Emphasis of spatial cues in the temporal fine structure during the rising segments of amplitude-modulated sounds

Mathias Dietz[a,1], Torsten Marquardt[a], Nelli H. Salminen[a,b], and David McAlpine[a]

[a]University College London Ear Institute, London WC1X 8EE, United Kingdom; and [b]NeuroImaging Brain and Mind Laboratory, Department of Biomedical Engineering and Computational Science, and MEG Core, Aalto NeuroImaging, Aalto University School of Science, FI-00076, Aalto, Finland

The ability to locate the direction of a target sound in a background of competing sources is critical to the survival of many species and important for human communication. Nevertheless, brain mechanisms that provide for such accurate localization abilities remain poorly understood. In particular, it remains unclear how the auditory brain is able to extract reliable spatial information directly from the source when competing sounds and reflections dominate all but the earliest moments of the sound wave reaching each ear. We developed a stimulus mimicking the mutual relationship of sound amplitude and binaural cues, characteristic to reverberant speech. This stimulus, named amplitude modulated binaural beat, allows for a parametric and isolated change of modulation frequency and phase relations. Employing magnetoencephalography and psychoacoustics it is demonstrated that the auditory brain uses binaural information in the stimulus fine structure only during the rising portion of each modulation cycle, rendering spatial information recoverable in an otherwise unlocalizable sound. The data suggest that amplitude modulation provides a means of "glimpsing" low-frequency spatial cues in a manner that benefits listening in noisy or reverberant environments.

spatial hearing | binaural processing | auditory system | psychoacoustics | auditory MEG

Human listeners are able to determine the location of a talker against a background of competing voices, even in rooms where walls generate reflections that, taken together, can be more intense than sounds arriving directly from the source. The dominant cues for localization in such complex sound fields are the interaural time differences (ITDs) conveyed in the temporal fine structure (TFS) of low-frequency (<1,500 Hz) sounds (1); normal-hearing listeners can discriminate ITDs as low as 10–20 μs in 500- and 1,000-Hz pure tones to judge the source location (2). In addition to source localization—the focus of the current study—sensitivity to ITDs is also reported to contribute to "spatial unmasking": different spatial configurations of the signal and background noise enable sources to be heard out, increasing their intelligibility (3, 4).

The majority of real-world sounds are strongly modulated in amplitude. Without these modulations humans are completely insensitive to the sound source location after its onset in reverberant environments [the Franssen effect (5)]. Human speech, for example, contains amplitude modulation (AM) rates ranging from those of syllables and phonemes to those conveying information about voice pitch (i.e., from 2 Hz up to about 300 Hz). These modulations act as potent grouping cues, enabling listeners to fuse sounds originating from a single talker, segregating them from competing talkers (6). Despite the importance of AM in real-world listening, however, behavioral measures of ITD sensitivity are commonly assessed for stimuli in which the amplitude is unmodulated. This is especially so when the focus of interest concerns ITDs conveyed in the TFS.

Fig. 1 illustrates the signals from a modulated, speech-like sound source arriving at the left and right ears in a listening environment comprising two reflecting walls. Whereas the direct sound arrives from the left-hand side, the two reflections arrive 4 and 8 ms later from the right. Despite the time-averaged interaural cues suggesting a sound originating from the right, listeners perceive this sound to originate from the left.* Extracting the ITD information from the TFS reveals that left-leading ITDs, corresponding to the true origin of the source, appear reliably, albeit briefly, following each energy minimum in the modulated waveform, whereas other segments, including the modulation peaks, are dominated by right-leading ITDs. Based simply on the average ITD, therefore, it is not self-evident why listeners robustly perceive this sound to originate from the left.

Although studies of spatial hearing suggest a perceptual dominance of early-arriving acoustic information (7), the non-deterministic (and ongoing) nature of speech waveforms considerably complicates assessment of which sound segments underpin its localization. Investigations into the related precedence effect (8) have tended to use pairs (or trains) of transient "clicks" (8, 9) to assess the dominance of the first wave front, but these stimuli have a very different temporal modulation compared with

Fig. 1. Acoustic properties of a spoken trilled <r>, low-pass–filtered at 800 Hz, to focus on temporal fine structure ITD. (*A*) Spatial configuration obtained by virtual acoustics. Direct sound presented from −30° (left leading), together with two delayed copies from the right (+65°; 4-ms delay and +130°, 8-ms delay). The signal has a negative direct-to-reverberant ratio [−4 dB; mean ILD: +3 dB (right louder); mean ITD: +140 μs (right leading)] and an even larger amplitude weighted mean ITD (+252 μs). It might, therefore, come as unexpected that informal tests revealed for all 10 participants a localization of the sound toward the left hand side, similar to the direct sound presented without reverberation. The enve-lope slope weighted mean ITD (−197 μs) appears to be a better indicator for perceived localization. (*B*) Acoustic signals at the left and right ear. The regular modulation pattern with ∼45-ms cycle duration is evident. Because the ITD cannot be identified by eye at this scale, the sign of the ITD is indicated by color-coding the Hilbert envelope (pale blue for left-leading ITDs and pale red for right-leading ITDs while within the human physiological range of ±700 μs). The ITD is consistently left-leading at the begin-

ning of each cycle but right-leading during the peak or end of the modulation cycle. (*C*) A magnified three-cycle segment from *B* is replotted together with the ITD (green). In this resolution, the fine-structure phase is visible. ITDs outside of the physiologic range occur during incoherent segments and are not plotted. (*D*) The same segment as in *C* but without any reflections; only the sound directly from the source arrives at the ears. A constant left-leading ITD is visible.
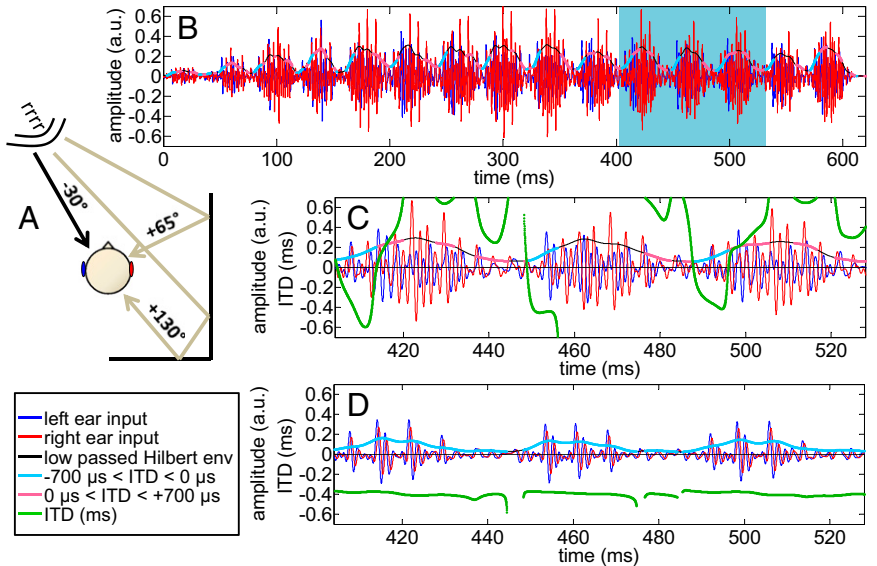
continuous speech signals. Here, we developed a class of stimuli [amplitude-modulated binaural beats (AMBBs)] designed to mimic the variation of ITD in the TFS for sounds heard in complex listening environments (such as that in Fig. 1). AMBBs enable a quantitative analysis of the time-dependent ITD sensitivity of ongoing sounds in the absence of spatial cues associated with the stimulus envelope. The unique property of the AMBB is that the rate at which the sound amplitude is modulated is identical to the binaural beat rate [i.e., each AM cycle accommodates a full 360° rotation in interaural phase difference (IPD)]. By these means, each position within the AM cycle corresponds to a spe-cific IPD (Fig. 2). This common rate of AM and binaural beating is referred to as the AMBB frequency.

The correlated change in IPD and the phase of the AM is typical of reverberant speech (e.g., Fig. 1) and is designed to mimic the rapid changes in ITD commonly encountered in reverberant listening. To this end, AMBBs provide highly controlled means for assessing those moments in time at which the ITD is read out from amplitude-modulated sounds.

Note that for beat rates above about 2 Hz, neither binaural beats nor AM alone elicits a lateralized percept in human listeners. The unmodulated binaural beat is perceived as having a diffuse in-tracranial image (10, 11), whereas diotic AM, in the absence of a binaural beat, obviously favors neither left nor right lateraliza-tion. With both binaural beating and AM in combination, how-ever, the resulting AMBB can elicit a compact intracranial image with a lateralization percept that depends on the phase relation-ship between the AM and binaural beat.

Results

Behavioral Assessment. To determine where within the modulation cycle ITD information dominates the spatial percept, listeners were asked to adjust the static IPD of a sinusoidally amplitude-modulated tone until it matched as closely as possible the spatial percept (including lateralization and compactness of the image) of a preceding AMBB. Despite each AM cycle of the AMBB cycle rotating through a full 360° of IPD, subjects were able, in the majority of cases, to match the intracranial spatial percept to that of a specific, static IPD (Fig. 3). This occurred despite the fact that AMBBs and matched static-IPD stimuli were, for most parameter settings, still distinguishable, including the reported

"compactness" of their intracranial image. In general, the static IPD that the listeners matched to the AMBBs corresponded to an IPD occurring during the early rising slope of the AMBB.

A three-way repeated-measures analysis of variance (ANOVA) on all ΔIPD (start IPD − matched static IPD) values with the factors AMBB frequency (five values), start IPD (eight values), and beat direction (two values) revealed no significant main effect of beat direction ($F_{1,28} = 0.01$; $P = 0.996$), as well as no signifi-cant interaction between beat direction and AMBB frequency ($F_{4,28} = 0.08$; $P = 0.945$) and no significant interaction between beat direction and start IPD ($F_{7,28} = 0.15$, $P = 0.941$). Because no significant influence of the beat direction was observed, data were combined across both beat directions by defining IPD with respect to beat direction (Fig. 3, filled and open symbols). This resulted in six repeats for each of the 40 conditions (5 modula-tion frequencies × 8 start IPDs).

Matching the intracranial percept of an AMBB comprising a particular "start IPD" to a static IPD indicates a constant IPD offset (ΔIPD) between the two. The fitted average ΔIPD is the cyclic mean of 48 IPD offsets derived from the 48 matches per AMBB frequency (three repetitions of two beat directions and eight start IPDs), and equivalent to the ordinate intercept at 0° start IPD. Its value indicates the position (phase) in the modu-lation cycle where the instantaneous IPD of the AMBB is equivalent to the matched static IPD and is visualized in Fig. 3 as diagonal line in the respective color. In each case, this appears to occur early in the modulation cycle, with across-subject mean values of 37°, 40°, 62°, 83°, and 115° for 4, 8, 16, 32, and 64 Hz, respectively. Although suggestive of a fixed delay, this pro-gression within the AM cycle, in fact, comprises a mixture of both time and phase delays (for quantitative analysis of this progression, see *Discussion*). Averaged across subjects, the SD of the ΔIPD values was 46°, 29°, 29°, 31°, and 37° for 4, 8, 16, 32, and 64 Hz, respectively. (For comparison, randomly distributed data would result in an average SD of 104°.) This attests to strong alignment of the data with a linear function of unity slope (the offset of these functions in Fig. 3 is equivalent to the average ΔIPD for that subject) but also indicates a decline in perfor-mance for the lowest and highest modulation frequencies. With the exception of subject (S) 4, all subjects showed greater vari-ability in their response when matching 4-Hz beats: whereas
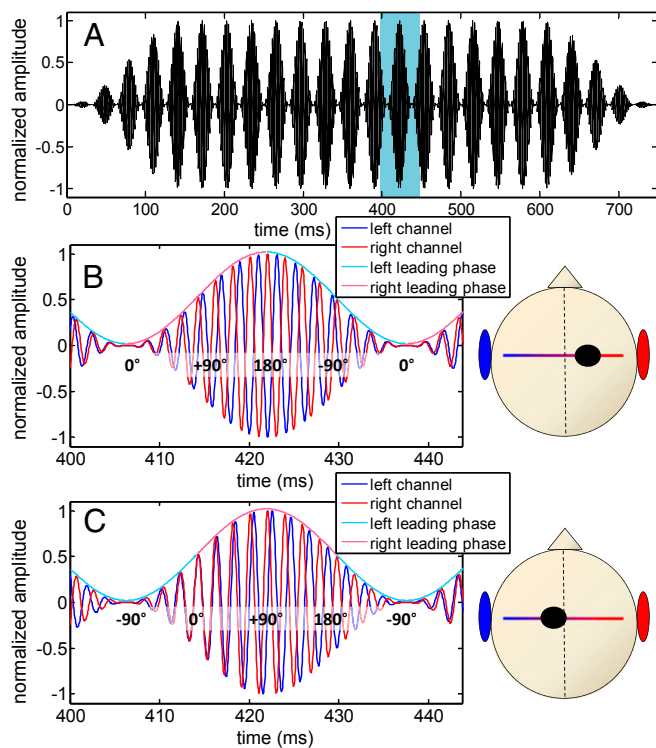
Dietz et al.

www.manaraa.com

Fig. 2. Waveform of the AMBB stimulus; 500-Hz sinusoidally amplitude-modulated tone with a 32-Hz beat frequency and AM rate. (*A*) Full waveform of one channel. (*B*) A 45-ms segment of the AMBB from *A*: $f_{left} < f_{right}$ with "start IPD" at modulation minimum $IPD_{Start} = 0$. The left channel is in blue, and the right channel is in red. The color of the envelope indicates the sign of the instantaneous IPD (blue, left leading; red, right leading). Some values of the instantaneous IPD are plotted over the signals. All subjects perceived such a stimulus clearly from the right as indicated by the dot within the cranial diagram. (*C*) Same format as *B* but now with $IPD_{Start} = -90°$. In this particular case, the IPD is right-leading when the amplitude is large (>0.5) and left-leading when the amplitude is small (<0.5). Most subjects perceive this stimulus slightly to the left even though the stimulus intensity is low during left-leading IPDs.

many matches still align with a slope of unity, others indicate that the percept of low-rate AMBBs is most similar to that of a more diffusely perceived 180°-IPD pointer. Note also that, for 64-Hz AMBBs, the data are less well-aligned compared with lower beat frequencies (for all subjects except S6). The data suggest that 4 and 64 Hz are close to the lower and upper-rate limits at which reliable IPD read out in the modulation cycle is possible. Overall, the data indicate that IPDs on the rising slope of each AM cycle dominate the lateralization percept, with a trend for the dominant segment to shift from near the modulation trough to a position just past the maximum slope (i.e., 90° after the modulation minimum) with increasing modulation frequency.

Magnetoencephalography. The behavioral data suggest an increased weighting of IPDs during the rising slope of the AM cycle and that this weighting is dependent on the modulation frequency. To assess the extent to which this perceptual weighting is evident in cortical responses, we undertook magnetoencephalography (MEG) recordings in human subjects, using an adaptation paradigm in which AMBB stimuli and AM tones with static IPDs were alternated every 500 ms, such that a continuous sound was delivered. Transitions in this continuous sound presentation generate auditory change responses ~100–200 ms following the transition (Fig. 4*A*) (12–14), and the amplitude of this change response is often inversely related to the degree of

similarity between the two sounds (14). Therefore, the weakest response was expected to occur for static-IPD values that perceptually matched the beating sound. This was first tested for 32-Hz AMBB stimuli with a start IPD of 0° at the modulation minimum alternating with (adaptor) AM tones of various static IPDs. This procedure was subsequently repeated with 64-Hz modulation. In contrast to the psychophysics, where eight different start IPDs were used for the AMBBs, only one AMBB (start IPD at modulation minimum, 0°) was assessed. For AMBB
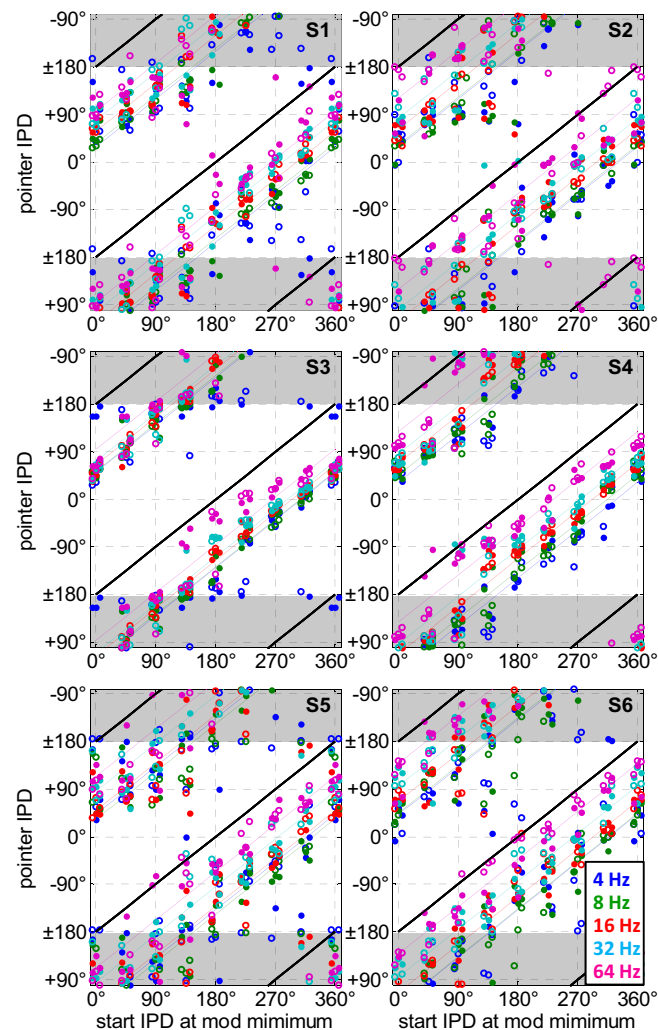

Fig. 3. Psychoacoustic results. Each graph shows data from one of the six subjects. Subjects adjusted the static IPD of sinusoidally AM tones to match the spatial images of the AMBB stimuli. The ordinate is plotted over 1.5 cycles to illustrate a continuous data distribution, with the result that each data point appearing in this additional gray-shaded area corresponds to a point within the white central cycle. Eight different start IPDs (abscissa), five different modulation frequencies (see legend for color code), and two different beat directions were tested (filled symbols: $f_{left} < f_{right}$; open symbols: $f_{left} > f_{right}$). Static IPD is defined with respect to the direction of the AMBB: positive IPD indicates sounds leading on the side with the higher carrier frequency. The black line indicates the expected marker positions if IPD weighting was based on stimulus intensity (ΔIPD = start IPD – matched static IPD = 180°). However, ΔIPD is between 0° and 180° for all subjects and modulation frequencies; thus, the diagonal lines, representing the mean ΔIPD for each AMBB frequency, are below the black line for all subjects and all modulation frequencies. Markers are jittered by up to 6° in the diagonal direction to avoid strong marker overlap caused by the discrete start IPD values.
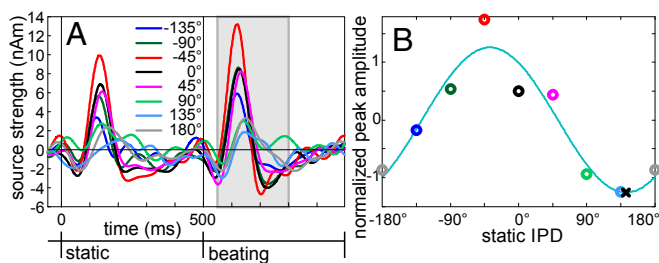
Fig. 4. Results of the MEG experiment. (A) Example source waveforms of one exemplary subject (S1) (same subject as S1 in Fig. 3). Continuous sound presentation alternated every 500 ms between a 32-Hz AMBB with start IPD = 0 and 32-Hz AM tones with various static IPD (see legend). The amplitude of the response to the transition from a static IPD to an AMBB was smallest if the static IPD corresponds to those occurring within the rising portion of an AMBB modulation cycle. (B) A sinusoidal function was fitted to the peak amplitudes corresponding to each static IPD value (circles correspond to peak amplitudes of S1) to estimate the IPD causing maximal adaptation. The minimum is marked with a cross.

frequencies below 32 Hz, the integration time for the change response [~40 ms (15)] is expected to be smaller than the cycle duration (e.g., 62.5 ms for 16 Hz). Therefore, it is not possible to extract the desired information from MEG responses with AMBB frequencies of 4, 8, or 16 Hz.

For most values of IPD, a prominent change response was generated by both transitions. However, only the response following the transition from the static IPD to the beating sound was analyzed, because the starting phase of the AMBB remains constant in all conditions (i.e., only the static IPD was varied). Differences between the peak amplitudes of the change response were tested using a repeated-measures ANOVA, with static IPD as the repeating factor. No evidence was found for differences between beat directions ($F_{1,63} = 0.47$; $P = 0.51$) and, therefore, data from the two conditions were combined, as in the psychoacoustics. The amplitude of the change response was systematically dependent on static IPD ($F_{7,63} = 4.22$; $P < 0.001$). The smallest responses occurred when an AM tone with a static IPD of 90° or 135° preceded the AMBB (i.e., when the static IPD matched an IPD within the rising slope of the AMBB cycle). The largest responses were recorded when the static IPD was −90° and −45° and matched the IPD within the modulation cycle of the AMBB at AM phases of 270° and 315° (i.e., on the falling slope of the modulation cycle). A post hoc analysis confirmed that these differences were statistically significant (Newman–Keuls tests: $P < 0.05$ for 90° vs. −45°, 90° vs. −90°, 135° vs. −45°, 135° vs. −90°).

To estimate the static IPD evoking the lowest change response with a precision better than the 45° measurement grid, a sinusoidal function was fitted to the responses and its minimum identified (cross in Fig. 4B). The minima fell within the range 47−146° static IPD (average 100°), indicating that, for all subjects, static IPD values that occurred during the rising slope of the AMBB evoked the smallest change response.

The experiment was repeated with a 64-Hz AMBB frequency with seven of the same subjects who took part in the 32-Hz condition. Similar to the 32-Hz AMBB, the amplitude of the change response depended on the static IPD (repeated-measures ANOVA: $F_{7,42} = 3.56$; $P < 0.01$). The smallest responses occurred when the static IPD was 135°, and the largest responses occurred when static IPD was 0° or −45° (Newman–Keuls test: $P < 0.05$ for 135° vs. 0°, 135° vs. −45°, 135° vs. 90°, and 135° vs. 45°). The minima of sinusoidal fits fell between 92° and 187° (average 138°). For each of the seven individual subjects taking part in both conditions, the minimum always corresponded to a position

later in the AMBB cycle for 64 Hz than for 32 Hz (dependent samples t test: $t_6 = 6.12$; $P < 0.001$).

Across-Method Comparison. All seven subjects whose MEG response had been recorded with both modulation frequencies showed an increase in the cycle positions that result in the strongest response adaptation when the AMBB frequency was increased from 32 to 64 Hz, consistent with the behavioral data. This increase was, on average, 29.4° ($\Delta IPD_{64} − \Delta IPD_{32}$) and relatively consistent across subjects (20–27° for five subjects and 41° and 53° for two others; SD across all seven subjects, 12.7°), as indicated by the alignment along the dashed diagonal in Fig. 5A that is shifted by the average 29.4° difference. For the behavioral assessment, each of the six subjects showed an increase in the dominant cycle position with increasing AMBB frequency. The average position that most strongly determined the spatial percept, at least in terms of the lateralized position, increased from 83.4° for 32 Hz to 115.2° for 64 Hz. This 31.8° increase (18–44°; SD, 10.5°) is very similar to the 29.4° observed in the MEG study.

The only subject for whom data were obtained using both methods and both frequencies was S2, with psychoacoustically obtained cycle positions of 82.4° and 120.8°, and MEG-obtained positions of 94° and 121°, respectively, for 32 Hz and 64 Hz AMBBs. A second subject who also took part in both methods, but for whom only 32-Hz MEG data were recorded, S1, showed cycle positions of 78.5° for 32 Hz in the behavioral assessment and 142° in the MEG.

Discussion

The observed dominance of ITDs during the early rising portion of a modulation cycle is crucial in complex listening environments: in realistic sound fields, segments that comprise a sharp amplitude rise usually contain the most reliable ITD information because this sound energy is most likely to arrive directly from the source. Later-arriving, reverberant sound often accumulates to higher energy than the direct sound and, as such, can overwhelm useful binaural cues, generating spurious spatial information (Fig. 1). This means that the direct-to-reverberant ratio is time-dependent and is most often maximal immediately following stimulus onset or after a modulation trough. Our data are consistent with the suggestion that reliable ITD information might be "glimpsed" throughout the ongoing stimulus (16, 17). To this end, enhanced sensitivity to low-frequency ITD cues conveyed in the TFS during the early rising segments of modulated sounds represents an effective decoding strategy for spatial hearing.
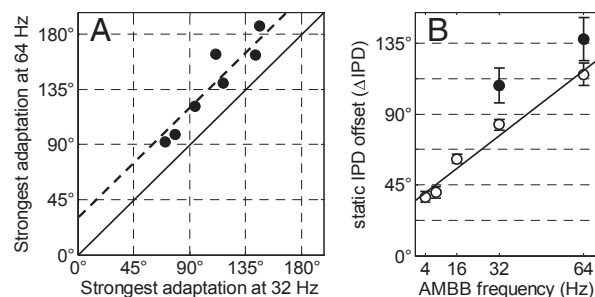
Fig. 5. Data comparisons. (A) MEG data plotted as 32-Hz minima vs. 64-Hz minima for the seven subjects for which 32- and 64-Hz data were collected. A diagonal with an offset equal to the average increase of 29.4° is plotted as a dashed line. (B) Comparison across AMBB frequencies and across methods. Error bars indicate the across subject SEM. The linear regression of the behavioral data (open symbols) reveals an ordinate intercept of 35° corresponding to the phase delay relative to AM minimum, and the slope of 1.3°/Hz corresponds to an additional time delay of 3.6 ms. The increase from 32 to 64 Hz in the MEG data (filled symbols) is similar to that observed in the behavioral data.

The regression fit in Fig. 5B reveals that the dominant cycle position contains a constant phase (35°), as well as a constant time (3.6-ms), portion. A possible explanation for the observed dependence is that the ITD sensitivity may be triggered by a slope level equivalent to ~35° cycle position, followed by a constant time window of high ITD sensitivity centered ~3.6 ms after its start (under the simple assumption of a symmetrical window, the read-out window would, therefore, be 7.2 ms in duration). Potential underlying mechanisms contributing to the generation of a read-out window could be some form of neural adaptation or inhibition that rapidly negates ITD sensitivity after these few milliseconds throughout the remainder of the modulation cycle and that is released in each modulation trough.

It is well established that the mean extent of lateralization of pure tones generally increases with IPD, reaching a maximum around 110–150° before declining again toward 180° (18). The same mean extent of lateralization can, therefore, be achieved by two different IPDs (e.g., one near 0° and another near 180°), with the latter perceived as a "less compact" intracranial image. The observation that, for modulation rates in the range 8–32 Hz, each static IPD can be mimicked with an AMBB shows that subjects not only match the center of gravity of the lateralization percept but also that they are able to disambiguate the two possible static IPDs by considering other perceptual differences such as compactness. The data also imply that AMBBs elicit the same maximum lateralization as static IPDs. This would not be possible if IPDs were averaged over an extended epoch of the AMBB cycle for the lateralization percept. At 64 Hz, some subjects do indeed avoid setting the static IPD pointer to the maximum lateralization interval 110–150°, indicating that, at this short cycle duration, the instantaneous IPD begins to be integrated over a larger portion of the modulation cycle, again in line with the estimated 7.2-ms duration for the read-out window.

The current data emphasize the importance of TFS information conveyed in the early, rising, portion of modulations in speech and other complex sounds. A number of recent studies (19, 20) provided comprehensive investigations of the importance of the modulated envelope of high-frequency sounds (i.e., for sound frequencies above those at which ITD information is conveyed in the TFS in humans). Nelson and Takahashi (19), for instance, examined brain mechanisms in the barn owl contributing to spatial listening in reverberant conditions. Complementary to the current study, they focused on the role of the modulated envelope per se in conveying spatial information. These studies reported that the rising portion of the ongoing envelope was critical to binaural sensitivity (20) or the ability to localize the direct source (19), consistent with the role of stimulus onsets in localization performance (8). Here, however, by using AM stimuli with a 500-Hz carrier frequency, a dominant frequency in human spatial hearing, we demonstrate that ITDs conveyed in the low-frequency TFS, in the absence of any envelope spatial cues, are strongly emphasized during the rising phases of the modulated waveform. Combined, these studies reveal the general dominance of interaural cues during the rising portions of stimulus envelopes, across species, frequency range, and domain.

Several behavioral studies have reported dominance of the ITD during the very first onset of a stimulus (21–24). Hafter and Dye (22) argued that onset dominance is related to the precedence effect (8) and probably caused by postonset adaptation. The large onset response of binaurally sensitive neurons in the auditory midbrain has recently been shown to account for robust sound-source localization in reverberant environments and was suggested as a mechanism for precedence processing (25). Although related to those studies, the current study demonstrates the importance of the spatial cues conveyed in the TFS of the ongoing portion of modulated sounds, arguably a more realistic comparison with temporally complex sounds such as speech. The

data, therefore, allow for a generalization of the importance of stimulus envelopes from onset to ongoing modulations and provide mechanistic insight into the nature of robust localization abilities in reverberant environments. In particular, the data demonstrate means by which the location of sounds with complex envelope structure can be reliably tracked over time when the source location is slowly varied. In contrast to ubiquitous click stimuli, the AMBB provides a realistic temporal structure for investigating the different relative importance for lateralization of a modulated complex signal (7).

The behavioral data suggest an earlier maximum of IPD weighting than that indicated by the MEG data, at least for S1 and for the mean data (Fig. 5B); for S2, both methods yielded very similar results. It is, therefore, hypothesized that a specific subpopulation of ITD-sensitive neurons encodes the perceived sound location by responding only to the earliest, rising portion of stimulus modulations to the exclusion of later-arriving information conveyed with more stimulus energy. We posit that these neurons are characterized by strong adaptation. The MEG data represent, then, the output of a mixture of such neurons and other more "sustained" neurons responding to ITDs throughout the full AM cycle, which, hypothetically, could code information about the auditory scene in general (for example, the degree of interaural coherence, which is a proxy for the room acoustics, source distance, and the number of sources).

In conclusion, lateralization is strongly dominated by ITDs during the rising flanks of amplitude-modulated sounds, providing a plausible explanation as to how sound-source localization in complex environments is achieved: stimulus-driven adaptation in the ascending auditory pathway is a simple and highly effective mechanism by which late-arriving, potentially spurious, spatial information can be suppressed, allowing sources, and their locations, to be heard out against a background of reflections.

Methods

Stimuli. The mean carrier frequency (f_c) of the AMBB stimuli is 500 Hz. To generate the binaural beating, one ear's f_c is reduced by half of the beat frequency while the other ear's f_c is increased by the same amount.

To depict both possible beat directions in the same coordinate system, IPD is defined with respect to the AMBB beat direction: positive IPD means leading on the side with the higher carrier frequency. This includes start IPD, instantaneous IPD, and the static IPD, compared with beat responses.

Behavioral Assessment. Six normal-hearing subjects (one female) with experience in binaural psychoacoustics participated in the study. All subjects provided written consent and were informed of the purpose of the study, potential risks, and their possibility to stop at any time by removing their headphones. Two setups were used: three subjects (S1, S3, and S5) were seated in triple-walled, sound-attenuating booth and listened over Sennheiser HDA 200 headphones to the stimuli presented at a level of 75-dB SPL (at the University College London (UCL) Ear Institute). The other three subjects were seated in a single-walled, sound-attenuating booth and listened over Sennheiser HD 650 headphones to the stimuli presented at a level of 70-dB SPL (Aalto University School of Electrical Engineering).

All stimuli were fully modulated. Each stimulus started and ended in a modulation minimum. Stimulus duration was 750 ms, and all stimuli with modulation frequencies >8 Hz were gated on with a squared sine ramp of 150-ms duration to minimize the influence of the stimulus onset.

After pressing a button, listeners were presented with two stimuli. The first was an AMBB stimulus with a carrier frequency of 500 Hz and an AMBB frequency of one of 4, 8, 16, 32, or 64 Hz. Both beat directions and eight different start IPDs (integer multiples of 45°) were used, resulting in a total of 80 conditions. After an interstimulus interval of 250 ms, the AMBB was followed by an AM-tone "pointer" with the same carrier and AM frequencies as the AMBB but with a static IPD. Using a slider on a computer screen, subjects were instructed to change the static IPD of the second interval until the intracranial spatial percept (including the center of gravity for the lateralization percept and its compactness) best matched the first interval. Subjects were made aware of the fact that, in some cases, two very different static IPDs might produce the same center of gravity (e.g., 0° and 180°). They were encouraged to scan across the 360° range of static IPDs for potentially

better matches in case their current lateralization match did not demonstrate a matching compactness. Subjects could replay the sequence of the AMBB and the static IPD pointer until they were satisfied with their match and continue with the next condition, which was randomly chosen. The experiment was repeated three times on 2–3 different days.

MEG. Fourteen volunteers (right-handed; one female; 20–42 y of age) took part in experiments performed at the MEG Core, Aalto University School of Science. All subjects gave written informed consent, and approval for the experiments was given by the Ethical Committee of Aalto University. All subjects reported normal hearing and no history of neural disorders. The data of four subjects were discarded, one because of a technical failure during recording and three because of the low signal-to-noise ratio of their MEG responses (see below for rejection criteria).

Sound presentation followed a paradigm in which a change is introduced to a continuous sound at regular intervals. Within each stimulation block, sound presentation alternated every 500 ms between an AMBB and an AM tone with a static IPD. The static IPDs were −135°, −90°, −45°, 0°, 45°, 90°, 135°, and 180°. Each static IPD was presented in its own stimulation block. The AMBB always had the start IPD of 0°. The AMBB frequency for all stimuli was either 32 or 64 Hz. The stimuli with 32- and 64-Hz frequencies were presented in separate experimental sessions. The beat direction was left for five, and right for the other five, subjects included in the final data analysis. All 10 subjects took part in the 32-Hz condition, and 7 took part in the 64-Hz condition.

Each stimulation block lasted for a minimum of 4 min and typically 200–250 (minimum 150) artifact-free transitions. The entire recording session took 70 min on average, including short breaks between stimulation blocks. The order of the stimulation blocks was counterbalanced across subjects. During recordings, subjects were instructed to sit still and to ignore the sound stimulation while reading a self-selected text.

Data were recorded with a whole-head MEG device (Vectorview; Elekta Neuromag) consisting of 102 recording elements, each containing two orthogonal planar gradiometers and a magnetometer. Data were recorded continuously at a sampling rate of 1,024 Hz and pass band of 0.01–300 Hz. Eye movements and blinks were monitored with horizontal and vertical electrooculograms (EOGs). To obtain information on head position, four head-position indicator coils were attached to the scalp of the subjects, and the positions of the coils and three anatomical landmark points (left and right preauricular points and nasion) were digitized before recordings. The head position was estimated at the beginning of each stimulation block and later used for modeling the brain responses with an equivalent current dipole (ECD) calculation.

Event-related fields were obtained by averaging offline epochs from 50 ms before to 500 ms after the transition from a static IPD to a beating sound. The analysis was focused on this transition and not the transitions from the beating sound to static IPD to gain better experimental control; transitions from beating to static IPDs could potentially reflect some effects of static IPD per se on the amplitude of the auditory cortical responses (for instance, larger response amplitudes for laterally, compared with centrally, perceived sounds). The transitions from static IPD to beating IPD, however, always represented the onset of the identical AMBB stimulus, and, therefore, variations in the amplitude of the change response can be attributed to adaptation to the static IPD stimulus before presentation of the AMBB. The epoch 400–500 ms following the transition from static to beating IPDs was used as a baseline. The baseline was chosen from a period following, rather than before, the transition of interest to exploit a time period during which the stimulation was identical in all conditions. The first 10 s of sound presentation was omitted from the analysis to exclude any effects of the stimulus onset. Epochs containing responses exceeding 150 μV in EOG or 3,000 fT/cm in gradiometers were regarded as artifacts and discarded. The remaining data were filtered between 1 and 10 Hz.

Activity at the source level was estimated with a single ECD in each cortical hemisphere (26). In ECD analysis, a spherical head model and a set of 22 gradiometer pairs over the temporal lobe were used, and the ECDs were obtained for left and right hemispheres separately. The ECD analysis was performed with Neuromag Xfit software (Elekta Neuromag). To ensure the model selection did not introduce any systematic bias, the location and orientation of the ECD model was derived based on a reference condition using the responses averaged over all transitions from one static IPD to another. First, ECDs were fitted without restrictions on the orientation, location, or the dipole moment over a time window of 50–300 ms at 1-ms intervals after the transition from one IPD to another. The ECD with the largest magnitude, and a goodness of fit value exceeding 75%, over this time window was chosen as the basis for the model. This ECD was then used to model the activity over the entire duration of the event-related field for each experimental condition. The magnitude was allowed to vary while the orientation and location were kept constant. In three subjects, the average level of goodness of fit failed to reach 60% in both hemispheres. This was related to the small amplitude of transition responses and weak signal-to-noise ratio already evident in the sensor space. These subjects were not included in further analysis. For most of the 10 remaining subjects, transition responses could be reliably identified in only one hemisphere for all experimental conditions. Therefore, only the data from the hemisphere showing larger response amplitudes, and thereby better goodness-of-fit values, were used in further analysis for each subject. The peak amplitude was identified from a time interval from 50 to 300 ms after the stimulus transition.

1. Wightman FL, Kistler DJ (1992) The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am* 91(3):1648–1661.
2. Zwislocki J, Feldman RS (1956) Just noticeable differences in dichotic phase. *J Acoust Soc Am* 28(5):860–864.
3. Licklider JCR (1948) The influence of interaural phase relations upon the masking of speech by white noise. *J Acoust Soc Am* 20(2):150–159.
4. Hawley ML, Litovsky RY, Culling JF (2004) The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *J Acoust Soc Am* 115(2):833–843.
5. Franssen NV (1960) Some considerations on the mechanism of directional hearing. PhD thesis (Technische Hogeschool, Delft, The Netherlands).
6. Bregman AS (1997) *Auditory Scene Analysis - The Perceptual Organization of Sound* (MIT Press, Cambridge, MA).
7. Haas H (1951) Über den Einfluss eines Einfachechos auf die Hörsamkeit von Sprache (On the influence of a single echo on the intelligibility of speech). *Acustica* 1(2):49–58, German.
8. Wallach H, Newman EB, Rosenzweig MR (1949) The precedence effect in sound localization. *Am J Psychol* 62(3):315–336.
9. Litovsky RY, Colburn HS, Yost WA, Guzman SJ (1999) The precedence effect. *J Acoust Soc Am* 106(4 Pt 1):1633–1654.
10. Licklider JCR (1950) On the frequency limits of binaural beats. *J Acoust Soc Am* 22(4):468–473.
11. Akeroyd MA (2010) A binaural beat constructed from a noise (L). *J Acoust Soc Am* 128(6):3301–3304.
12. Ross B, Tremblay KL, Picton TW (2007) Physiological detection of interaural phase differences. *J Acoust Soc Am* 121(2):1017–1027.
13. Halliday R, Callaway E (1978) Time shift evoked potentials (TSEPs): Method and basic results. *Electroencephalogr Clin Neurophysiol* 45(1):118–121.
14. Sams M, Hämäläinen M, Hari R, McEvoy L (1993) Human auditory cortical mechanisms of sound lateralization: I. Interaural time differences within sound. *Hear Res* 67(1-2):89–97.
15. Joutsiniemi SL, Hari R, Vilkman V (1989) Cerebral magnetic responses to noise bursts and pauses of different durations. *Audiology* 28(6):325–333.
16. Dietz M, Ewert SD, Hohmann V (2011) Auditory model based direction estimation of concurrent speakers from binaural signals. *Speech Commun* 53(5):592–605.
17. Day ML, Koka K, Delgutte B (2012) Neural encoding of sound source location in the presence of a concurrent, spatially separated source. *J Neurophysiol* 108(9):2612–2628.
18. Sayers BM (1964) Acoustic-image lateralization judgments with binaural tones. *J Acoust Soc Am* 36(5):923–926.
19. Nelson BS, Takahashi TT (2010) Spatial hearing in echoic environments: The role of the envelope in owls. *Neuron* 67(4):643–655.
20. Klein-Hennig M, Dietz M, Hohmann V, Ewert SD (2011) The influence of different segments of the ongoing envelope on sensitivity to interaural time delays. *J Acoust Soc Am* 129(6):3856–3872.
21. Perrott DR, Baars BJ (1974) Detection of interaural onset and offset disparities. *J Acoust Soc Am* 55(6):1290–1292.
22. Hafter ER, Dye RH, Jr. (1983) Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number. *J Acoust Soc Am* 73(2):644–651.
23. Akeroyd MA, Bernstein LR (2001) The variation across time of sensitivity to interaural disparities: Behavioral measurements and quantitative analyses. *J Acoust Soc Am* 110(5 Pt 1):2516–2526.
24. Stecker GC, Brown AD (2010) Temporal weighting of binaural cues revealed by detection of dynamic interaural differences in high-rate Gabor click trains. *J Acoust Soc Am* 127(5):3092–3103.
25. Devore S, Ihlefeld A, Hancock K, Shinn-Cunningham B, Delgutte B (2009) Accurate sound localization in reverberant environments is mediated by robust encoding of spatial cues in the auditory midbrain. *Neuron* 62(1):123–134.
26. Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993) Magnetoencephalography – theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev Mod Phys* 65(2):413–497.

Dietz et al.

www.manaraa.com