19th INTERNATIONAL CONGRESS ON ACOUSTICS
MADRID, 2-7 SEPTEMBER 2007

AUDITORY SCENE ANALYSIS AND THE PRECEDENCE-EFFECT WITH COCHLEAR IMPLANTS – PREDICTIONS FROM SIMULATIONS

PACS: 43.66.Pn

Seeber, Bernhard U.
Auditory Perception Lab, Dept. of Psychology, University of California at Berkeley, Berkeley, CA
Now: MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, UK;
seeber@ihr.mrc.ac.uk

ABSTRACT

The analysis of auditory scenes is based on the evaluation of monaural spectral and temporal cues as well as on binaural cues. Using current devices cochlear implant (CI) subjects have very limited access to interaural time differences (ITDs) and spectral information while interaural level cues (ILDs) and the temporal envelope are relatively well reproduced. Localization in quiet is often possible but the presence of other sounds disturbs it. We previously showed that the precedence-effect was disturbed in a noise-band vocoder CI-simulation for ongoing sounds and lead and lag images were heard in separation. The present study shows that temporal quantization of the envelope in the vocoder has only minor effect on precedence. Further study varied the interaural match of the carrier frequencies in a sinusoidal vocoder. Precedence broke down if carriers were not matched in frequency. However, for zero or small frequency offsets some subjects showed precedence and fusion into one image. Interestingly, the interaural phase of the carrier played only a limited role which shows that precedence of ongoing sounds can solely be based on ILDs and envelope-ITDs provided that spectral cues are matched across the ears. The results suggest that proper place matching of CI-electrodes would help the analysis of concurrent sounds.

INTRODUCTION

The auditory system possesses the amazing ability to analyze the auditory scene into separate objects that are present simultaneously, often at different locations. The information for this is taken from the spectrum, from temporal relationships of components, and from binaural cues in time and intensity. In impaired hearing and with CIs in particular, some of this information is missing. Current generations of CIs do not encode ITDs in the carrier pulses and access to spectral information (pitch) is very limited. For coping with situations of hearing in noise or reverberation, however, previous studies found that carrier-ITDs are essential, at least for the localization of longer duration sounds [1-5]. An explicit way to study the auditory system's ability to deal with reverberation is to investigate precedence or summing localization: Subjects localize a sound that is accompanied by its echo. Localization will be at the leading sound for short delay times and the echo will not be audible by itself. When we previously simulated CI-listening with a noise-band vocoder we found that precedence breaks down, i.e. that the echo is not suppressed, but separately audible [6]. The following two mechanisms which we will address in the present paper might contribute to the breakdown of precedence:

a) The information present in the envelope might decorrelate signals on both ears so far that two images appear in both ears even if a correlated carrier is used. We investigated the influence of envelope quantization on precedence. The quantization should better simulate the pulsatile stimulation of CIs compared to low-pass filtering of the envelope which we previously used. Similar quantization on both ears should also reduce the availability of envelope-ITDs and increase the interaural correlation of the envelope.
b) We wonder if the breakdown of precedence is caused by an absence of auditory object grouping in and across the ears. Auditory scene analysis is predominantly influenced by spectral grouping. The occurrence of separate images in both ears might thus be based on the lack of spectral grouping across the ears. We studied this hypothesis by varying carrier frequencies in a sinusoidal vocoder.

METHODS

Experimental paradigms

The precedence effect was investigated in a localization dominance experiment. The lead and lag sounds were played from virtual directions ±30° with a probability of 0.5 that the lead would be on the left. The lead-lag delay was varied within 0-60 ms, with the upper limit depending on the stimulus. In one experimental session subjects were instructed to localize the leftmost of the sound images if they perceived more than a single image. Randomizing the side of lead and lag sounds on every trial thus meant that subjects responded to the lead on one half of the trials and the lag on the other half without having to determine which of the sounds came first. Pointing biases were reduced by localizing the rightmost sound in separate sessions. In other conditions subjects were asked to localize the most dominant or the weakest image. The following experimental paradigms were studied:

(1) Baselines were gained for precedence in the free field with two loudspeakers at ±30° and with virtual acoustics. Results were discussed in [6].
(2) The precedence-stimuli were processed through a binaural noise-band vocoder to simulate CI-processing. The results showed a breakdown of precedence, i.e. the lag was audible [6].
(3) A CI-simulation based on the noise-band vocoder of paradigm 2 was used, but channel envelopes were quantized in 1.5 ms steps before being applied to the carrier noise. This was done to reduce the impact of envelope-ITDs and to render the simulation more realistic.
(4) The CI-simulation was modified to use sinusoids as the carrier instead of noise bands. The frequency of the sinusoids was varied randomly separately on each ear within a restricted range. The influence of the frequency range and of interaural carrier phase was studied.

Vocoder-simulation of cochlear implants

A 16-channel vocoder was used to simulate CI-processing. The precedence sounds were first filtered with individually selected head-related transfer functions (HRTFs) [7]. In the vocoder, the HRTF-processed sound was band-pass filtered in 16 logarithmically spaced channels in 300 Hz - 8 kHz. The channel-envelopes were computed by rectification and low-pass filtering at 200 Hz. Envelopes were computed independently for both ears. In experiments 2 and 3 envelopes were applied to noise bands (carrier) of varying interaural correlation. In experiment 3 the envelope was quantized in 1.5 ms steps and low-pass filtered with a cut-off at 1 kHz to reduce spectral splatter before being used to modulate the noise. In experiment 4 sinusoids were modulated by the envelope. The frequency of each carrier sinusoid was chosen randomly in each trial within a range around the center frequency of the analysis channel. The frequency range was varied, but carrier frequency was always limited to the frequencies covered by the analysis filter of that channel. This resulted in different carrier frequencies in each ear which roughly lay within the same critical band. Carrier phase was chosen at random. Interaural phase was undefined, except for one condition in which phase and frequency were interaurally matched. In another condition carrier phase was left at random, but carrier frequency was identical on each ear. Spectral splatter due to envelope modulation was limited by bandpass-filtering the modulated carriers with filters three times the bandwidth of the analysis filters.

Subjects and Stimuli

Table 1: Overview of stimuli.

Stimulus[1]	Bandwidth [Hz]	Envelope	Duration [ms]	Level [dB(A)]
WBN Burst	300-10000	1 ms Gaussian	10	60
LPN Burst 770 Hz	300-770	1 ms Gaussian	10	60
“Shape”	Speech, CVC word		800	55

1 WBN: Wide-band noise (Gaussian noise); LPN: Low-pass noise.

Table 1 gives an overview of all stimuli regarding their type, bandwidth, envelope structure, duration, and level. Four subjects, age 19-29 years, participated in experiments 3 and 4. Results are presented of one subject (male, age 21). All subjects had normal hearing thresholds within 300 Hz – 10 kHz as assessed with a Bèkèsy-tracking procedure. Subjects received payment for their participation. The study protocol was approved by the ethics committee at the University of California at Berkeley.

Experimental Procedures

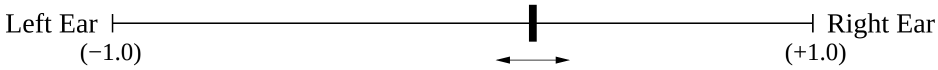


Figure 1: Line dissection method: The line was projected on a screen in 2m distance in front of the subject. The bar could be adjusted with a trackball to the position of the lateralized sound image.

Lateralization of the precedence stimuli was studied with a line-dissection method (Figure 1). A white line was projected on an otherwise dark screen in 2m distance in front of the subject. It covered a visual angle of approx. ±25° and the endpoints were marked with vertical bars and the words “left ear” or “right ear”. The subject adjusted a red bar to the lateralized sound position with a trackball. For data analysis the left ear position is assigned -1, the right ear +1.

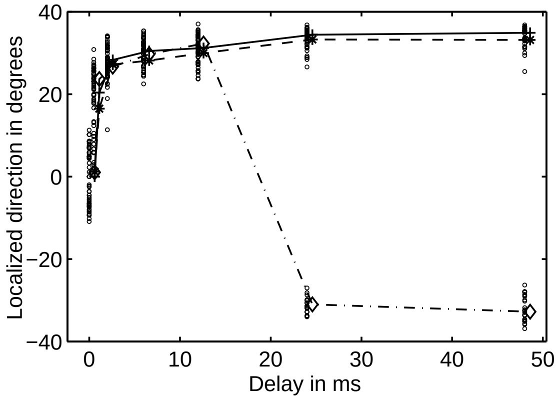
A fully factored design was used to investigate precedence for different interaural correlations of the carrier noise (exp. 2 and 3) or for different widths of the range of carrier frequencies (exp. 4). Ten trials were collected for each condition (3 sounds * 2 lead/lag directions * 6 lead-lag delays * 5 correlations/frequency ranges * 10 trials = 1800 trials total). The 1800 trials were administered in random order divided into 15 runs. Level was roved randomly in 2dB-steps within ±6dB of the target sound level (Table 1). Each run took about 8 min to complete and subjects had to take short breaks between runs. Subjects received training prior to data collection which consisted of 2 runs in which all conditions for all sounds were presented twice. The experiment was repeated three times with different instructions: a) point to the left image, b) point to the right image, c) point to the dominant image if you hear two or more sounds.

In a single trial the sound was presented followed by a pause of 0.5 sec after which the adjustable bar appeared in the middle of the line. The subject moved the bar to the lateralized sound position and confirmed this by pressing a button on the trackball. The left button coded hearing a single image while subjects were instructed to press the right button if two or more images were perceived. The bar disappeared and after 0.5 sec the next sound was presented. The line was visible throughout the experiment. At the beginning of each experimental run five uncounted random trails were presented for accommodation. Experiments utilized the Simulated Open-Field Environment in the otherwise darkened anechoic chamber [8]. Stimuli were presented through a calibrated, diffuse-field equalized Sennheiser HD580 headphone.

RESULTS AND DISCUSSION

Figure 1 gives baseline precedence results for the subject of the current study for sounds presented in the free field. For delays up to 12 ms precedence is strong and a single image is reported at the lead while no responses appear at the lag. At 24 ms delay the echo threshold is exceeded and two images are reported, one at the lead and one at the lag. Figure 2 shows the corresponding results obtained with a noise-band vocoder with a correlated carrier (experiment 2). At all delay times responses appear at the lag, indicating an incomplete suppression of the lag. Nevertheless, the subject responded a few more times to the lead resulting in medians located at the lead. The carrier consisted of correlated noise which by itself is localized centered in the head. The binaural information in the envelope, ILDs and envelope-ITDs, was sufficient to widen the image and even to suggest two separate images. Full lag suppression, however, did not occur which we previously termed the breakdown of precedence [6]. The following experiments investigated, a) if the breakdown could be altered by reducing the temporal information in the envelope that might suggest two images, or b) by changing interaural grouping mechanisms through using sinusoidal carriers [1, 6].

The influence of envelope information was tested in experiment 3. The low-pass filtered envelope was additionally quantized every 1.5 ms which simulates a stimulation rate of 667pps. Figure 3 shows results for again the same subject for the word “shape”. The CI-vocoder used for Figure 2 and the one for Figure 3 differed only by the envelope quantization and both used a correlated carrier. Results look almost identical, suggesting that envelope quantization has only minor effects on precedence grouping. This might be in part due to the low cut-off frequency of 200 Hz used for envelope extraction. However, informal listening found a similar breakdown of precedence for higher cut-off frequencies, at least when spectral splatter was contained to a



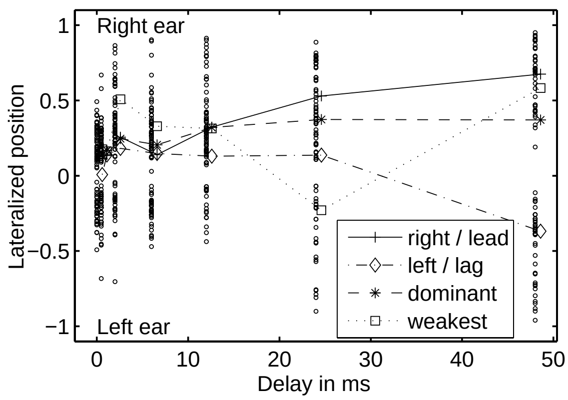


Fig. 1 (left): Results of experiment 1 on precedence in the *free-field*. Data from one subject are presented for the CVC-word *"shape"*. Scattered localization responses and medians are shown as a function of the lead-lag delay time in the precedence experiment. In different sessions the subject was instructed to respond either to the image on the rightmost (+, here: leading), the leftmost (◊, lagging), or the dominant (*) sound image if two images were heard (see legend in Figure 2). In the experiment the lead was played randomly from ±30°, but for clarity in the picture the lead is plotted at +30° and the lag at -30°. Data plotted at the lead (+) were combined from data for pointing to the rightmost image if the lead was on the right and from side-inverted data for pointing to the leftmost image if the lead was on the left. Data for the lag (◊) were combined in a similar way.

Fig. 2 (right): Results of experiment 2 for precedence with *CI-simulation* with interaurally *correlated carrier noise* for the word *"shape"*. Lateralization was measured with the ears depicted by ±1. Results for pointing to the lead/lag in either the left or rightmost sound image were combined and plotted analog to Fig. 1 with the lead at +1. In one additional session the subject was instructed to point to weakest (□) sound image if two images were heard. The results show the breakdown of precedence: Despite the correlation of the carrier with its centralizing effect images are reported on the lag side.

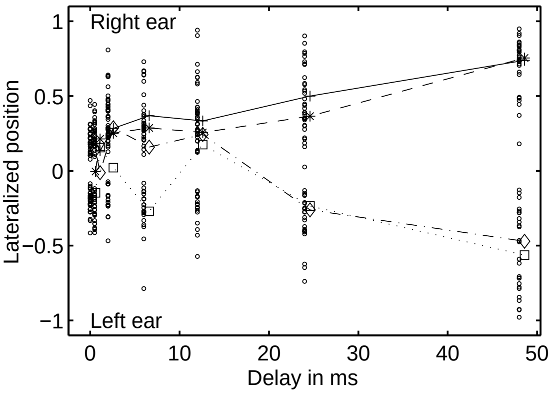


Fig. 3: Results of experiment 3 for precedence with a CI-simulation with *temporal quantization* of the envelope. Interaurally correlated carrier noise was used. The stimulus was the word "shape". The data presentation is analog to Figure 2 and the only difference resides in the envelope quantization. As results are highly similar, envelope quantization seems to have only minor effects on binaural grouping.

reasonable frequency range by using an output-filterbank. Reducing envelope information does not lead to increased fusion of both images, not even when a correlated carrier is used.

Experiment 4 addressed hypothesis 2 that the breakdown of precedence is less a binaural phenomenon but rather due to disturbed auditory scene analysis. The idea is that the carrier noise is detrimental to interaural fusion compared to sinusoidal carriers because it does not elicit a clear pitch. Further, sinusoidal carriers might enhance monaural grouping, provided there is a harmonic relationship between them. Figures 4 and 5 display selected results of experiment 4. For unmatched interaural carrier frequencies precedence breaks down and the lag is clearly audible (Figure 4). Lead and lag images are perceived close to both ears. This result is very similar to what was previously obtained with uncorrelated noise as a carrier [6].

The results in Figure 5 stem from a condition in which the carrier frequencies were identical on both ears, but interaural phases were random. For this subject ideal precedence was observed, characterized by localization of a single image at the lead location at short delays and the complete absence of responses at the lag. In comparison to Figure 1 which presents data from the free-field similar echo thresholds are found. Unlike free-field results responses are split in

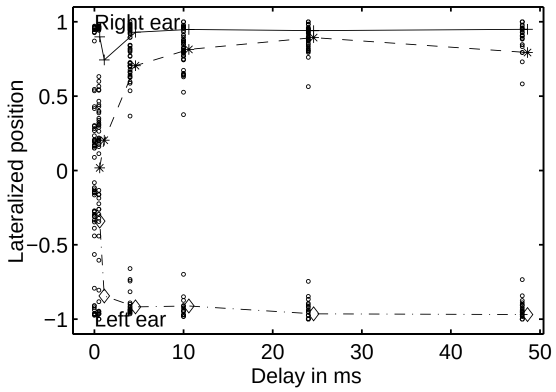


Fig. 4: Results of experiment 4 on precedence with *sinusoidal carrier* CI-simulation with *carrier frequencies chosen randomly* on both ears within the full analysis frequency interval. Data are presented for the same subject as above for the word "shape" in identical fashion to previous figures. With unmatched interaural carrier frequencies that differ roughly no more than a critical band, precedence breaks down similar to an uncorrelated noise carrier [6]. Two separate images are localized at both ears without the centralizing effect of the correlated carrier seen in Fig. 2.

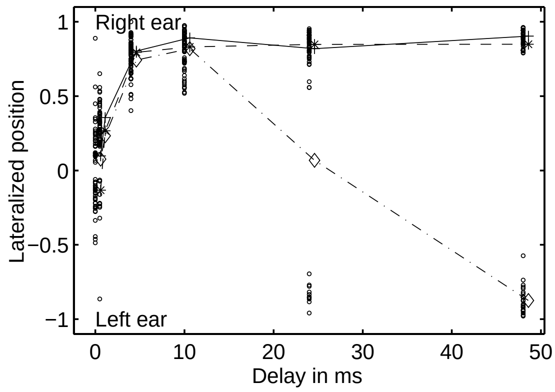


Fig. 5: Results of experiment 4 on precedence with *sinusoidal carrier* CI-simulation with *carrier frequencies identical* on both ears, but *random interaural phase* in the carrier. Subject, stimulus and way of presentation are identical to Figure 4. The subject shows perfect precedence almost comparable to the free-field results despite the random interaural carrier phase. This shows that even for a wide-band stimulus of longer duration ITD- and ILD-information in the envelope may be sufficient for precedence.

two nearby images for no delay (0 ms) and they are more centered at 0.5 ms delay. The latter might be caused by the low-pass filtering in the envelope computation. Although the selected subject shows ideal precedence some other subjects still give some responses at the lag. Common to all subjects is an increase in lag-suppression and a better tendency to form only a single image when the separation of interaural carrier frequencies decreases.

Precedence with CI-simulation was shown to be possible by using a vocoder with sinusoidal carriers of equal frequency in both ears. The surprising result is that near-perfect precedence could be obtained with random interaural phase in the carrier. Precedence must have been based entirely on the ITDs and ILDs in the envelope while random ITDs in the carrier were ignored. To our knowledge, this altered cue weighting has rarely been seen for natural wide-band sounds of long duration. Several studies have previously found that those sounds are localized on the basis of ITDs at low frequencies (in the carrier). Particularly in situations with repeated, brief sounds the envelope-ITD in the onset plays an increased role, but for sounds of longer duration, longer rise-times, more repetitions or sparser repetition intervals, or non-identical repetitions the ITD is evaluated from the ongoing part [2, 3, 5]. A speech sound with components down to 300 Hz is thus most likely be localized on the basis of carrier-ITDs and the present ignorance of those cues is surprising. The precedence condition must have lead to a cue weighting in which ILDs and envelope-ITDs were enhanced. We can speculate if the relatively sparse spectral composition of the sound and the inharmonic relationship of the carriers might have had an influence.

Precedence was observed more strongly when the interaural frequency difference in the carrier decreased. In agreement with our second hypothesis this may be based on increased fusion of spectral components presented to both ears.

The present results have a wide-ranging impact on CIs. Carrier pulse trains used with current CIs do not encode phase-ITDs and the current spread in the cochlea leads to a weak representation of pitch. The present results show that the auditory system can deal with multiple sounds at a time without receiving sound-specific information in the carrier. However, carriers need to be frequency-matched across ears – a prerequisite that is hard to fulfill with current CIs. Since the cochlea is stimulated over a wide region with CIs it is also questionable if this rather corresponds to the usage of sinusoids in the simulation, or if noise-bands which excite the cochlea over a wider range would be more appropriate in a simulation. However, in the simulation with a noise-carrier precedence was not as pronounced, suggesting that a narrow focus for the place of stimulation is required.

CONCLUSIONS

The current study provides evidence that the auditory system may suppress reflections of ongoing natural sounds on the basis of ILDs and ITDs in the envelope alone while previous studies instead emphasized the importance of ITD-information in the carrier for localization of ongoing stimuli. The present study showed that carrier information including phase-ITDs might be disregarded for the precedence effect if carrier frequencies are interaurally matched and interaural carrier phases are random. The results have a wide impact on cochlear implants as they suggest that proper place matching may enable patients to deal better with reflections or noisy environments in general.

ACKNOWLEDGEMENTS

I am indebted to Prof. Ervin Hafter for his extensive support. This work was financed by NIH RO1 DCD 00087. Parts of the data analysis and the write-up of the manuscript occurred after BS joined the MRC.

References:

[1] J. F. Culling and C. J. Darwin, "Perceptual separation of simultaneous vowels: Within and across-formant grouping," *J. Acoust. Soc. Am.*, vol. 93, pp. 3454-3467, 1993.

[2] R. L. Freyman and P. M. Zurek, "Onset dominance in lateralization," *J. Acoust. Soc. Am.*, vol. 101, pp. 1649-1659, 1997.

[3] Z. M. Smith, B. Delgutte, and A. J. Oxenham, "Chimaeric sounds reveal dichotomies in auditory perception," *Nature*, vol. 416, pp. 87-90, 2002.

[4] T. N. Buell and E. R. Hafter, "Discrimination of interaural differences of time in the envelopes of high-frequency signals: Integration times," *J. Acoust. Soc. Am.*, vol. 84, pp. 2063-2066, 1999.

[5] E. Hafter, "Binaural Adaptation and the Effectiveness of a Stimulus beyond Its Onset," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R. H. Gilkey and T. R. Anderson, Eds. Mahwah, New Jersey: Lawrence Erlbaum Ass., 1996, pp. 211-232.

[6] B. Seeber and E. Hafter, "Precedence-effect with cochlear implant simulation," in *Hearing - from basic research to applications (Int. Symposium on Hearing 2006)*: Springer, 2006.

[7] B. U. Seeber and H. Fastl, "Subjective Selection of Non-Individual Head-Related Transfer Functions," in *Proc. 9th Int. Conf. on Aud. Display*, E. Brazil and B. Shinn-Cunningham, Eds. Boston, USA: Boston University Publications Prod. Dept., 2003, pp. 259-262.

[8] E. Hafter and B. Seeber, "The Simulated Open Field Environment for auditory localization research," in *Proc. ICA 2004, 18th Int. Congress on Acoustics, Kyoto, Japan, 4.-9.04.2004*, vol. V: Int. Commission on Acoustics, 2004, pp. 3751-3754.