACE would provide clearer ITD cues.

References

- Ching, T., Incerti, P., and Hill, M. (2004). “Binaural benefits for adults who use hearing aids and cochlear implants in opposite ears.”, *Ear Hear* **25**, 9–21.
- Ching, T., van Wanrooy, E., and Dillon, H. (2007). “Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: a review.”, *Trends Amplif* **11**, 161–92.

- Dunn, C., Tyler, R., and Witt, S. (2005). “Benefit of wearing a hearing aid on the unimplanted ear in adult users of a cochlear implant.”, *J Speech Lang Hear Res* **48**, 668–80.
- Feddersen, W., Sandel, T., Teas, D., and Jeffress, L. A. (1957). “Localization of high-frequency tones”, *J Acoust Soc Am* **29**, 988–991.
- Francart, T. (2008). “Perception of binaural localization cues with combined electric and acoustic hearing”, Ph.D. thesis, K.U.Leuven.
- Francart, T., Brokx, J., and Wouters, J. (2008a). “Sensitivity to interaural level difference and loudness growth with bilateral bimodal stimulation”, *Audiol Neurotol* **13**, 309–319.
- Francart, T., Brokx, J., and Wouters, J. (2009a). “Sensitivity to interaural time differences with combined cochlear implant and acoustic stimulation”, *J Assoc Res Otolaryngol* **10**, 131–41.
- Francart, T., Lenssen, A., and Wouters, J. (2011). “Sensitivity of bimodal listeners to interaural time differences with modulated single- and multiple-channel stimuli”, *Audiol Neurotol* **16**, 82–92.
- Francart, T., Van den Bogaert, T., Moonen, M., and Wouters, J. (2009b). “Amplification of interaural level differences improves sound localization in acoustic simulations of bimodal hearing”, *J Acoust Soc Am* **126**, 3209–13.
- Francart, T., van Wieringen, A., and Wouters, J. (2008b). “APEX 3: a multi-purpose test platform for auditory psychophysical experiments”, *J Neurosci Methods* **172**, 283–293.
- Francart, T. and Wouters, J. (2007). “Perception of across-frequency interaural level differences”, *J. Acoust. Soc. Am.* **122**, 2826–2831.
- Grantham, D., Ricketts, T., Ashmead, D., Labadie, R., and Haynes, D. (2008). “Localization by postlingually deafened adults fitted with a single cochlear implant.”, *Laryngoscope* **118**, 145–51.
- Keidser, G., Rohrseitz, K., Dillon, H., Hamacher, V., Carter, L., Rass, U., and Convery, E. (2006). “The effect of multi-channel wide dynamic range compression, noise reduction, and the directional microphone on horizontal localization performance in hearing aid wearers.”, *Int J Audiol* **45**, 563–79.

Macpherson, E. and Middlebrooks, J. (2002). “Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited.”, *J Acoust Soc Am* **111**, 2219–36.

McKay, C., Henshall, K., Farrell, R., and McDermott, H. (2003). “A practical method of predicting the loudness of complex electrical stimuli.”, *J Acoust Soc Am* **113**, 2054–63.

Moore, B. (2003). *An introduction to the Psychology of Hearing, 5th edition*, chapter 7, 233–267 (Elsevier Science).

Moore, B. and Glasberg, B. (2004). “A revised model of loudness perception applied to cochlear hearing loss.”, *Hear Res* **188**, 70–88.

Moore, B., Glasberg, B., and Schlueter, A. (2010). “Detection of dead regions in the cochlea: relevance for combined electric and acoustic stimulation.”, *Adv Otorhinolaryngol* **67**, 43–50.

Musicant, A. and Butler, R. (1984). “The influence of pinnae-based spectral cues on sound localization.”, *J Acoust Soc Am* **75**, 1195–200.

Seeber, B., Baumann, U., and Fastl, H. (2004). “Localization ability with bimodal hearing aids and bilateral cochlear implants.”, *J Acoust Soc Am* **116**, 1698–709.

Seeber, B. and Fastl, H. (2008). “Localization cues with bilateral cochlear implants.”, *J Acoust Soc Am* **123**, 1030–42.

Shaw, E. (1974). “Transformation of sound pressure level from the free field to the eardrum in the horizontal plane.”, *J Acoust Soc Am* **56**, 1848–61.

Wightman, F. and Kistler, D. (1989). “Headphone simulation of free-field listening. i: Stimulus synthesis.”, *J Acoust Soc Am* **85**, 858–67.

Wightman, F. and Kistler, D. (1992). “The dominant role of low-frequency interaural time differences in sound localization.”, *J Acoust Soc Am* **91**, 1648–61.

Wightman, F. and Kistler, D. (1997). “Monaural sound localization revisited.”, *J Acoust Soc Am* **101**, 1050–63.

Yost, W. and Dye, Jr, R. (1988). “Discrimination of interaural differences of level as a function of frequency.”, *J Acoust Soc Am* **83**, 1846–51.

TABLE I. Subject information: “Age” is in years at the time of testing. “M of use” is the number of months of implant use at the time of testing. “CI side” is left (L) or right (R), the HA was on the other side. ”HA type” is the type of HA the subjects used at the time of testing.

Subject	Age	M of use	CI side	HA type	Etiology
S02	68	78	R	GNReSound Live	Noise exposure
S07	42	100	R	Phonak PowerMAXX 411	Meningitis
S09	34	99	L	Phonak PicoForte3	Auto-immune
S12	67	34	L	Phonak Perseo 311dAZ	Genetic (DFNA9)
S13	37	92	R	Phonak SuperFront PP-C-L-4	Unknown
S23	20	89	R	Phonak F3	Genetic

TABLE II. Results of the balancing experiments. “Slope” and “Intercept” indicate the coefficients of the line $ac = slope * el + intercept$ fitted though the results from the balancing procedure. R^2 indicates the goodness of the linear fit. “Transducer” indicates the transducer used for that particular subject, either the Resound Viking HA with audio input or the Etymotic ER-3A insert phones. “Cut-off” indicates the cut-off frequency of the residual hearing used in the loudness model.

Subject	Slope	Intercept (dB)	R^2	Transducer	Cutoff (Hz)
S02	0.65	0.80	0.85	viking	3000
S07	0.72	-30.89	0.94	viking	1000
S09	0.60	-26.00	0.78	viking	1000
S12	0.75	-24.38	0.95	er-3a	8000
S13	0.63	5.80	0.83	viking	2000
S23	0.77	2.50	0.72	er-3a	2000

List of Figures

FIG. 1	Unaided pure tone audiograms of the non-implanted ear for the subjects that participated in the current study. Note that the vertical axis starts at 60 dBHL.	7
FIG. 2	Broadband ILD present in the telephone alerting signal after filtering with HRTFs. In the conditions labelled 1000 and 2000 Hz, the left-ear signal was filtered with a low pass (LP) filter with these respective cut-off frequencies. This is a simulation of high-frequency hearing loss at the acoustically stimulated side (left).	9
FIG. 3	ILD-versus-angle functions for a white noise stimulus with a BTE HA shell and with the Freedom device, for the telephone alerting signal used in the current study with the Freedom device, and for the artificial ILD function used in the current study. All measurements were made on the same artificial head.	11
FIG. 4	Level-versus-angle function for a white noise sent through the Freedom device, derived from the same data as shown in figure 3, and artificial levels used in the current study.	12
FIG. 5	Spectrum of the telephone alerting signal used.	14
FIG. 6	Loudness balancing results for S12. The circles indicate individual results from single runs of the loudness balancing procedure. The full line indicates average results per electric level and the dashed line is the result of a linear fit.	17
FIG. 7	Localization results for each subject in the different conditions. The average response as a function of angle of incidence is shown and the total length of each error bar indicates twice the standard deviation of the response.	20

FIG. 8	Average absolute error for each subject and condition. The total length of each error bar indicates twice the standard deviation. The bar labelled [Pleaseinsertintopreamble]NH[Pleaseinsertintopreamble] shows the average result of 23 runs of the SF condition in NH subjects (Francart, 2008).	22
FIG. 9	Difference in absolute error per angle between the VF-NoEnh and VF-Enh conditions, averaged over all subjects. Positive values indicate that performance was better in the VF-Enh condition. The total length of each error bar indicates twice the standard deviation.	23