INVITED PAPER

Sound Processing for Better Coding of Monaural and Binaural Cues in Auditory Prostheses

Advances in signal processing technologies for modern hearing aids and cochlear implants with improved sound perception are discussed in this paper.

By Jan Wouters, Simon Doclo, Raphael Koning, and Tom Francart

ABSTRACT | Despite many considerable technical advances in the field of hearing aids and cochlear implants, people using auditory prostheses still have major problems with speech understanding in the presence of interfering sounds and with directional hearing. Both abilities are dependent on sound stream segregation in real-world listening environments. In this paper, two timely and important issues related to sound stream segregation in auditory prostheses are addressed, namely, the coding of monaural and binaural cues. Several state-of-the-art signal processing algorithms used in cochlear implants (CIs) and in hearing aids (HAs) are introduced. A review is given of some recent proposals to improve temporal coding in monaural CIs, and of recent work to improve the transmission of binaural cues in both HAs, CIs, and combined acoustic and electric hearing (bimodal hearing). The ultimate aim is to improve speech and music perception, and, additionally, the preservation of binaural cues to preserve directional hearing.

KEYWORDS | Auditory prostheses; binaural cues; signal processing; temporal coding

I. INTRODUCTION

The technology of auditory prostheses has progressed tremendously during the last two decades. Hearing aids (HAs) and cochlear implants (CIs) are the most common examples of auditory prostheses. In many people with a mild to moderate hearing loss, the auditory perception of speech and music can be restored with behind-the-ear or in-the-ear HAs, typically consisting of one or two microphones, signal processing, amplification, and a transducer to present the processed sound to the ear. More than 95% of all HAs sold in the United States today contain digital signal processing (DSP) technology. CIs are auditory prostheses that can restore hearing in people with severe to profound hearing impairment (or deafness) through electrical stimulation of the auditory nerve. CIs consist of an external sound processor, wirelessly connected to the implanted internal part. The internal part includes a decoder chip and an array of about 20 stimulation electrodes that are surgically inserted into the cochlea. The sound processor converts the received acoustical signals to trains of current pulses that are delivered in the internal part to the electrodes.

Auditory prostheses have been shown to be very successful in restoring auditory perception. As an example,

Manuscript received August 10, 2012; revised January 13, 2013; accepted March 9, 2013. Date of publication July 29, 2013; date of current version August 16, 2013. This work was supported by a project of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT), by the European Union (EU) Marie Curie ITN Project "Digital Signal Processing in Audiology" (AUDIS), by the Research Unit Individualized Hearing Acoustics and the Cluster of Excellence Hearing4All of the German Research Foundation (DFG), and was approved by the Local Research Ethics Committees of the University of Leuven (KU Leuven) and the University of Oldenburg. The work of R. Koning was supported by a fellowship of the EU Marie Curie ITN Project "Digital Signal Processing in Audiology" (AUDIS). The work of T. Francart was supported by postdoctoral fellowships of the Fund for Scientific Research of the Flemish Government and a Marie Curie International Outgoing Fellowship of the EU.
J. Wouters, **R. Koning**, and **T. Francart** are with Experimental Oto-rhino-laryngology (ExpORL), Department of Neurosciences, University of Leuven, Leuven 3000, Belgium (e-mail: jan.wouters@med.kuleuven.be; raphael.koning@med.kuleuven.be; tom.francart@med.kuleuven.be).
S. Doclo is with the Signal Processing Group, Department of Medical Physics and Acoustics, University of Oldenburg, 26111 Oldenburg, Germany (e-mail: simon.doclo@uni-oldenburg.de).

Digital Object Identifier: 10.1109/JPROC.2013.2257635

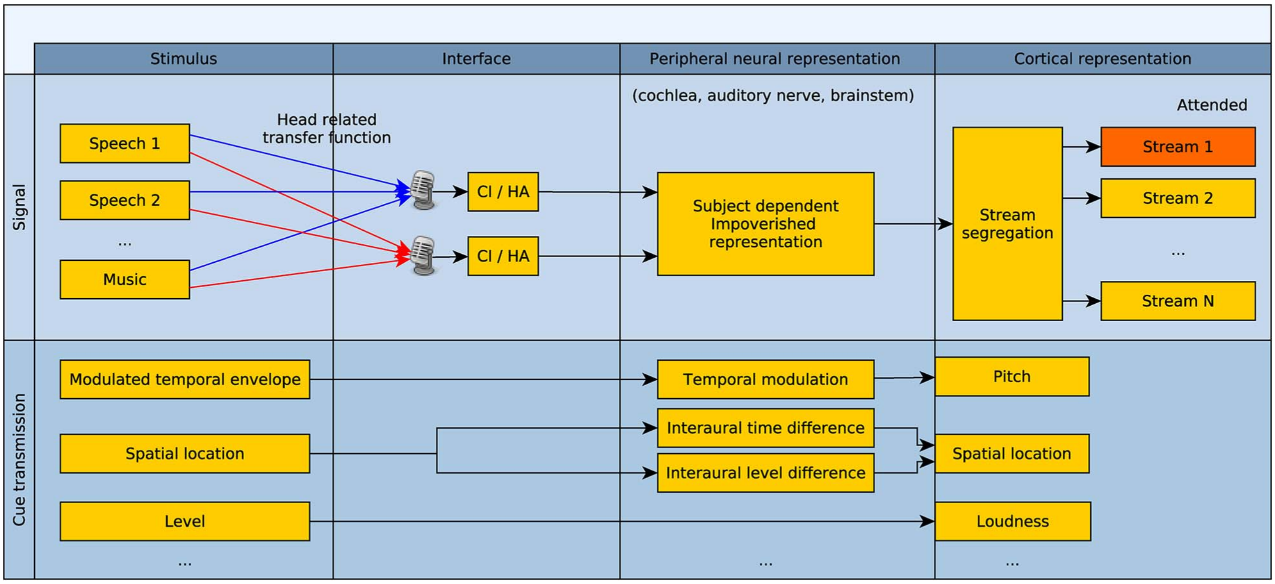

Fig. 1. ***Illustration of auditory stream segregation. The top row illustrates how different sound sources are transmitted from the acoustical domain to the neural/cortical representation. The bottom row illustrates some of the cues that are present in the acoustical domain and their neural counterparts.***

in countries with neonatal hearing screening programs in place, most hearing loss in children can today be detected at a very early stage. With a well-adjusted hearing instrument (e.g., a CI for the severely hearing impaired), these children can receive mainstream education, albeit with continuing audiological and rehabilitative follow-up. The number of people with deafness, from a young child to a senior adult, aided with a CI has significantly increased in the last decade. There are now more than 300 000 people implanted with a CI, and of the newly implanted population 50% are children and 50% have residual hearing due to relaxed implantation criteria.

However, despite many considerable technical advances in the field of auditory prostheses, the progress did not kindle a more widespread use of HAs in people with sensorineural hearing impairment, resulting in more nonusers than users. Today, many hearing problems of the hearing impaired are still unresolved. Whereas speech perception of hearing instrument users is very good in quiet acoustical environments and in one-to-one scenarios, it seriously deteriorates in more challenging listening scenarios. Even when loudness is adequately restored across frequency, i.e., when all sounds are made audible by appropriate fitting of the auditory prostheses to the individual, people using auditory prostheses still have problems with speech understanding in the presence of interfering background sounds, competing talkers or reverberation, and with directional hearing. These situations are abundant in our daily active lives. Hearing-impaired people with appropriately fitted auditory prostheses can hear but not necessarily understand speech. Whereas for normal-hearing listeners, the speech-to-noise ratio (SNR) for 50% speech understanding is about -5 to -10 dB for normal conversation, for most CI users, this SNR is increased by 10 to 20 dB, and for HA users, to a level in between. Additionally, the sound coding in current CIs does not allow proper perception of music and pitch. Also, directional hearing is often very limited. The coded sounds are in many cases perceived as very diffuse or coming from incorrect directions, severely restricting spatial awareness.

This inadequate auditory scene analysis can be due to the sound processing in the auditory prostheses and/or to limited neurophysiological abilities, and/or, in the case of CI, to the nonoptimal signal interface between the electrodes and the auditory system. Auditory prostheses cannot turn the dysfunctioning auditory system to normal, but the aim is to transform the input sounds to stimulation patterns and presentation of auditory cues that the human brain can use to build up the auditory information streams. Additionally, noise reduction algorithms aim to increase the SNR to alleviate the need for good stream segregation. Furthermore, as the auditory profile may vary widely among people, the optimal signal processing in the auditory prosthesis depends on the individual.

To understand why users of auditory prostheses still have the problems outlined above, we need to consider the processes that take place with normal hearing. To segregate simultaneous sound streams, the auditory system uses a process called auditory scene analysis, as illustrated in Fig. 1. Good stream segregation leads to good speech perception in noise [1]. Sound sources are segregated according to their physical location, loudness, temporal structure, and other properties. Each property gives rise to a number of acoustic cues. These cues are transmitted via the auditory periphery to the brain, where they are interpreted and mapped back to the original sound source properties, which are then used to segregate sound sources. In what follows, we will describe several processes that are a part of this mechanism. We will distinguish between monaural cues, which are available at each ear separately, and binaural cues, which are obtained by combining information from both ears.

An important monaural cue is the temporal structure of the input sound signal. The temporal structure of speech consists of three main time scales based on the dominant fluctuation rate: the temporal envelope related to (supra)-segmental variation, the periodicity related to fundamental frequency (F0), and the temporal fine structure [2]. The temporal fine structure are the fast fluctuations in a signal that can be used by normal-hearing listeners to perceive pitch, to localize sounds, and to binaurally segregate different sound sources. The fine structure is modulated in amplitude by the temporal envelope and periodicity. In current signal processing of CIs, only the envelope cues are well encoded in the electrical stimulation waveform, and the periodicity is only encoded in a very restricted way. Monaural signal processing schemes that emphasize temporal coding, based on findings in auditory neurophysiology and human perception, may improve speech and music perception, as will be shown in Sections II–IV.

Bilateral input can generate an important binaural advantage. Genuine binaural hearing is based on the perception of binaural cues of the sounds arriving at the two ears: the interaural time (ITD) and level differences (ILDs). ILDs are due to the acoustic shadow of the head. ILDs are mainly present at higher frequencies and range up to about 20 dB. ITDs are caused by the time difference of arrival to the two ears of the sound waves. This means that, e.g., for a sound coming from the right side, its level will be softer at the left ear and it will arrive later at the left ear than at the right ear. Useful ITDs are present at lower frequencies ($<$ 1500 Hz) and in the envelope of higher frequency sounds [3], [4]. Normal-hearing listeners can use these binaural cues to localize sound sources, binaural unmask speech in background sounds, and experience spatial awareness. Normal-hearing listeners are sensitive to ITDs in both temporal fine structure and the envelope, and their sensitivity is much better for fine-structure ITDs [5]. The human sensitivity can be expressed by the just noticeable differences in ITD and ILD. These can be as low as 10 μs and 1 dB in normal hearing, respectively, and relate to a maximal ITD of about 700 μs and ILD of up to 20 dB in high frequencies, for sounds coming from the side of the head.

In hearing-impaired listeners, usually the main problem lies at the interface between sound wave and auditory nerve. While auditory prostheses can shape the signals to return audibility, the interface is still impaired, so only a fraction of the information transmission possible in normal hearing remains. Therefore, when developing signal processing, the important cues in the input sound signal need to be adequately coded and transmitted to allow suitable presentation to the (impaired) auditory system and to allow appropriate perception.

Moreover, usually hearing instruments for left and right ears are fitted and operate independently (i.e., a so-called bilateral system), and do not necessarily preserve ITDs and ILDs [6]. The binaural benefit with bilateral prostheses compared to unilateral prostheses can be very limited compared to the difference between one and two ears in normal hearing. The independence of the devices adversely impacts binaural cues transmission, which impacts the auditory scene analysis and forthcoming auditory stream segregation. With the advent of wireless ear-to-ear connections, a variety of novel front–end signal processing techniques are being studied, e.g., binaural noise reduction. In addition, binaural cue transmission may be obtained with (cooperating) binaural HAs or CIs, leading to binaural unmasking of speech in noise and spatial hearing. Major interests in this field are front–end processing to improve SNR without distorting binaural cues, and sound processing for CIs to emphasize and properly transmit binaural cues. Several such schemes will be discussed in Sections V–VII.

Due to the aforementioned limitations in the coding of monaural and binaural cues, there is an important societal need for advances in hearing instrument technology.

In general, the signal processing path in most auditory prostheses (HA and CI) consists of the following modules: 1) one or two microphones per device that pick up the sound; 2) filtering into a number of spectral bands, which allows to vary the processing across frequency or stimulation channel; 3) preprocessing aimed at improving the SNR and reducing reverberation; 4) compression to compensate for the reduced dynamic range; 5) amplification in HAs to take into account the higher detection thresholds due to the hearing loss; and 6) in HAs a sound transducer, and in CIs, the transformation by a stimulation algorithm to trains of electrical pulses on the separate electrodes. In the current design of these signal processing modules in commercial auditory prostheses, issues of filter bank structure (e.g., bandwidths) and loudness (from detection threshold to uncomfortably loud) are the main aspects of auditory perception and hearing loss taken into account.

It is clear that new developments in auditory prostheses have to be based on further inclusion of knowledge of the (impaired) human auditory system and perception. As an illustration, several studies have demonstrated that some noise reduction approaches with detrimental effects in sound quality for normal-hearing listeners can yield immediate speech perception improvements in CI users, without noticeable quality differences [7].

Furthermore, in the evaluation phase, new signal processing algorithms for auditory prostheses need to be evaluated behaviorally with hearing-impaired listeners using an appropriate set of perceptual performance measures. Measures can be speech reception thresholds (SRTs; the speech level in quiet or SNR where 50% of the speech is understood), just noticeable differences (or sensitivity to certain processed sound cues), listening effort, cognitive load, or error in localization of sound direction. A recent development is that increasingly objective (or instrumental) performance measures based on human auditory perception are being used [8] to predict outcomes in the design phase of signal processing techniques, as well as perceptually or physiologically validated models of (aspects of) the impaired auditory system [9]–[11]. However, while this can be very useful during design, it will most probably never replace proper psychophysical/behavioral validations. Moreover, no methods exist yet to accurately predict performance with a CI.

Instead of presenting a general treatise on signal processing in auditory prostheses, the authors have chosen to review in this paper a number of recent and perceptually relevant signal processing schemes, which have been perceptually evaluated and have demonstrated a benefit for the hearing impaired. First, different areas related to sound stream segregation and auditory scene analysis in auditory prostheses will be addressed and some state-of-the-art signal processing schemes used for stimulation in CIs and in HAs will be introduced. Second, and more particularly, some recent proposals to improve temporal coding in the monaural electric stimulation of CIs are reviewed in Sections II–IV. Third, recent work to improve the transmission of binaural cues in both acoustic HAs and CIs, as well as in bimodal combinations of acoustic and electric hearing, is reviewed in Sections V–VII. Both for temporal coding of monaural cues and transmission of binaural cues, the problems and possible solutions are discussed.

II. SPEECH ENVELOPE ENHANCEMENT

The speech envelope is important for speech understanding [12], and it is the main contributor to speech intelligibility in CI. Many speech modification approaches, such as offline variations of vowel–consonant level ratios, as well as real-time speech processing algorithms have been studied through the years, aiming at the enhancement of speech intelligibility via speech envelope manipulations. Recent findings from cochlea-scaled entropy calculations [13] corroborate that rapidly changing parts of the speech signal like consonant–vowel transitions, onsets and offsets carry most information for speech intelligibility. While cochlea-scaled entropy-based measures show that the formant transitions carry enough information to contribute to speech intelligibility, the rapid changing consonant–vowel transitions are the main contributor to speech intelligibility in adverse listening conditions [14].

Moreover, in CIs, a neurophysiological rationale exists for the enhancement of transient parts of the speech envelope. The rapid adaptation effect of the auditory nerve synapses that leads to a natural emphasis of onsets in the speech envelope is not present, because it is bypassed by the direct electrical stimulation of the auditory nerve. This transient emphasis aids in segregating different sound sources [15].

In HAs, the spectral changes in speech can be emphasized to circumvent the reduced frequency resolution of hearing-impaired listeners, while in CIs, the frequency resolution is fixed due to the number of electrodes. Therefore, envelope enhancement strategies were developed that enhance the perception of spectral changes for hearing-impaired listeners to help to segregate adjacent frequencies [16]. A benefit in terms of speech intelligibility in noisy conditions was obtained but the speech quality of the unenhanced condition was preferred.

In CIs, the envelope enhancement continuous interleaved sampling (EECIS) [17] and the transient emphasis spectral maxima (TESM) [18] strategies were developed to emphasize transient parts of the speech signal. These algorithms intentionally distort the speech envelope to improve speech intelligibility. The rationale is that emphasizing parts rich in information leads to a better contrast between the target signal and the background sound, which should improve speech intelligibility for hearing-impaired listeners in adverse listening conditions. Although results of these first studies showed only a marginal benefit, envelope enhancement algorithms that focus on the transient parts of the speech signal have recently gained renewed interest, either in combination or not in combination with noise reduction schemes. In Fig. 2, the envelope enhancement for the Dutch sentence "Morgen gaan wij naar de stad" (Tomorrow we will go to the city) processed with a real-time algorithm is shown in red, for a HA (top, waveform) and a CI (bottom, electrodogram, the analogon of spectrogram in CI) [19]. The black line represents the unenhanced condition. Improvements of speech perception in certain background sound scenarios have been observed for envelope enhancement algorithms in normal hearing subjects listening to noise vocoded speech [19] and in hearing-impaired and CI subjects. This suggests better segregation of the target signal from the interfering background due to the increased contrast between the two.

III. TRANSMISSION OF TEMPORAL FINE STRUCTURE

Current commercial sound processors for CI preserve the temporal envelope quite well, but discard or distort the temporal fine structure. We will briefly describe the function of the three most commonly used commercial sound processing schemes for CIs: ACE (Cochlear), HiRes (Advanced Bionics), and FS4 (Med-El).

In ACE [20], [21], the signal is sent through a filter bank, then envelopes are computed using full-wave

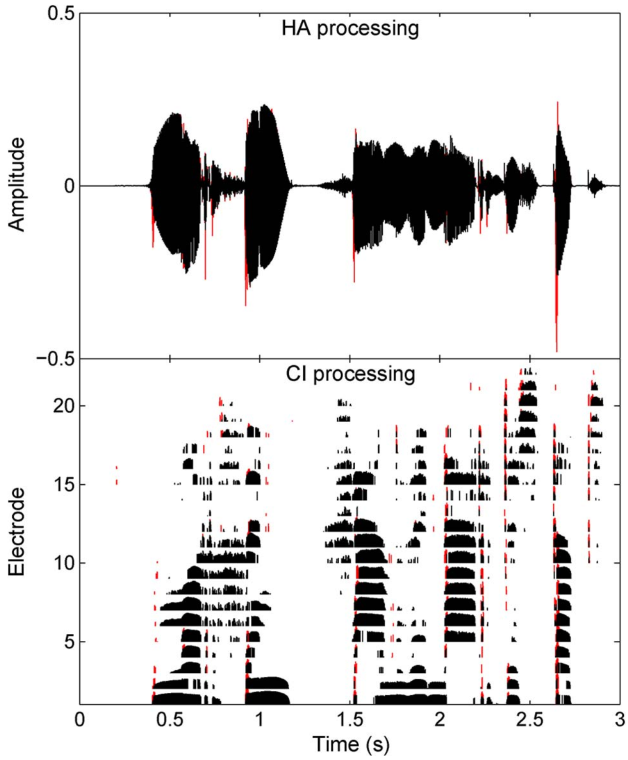

Fig. 2. *Example of an envelope enhancement algorithm that emphasizes the onsets of the speech envelope for the Dutch sentence "Morgen gaan wij naar de stad" (Tomorrow, we will go to the city). The output of the envelope enhancement algorithm is shown for HAs (top) and CIs (bottom).*

rectification and low-pass filtering, with a cutoff frequency corresponding to the channel bandwidth. In each time frame, the N channels with maximal amplitude are selected for stimulation, with usually $N = 8$. A fixed-rate carrier of around 900 pulses per second (pps) is then modulated with the resulting envelopes. Clearly ACE discards all temporal fine structure, and, additionally, all temporal information is quantized to the frame repetition period, which corresponds to the channel stimulation period (1.1 ms, corresponding rate of 900 pps).

HiRes uses a filter bank, followed by envelope detection using half-wave rectification and a low-pass filter with a fairly high cutoff frequency, thereby retaining part of the temporal fine structure. The resulting envelopes are used to modulate fixed rate pulse trains with fairly high rates (around 2900 pps).

FS4 [22] uses different processing for the two to four most apical electrodes and the rest. On the non-apical electrodes, a standard filter bank plus envelope detection (using the Hilbert transform) is used. On the apical electrodes, the channel-specific sampling sequence (CSSS) strategy is applied. In CSSS, a pulse burst is initiated at each positive zero crossing with amplitudes corresponding to the peak amplitudes of the previous half-wave segment. While in this way the periodicity in the fine structure is coded, its exact timing information is lost.

There is no evidence that users of any of these strategies would be able to perceive timing cues in the temporal fine structure in the conventional sense.

IV. F0 CODING

Normal-hearing listeners use three different mechanisms to perceive pitch: place of stimulation in the cochlea (varies with frequency content), temporal fine structure, and periodicity in the envelope. The most salient cue is in the temporal fine structure.

As CI users usually do not have access to the temporal fine structure, they can only use place cues and envelope periodicity. Due to the spread of excitation, placed pitch perception through CIs is quite poor. Salience of envelope periodicity cues depends on modulation depth and across channel synchronization.

The modulation depth of a pulse train on a certain electrode is determined by the modulation depth of the input signal and the sound pressure level of the input signal. The maximal pulse current on each electrode is limited to a fixed value, so with increasing intensity the modulation depth will be reduced until there remains an unmodulated pulse train at maximal current. This is illustrated in Fig. 3 for a speech spectrum weighted click train with an F0 of 100 Hz, at different input levels, processed by ACE. While this could be avoided using an appropriate compression system, this would have undesirable side effects, such as degrading the perception of

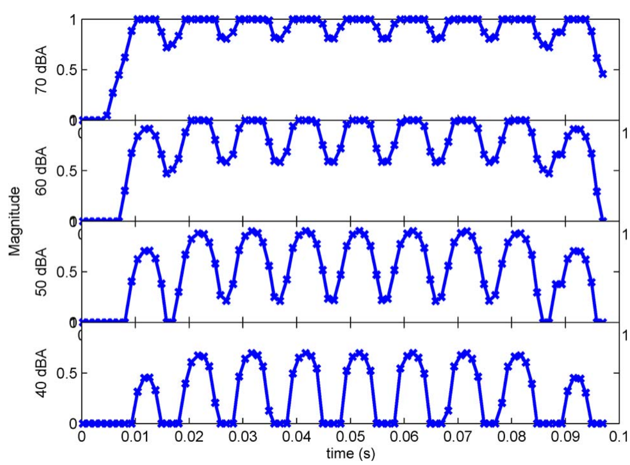

Fig. 3. *Illustration of reduced modulation depth with increasing input sound pressure level. The input signal was a speech weighted click train with an F0 of 100 Hz, at different input levels. The pulses on the fourth most apical electrode are shown. Magnitudes between 0 and 1 are mapped to current units between threshold and comfortable loudness.*

overall loudness variations, and possibly reducing spectral details.

Spread of excitation means that the number of independent perceptual channels is limited. For instance, if two adjacent electrodes carry a modulated pulse train and the modulations are out of phase, the overall neural excitation might be a combination of the two, with reduced modulation depth and possibly altered periodicity. Several sound processing strategies have been developed to address these problems by explicitly enhancing the speech envelope modulations.

In the F0mod scheme [23], existing modulations in the channel envelopes are removed using a low-pass filter, the F0 of the input signal is determined using an autocorrelation method, and all channels are modulated in phase with a sinusoidal modulator based on the detected F0. Modulation is only applied in voiced parts of the signal; for unvoiced parts the scheme falls back to standard ACE.

The MEM strategy [24] also removes existing modulations, but its modulator is an enhanced version of the envelope of the broadband signal. In this way, deep and synchronous modulation of all channels is also achieved.

The eTone strategy [25] estimates F0 using a harmonic sieve, and modulates each time frame to an extent proportional to estimated harmonicity of the time frame.

In laboratory experiments with single F0 sources, all these strategies have yielded improved performance on tasks, such as pitch discrimination, pitch ranking, or melody recognition, and even indications of improvements of tonal language speech perception in noise.

Other strategies that may improve F0 perception are those initially proposed to improve transmission of temporal fine structure or binaural cues.

V. BILATERAL CIs

There is ample evidence that bilateral CIs yield better performance than unilateral ones for a number of tasks: speech perception in noise when speech and noise are spatially separated, sound source localization, and general comfort of use due to a more balanced sensation. This is mainly due to perception of ILDs, increased redundancy, or the better ear effect.

A. Binaural Cue Perception

Bilateral CI listeners are sensitive to ILDs with JNDs in the order of 1–5-dB input level with typical commercial sound processors [26], or 0.17–0.68-dB change in electric current [27]. For normal-hearing listeners just noticeable differences are in the order of 1 dB. Therefore, sensitivity to interaural level differences *per se* should not preclude perception.

In laboratory experiments with synchronized devices and well-controlled stimuli, it has been found that bilateral CI users can be sensitive to ITDs, but less so than normal-hearing listeners and with large intersubject variability. Just noticeable differences for the best performers are in the order of 100 μs [28], [29], which is comparable to normal-hearing listeners' sensitivity to ITDs in the envelope. The worst performers are not sensitive to ITDs at all. Several studies utilizing experimental devices have also shown binaural masking level differences: improved detection thresholds if the signal but not the masker is interaurally out of phase [30]. This is an indication that binaural unmasking of speech in noise could be possible with improved sound processing.

For sound source localization, however, bilateral CI users mainly seem to use ILDs, and there is no direct evidence of true binaural unmasking using interaural timing cues [31], [32]. When a listener cannot perceive ITDs, either due to perceptual limitations or due to the signal processing in the devices, the ILD is the only remaining binaural cue. In this case, there is an additional problem: nonmonotonicity of the ILD versus angle function. Beyond approximately 45° to the left or right, the ILD versus angle function flattens, such that it becomes impossible to distinguish between angles larger than 45° on one side. Normal-hearing listeners cope with this using ITD cues.

B. Binaural Cue Transmission

While binaural cues can be physically present in the acoustic signal and the binaural system might be able to perceive such cues, this does not mean that they are transmitted by the devices used.

For normal-hearing listeners, the most salient ITD cues are those in the temporal fine structure, which is not well transmitted by current commercial sound processors. Cochlear's ACE strategy uses a fixed carrier rate and completely discards fine structure. While strategies such as HIRes (Advanced Bionics) and FS4 (Med-El) transmit fine structure to some extent, which may be useful for pitch perception, there are still quantization issues that may distort the very short time differences that can be used by the binaural system.

Most commercial strategies transmit ITDs in the envelope, but across-channel timing issues and limited modulation depth might limit their usefulness. Several sound processing schemes have been proposed to address these problems.

The peak-derived timing (PDT) strategy [27] operates by synchronizing stimulation pulses with amplitude peaks in the fine structure of the signals in the different channels of the filter bank. The FAST strategy [33] operates in a similar manner, but aligns pulses with peaks in the envelope in each channel. The FSP strategy outputs pulse bursts on positive zero crossings of the filtered signals in the lowest one to four channels.

So far no or limited practical benefit of these strategies has been shown. This could be due to perceptual limitations, inability of the strategies to cope with complex everyday sounds, or lack of acclimatization of the listeners.

VI. BIMODAL STIMULATION

In the past, patients were only considered for cochlear implantation when they were profoundly deaf in both ears. Today, however, it is seen that cochlear implantees often have better speech perception than severely hearing-impaired listeners with HAs. Therefore, implantation criteria are changing, and currently most of the newly implanted patients have residual hearing to some degree, in either the implanted or nonimplanted ear [34]. If there is residual hearing in the nonimplanted ear, it can be stimulated using an HA. We call this combination bimodal stimulation. While the residual hearing in the nonimplanted ear ranges from normal hearing to severely hearing impaired, most CI users only have residual hearing at lower frequencies, e.g., up to 2 kHz.

With bimodal stimulation, the two ears are stimulated, so binaural cues (ITD and ILD) can potentially be perceived, which is obviously not possible with unilateral CI stimulation. In what follows, we will first discuss binaural cue perception through bimodal stimulation, then the effect of current signal processing systems on binaural cue transmission, and, finally, a number of novel signal processing schemes designed to improve binaural cue transmission.

A. Binaural Cue Perception

Electrical (via CI) and acoustic auditory stimulations lead to very different stimulation patterns at the auditory nerve level. It is not obvious whether the human binaural system is able to combine those different inputs to extract binaural cues. Therefore, a number of basic psychophysical studies were conducted to investigate bimodal binaural cue perception under laboratory conditions, with specially crafted stimuli designed to transmit maximal binaural information.

Bimodal listeners are sensitive to ILDs, with an average just noticeable difference of 1.7-dB change at the acoustically stimulated side [35], but with large intersubject differences. Similarly to bilateral CI listeners, sensitivity to ILDs *per se* should, therefore, not preclude perception. There is, however, another issue when considering real-life signals: a lack of high-frequency residual hearing. ILDs in real signals are largest for higher frequencies (> 1500 Hz). Unfortunately, most bimodal listeners have no or very limited residual hearing at higher frequencies. Therefore, only very small low-frequency interaural level differences are available to the binaural system, with their magnitude often even below the subject's sensitivity.

Bimodal listeners can be sensitive to ITDs in the envelope if the thresholds of their residual hearing at 1000 and 2000 Hz are better than the 100-dB sound pressure level (SPL) [36]. Bimodal listeners are not sensitive to ITDs in temporal fine structure [37]. A large intersubject variability was observed in these studies, but best thresholds were in the order of 100 μs, while in ecological signals, ITDs up to 700 μs can occur.

When a bimodal listener cannot perceive ITD cues, the same problem of nonmonotonicity of the ILD versus angle function occurs as with bilateral CIs (see Section V).

B. Binaural Cue Transmission Through Commercial Devices

Differences in loudness growth between modalities can distort perception of ILD cues. For normal-hearing listeners, the term ILD can be used to indicate the physical cue as well as the perceptual effect. For listeners with asymmetric hearing losses, of which bimodal listeners are an extreme example, it makes more sense to consider interaural loudness differences, i.e., the perceptual effect of ILDs. If loudness growth is different across the ears, interaural loudness differences are not consistent across stimulation levels. It is, for example, possible for a signal at 60-dB SPL to be centered in the middle of the head, but for the same signal at 50-dB SPL to be shifted to one side. This is illustrated in Fig. 4. For a 1-kHz tone and a speech weighted noise signal, interaural loudness differences were predicted using models of loudness [38], [39] configured with the parameters of a typical bimodal listener. While for a normal-hearing listener the interaural loudness difference is invariant with level, this is not the case for a bimodal listener, and, additionally, there is a dependence on frequency content of the signal and even a nonmonotonicity at low levels.

ITDs can be present in the fine structure and temporal envelope of the acoustic signal. Fine structure ITDs are

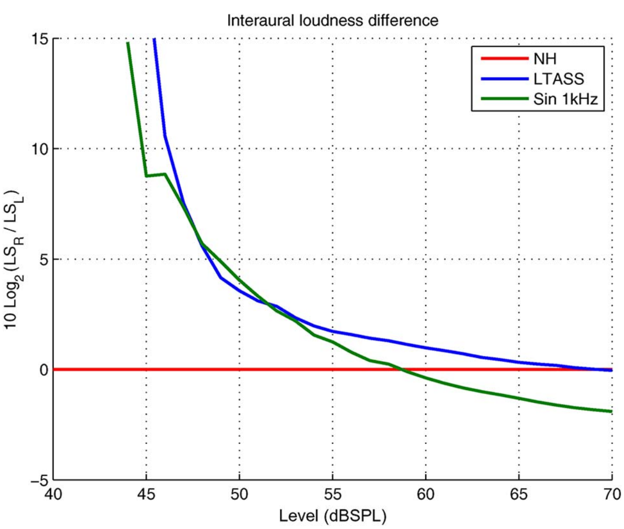

Fig. 4. ***Interaural loudness differences for either a normal-hearing listener (for any signal) and for a bimodal listener for a long-term-average-speech-spectrum weighted noise and a 1-kHz sinusoid. The interaural loudness difference is plotted here as ten times the base 2 logarithm of the ratio of the left and right loudness in tones.***

either not transmitted at all or seriously distorted by most current commercial sound processing schemes (see above). Envelope ITDs can be transmitted, but their perceptual usefulness depends on across-channel timing and channel modulation depth, which depend on the phase response of the filter bank used and the intensity of the input signal.

C. Novel Signal Processing for Improved Binaural Cue Transmission

The SCORE bimodal strategy [40] uses loudness models to normalize loudness perception for both electric and acoustic stimulation, and thus yields equal loudness growth functions on both sides, which leads to improved binaural balance and transmission of interaural loudness cues. The loudness of the microphone signals for a normal-hearing listener is estimated, the loudness of the signals after CI and HA processing is estimated using loudness models for impaired hearing, and level adjustments are calculated to return the loudness to normal.

Bilateral CI strategies are often not useful for bimodal stimulation because they introduce cues (temporal features) that are not available in the acoustical signal. For good ITD perception with bimodal stimulation, we need deep modulations, synchronized across channels, and long dead times between stimulation periods [41], [42].

The modulation enhancement strategy (MEnS) imposes a deeply modulated envelope on all frequency channels simultaneously [43], similarly to CI F0 perception improvement strategies, as described in Section IV. This results in deep and synchronous modulations of the pulse trains delivered to all electrodes. In preliminary experiments, improved ITD detection thresholds were found compared to the commercial ACE processing in five bimodal listeners.

VII. BILATERAL AND BINAURAL NOISE REDUCTION

To improve speech perception in noisy environments, several noise reduction techniques for HAs and CIs have been developed, with the aim to improve the SNR while preserving the speech quality. Hereby, it should be realized that an SNR improvement of 1 dB around the SRT can generate an increase in speech understanding of up to 15% in everyday communication [44].

Noise reduction algorithms can be broadly classified into single-microphone and multimicrophone algorithms. Although single-microphone noise reduction algorithms provide an SNR improvement, this typically comes at the price of speech distortion and other artifacts, such as so-called musical noise. For HAs, current single-microphone noise reduction algorithms have not been found to yield any SRT improvement [45], [46], although they attenuate the overall noise level and may reduce the listening effort. For CIs, small SRT improvements can be obtained using single-microphone algorithms [47], [48]. In comparison with single-microphone algorithms, which can only use spectral and temporal information, multimicrophone algorithms can additionally exploit the spatial information of the sound sources, by combining different microphone signals. This generally results in a larger SNR improvement than single-microphone algorithms, especially when the speech and the noise sources have different spatial characteristics. Many state-of-the-art HAs and CIs currently contain two or three closely spaced microphones, enabling the use of multimicrophone noise reduction algorithms [45], [49], [50].

In principle, applying hearing instruments bilaterally can generate an important binaural advantage. The auditory system can use binaural cues and the signal processing can use information from multiple microphones at both hearing instruments. However, in a bilateral system where both hearing instruments work independently, this potential is not fully exploited since not all microphone signals from both hearing instruments are combined. Moreover, it has been shown that the localization performance of bilateral HA users may even be compromised when the multimicrophone noise reduction is switched on, due to distortion of the binaural cues [6].

To achieve true binaural processing, both hearing instruments need to cooperate with each other and exchange information or signals, e.g., through a wireless link [51]. In current state-of-the-art binaural devices, one microphone signal can be transmitted from one device to the other, in half-duplex mode. It is expected that in future devices it will also be possible to exchange one (or even multiple) microphone signals in full-duplex mode. These systems would allow for binaural multimicrophone noise reduction algorithms, where microphone signals from both devices are processed and combined in each device.

The objective of a binaural noise reduction algorithm is not only to selectively extract the target speech and to suppress background noise, but also to preserve the binaural cues of the sound sources, so as to preserve the auditory impression of the acoustic scene. Two paradigms are typically adopted for binaural algorithms. In the first paradigm, two microphone signals (i.e., one on each device) are filtered with the same real-valued spectral gain, hence automatically guaranteeing binaural cue preservation (cf., Section VII-A). In the second paradigm, all microphone signals from both devices are processed by different complex-valued spatial filters, e.g., using fixed or adaptive beamforming (cf., Section VII-B) or using binaural multichannel Wiener filter (MWF) approaches (cf., Section VII-C). Although the second paradigm allows for more degrees of freedom to achieve noise reduction, there is necessarily a tradeoff between noise reduction performance and binaural cue preservation. In the next sections, we will describe recent advances for both paradigms and discuss their benefit in terms of SRT improvement and localization performance.

A. Spectral Postfiltering Techniques

In spectral postfiltering techniques, an identical real-valued spectrotemporal gain is applied to one microphone signal of each hearing instrument, where a gain close to one is applied when the time–frequency bin should be retained (target speech), and a gain close to zero is applied when the time–frequency bin should be suppressed (background noise).

In [52], this spectral gain has been computed by comparing the estimated binaural properties, such as the interaural coherence, in each frequency bin with the expected properties of the target speech (e.g., assuming that the target speech arrives from the frontal direction with ITD and ILD values close to 0 μs and 0 dB). Another coherence-based technique has been proposed in [53], where the fluctuation of the interaural phase difference is used as a measure for interaural coherence. Another approach to compute the spectral gain is based on the output signal of a fixed or adaptive beamformer [54], [55].

Although these spectral postfiltering techniques preserve the binaural cues of the sound sources, in essence, they can be viewed as single-microphone noise reduction techniques, hence introducing some speech distortion and exhibiting the typical musical noise artefacts, especially at low input SNRs.

B. Fixed and Adaptive Beamforming

Well-known multimicrophone noise reduction techniques in unilateral and bilateral HAs and CIs are based on fixed and adaptive beamforming. The objective of a fixed beamformer is to obtain spatial focusing on the target speech source, thereby reducing background noise not coming from the direction of the speech source. For the design of fixed beamformers, the direction of the speech source and the complete microphone configuration need to be known. Hence, fixed beamformers have mainly been used for unilateral HAs, although for binaural HAs fixed beamformers have also been proposed that aim to combine spatial selectivity with the preservation of the binaural cues of the speech source [56].

Since in practice the background noise is unknown and can change both spectrally and spatially, information about the noise field needs to be adaptively estimated. Adaptive beamformers combine the spatial focusing of fixed beamformers with adaptive noise suppression, hence generally exhibiting a higher noise reduction performance. A very commonly used adaptive beamforming technique in unilateral hearing instruments is the so-called adaptive directional microphone (ADM) [57], adaptively forming a null in the direction of the strongest interferer. A more general approach is the generalized sidelobe canceler (GSC), which consists of a spatial preprocessor, i.e., a fixed beamformer and a blocking matrix, combined with a (multichannel) adaptive noise canceler. The GSC can be considered as the current state-of-the-art solution for unilateral hearing instruments. A two-microphone implementation was indeed shown to achieve a considerable SRT improvement for CI users (about 2–3-dB improvement compared to a hardware directional microphone for three babble noise sources) [49].

In an effort to combine adaptive noise reduction with binaural processing, adaptive beamforming techniques producing a binaural output signal have been proposed. In [58], the microphone signals are divided into low- and high-frequency components ($f_c = 800$ Hz), where the low-frequency components are passed through unprocessed in order to preserve the ITD cues of the speech source and adaptive noise reduction is performed only for the high-frequency components. When using this approach with hearing-impaired subjects, an SRT improvement of about 2 dB was obtained (for a single noise source). However, the binaural cues are preserved only for the target speech but not for the noise, and only when the speech source is arriving from the frontal direction [59].

C. Multichannel Wiener Filter

A more recent class of multimicrophone noise reduction techniques is based on the MWF, where the objective is to obtain a minimum mean square error (MMSE) estimate of the speech component in a reference microphone signal. To provide an explicit tradeoff between speech distortion and noise reduction, the speech distortion weighted multichannel Wiener filter (SDW–MWF) has been proposed [60]. A benefit over the GSC is that, in principle, no *a priori* knowledge about the acoustic environment, target speech location, or microphone characteristics is required. The SDW–MWF is uniquely based on estimates of the second-order statistics of the speech and the noise signals, and hence requires a voice activity detector (VAD), determining time–frequency regions where the desired speech is dominant and time–frequency regions where the noise is dominant. In [45], a three-microphone MWF implementation for a unilateral HA was evaluated at different test sites, and compared with other single-microphone and multimicrophone noise reduction techniques. In this study, it was shown that overall the MWF achieved the largest SRT improvements (up to 7 dB), even in highly reverberant environments.

The MWF can be straightforwardly extended into a binaural version producing a binaural output, by estimating the speech component in two reference microphone signals, namely one on each device. In [61], it was shown both mathematically and using physical evaluations that the binaural MWF perfectly preserves the binaural cues of the speech component but changes the binaural cues of the noise component into those of the speech component. To optimally benefit from spatial release from masking and to optimize the spatial awareness of the HA user, it would be beneficial to also preserve the binaural cues of the noise component. Hence, several extensions for the binaural MWF have been proposed [61], either using partial noise estimation (MWF–N) or by extending the MWF cost

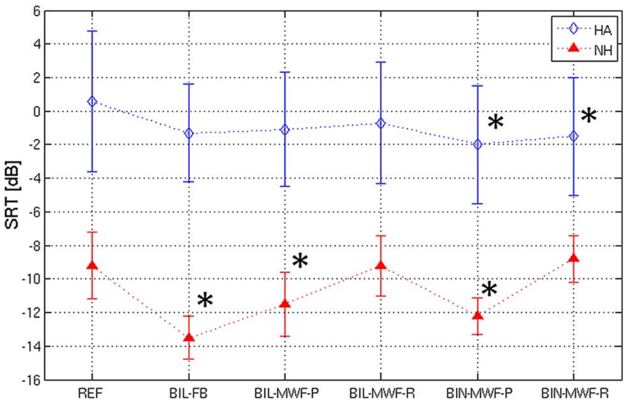

Fig. 5. ***Average SRT results of ten normal-hearing subjects (NH) and eight hearing aid users (HA), for the cafeteria scenario. Standard deviations are indicated by error bars. Significant SRT improvements compared to the unprocessed (REF) condition are marked by *. The algorithms are denoted by: bilateral beamformer (BIL–FB), bilateral MWF with perfect and real VAD (BIL–MWF–P, BIL–MWF–R), and binaural MWF with perfect and real VAD (BIN–MWF–P, BIN–MWF–R).***

function with terms related to the interaural transfer function of the noise component.

The performance of the binaural MWF and its extensions has also been perceptually evaluated, both in terms of SRT improvement and localization performance. In [59] and [62], offline bilateral and binaural MWF implementations using a perfect VAD were evaluated using ten normal-hearing subjects for different spatial speech and noise scenarios in a realistic environment ($T_{60} = 0.61$ s). First, it was shown that the binaural MWF achieved significant SRT improvements compared to the bilateral MWF and the bilateral ADM. This demonstrates that transmitting and processing contralateral microphone signals can result in a significant gain in noise reduction, especially for multiple noise source scenarios. Second, using a localization experiment in the frontal horizontal hemisphere, it was shown that the binaural MWF and MWF–N have advantages in terms of spatial awareness for the HA user in comparison with a bilateral ADM. In contrast with the bilateral ADM, the binaural MWF preserved the location of the target speech source independently of its angle of arrival, even though in some conditions the subjects located the noise source at the place of the speech source (as theoretically predicted in [61]). However, when using the binaural MWF–N, the subjects correctly localized both the speech and the noise sources. Third, it was demonstrated that for the binaural MWF–N, using a partial noise estimate that is large enough to sufficiently restore spatial awareness, the speech intelligibility performance was only slightly affected. In some conditions, the MWF–N was even able to outperform the MWF, which may be due to improved spatial release from masking.

In [63], online adaptive bilateral and binaural MWF implementations using a real VAD were evaluated using ten normal-hearing and eight hearing-impaired subjects for different spatial speech and noise scenarios, including a challenging cafeteria scenario with highly nonstationary noise (i.e., interfering speakers). The average SRT results are depicted in Fig. 5. For the group of HA users only, the binaural MWF achieved a significant SRT improvement (2.2 dB with real VAD). Moreover, no significant performance degradation was observed when using a real VAD instead of a perfect VAD.

VIII. CONCLUSION

The appropriate coding and transmission of monaural and binaural cues in the incoming sound signal are a prerequisite for neural auditory representations that lead to improved perception of speech and music, and directional hearing.

In this paper, a number of state-of-the-art signal processing schemes are described that have been recently proposed, perceptually evaluated, and shown to yield some benefit as compared to current technology used in monaural or binaural auditory prostheses. A review has been given of some monaural applications such as the emphasis of onsets in the speech envelope and the explicit coding of the fundamental frequency, which result in better speech and music perception. Several binaural processing strategies have been reviewed as well. A more faithful transmission of binaural cues leads to better ITD and ILD perception in binaural as well as bimodal HA and CI systems. The combination of directional hearing with noise reduction is an additional challenge, but progress is made. Sound localization and speech perception have been evaluated and clinically relevant improvements have been demonstrated. Some of these new strategies are very promising and can lead to implementations in future commercial devices.

Additionally, the importance of behavioral validation with user groups for the development of adequate signal processing for auditory prostheses has been emphasized. ■

REFERENCES

[1] A. S. Bregman, *Auditory Scene Analysis*. Cambridge, MA, USA: MIT Press, 1990.

[2] S. Rosen, "Temporal information in speech: Acoustic, auditory and linguistic aspects," *Phil. Trans. Roy. Soc. Lond. B, Biol. Sci.*, vol. 59, pp. 367–373, Jul. 1992.

[3] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localisation*. Cambridge, MA, USA: MIT Press, 1983.

[4] A. Bronkhorst and R. Plomp, "The effect of head-induced interaural time and level differences on speech intelligibility in noise," *J. Acoust. Soc. Amer.*, vol. 83, pp. 1508–1516, Apr. 1988.

[5] D. McFadden and E. G. Pasanen, "Lateralization of high frequencies based on interaural time differences," *J. Acoust. Soc. Amer.*, vol. 59, no. 3, pp. 634–639, 1976.

[6] T. van den Bogaert, T. J. Klasen, M. Moonen, L. Van Deun, and J. Wouters, "Horizontal localisation with bilateral hearing aids: Without is better than with," *J. Acoust. Soc. Amer.*, vol. 119, pp. 515–526, Jan. 2006.

[7] O. Qazi, B. van Dijk, M. Moonen, and J. Wouters, "Speech understanding performance of cochlear implant subjects using time-frequency masking-based noise reduction," *IEEE. Trans. Biomed. Eng.*, vol. 59, no. 5, pp. 1364–1373, May 2012.

[8] K. Eneman, A. Leijon, S. Doclo, A. Spriet, M. Moonen, and J. Wouters, "Auditory-profile-based physical evaluation of multimicrophone noise reduction techniques in hearing instruments," in *Advances in Digital Speech Transmission*. New York, NY, USA: Wiley, 2008, pp. 431–458.

[9] M. L. Jepsen and T. Dau, "Characterizing auditory processing and perception in individual listeners with sensorineural hearing loss," *J. Acoust. Soc. Amer.*, vol. 129, pp. 262–281, 2011.

[10] M. S. Zilany and I. C. Bruce, "Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," *J. Acoust. Soc. Amer.*, vol. 120, pp. 1446–1466, Sep. 2006.

[11] T. Dau, D. Püschel, and A. Kohlrausch, "A quantitative model of the 'effective' signal processing in the auditory system: II. simulations and measurements," *J. Acoust. Soc. Amer.*, vol. 99, pp. 3623–3631, 1996.

[12] R. V. Shannon, F.-G. Zeng, V. Kamath, J. Qygonski, and M. Ekelid, "Speech recognition with primarily temporal cues," *Science*, vol. 270, pp. 303–304, 1995.

[13] C. E. Stilp and K. R. Kluender, "Cochlea-scaled entropy, not consonants, vowels, or time, best predicts speech intelligibility," *Proc. Nat. Acad. Sci. USA*, vol. 107, no. 27, pp. 12 387–12 392, 2010.

[14] F. Chen and P. C. Loizou, "Contribution of cochlea-scaled entropy and consonant-vowel boundaries to prediction of speech intelligibility in noise," *J. Acoust. Soc. Amer.*, vol. 131, pp. 4104–4113, 2006.

[15] B. Delgutte and N. Y. S. Kiang, "Speech coding in the auditory nerve: Iv. sounds with consonant-like dynamic characteristics," *J. Acoust. Soc. Amer.*, vol. 75, pp. 897–907, 1984.

[16] J. Chen, T. Baer, and B. C. J. Moore, "Effect of enhancement of spectral changes on speech intelligibility and clarity preferences for the hearing impaired," *J. Acoust. Soc. Amer.*, vol. 131, pp. 4104–4113, 2012.

[17] L. Geurts and J. Wouters, "Enhancing the speech envelope of continuous interleaved sampling processors for cochlear implants," *J. Acoust. Soc. Amer.*, vol. 105, pp. 2476–2484, 1999.

[18] A. Vandali, "Emphasis of short-duration acoustic speech cues for cochlear implant users," *J. Acoust. Soc. Amer.*, vol. 109, pp. 2049–2061, 2001.

[19] R. Koning and J. Wouters, "The potential of onset enhancement for increased speech intelligibility in auditory prostheses," *J. Acoust. Soc. Amer.*, vol. 132, no. 4, pp. 2569–2581, 2012.

[20] M. McKay, A. E. Vandali, H. McDermott, and G. M. Clark, "Speech processing for multichannel cochlear implants: Variations of the spectral maxima sound processor strategy," *Acta Otolaryngol.*, vol. 114, no. 1, pp. 52–58, 1994.

[21] A. E. Vandali, L. A. Whitford, K. L. Plant, and G. M. Clark, "Speech perception as a function of electrical stimulation rate: Using the nucleus 24 cochlear implant system," *Ear Hear.*, vol. 21, pp. 608–624, 2000.

[22] I. Hochmair, P. Nopp, C. Jolly, M. Schmidt, H. Schosser, C. Garnham, and I. Anderson, "Med-el cochlear implants: State of the art and a glimpse into the future," *Trends Amplif.*, vol. 10, pp. 201–219, 2006.

[23] M. Milczynski, J. Wouters, and A. van Wieringen, "Perception of mandarin Chinese with cochlear implants using enhanced temporal pitch cues," *Hear Res.*, vol. 285, pp. 1–12, 2012.

[24] A. Vandali, C. Sucher, D. Tsang, C. McKay, J. Chew, and H. McDermott, "Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies," *J. Acoust. Soc. Amer.*, vol. 117, pp. 3126–3138, 2005.

[25] A. Vandali and R. van Hoesel, "Enhancement of temporal cues to pitch in cochlear implants: Effects on pitch ranking," *J. Acoust. Soc. Amer.*, vol. 132, pp. 392–402, 2012.

[26] B. Laback, S.-M. Pok, W.-D. Baumgartner, W. A. Deutsch, and K. Schmid, "Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors," *Ear Hear.*, vol. 25, pp. 488–500, Oct. 2004.

[27] R. J. M. van Hoesel and R. S. Tyler, "Speech perception, localization, and lateralization with bilateral cochlear implants," *J. Acoust. Soc. Amer.*, vol. 113, no. 3, pp. 1617–1630, 2003.

[28] C. J. Long, D. K. Eddington, H. S. Colburn, and W. M. Rabinowitz, "Binaural sensitivity as a function of interaural electrode position with a bilateral cochlear implant user," *J. Acoust. Soc. Amer.*, vol. 114, pp. 1565–1574, Sep. 2003.

[29] R. J. M. van Hoesel, "Sensitivity to binaural timing in bilateral cochlear implant users," *J. Acoust. Soc. Amer.*, vol. 121, pp. 2192–2206, Apr. 2007.

[30] L. Van Deun, A. van Wieringen, T. Francart, A. Büchner, T. Lenarz, and J. Wouters, "Binaural unmasking of multi-channel stimuli in bilateral cochlear implant users," *J. Assoc. Res. Otolaryngol.*, vol. 12, pp. 659–670, 2011.

[31] T. Ching, E. van Wanrooy, and H. Dillon, "Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: A review," *Trends Amplif.*, vol. 11, pp. 161–192, 2007.

[32] B. U. Seeber and H. Fastl, "Localization cues with bilateral cochlear implants," *J. Acoust. Soc. Amer.*, vol. 123, pp. 1030–1042, 2008.

[33] Z. Smith, "Speech coding of interaural time differences for bilateral cochlear implants," in *Proc. Conf. Implantable Auditory Prostheses*, 2009.

[34] C. A. von Ilberg, U. Baumann, J. Kiefer, J. Tillein, and O. F. Adunka, "Electric-acoustic stimulation of the auditory system: A review of the first decade," *Audiol. Neurootol.*, vol. 16, no. Suppl. 2, pp. 1–30, 2011.

[35] T. Francart, J. Brokx, and J. Wouters, "Sensitivity to interaural level difference and loudness growth with bilateral bimodal stimulation," *Audiol. Neurootol.*, vol. 13, pp. 309–319, 2008.

[36] T. Francart, J. Brokx, and J. Wouters, "Sensitivity to interaural time differences with combined cochlear implant and acoustic stimulation," *J. Assoc. Res. Otolaryngol.*, vol. 10, pp. 131–141, 2009.

[37] A. Lenssen, T. Francart, J. Brokx, and J. Wouters, "Bimodal listeners are not sensitive to interaural time differences in unmodulated low-frequency stimuli," *J. Acoust. Soc. Amer.*, vol. 129, pp. 3457–3460, 2011.

[38] B. C. J. Moore, B. R. Glasberg, and T. Baer, "A model for the prediction of thresholds, loudness, and partial loudness," *J. Audio Eng. Soc.*, vol. 45, no. 4, pp. 224–240, 1997.

[39] C. M. McKay, K. R. Henshall, R. J. Farrell, and H. McDermott, "A practical method of predicting the loudness of complex electrical stimuli," *J. Acoust. Soc. Amer.*, vol. 113, no. 4, pt. 1, pp. 2054–2063, 2003.

[40] T. Francart and H. McDermott, "Development of a loudness normalisation strategy for combined cochlear implant and acoustic stimulation," *Hear Res.*, vol. 294, pp. 114–124, Dec. 2012.

[41] T. Francart, A. Lenssen, and J. Wouters, "Sensitivity of bimodal listeners to interaural time differences with modulated single- and multiple channel stimuli," *Audiol. Neurootol.*, vol. 16, no. 2, pp. 82–92, 2011.

[42] T. Francart, A. Lenssen, and J. Wouters, "The effect of interaural differences in envelope shape on the perceived location of sounds (l)," *J. Acoust. Soc. Amer.*, vol. 132, pp. 611–614, Aug. 2012.

[43] T. Francart, A. Lenssen, and J. Wouters, "Modulation emphasis improves perception of interaural time differences in vowels with bimodal stimulation," *PLoS One*, under review.

[44] R. Plomp and A. M. Mimpen, "Improving the reliability of testing the speech reception threshold for sentences," *Audiology*, vol. 18, no. 1, pp. 43–52, 1979.

[45] H. Luts, K. Eneman, J. Wouters, M. Schulte, M. Vormann, M. Buechler, N. Dillier, R. Houben, W. A. Dreschler, M. Froehlich, H. Puder, G. Grimm, V. Hohmann, A. Leijon, A. Lombard, D. Mauler, and A. Spriet, "Multicenter evaluation of signal enhancement algorithms for hearing aids," *J. Acoust. Soc. Amer.*, vol. 127, pp. 2054–2063, Mar. 2010.

[46] P. C. Loizou and G. Kim, "Reasons why current speech-enhancement algorithms do not improve speech intelligibility and suggested solutions," *IEEE Trans. Audio Speech Lang. Process.*, vol. 19, no. 1, pp. 47–56, Jan. 2011.

[47] P. W. Dawson, S. J. Mauger, and A. A. Hersbach, "Clinical evaluation of signal-to-noise ratio-based noise reduction in nucleus cochlear implant recipients," *Ear Hear.*, vol. 32, no. 3, pp. 382–390, 2011.

[48] S. J. Mauger, P. W. Dawson, and A. A. Hersbach, "Perceptually optimized gain function for cochlear implant signal-to-noise ratio based noise reduction," *J. Acoust. Soc. Amer.*, vol. 131, no. 1, pp. 327–336, 2012.

[49] A. Spriet, L. Van Deun, K. Eftaxiadis, J. Laneau, M. Moonen, B. van Dijk, A. van Wieringen, and J. Wouters, "Speech understanding in background noise with the two-microphone adaptive beamformer beam in the nucleus freedom cochlear implant system," *Ear Hear.*, vol. 28, pp. 62–72, Feb. 2007.

[50] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic beamforming for hearing aid applications," in *Handbook on Array Processing and Sensor Networks*. New York, NY, USA: Wiley, 2010, pp. 269–302.

[51] V. Hamacher, U. Kornagel, T. Lotter, and H. Puder, "Binaural signal processing in hearing aids: Technologies and algorithms," in *Advances in Digital Speech Transmission*. New York, NY, USA: Wiley, 2008, pp. 401–429.

[52] T. Wittkop and V. Hohmann, "Strategy-selective noise reduction for binaural digital hearing aids," *Speech Commun.*, vol. 39, no. 1–2, pp. 111–138, 2003.

[53] G. Grimm, V. Hohmann, and B. Kollmeier, "Increase and subjective evaluation of feedback stability in hearing aids by a binaural coherence-based noise reduction scheme," *IEEE Trans. Audio Speech Lang. Process.*, vol. 17, no. 7, pp. 1408–1419, Sep. 2009.

[54] T. Lotter and P. Vary, "Dual-channel speech enhancement by superdirective beamforming," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 175–175, Jan. 2006.

[55] T. Rohdenburg, "Development and objective perceptual quality assessment of monaural and binaural noise reduction schemes for hearing aids," Ph.D. dissertation, Dept. Med. Phys., Univ. Oldenburg, Oldenburg, Germany, 2008.

[56] J. Desloge, W. Rabinowitz, and P. Zurek, "Microphone-array hearing aids with binaural output—Part I: Fixed-processing systems," *IEEE Trans. Speech Audio Process.*, vol. 5, no. 6, pp. 529–542, Nov. 1997.

[57] F. L. Luo, J. Y. Yang, C. Pavlovic, and A. Nehorai, "Adaptive null-forming scheme in digital hearing aids," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1583–1590, Jul. 2003.

[58] D. Welker, J. Greenberg, J. Desloge, and P. Zurek, "Microphone-array hearing aids with binaural output—Part II: A two-microphone adaptive system," *IEEE Trans. Speech Audio Process.*, vol. 5, no. 6, pp. 543–551, Nov. 1997.

[59] T. van den Bogaert, S. Doclo, J. Wouters, and M. Moonen, "The effect of multimicrophone noise reduction systems on sound source localization by users of binaural hearing aids," *J. Acoust. Soc. Amer.*, vol. 124, pp. 484–497, Jul. 2008.

[60] S. Doclo, A. Spriet, J. Wouters, and M. Moonen, "Frequency-domain criterion for speech distortion weighted multichannel Wiener filter for robust noise reduction," *Speech Commun.*, vol. 49, pp. 636–656, Jul.–Aug. 2007.

[61] B. Cornelis, S. Doclo, T. van den Bogaert, J. Wouters, and M. Moonen, "Theoretical analysis of binaural multi-microphone noise reduction techniques," *IEEE Trans. Audio Speech Lang. Process.*, vol. 18, no. 2, pp. 342–355, Feb. 2010.

[62] T. van den Bogaert, S. Doclo, J. Wouters, and M. Moonen, "Speech enhancement with multichannel Wiener filter techniques in multimicrophone binaural hearing aids," *J. Acoust. Soc. Amer.*, vol. 125, pp. 360–371, Jan. 2009.

[63] B. Cornelis, M. Moonen, and J. Wouters, "Speech intelligibility improvements with hearing aids using bilateral and binaural adaptive multichannel Wiener filtering based noise reduction," *J. Acoust. Soc. Amer.*, vol. 131, pp. 4743–4755, Jun. 2012.

ABOUT THE AUTHORS

Jan Wouters was born in 1960. He received the M.S. and Ph.D. degrees in physics from the University of Leuven (KU Leuven), Leuven, Belgium, in 1982 and 1989, respectively, with intermission for officer military service.

From 1989 until 1992, he was a Postdoctoral Research Fellow with the Belgian National Fund for Scientific Research (FWO), Institute of Nuclear Physics (UCL Louvain-la-Neuve) and at NASA Goddard Space Flight Center, Greenbelt, MD, USA. Since 1993, he has been a Professor at the Neurosciences Department, KU Leuven (Full Professor since 2005). He is responsible for the hearing research group at Experimental Oto-rhino-laryngology (ExpORL), Department of Neurosciences, and teaches five physics and audiology courses at the Faculty of Medicine, KU Leuven. His research activities center around audiology and the auditory system, signal processing for cochlear implants and hearing aids. He is an author of about 200 articles in international peer-reviewed journals, and he is a reviewer for several international journals.

Dr. Wouters received an Award of the Flemish Government in 1989, a Fulbright Award and a NATO Research Fellowship in 1992, the Flemish VVL Speech Therapy—Audiology Award in 1996, and an award from the Deutsche Gesellschaft für Audiologie (DGA) in 2007. He is on the editorial board of the *International Journal of Audiology*, the *Journal of Communication Disorders*, and the journal *B-ENT*. He is the President of the European Federation of Audiology Societies (EFAS), member of the International Collegium for ORL (CORLAS), and a board member of the International Collegium for Rehabilitative Audiology (ICRA).

Simon Doclo received the M.Sc. degree in electrical engineering and the Ph.D. degree in applied sciences from the University of Leuven (KU Leuven), Leuven, Belgium, in 1997 and 2003, respectively.

From 2003 to 2007, he was a Postdoctoral Fellow with the Research Foundation—Flanders at KU Leuven, and McMaster University, Hamilton, ON, Canada. From 2007 to 2009, he was a Principal Scientist with NXP Semiconductors at the Sound and Acoustics Group in Belgium. Since 2009, he has been a Full Professor at the University of Oldenburg, Oldenburg, Germany, and scientific advisor for the project group Hearing, Speech and Audio Technology of the Fraunhofer Institute for Digital Media Technology. His research activities center around signal processing for acoustical and biomedical applications, more specifically microphone array processing, active noise control, acoustic sensor networks, and hearing aid processing.

Prof. Doclo received the Master Thesis Award of the Royal Flemish Society of Engineers in 1997 (with E. De Clippel), the Best Student Paper Award at the International Workshop on Acoustic Echo and Noise Control in 2001, the EURASIP Signal Processing Best Paper Award in 2003 (with M. Moonen), and the IEEE Signal Processing Society 2008 Best Paper Award (with J. Chen, J. Benesty, and A. Huang). He is a member of the IEEE Signal Processing Society Technical Committee on Audio and Acoustic Signal Processing (2008–2013). He has been Secretary of the IEEE Benelux Signal Processing Chapter (1998–2002), has served as a Guest Editor for the *EURASIP Journal on Advances in Signal Processing* and *Elsevier Signal Processing* and was the Technical Program Chair for the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) in 2013.

Raphael Koning was born in 1985 in Hattingen, Germany. He received the Dipl.-Ing. degree in electrical engineering from the Ruhr University Bochum, Bochum, Germany, in 2009.

He became a member of the research group Experimental Oto-rhino-laryngology (ExpORL), Department Neurosciences, University of Leuven (KU Leuven), Leuven, Belgium, and received a Marie Curie Scholarship for Doctoral studies under the European Union (EU) Marie Curie ITN Project ''Digital Signal Processing in Audiology'' (AUDIS) in 2009. His general research interests include digital speech signal processing and speech enhancement. In particular, he is interested in the development and application of speech enhancement algorithms in hearing aids and cochlear implants.

Tom Francart was born in 1981. He received the M.S. and Ph.D. degrees in engineering from the University of Leuven (KU Leuven), Leuven, Belgium, in 2004 and 2008, respectively.

From 2010 to 2013, he was a Postdoctoral Fellow of the Belgian Fund for Scientific Research (FWO) and he holds a Marie Curie International Outgoing Fellowship from the European Commission (EC), which funded his stay from 2010 to 2012 at the Bionics Institute, Melbourne, Vic., Australia.