Influence of envelope fluctuation on the lateralization of interaurally delayed low-frequency stimuli

Jörg Encke and Mathias Dietz

Citation: The Journal of the Acoustical Society of America **150**, 3101 (2021); doi: 10.1121/10.0006571
View online: https://doi.org/10.1121/10.0006571
View Table of Contents: https://asa.scitation.org/toc/jas/150/4
Published by the Acoustical Society of America

JASA ARTICLE

Influence of envelope fluctuation on the lateralization of interaurally delayed low-frequency stimuli

Jörg Encke[a),b)] and Mathias Dietz[b)]
Department für Medizinische Physik und Akustik, Universität Oldenburg, 26111 Oldenburg, Germany

ABSTRACT:
Disregarding onset and offset effects, interaurally delaying a 500 Hz tone by 1.5 ms is identical to advancing it by 0.5 ms. When presented over headphones, humans indeed perceive such a tone lateralized toward the side of the nominal lag. Any stimulus other than a tone has more than one frequency component and is thus unambiguous. It has been shown that phase ambiguity can be resolved when increasing the stimulus bandwidth. This has mostly been attributed to the integration of information across frequencies. Additionally, interaural timing information conveyed in the stimulus envelope within a single frequency channel is a second possible cue that could help to resolve phase ambiguity. This study employs stimuli designed to differ in the amount of envelope fluctuation while retaining the same power spectral density as well as interaural differences. Any difference in lateralization must thus be a result of the difference in envelope. The results show that stimuli with strong envelope fluctuation require significantly smaller bandwidths to resolve phase ambiguity when compared to stimuli with weak envelope fluctuation. This suggests that within-channel information is an important cue used to resolve phase ambiguity.
 https://doi.org/10.1121/10.0006571

(Received 18 March 2021; revised 6 September 2021; accepted 16 September 2021; published online 26 October 2021)
[Editor: G. Christopher Stecker] Pages: 3101–3108

I. INTRODUCTION

Spatial hearing—listening with two ears—provides a spatial representation of the acoustic surrounding. Unlike in the visual system, where two dimensions are coded at the sensory level, the auditory system has to generate this space. The physical cues that are used to localize sound sources in the lateral dimension are differences in the arrival time and in the intensity between the signals at the two ears. The neurons that are involved in encoding interaural time differences (ITDs) are highly specialized for this purpose and are considered the most temporally precise in the mammalian nervous system (Grothe *et al.*, 2010). This indicates the high evolutionary pressure on, and thus the importance of, ITD sensitivity. Even in unnatural man-made environments, such as urban traffic, we can rely on our ability to localize sound sources in order to warn us of a vehicle approaching from a certain direction. As a consequence of this superb sound localization ability, one might speculate that a specific physical cue, such as ITD, could be directly mapped to the activity of specific neurons, similar to the 2D image projection on the retina. The classic understanding of ITD encoding comprises binaural coincidence detecting neurons along an array of different input delays (Jeffress, 1948). Each specific delay leads to activity in a corresponding neuron whose internal relative left–right latency difference compensates for the external ITD. This arrangement would indeed resemble a place coding of azimuthal direction similar to the visual system. Several psychoacoustic studies, however, report that two stimuli with the same physical ITD and no interaural intensity difference can be perceived as either coming clearly from the left or from the right, depending on the bandwidth of the respective stimulus [Trahiotis and Stern, 1989; see Fig. 1(a)]. In addition, it was also shown that lateralization can change with stimulus level (Dietz *et al.*, 2009).

The effect of bandwidth-dependent lateralization of stimuli with the same ITD can be understood in light of the periodicity of a sinusoidal stimulus. Disregarding onset/offset effects, delaying a 500 Hz sinusoid by 1.5 ms is equivalent to phase shifting it by 1.5π. Due to its cyclic nature, this is identical to a phase shift of -0.5π. As a consequence, introducing an ITD of 1.5 ms into the ongoing part of a 500 Hz tone is equivalent to an ITD of -0.5 ms. In the experiment of Trahiotis and Stern (1989), 500 Hz centered noise tokens with different bandwidths were presented via headphones. At narrow bandwidths, the subjects perceived the noises lateralized to the nominal lag, consistent with an ITD of -0.5 ms. With increasing bandwidth, however, the percept moved to the correct (leading) side [see Figs. 1(a) and 1(b)]. Similar results for a larger number of experienced as well as inexperienced subjects were reported by Yost *et al.* (2007). The change in lateralization suggests that subjects were able to resolve phase ambiguity in stimuli with larger bandwidth. Two different mechanisms that could underlie this effect have since been proposed.

[a)]Electronic mail: joerg.encke@uol.de, ORCID: 0000-0002-6067-1602.
[b)]Also at: Cluster of Excellence "Hearing4all", Universität Oldenburg, 26111 Oldenburg, Germany, ORCID: 0000-0002-1830-469X.

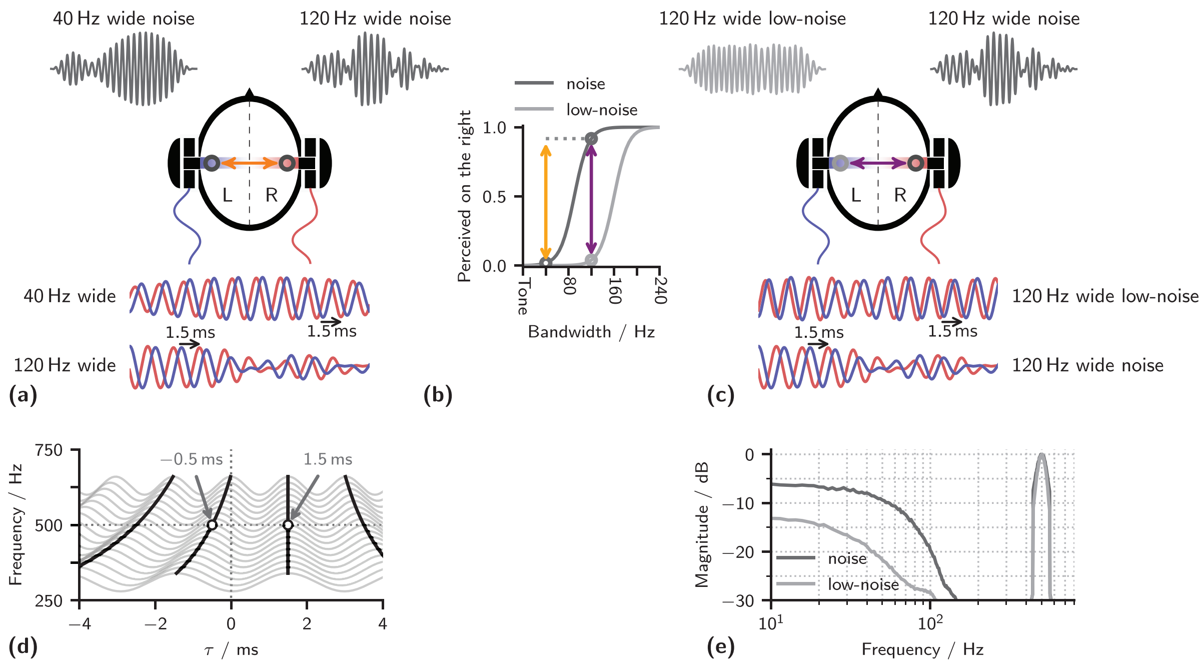

FIG. 1. (Color online) (a) Illustration of the effect of changing bandwidth on the lateralization of a 1.5 ms delayed noise stimulus. Noise stimuli with a low bandwidth are lateralized to the left while stimuli with a high bandwidth are lateralized to the right. (b) Schematic plot of the fraction of stimuli perceived as lateralized to the right as a function of noise bandwidth. The dark gray line depicts results for band-limited white noise, the light gray line results for low-noise noise, a special type of noise that was optimized in order to reduce envelope fluctuation while keeping the magnitude spectrum of normal noise. The orange arrow marks the difference between the two noise bandwidths illustrated in (a); the purple arrow marks the difference between the noise and the low-noise condition as illustrated in (c). (c) Illustration of the hypothesized effect of envelope fluctuation on stimulus lateralization. (d) Cross correlogram for a 320 Hz wide noise stimulus centered at 500 Hz with an ITD of 1.5 ms. The black lines indicate the location of local maxima. The sinusoids illustrate the underlying periodicity. (e) Ensemble average of the magnitude spectra of 120 Hz wide noise and low-noise after applying a 79 Hz gamma-tone filter and one-way rectification. Magnitudes are shown relative to the value at 500 Hz.

The first mechanism was proposed by Stern *et al.* (1988), who concluded that subjects had to combine information across frequencies. In each frequency channel, the coincidence detecting neurons along a delay line effectively cross correlate the two inputs and show response peaks when the internal delay compensates for the ITD of 1.5 ms [see Fig. 1(d)]. Additional peaks appear at internal delays that are offset from the central peak by one or more periods. The period duration depends on the frequency, yielding curved trajectories for all additional peaks. Multiplying the cross correlation functions of multiple channels along the frequency axis thus attenuates all curved trajectories while amplifying the straight 1.5 ms trajectory. The amplification of the consistent peak position across-frequency bands was coined "straightness" weighting.

A second possible mechanism is to rely on information that is available within-frequency channels. Dietz *et al.* (2009) proposed a model that leveraged the demodulating properties of the auditory periphery. As soon as some form of nonlinear processing, such as half-wave rectification is encountered prior to binaural interaction; the envelope of the noise is demodulated and appears as low-frequency components. The highest frequency of these demodulated components is restricted by the peripheral bandwidth so that their period is much larger than the 1.5 ms. As a consequence, they are not affected by periodicity, and their ITD can thus be used as an unambiguous cue, which could either be used on its own or in combination with the ambiguous phase cue. Increasing the stimulus bandwidth increases the amount of envelope fluctuation and thus the magnitude of the low-frequency components. The use of the demodulated envelope as an ITD cue is well established at high frequencies, where the auditory system is not sensitive to the fine-structure so that the auditory system relies on envelope ITD information (Bernstein and Trahiotis, 2002, 2003; Henning, 1980; Henning, 1974). For models that do not use straightness, within-channel cues are of central importance (Dietz *et al.*, 2009; Stecker *et al.*, 2019).

If within-channel information is indeed used to disambiguate the ITD, this process should be influenced by the amount of envelope fluctuation. Two stimuli that share the same power density spectrum but differ in the amount of envelope fluctuation would also show differences in the magnitude of the demodulated envelope. This is clearly visible in the low-frequency region of Fig. 1(e), which shows the ensemble average of magnitude spectra for two noise tokens that differ in their amount of envelope fluctuation. If the demodulated envelope is used, these two stimuli could thus be lateralized differently [see Figs. 1(b) and 1(c)]. The aim of this study was to investigate whether and to which degree within-channel information is employed by the auditory system to resolve phase ambiguity within the 500 Hz, 1.5 ms ITD stimuli. To this end, this study used a set of stimuli with different amounts of envelope fluctuation but identical interaural differences and power density spectra.

Two experiments, with two different stimulus types, that both have these properties, were conducted. The first experiment, which was termed the "*band*" experiment, consisted of bands of noise. Two conditions were tested: a standard Gaussian *noise* condition and a *low-noise* noise condition where the noise tokens were designed to reduce the amount of envelope fluctuation. The spectral content of the stimuli in this experiment are visualized by the horizontal gray lines in Fig. 2(a). The second experiment, which will be referred to as "*complex*", consisted of harmonic tone complexes with a fundamental frequency $f_0 = 20$Hz as visualized by the vertical bars in Fig. 2(a). Within this experiment, three conditions, that only differed in the phase relation of the harmonic components, were compared. In the *cosine* condition, all components were cosines starting with zero phases, maximizing envelope fluctuation. In the *random* condition, starting phases were picked randomly between π and $-\pi$. In the *Schröder* condition, the starting phases were arranged to minimize the amount of envelope fluctuation using the method introduced by Schroeder (1970) with a negative phase curvature. Short example stimuli of all five conditions are shown in Fig. 2(b). Importantly, the amount of envelope fluctuations can be influenced by the bandpass characteristic of the auditory periphery. Illustrating this influence, the same stimuli are also shown after applying a gammatone bandpass filter [Fig. 2(c), bottom row].

If within-channel envelope-ITD information is used to disambiguate the phases in these stimuli, then any improvement in how well the envelope is encoded should also improve the weight of this cue. Previous studies found that monaural modulation detection improved with stimulus level (Kohlrausch *et al.*, 2000). Additionally, Dietz *et al.* (2009) reported that the lateralization of amplitude-modulated tones with conflicting phase and envelope ITDs were influenced by stimulus level with higher levels, resulting in a shift of lateralization in favor of the envelope-ITD. To test if a level dependency in envelope ITD encoding also influences the ability of the auditory system to resolve phase ambiguity within the 500 Hz, 1.5 ms stimulus, a third experiment referred to as "*level*" was conducted. In this experiment, the *noise* stimulus from the *band* experiment was presented at different levels.

To summarize, this study used three experiments to investigate the influence of envelope fluctuation on the lateralization of low-frequency stimuli with identical bandwidths. The first two experiments used stimuli that only differ in the amount of envelope fluctuation while keeping the same amplitude and IPD spectra. The underlying hypothesis was that any difference in the lateralization should be due to the difference in their envelope. In the third experiment, the influence of sound level was tested with the hypothesis that an increase in stimulus level should improve neuronal envelope coding and thus also, the use of binaural envelope cues.

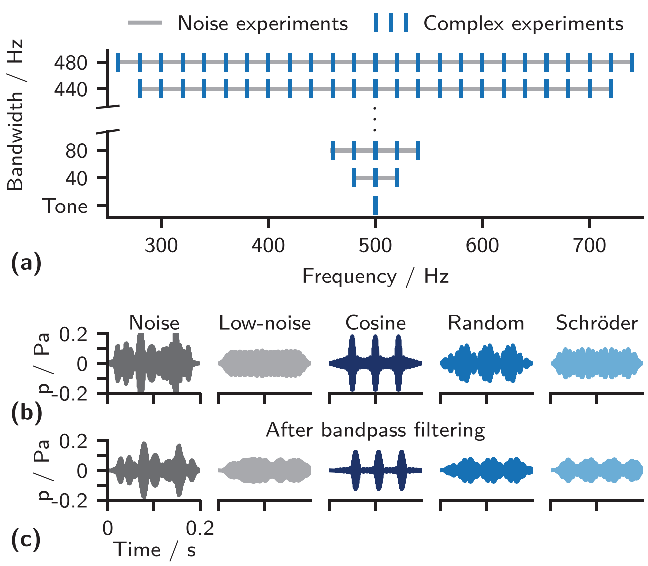

FIG. 2. (Color online) (a) Visualization of the frequency content of the stimuli used in the experiments. The gray lines visualize the frequency content of the *band* stimuli. Vertical bars indicate the frequency components of the *complex* experiment. (b) 200 ms long example waveforms for the stimuli in the *band* and *complex* group (amplitudes for the *complex* conditions were scaled for better visualization). (c) Same stimuli as in (b) but after applying a 79 Hz wide gammatone filter centered at 500 Hz to simulate the influence of peripheral frequency selectivity.

II. METHODS

The three experiments discussed in the introduction as well as their respective stimulus conditions, are summarized in Table I.

A. Stimuli and apparatus

All stimuli were centered around 500 Hz with bandwidths ranging from 40 to 480 Hz in 40 Hz increments. Stimuli were 700 ms long and gated using 100 ms long cosine ramps. Long ramp times were chosen to minimize potential onset-offset effects. The stimulus duration was chosen to ensure that stimuli in the *complex* experiment contained a sufficiently large number of envelope repetition cycles (50 ms period).

For the *noise* condition, unique Gaussian noise tokens were generated for every stimulus presentation. The tokens were bandpass filtered in the frequency domain by setting the magnitude of frequency components outside the desired frequency range to zero. ITDs of 1.5 ms were introduced prior to bandpass filtering and gating by cyclically shifting the noise token of the left channel by the corresponding number of samples. The *low-noise* signals were generated based on noise tokens with the same parameters. The

TABLE I. Overview over the experiments and stimulus conditions used in the study.

Experiment	Conditions
Band	noise, low-noise
Complex	cosine, random, Schröder
Level	noise at 38, 44, 50, 56 dB

amount of envelope fluctuation at each bandwidth was then minimized following the procedure described by Kohlrausch *et al.* (1997) (10 repetitions). In the *noise* and *low-noise* conditions, all tokens were presented with a constant spectrum level of 50 dB Hz^{-1}. The *level* experiment used the same stimulus as the *noise* experiment but with the spectrum level set according to the condition to be tested. For all *complex* conditions, stimuli were created by summing over a set of tones with the desired frequencies. ITDs were introduced prior to applying the gating window by adding a phase shift equivalent to an ITD of 1.5 ms to each of the frequency components comprising the left channel signal. The stimuli were created so that the sound pressure level of each frequency component was equivalent to 50 dB. For the single tone condition used in all experiments, a single 500 Hz tone was either presented at a sound pressure level of 50 dB (*band* and *complex*) or at the sound pressure level indicted by the experimental condition (*level*).

All stimuli were created digitally in Matlab R2016b (MathWorks, Natick, Massachusetts, USA) with a sampling rate of 192 kHz. The signals were then transmitted to an external audio interface (ADI-2 DAC FS, RME, Heimhausen, Germany) set to a D/A-converter word-length of 16 bit and presented using circumaural headphones (HD650, Sennheiser electronic GmbH, Wedemark, Germany) which were calibrated by scaling the outputs according to the response of each headphone at 500 Hz. Subjects were located in a sound-isolated chamber and used the two buttons (left/right) of a wireless presenter (R400; Logitech Europe S.A.,195 Lausanne, Switzerland) for their response.

B. Procedure and data analysis

A two-alternative, forced-choice procedure was used throughout the experiments. Two stimuli, one with −1.5 ms and one with +1.5 ms ITD, were presented to the subjects in random order separated by a 200 ms inter-stimulus interval. By pressing one of the two buttons of the presenter, the subjects had to indicate whether they perceived the second stimulus left or right of the first stimulus. No feedback about their decision was provided to the subject. Before starting data collection, subjects were able to familiarize themselves with the task by conducting half of the repetitions of the *noise* condition.

The method of constant stimuli was used with each stimulus presented 40 times. In the *complex* and *level* experiments, presentations were randomized across both conditions and bandwidths. For the *band* experiment, individual presentations were randomized across bandwidths only. For each condition, subject and bandwidth, the fraction of responses at which the subject reported the stimulus with an ITD of +1.5 ms to be located to the right of the stimulus with −1.5 ms (fraction "right") was calculated. For each subject and condition, logistic functions were fitted to these values as a function of bandwidth by using a maximum log-likelihood approach (Wichmann and Hill, 2001). Data analysis was performed in Python 3.7.3 using the "Scipy" package (Virtanen *et al.*, 2020). Statistics were calculated using the "Pingouin" package (Vallat, 2018).

C. Subjects

The data of seven subjects (five female, two male) with an age ranging from 19 to 26 y (median 24 y) were analyzed in this study. All subjects were without known hearing difficulties and showed monaural pure-tone thresholds below 15 dB HL at all studied frequencies up to 8 kHz. Thresholds between the left and the right ear differed by not more than 5 dB at any frequency and at 4 kHz, they were less or equal to 7.5 dB HL. Only one of the listeners (subject 3) was familiar with the design or aim of the experiment. The subjects participated voluntarily and were reimbursed for their participation. Data from one additional subject that took part in the experiments were discarded as this subject could not reliably solve the task for the 500 Hz tone condition for which the subject reported a fraction "right" value of around 0.2. The experimental procedures were approved by the ethics committee of the University of Oldenburg.

III. RESULTS

Symbols in Fig. 3(a) show the individual results for all experimental conditions. Each row represents results for one experiment, while each column shows the result of a single subject. The result for the 500 Hz tone was only measured once and is identical in all conditions. Also, the 50 dB results in the *level* experiment are the replotted data from the *noise* experiment.

Irrespective of the stimulus condition, at low bandwidth and for the 500 Hz tone, all subjects consistently reported the stimuli with +1.5 ms ITD to be lateralized left of stimuli with −1.5 ms ITD resulting in a fraction "right" close to zero. With increasing bandwidth, the relative lateralization flipped, and at 480 Hz, all subjects consistently reported the +1.5 ms ITD stimuli to be lateralized to the right of the −1.5 ms ITD stimuli. The bandwidth resulting in a fraction "right" of 0.5 was defined as the changeover bandwidth, which was determined by fitting logistic functions. Fits are shown as lines in Fig. 3(a).

Boxplots over the individual changeover bandwidths in each condition are shown in Fig. 3(b). To test the influence of different conditions within each group of experiments, the one-way repeated-measures ANOVA with Greenhouse–Geisser correction was calculated. If these resulted in a significant difference, *post hoc* Bonferroni corrected pairwise t-tests (one-tailed) and the effect size Cohen's d were calculated. A significance criterion of $\alpha = 0.05$ was used in all tests.

For the *band* group of stimuli, all subjects showed an increase in the changeover bandwidth from *noise* to *low-noise* In the *complex* group, six of the seven subjects (subjects 2–7) showed a consistent increase from *cosine* to *random* to *Schröder*. Subject 1 had its lowest changeover point for *random* and the highest for *cosine*.

The ANOVA resulted in a significant difference between the conditions within the *band* experiment

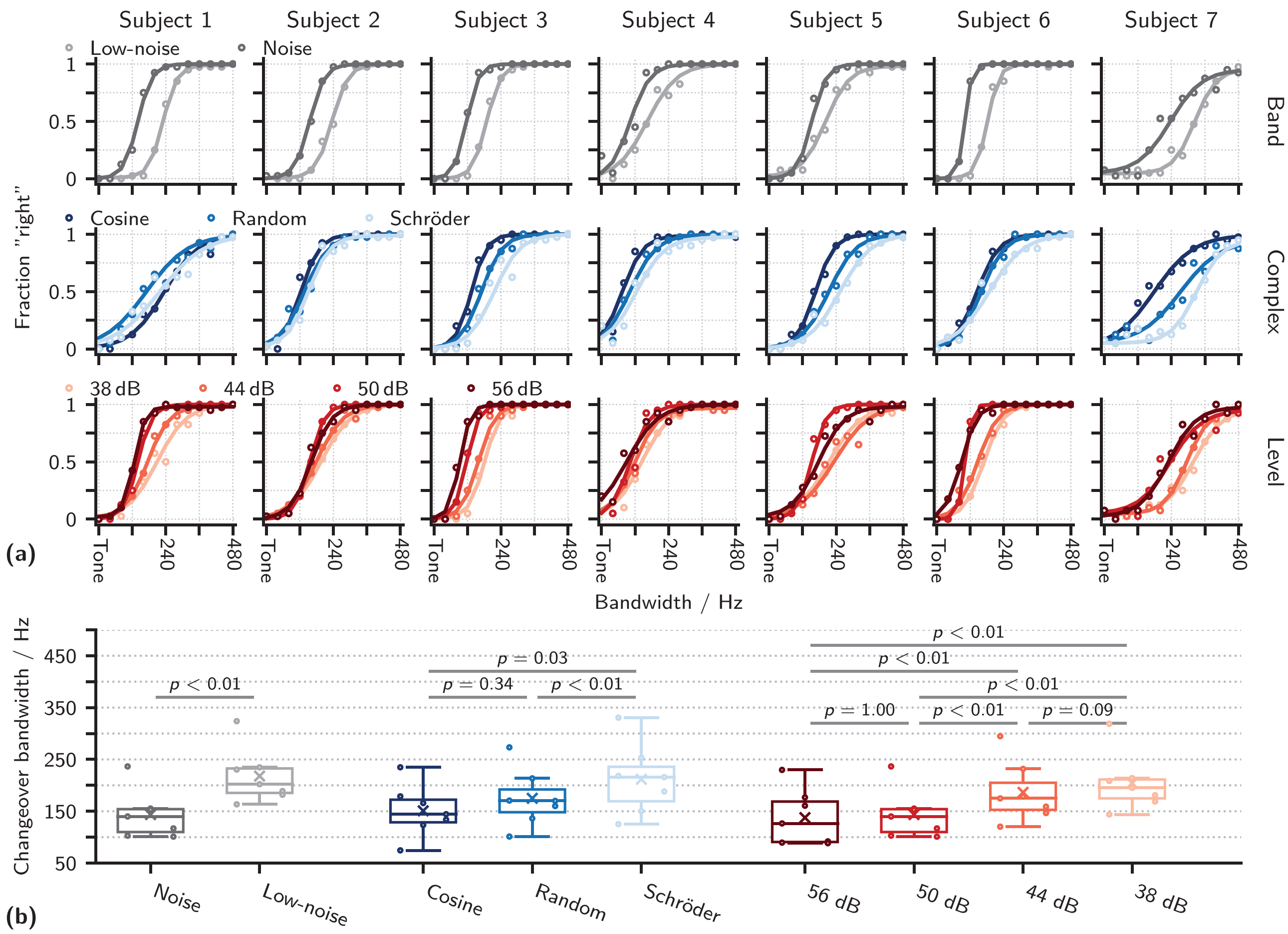

FIG. 3. (Color online) (a) Individual results of the experiments. Each row contains the results of one experimental group and each column of one subject. Stimulus conditions are plotted in different brightness. Symbols indicate the mean fraction of "right" responses at a given bandwidth; lines are logistic functions that were fitted to these values. (b) Boxplots for the changeover bandwidths defined as the bandwidth at which the fits reached a fraction "right" of 0.5. Small circles indicate individual changeover points while crosses indicate the inter-individual means. Probability values (p) are the results of pairwise t-tests.

$[F(1,6) = 174.40,\ p < 0.01]$ and the conditions within the *complex* experiment $[F(2,12) = 8.21,\ p = 0.03]$. *Post hoc* t-tests revealed significantly higher changeover bandwidths $[t(6) = 13.21,\ p < 0.01,\ d = 1.49]$ of the *low-noise* compared to the *noise* condition. They also showed significantly higher changeover bandwidths of the *Schröder* compared to the *cosine* $[t(6) = 3.03,\ p = 0.03,\ d = 1.03]$ and compared to the *random* $[t(6) = 6.39,\ p < 0.01,\ d = 0.54]$ condition. No significant increase was found from the *cosine* to the *random* condition $[t(6) = 1.34,\ p = 0.34,\ d = 0.52]$.

In the *level* group of stimuli, all subject showed an increase in changeover bandwidths from the 56 dB to both 44 dB and 38 dB. Trends across other conditions were less systematic. The ANOVA resulted in a significant difference between *level* conditions $[F(3,18) = 42.00,\ p < 0.01]$. The *post hoc* test found the changeover bandwidth of the 56 dB condition to be significantly lower than those of the 44 dB $[t(6) = 7.03,\ p < 0.01,\ d = 0.87]$ and the 38 dB condition $[t(6) = 7.33,\ p < 0.01,\ d = 1.21]$. No significant differences were found between 56 dB and 50 dB $[t(6) = 0.99,\ p = 1.00,\ d = 0.12)$. The changeover points at 50 dB were significantly lower than those at 44 dB $[t(6) = 5.4,\ p < 0.01,\ d = 0.8]$ and at 38 dB $[t(6) = 10.49,\ p < 0.01,\ d = 1.18]$. No significant difference was found between 44 dB and 38 dB $[t(7) = 2.87,\ p = 0.09,\ d = 0.31]$.

IV. DISCUSSION

In both, the *band* and *complex* group, stimuli with stronger envelope fluctuation resulted in significantly lower changeover bandwidth [see Fig. 3(b)] supporting the hypothesis that within-frequency channel information is used to resolve ambiguous phase differences.

Previous studies that used the same stimulus as the *noise* condition did not directly measure the changeover bandwidth, but instead used an acoustic pointer as a proxy for lateralization. Changeover bandwidths, however, can be estimated from the zero crossings of this lateralization data (e.g., Trahiotis and Stern, 1989), which result in values similar to those reported in the present study. Changeover bandwidths within the *complex* experiment showed considerably more inter-subject variability than within the *band* experiment and only six of seven subjects did show an increase in changeover bandwidth from the *cosine* to the *random* condition so that no significant increase was found. The comparison still resulted in a medium effect size of $d = 0.52$. One explanation for the increased variability could be that stimulus presentations within the *complex* experiment were randomized across all three conditions while the two conditions within the *noise* experiment were collected separately. Another reason could be the low envelope repetition rate of 20 Hz, which can also be perceived in the form of fluctuations in stimulus intensity (Fastl, 1983). The low envelope repetition rate offers a limited number of glimpses making it harder to lateralize these stimuli when compared to the more continuous *band* stimuli which do not fluctuate perceptibly.

Changeover-bandwidths for the *noise* and the *cosine* condition did not differ significantly [$t(6) = 0.37$, $p = 0.36$, $d = 0.15$]. This might be surprising considering the strong modulations of the *cosine* stimulus when compared to the *noise* stimulus [see Fig. 2(b)]. To quantify the amount of envelope fluctuation in the different conditions, the crest-factor C, which is the ratio of the waveform's peak value to its root mean square [$C = |p_{\mathrm{peak}}|/p_{\mathrm{rms}}$], was calculated. Results for each condition are shown in Fig. 4. In both, the *low-noise* and the *Schröder* condition phase relations are optimized in order to minimize the amount of envelope fluctuation. This optimization is based on the full bandwidth of the stimulus. If the bandwidth is changed after optimization, for example, due to the bandpass filtering properties of the auditory periphery, the phase relation becomes non-optimal resulting in increased envelope fluctuations. Furthermore, the signal phase is also changed due to the filters' phase response, which can result in an additional increase in fluctuation. To estimate the impact of bandpass filtering, the crest-factor was also calculated after applying a 500 Hz centered, gammatone bandpass filter with an equivalent rectangular bandwidth (ERB) of 79 Hz [see Fig. 4(b)]. As expected, the crest-factor of both the *Schröder* condition and *low-noise* condition increased with bandwidth after bandpass filtering. The crest-factor of the *cosine* condition increased continuously before bandpass filtering as increasingly more harmonic components interfered constructively. After bandpass filtering, C only increased until the additional harmonic components fell outside the bandpass filter. As a result, the crest-factor of the *cosine* condition was not much larger than the one calculated for the *noise* condition. The crest-factor also varied with the filters' center frequency. Figure 4(c) shows C calculated for filters of different center frequencies where the ERB was set according to Glasberg and Moore (1990) and stimulus bandwidths were set to the changeover bandwidths as determined in the experiment. For the *Schröder*, *random*, and *noise* conditions, C did not change considerably with center frequency. For the *cosine* condition, however, C decreased with increasing distance to the center frequency while it increased for *low-noise* condition, indicating that off-center frequency channels might carry different amounts of within-channel information. More importantly, however, the crest-factors at 500 Hz were different across conditions. This indicates that the crest-factor, as calculated here, is not sufficient in order to explain the differences in changeover bandwidths. One explanation could be the change in crest-factor in off-frequency channels discussed above. Another explanation could be the use of a gammatone filter as the only pre-processing step which neglects any non-linear processing of the auditory periphery. Furthermore, the phase response of the gammatone filter deviates significantly from the auditory periphery (Oxenham and Dau, 2001) which will influence the amount of envelope fluctuations for all non-random phase conditions. Replacing the simple gammatone filter with more detailed "off-the-shelf" models of the auditory periphery such as the model of Zilany *et al.* (2014) or Lyon (2011), did also not result in better predictions of the changeover bandwidth. As a consequence, a combination of a more complex peripheral model and possibly an envelope-ITD processing model, such as the one proposed by Klug *et al.* (2020), might be necessary to model the differences in changeover-bandwidth.

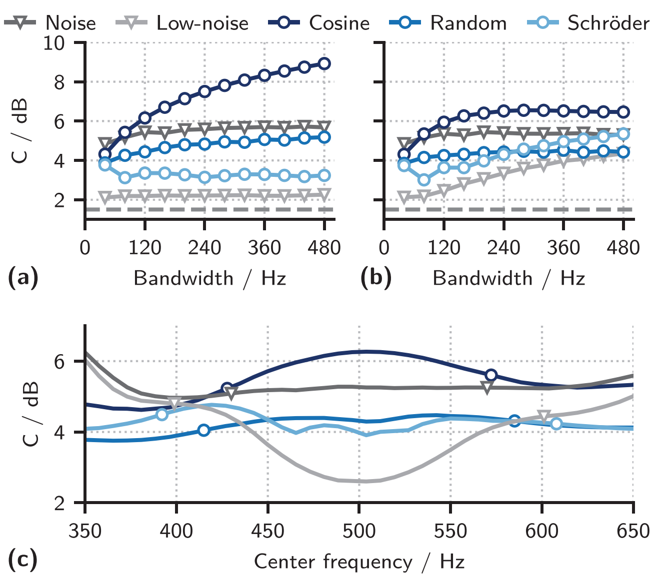

FIG. 4. (Color online) (a) The crest-factor C which is used to quantify how the amount of envelope fluctuation depends on bandwidth and condition. The crest-factor of a tone is indicated by a dashed line. (b) Same as (a) but after bandpass filtering to simulate frequency selectivity of the auditory periphery. For the *cosine* condition, C is limited by the bandwidth of the gammatone filter. For the *Schröder* and *low-noise* condition, C increases with bandwidth due to non-optimal phase alignment after bandpass filtering. (c) Crest-factor calculated for the changeover bandwidth after peripheral filtering. The abscissa indicates the filter center frequency. Symbols indicate the stimulus' frequency range.

Both the *noise* and the *cosine* experiment directly investigated the influence of envelope fluctuations on the changeover-bandwidth. The *level* experiment was designed to test if increases in stimulus level, which were previously shown to improve monaural amplitude modulation (AM) detection thresholds, also improved the ability of the auditory system to use the resolved phase ambiguity in the 1.5 ms delayed stimulus. Decreasing the sound level of the *noise* stimulus resulted in significantly higher changeover bandwidths. This is in line with the finding of Dietz *et al.*

(2009), who reported similar results for SAM tones. It is also in line with monaural studies indicating improved amplitude modulation (AM) detection thresholds at higher levels (Kohlrausch *et al.*, 2000). Improved monaural AM encoding can be expected to promote the use of binaural envelope cues. Furthermore, an additional effect is expected due to slightly increased peripheral bandwidth at higher stimulus levels. An increase in bandwidth results in higher modulation frequencies to be included within the frequency channels. This increases the availability of within-channel cues resulting in a decrease in changeover-bandwidth.

The stimulus conditions used in the *band* and *complex* experiments of this study were designed to reveal the same power density spectrum while differing only in the monaural phase relation of their spectral components. As a consequence of this stimulus design, a pure across-frequency mechanism, such as the peak-based straightness approach proposed by Stern *et al.* (1988) or a group delay approach based on the slope of the mean interaural phase differences across frequencies, could not explain the results of this experiment.

In praxis, however, the straightness approach is inherently linked to the cross correlation functions used in delay-line models, which are usually combined with a non-linear periphery. As discussed in the introduction, non-linear processing, such as rectification, will lead to demodulation which delay-line models rely on to create envelope-ITD sensitivity for frequencies where fine-structure information is not available (Bernstein and Trahiotis, 2002, 2003). The demodulated frequency components do, however, also influence the cross correlation function at frequencies where fine-structure information is available. This is visualized in Fig. 5(a), which shows parts of the cross correlation function for *noise* and *low-noise* stimuli, after 500 Hz bandpass filtering (gammatone filter 79 Hz ERB) and half-wave rectification. To isolate the impact of the demodulated envelope, Fig. 5(b) shows the cross correlation function for the same signal but calculated after removing frequencies above 200 Hz. While both functions of Fig. 5(b) have their maximum at 1.5 ms, the function for *noise* decays much quicker than the one for *low-noise*.

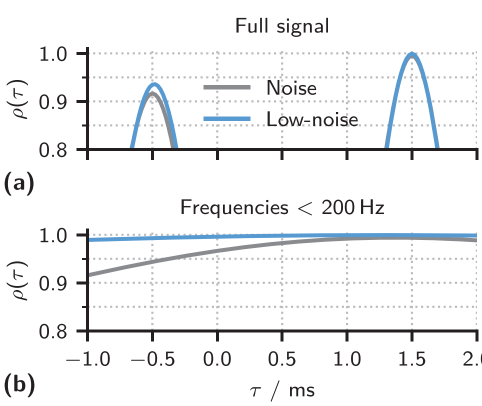

FIG. 5. (Color online) (a) Parts of the cross correlation function for a *noise* and *low-noise* stimulus after applying a 79 Hz wide bandpass filter (ERB) centered at 500 Hz. (b) Cross correlation function for the same signals as in (a) but after removing all frequency components above 200 Hz.

This is a consequence of the increased modulation energy in the high-frequency region of the *noise* condition [see also Fig. 1(e)]. The cross correlation functions of the full stimulus show their maximum at 1.5 ms with the second maximum located at −0.5 ms. This second maximum is higher in the *low-noise* than in the *noise* condition. Similarly, a *noise* stimulus with narrow bandwidth would show a higher peak at −0.5 ms than a *noise* stimulus with larger bandwidth. Trahiotis and Stern (1989), however, identified that this effect of within-channel information was not sufficient to explain bandwidth-dependent lateralization, which is why straightness weighting was introduced.

Alternatively, fast-acting neuronal adaptation as visible in auditory nerve responses could increase the effect of within-channel cues on the cross correlation function by increasing the amount of envelope fluctuations and thus, amplifying the effect of the demodulated envelope (Stecker *et al.*, 2019). Similarly, calculating a running short-term cross correlation instead of a single long-term cross correlation could potentially result in an increased effect of within-channel cues for time-segments which contain strong fluctuations compared to those with low fluctuations. Crucially, amplified within-channel information could be leveraged independently from the assumed binaural interaction mechanism, either with a delay line (Shackleton *et al.*, 1992) or without (Dietz *et al.*, 2009). Additional support for the importance of fast-acting adaptation in envelope processing can be found when again considering high-frequency ITD sensitivity: Envelope-ITD sensitivity can differ vastly for stimuli with identical envelope power spectra but different envelope shape. For example, a steep modulation onset and shallow decay generates a much higher ITD sensitivity than temporally inverted shallow onsets and steep offsets (Greenberg *et al.*, 2017; Klein-Hennig *et al.*, 2011). So far, only detailed models that did include fast-acting adaptation were able to reproduce this effect based on within-channel information (Klug *et al.*, 2020).

The main conclusion from the present data is that within-channel cues do influence the lateralization of low-frequency stimuli. The data, however, do not refute the contribution of across-frequency cues especially in cases where changeover-bandwidths are clearly larger than the ERB of the auditory filter.

V. DATA SHARING

The underlying raw data as well as the derived changeover bandwidths are openly available and can be found online (Encke and Dietz).

ACKNOWLEDGMENTS

The authors thank Helen Heinermann for her support in conducting the psychoacoustic experiments, Torsten Marquardt as well as David Mcalpine for comments on an earlier version of this manuscript. This work was supported by the Deutsche Forschungsgemeinschaft DFG (SFB 1330 - 352015383 Project B4).

Bernstein, L. R., and Trahiotis, C. (**2002**). "Enhancing sensitivity to interaural delays at high frequencies by using 'transposed stimuli'," J. Acoust. Soc. Am. **112**(3), 1026–1036.
Bernstein, L. R., and Trahiotis, C. (**2003**). "Enhancing interaural-delay-based extents of laterality at high frequencies by using 'transposed stimuli'," J. Acoust. Soc. Am. **113**(6), 3335–3347.
Dietz, M., Ewert, S. D., and Hohmann, V. (**2009**). "Lateralization of stimuli with independent fine-structure and envelope-based temporal disparities," J. Acoust. Soc. Am. **125**(3), 1622–1635.
Encke, J., and Dietz, M. "Supplemental data for the article: Influence of envelope fluctuation on the lateralization of interaurally delayed low-frequency stimuli." https://doi.org/10.5281/zenodo.5547578
Fastl, H. (**1983**). "Fluctuation strength of modulated tones and broadband noise," in *HEARING - Physiological Bases and Psychophysics*, edited by R. Klinke and R. Hartmann (Springer, Berlin Heidelberg), pp. 282–288. doi:10.1007/978-3-642-69257-4_41.
Glasberg, B. R., and Moore, B. C. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**(1-2), 103–138.
Greenberg, D., Monaghan, J. J. M., Dietz, M., Marquardt, T., and McAlpine, D. (**2017**). "Influence of envelope waveform on ITD sensitivity of neurons in the auditory midbrain," J. Neurophysiol. **118**(4), 2358–2370.
Grothe, B., Pecka, M., and McAlpine, D. (**2010**). "Mechanisms of sound localization in mammals," Physiological Rev. **90**(3), 983–1012.
Henning, B. (**1980**). "Some observations on the lateralization of complex waveforms," J. Acoust. Soc. Am. **68**(2), 446–454.
Henning, G. B. (**1974**). "Detectability of interaural delay in high-frequency complex waveforms," J. Acoust. Soc. Am. **55**(1), 84.
Jeffress, L. A. (**1948**). "A place theory of sound localization," J. Comp. Physiol. Psychol. **41**(1), 35–39.
Klein-Hennig, M., Dietz, M., Hohmann, V., and Ewert, S. D. (**2011**). "The influence of different segments of the ongoing envelope on sensitivity to interaural time delays," J. Acoust. Soc. Am. **129**(6), 3856–3872.
Klug, J., Schmors, L., Ashida, G., and Dietz, M. (**2020**). "Neural rate difference model can account for lateralization of high-frequency stimuli," J. Acoust. Soc. Am. **148**(2), 678–691.
Kohlrausch, A., Fassel, R., and Dau, T. (**2000**). "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers," J. Acoust. Soc. Am. **108**(2), 723–734.
Kohlrausch, A., Fassel, R., van der Heijden, M., Kortekaas, R., van de Par, S., Oxenham, A. J., and Püschel, D. (**1997**). "Detection of tones in low-noise noise: Further evidence for the role of envelope fluctuations," Acta Acust. United Acust. **83**(4), 659–669.
Lyon, R. F. (**2011**). "Cascades of two-pole-two-zero asymmetric resonators are good models of peripheral auditory function," J. Acoust. Soc. Am. **130**(6), 3893–3904.
Oxenham, A. J., and Dau, T. (**2001**). "Reconciling frequency selectivity and phase effects in masking," J. Acoust. Soc. Am. **110**(3), 1525–1538.
Schroeder, M. (**1970**). "Synthesis of low-peak-factor signals and binary sequences with low autocorrelation (corresp.)," IEEE Trans. Inf. Theory **16**(1), 85–89.
Shackleton, T. M., Meddis, R., and Hewitt, M. J. (**1992**). "Across frequency integration in a model of lateralization," J. Acoust. Soc. Am. **91**(4), 2276–2279.
Stecker, G. C., Dietz, M., and Stern, R. M. (**2019**). " 'Straightness' versus 'briefness' in binaural cue extraction," J. Acoust. Soc. Am. **145**(3), 1759–1759.
Stern, R. M., Zeiberg, A. S., and Trahiotis, C. (**1988**). "Lateralization of complex binaural stimuli: A weighted-image model," J. Acoust. Soc. Am. **84**(1), 156–165.
Trahiotis, C., and Stern, R. M. (**1989**). "Lateralization of bands of noise: Effects of bandwidth and differences of interaural time and phase," J. Acoust. Soc. Am. **86**(4), 1285–1293.
Vallat, R. (**2018**). "Pingouin: Statistics in python," J. Open Source Softw. **3**(31), 1026.
Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., and van Mulbregt, P. and SciPy 1.0 Contributors. (**2020**). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," Nat. Methods **17**, 261–272.
Wichmann, F. A., and Hill, N. J. (**2001**). "The psychometric function: I. fitting, sampling, and goodness of fit," Percept. Psychophys. **63**(8), 1293–1313.
Yost, W. A., Dye, R. H., and Sheft, S. (**2007**). "Interaural time difference processing of broadband and narrow-band noise by inexperienced listeners," J. Acoust. Soc. Am. **121**(3), EL103–EL109.
Zilany, M. S. A., Bruce, I. C., and Carney, L. H. (**2014**). "Updated parameters and expanded simulation options for a model of the auditory periphery," J. Acoust. Soc. Am. **135**(1), 283–286.