Localization of Amplitude-Panned Virtual Sources II: Two- and Three-Dimensional Panning*

VILLE PULKKI, *AES Member*

Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology, FIN-02015 HUT, Finland

The localization of amplitude-panned virtual sources is studied with loudspeaker pairs and triplets in arbitrary setups using listening tests and model-based analysis. The relation between panning direction and perceived direction is studied. The panning angle predicts well the angle between a virtual source and the median plane when loudspeakers are near the median plane. When loudspeakers are in lateral directions, the perceived direction is biased toward the median plane. The direction of virtual sources within the median plane is perceived individually. Various directional cues for elevation decoding are discussed

0 INTRODUCTION

Former studies in amplitude panning have mostly focused on setups in which virtual sources are generated with two sound sources placed in the horizontal plane in front or on the side of the listener. Multichannel reproduction systems and research on virtual acoustic environments have made it more common to use unconventional systems for sound reproduction [1], [2]. Questions have been raised about virtual source perception in arbitrary setups. Earlier results on virtual source localization therefore cannot be applied in several such cases.

The perceptual quality of virtual sources in horizontal setups is known quite well. The main localization cues, the interaural time difference (ITD) and the interaural level difference (ILD), explain all major phenomena of frequency-dependent localization. This is possible because the ITD and ILD decoding mechanisms are known well enough, and the effect of other cues is minor in horizontal direction decoding.

In arbitrary setups the localization cannot be explained using the main localization cues only, since cues that decode elevation are also needed. This makes research more demanding since the use of these cues is not known thoroughly. Most researchers agree that the spectral modifications caused by the pinna, head, and torso carry information regarding sound source elevation.

An important special case is when two sound sources are placed in the median plane. In this case panning between sound sources does not affect the main auditory cues at all. The localization of virtual sources, if it does occur, is based on spectral cues only. The virtual source cues may be different from cues produced by any real source. The way in which subjects perceive these directional cues reflects on the functioning of human auditory mechanisms and thus may reveal new information regarding spatial hearing.

In three-dimensional amplitude panning sound is applied to three sound sources that form a triangle from the listener's viewpoint. The localization of amplitude-panned virtual sources in such setups has not even been studied theoretically.

In a companion paper [3] the localization of virtual sources in stereophonic listening was discussed. In the present paper the virtual sources created with loudspeaker pairs and triplets in arbitrary setups are studied. Human spatial hearing and amplitude panning are reviewed in Section 1. A model of the perception of the direction of virtual sources is presented in Section 2. The methodology of conducting listening tests and simulations is presented in Section 3, and the test results are reported in Sections 4 and 6. A cue for elevation direction decoding is proposed in Section 5. The localization of broad-band virtual sources in different conditions is simulated in Section 7.

---

* Manuscript received 2001 October 10; revised 2001 May 17.

1 THEORY

1.1 Spatial Hearing

The duplex theory of sound localization states that the two frequency-dependent main cues of sound source localization are the interaural time difference (ITD) and the interaural level difference (ILD). They are caused by the wave propagation time difference (primarily below 1.5 kHz) and the shadowing effect of the head (primarily above 1.5 kHz), respectively [4], [5]. The cues specify a cone of confusion where the sound source is located. A cone of confusion can be approximated with a cone that has a symmetry axis along a line passing through the listener's ears, as shown in Fig. 1. The direction perception inside a cone of confusion is refined using other cues.

1.1.1 Coordinate Systems

Conventionally the spherical coordinates are denoted by azimuth θ, elevation ϕ, and range r. If elevated sound sources are also applied in auditory research, this formulation is inconvenient, since a cone of confusion cannot be easily formulated. An alternative spherical coordinate system has been used by Duda [6]. The sound source location is specified by defining the cone of confusion in which it lies, and further by the direction within the cone of confusion. The cone of confusion is defined by an angle θ_{cc} between the median plane and the cone (Fig. 1). The variable θ_{cc} may have values between $-90°$ and $90°$. The direction within the cone of confusion is denoted by ϕ_{cc}, and it varies between $-180°$ and $180°$. These coordinates are used throughout this paper. Standard nomenclature for these variables is lacking, though it would be beneficial. In some studies θ_{cc} is referred to as the left/right (L/R) direction.

1.1.2 Perception of ϕ_{cc} Direction

In pinna-occlusion experiments it has been demonstrated that the localization ability in the median plane deteriorates when the pinna cavities are filled [7], [8]. Shaw and Teranishi investigated the resonances of pinna cavities [9], [10]. They found two modes that were vertical dipoles and three modes that were horizontal dipoles. Typically these modes create a strong response to elevated sound sources and a weak response to frontal sound sources at frequencies between 5 and 10 kHz, and vice versa at frequencies higher than 12 kHz. There also exist pronounced peaks or notches in HRTFs, the frequencies of which are dependent on ϕ_{cc}. The pinna effects appear at frequencies above 4–5 kHz [10]. Algazi and Avendano [11] have shown that the effect of shoulder reflection carries information of the ϕ_{cc} direction outside the median plane also at lower frequencies.

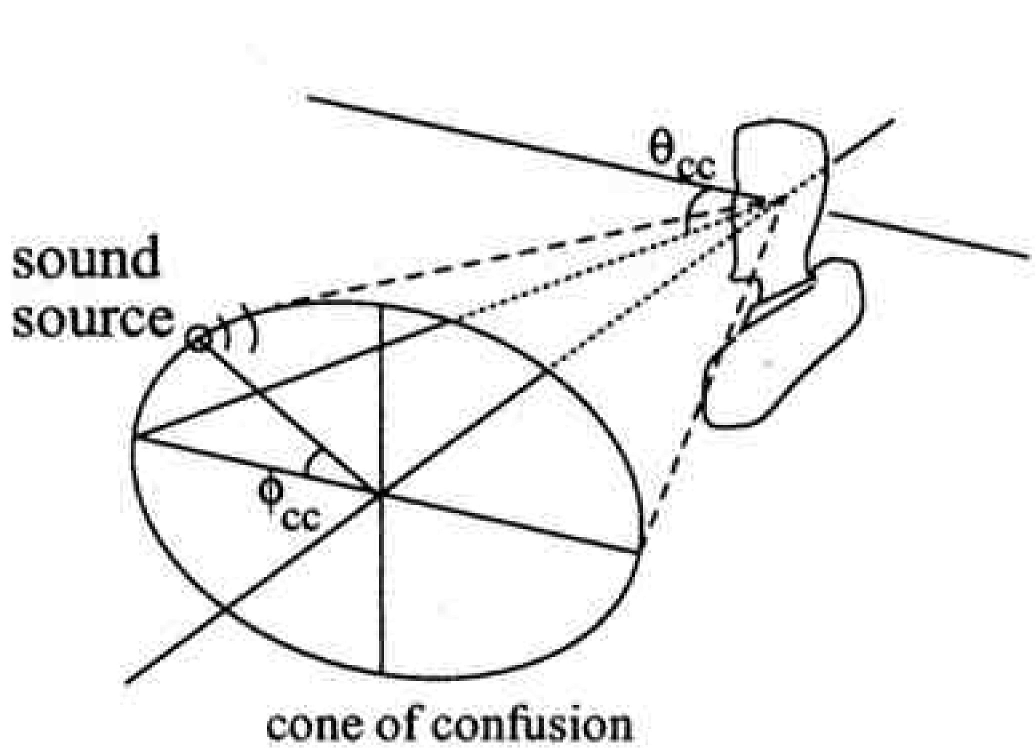


Fig. 1. Cone of confusion. Spherical coordinate system used in directional hearing.

Blauert [12] found that in the median plane one-third-octave noise is localized as a function of its frequency and not of its actual ϕ_{cc} direction. He found that most listeners localized the signal upward in direction when its center frequency was near 8 kHz. Middlebrooks [13] conducted tests in which he found that one-sixth-octave band-pass noise bursts at 6 kHz are localized upward and 8- and 10-kHz sounds are localized downward, independent of the position of a real source. 12-kHz sounds were localized quite consistently in the horizontal plane. The spectral contents of the ear signals thus include directional information, although it is not known which features (cues) in the spectra are used in decoding.

It has been shown that with a broad-band stimulus with constant level a neural network learned to decode ϕ_{cc} from monaural spectra with greater accuracy than humans [14]. The feature vector used in training was derived from the broad-band magnitude spectrum. Unfortunately it was reported that this network gave consistent results only with the stimuli used in training the network. It can be assumed that cues that are less dependent on stimulus spectra are used in ϕ_{cc} direction decoding by humans.

Since the black-box approach of using neural network decoding does not reveal the underlying mechanisms, more structured parametric modeling is desired. Blauert [12] suggested that the frequencies of spectral peaks are a prominent cue in ϕ_{cc} direction decoding. In some studies the frequencies of spectral notches were found to be prominent cues [15], whereas other studies state that the frequencies of both notches and peaks are important [16]. Middlebrooks [13] calculated the correlations between the shapes of HRTFs and the shapes of sound spectra appearing at the ears. Zakarauskas and Cynader [17] suggested that the second finite difference of the spectrum could be used as a cue. The second finite difference corresponds to the second frequency derivative of the signal spectrum. It has been suggested that the spectral modifications could be decoded from the ILD spectrum [6] when the sound source is outside the median plane. It has also been shown that interaural cross correlation conveys elevation and front–back discrimination ability in the median plane [18] when neural networks are used in elevation decoding.

1.2 Amplitude Panning

In amplitude (intensity) panning two or more sound sources are placed in different directions at equal distances from the listener. The same sound signal, but with different amplitudes, is applied to the sound sources. The sound signals from each source arrive at each ear. The signals are summed at the ear canals, forming new signals the attributes of which specify the perceived localization. This is called summing localization and the perceived auditory object is called the amplitude-panned virtual source. In this paper only the directional localization of virtual sources is discussed. The perception of distance is omitted in this study.

In two-dimensional amplitude panning two sound sources (loudspeakers) are in the same plane with the listener. In three-dimensional amplitude panning three loudspeakers form a triangle from the listener's point of view; in other words, the loudspeakers and the listener are not coplanar. The terms two- and three-dimensional thus are used in their natural meanings as spatial dimensionalities. The setups are called loudspeaker pairs or triplets and, more generally, loudspeaker sets. The perceptions of the θ_{cc} and ϕ_{cc} directions of virtual sources in two- and three-dimensional amplitude panning are considered separately.

1.2.1 Perception of θ_{cc} Direction in Two-Dimensional Panning

The localization of the θ_{cc} direction of amplitude-panned virtual sources is often estimated based on the low-frequency ITD, which is an important cue in sound localization. The ITD generated by amplitude panning can be approximated using a simplified model of a listener. In the model the listener is approximated by two spatially separated ears with no acoustic shadow from the head [19]. The model is valid at frequencies roughly below 400 Hz. The ILD or the high-frequency ITD cannot be estimated with this model.

Using this model and phasor analysis it has been shown that the amplitude difference of loudspeakers generates an ITD between the ears at low frequencies [20]. It can also be shown that amplitude panning creates virtual source ITDs ranging between the ITDs that the loudspeakers would produce alone. Thus the θ_{cc} direction of the virtual source is always between the θ_{cc} values of the loudspeakers.

Two laws for virtual source localization have been derived for stereophonic listening using the simplified model of a listener. In stereophonic listening two loudspeakers are positioned in the horizontal plane azimuth directions θ_0 and $-\theta_0$, where θ_0 is called the stereophonic base angle, and is typically 30°. The base angle value has also an effect on the perception of the virtual sources, although in this study it is not investigated.

In this study we are searching for generalizations of these laws. They can be found when it is noted that the only information needed in the derivation of the laws is the distance from the loudspeakers to the ears. Changing the ϕ_{cc} directions of loudspeakers does not affect the derivation. Although the laws have been traditionally applied to loudspeaker pairs in a horizontal setup, they are also valid for pairs having loudspeakers equally distant from the median plane with arbitrary ϕ_{cc} directions, as seen in Fig. 2. The θ_{cc} angle between the loudspeakers and the median plane is θ_{cc_0}, and is called the stereophonic base θ_{cc} angle.

The laws are now rewritten using the θ_{cc} directions. The sine law by Bauer [20] is formulated as

$$\frac{\sin\theta_{cc_s}}{\sin\theta_{cc_0}} = \frac{g_1 - g_2}{g_1 + g_2}. \qquad (1)$$

The ϕ_{cc} directions of real or virtual sources are not restricted by this law. In the equation θ_{cc_s} is the predicted θ_{cc} direction perception of a virtual source, and g_1 and g_2 are the gain factors (see Fig. 2). The gain factor values are between 0 and 1 and control the amplitudes of the sound signals emanating from the respective loudspeakers.

Bennett et al. [21] derived a panning law by improving the head model used in the derivation of the sine law by approximating the propagation path from the contralateral loudspeaker to the ear with a curved line around the head. It is rewritten as

$$\frac{\tan\theta_{cc_T}}{\tan\theta_{cc_0}} = \frac{g_1 - g_2}{g_1 + g_2} \qquad (2)$$

where θ_{cc_T} is the predicted θ_{cc} direction perception of a virtual source and the other variables are the same as in the sine law. In this law the θ_{cc} directions are defined as angles from the median plane of the listener prior to head rotation.

Although the model of the listener is valid only at frequencies below roughly 400 Hz, in the companion paper [3] it was shown that in a horizontal setup the laws are approximately correct up to 1100 Hz. Also, the high-frequency ILD cues are also roughly consistent with the low-frequency ITD cues. There are, however, some deviations that cause some frequencies to be localized differently.

No laws have been derived for pairs that are not symmetric with the median plane. In this paper the virtual source perceptions are simulated in such setups.

1.2.2 Perception of ϕ_{cc} Direction in Two-Dimensional Panning

As stated earlier, the ϕ_{cc} direction perception is based mostly on decoding the spectral modifications caused by the pinna, head, and torso. When the loudspeakers are in the same ϕ_{cc} direction, the spectral cues produced by

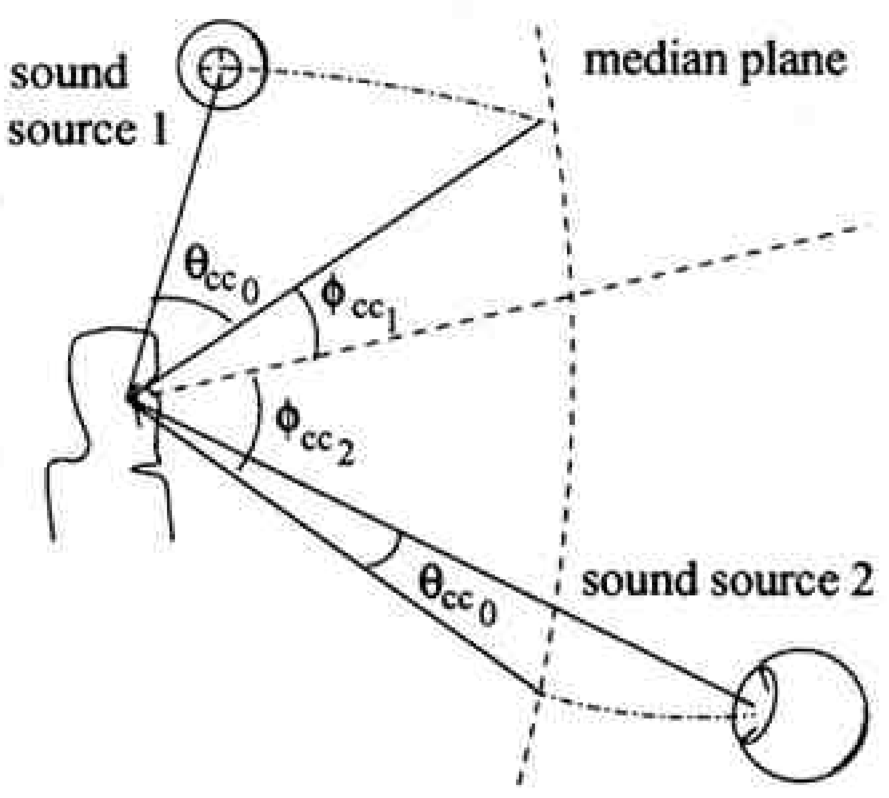


Fig. 2. θ_{cc} angle between median plane and both loudspeakers is the same, but ϕ_{cc} directions are different.

the loudspeakers alone may be similar. When amplitude panning is applied, the spectrum of the virtual source is colored due to summing localization. The virtual source may thus be perceived in a different ϕ_{cc} direction than the ϕ_{cc} direction of the loudspeakers. However, in most cases the virtual source is perceived to have the same ϕ_{cc} direction as the loudspeakers.

More problematic cases can be found when the ϕ_{cc} directions of the loudspeakers are different. An important special case is when the loudspeakers have the same θ_{cc} direction value. In these cases the spectra of the signals arriving from the loudspeakers differ between each other. The spectral content of the summed sound signal should then propose ϕ_{cc} directions that are between the ϕ_{cc} directions of the loudspeakers. There is no a prior knowledge regarding what the perceived virtual source ϕ_{cc} direction is.

1.2.3 Three-Dimensional Panning

This paper originated from the problem of evaluating the localization of virtual sources created using the vector base amplitude panning (VBAP) method [22]. VBAP is a reformulation of the tangent law, which can be generalized to loudspeaker triplets.

In three-dimensional VBAP a loudspeaker triplet is specified with vectors, as in Fig. 3. Cartesian unit-length vectors $\boldsymbol{l}_m$, $\boldsymbol{l}_n$, and $\boldsymbol{l}_k$ point from a listening position to the loudspeakers. The direction of the virtual source is presented with a unit-length vector $\boldsymbol{p}$, expressed as a weighted sum of the loudspeaker vectors,

$$\boldsymbol{p} = g_m \boldsymbol{l}_m + g_n \boldsymbol{l}_n + g_k \boldsymbol{l}_k \, . \tag{3}$$

Here g_m, g_n, and g_k are the gain factors of the respective loudspeakers. The gain factors can be solved as

$$\boldsymbol{g} = \boldsymbol{p}^{\mathrm{T}} \boldsymbol{L}_{mnk}^{-1} \tag{4}$$

where T denotes matrix transposition, $\boldsymbol{g} = [g_m \;\; g_n \;\; g_k]^{\mathrm{T}}$, and $\boldsymbol{L}_{mnk} = [\boldsymbol{l}_m \;\; \boldsymbol{l}_n \;\; \boldsymbol{l}_k]$. The calculated factors are used in amplitude panning as gain factors of the signals applied to the respective loudspeakers after suitable normalization; for example, $\|\boldsymbol{g}\| = 1$. VBAP can be controlled using (azimuth, elevation) or (θ_{cc}, ϕ_{cc}) coordinates. Spherical coordinates are transformed to Cartesian coordinates prior to calculating the gain factors.

Although VBAP is formulated to be used with triplets that have loudspeakers in arbitrary directions, no evidence exists that the virtual sources would appear in the directions that VBAP suggests. However, a regularity for the perception of the θ_{cc} direction can be formulated at low frequencies in analogy with two-dimensional panning. The θ_{cc} directions of loudspeakers 1, 2, and 3 in a triplet are denoted by θ_{cc_1}, θ_{cc_2}, and θ_{cc_3}. If loudspeaker 3 is not used, loudspeakers 1 and 2 produce summed ear signals that can be produced by a loudspeaker in the direction $\theta_{cc_{12}}$. The following then holds:

$$\min_i(\theta_{cc_i}) \leqslant \theta_{cc_{12}} \leqslant \max_j(\theta_{cc_j}) \, , \qquad i, j \in \{1, 2\} \, . \tag{5}$$

When the third loudspeaker is applied, the summing localization can be thought to happen between the virtual sound source at the $\theta_{cc_{12}}$ direction and the loudspeaker at the θ_{cc_3} direction. The listener will perceive a virtual source at the direction $\theta_{cc_{123}}$, for which, naturally,

$$\min_i(\theta_{cc_i}) \leqslant \theta_{cc_{123}} \leqslant \max_j(\theta_{cc_j}) \, , \qquad i, j \in \{1, 2, 3\} \, . \tag{6}$$

This means that with amplitude panning the virtual source θ_{cc} direction can vary between the minimum and maximum of the θ_{cc} directions of the loudspeakers. If a θ_{cc} direction of 90° is located inside the triangle, there exists a circle around this direction inside of which the virtual source cannot be positioned. This follows since the absolute values of the loudspeaker's θ_{cc} directions are smaller than 90° and from Eq. (6). This can be demonstrated with a triplet, as shown in Fig. 4. The directions where the virtual sources could theoretically be perceived are shaded. The virtual sources are therefore not necessarily located inside the triangle that the three loudspeakers define. However, human localization accuracy is low in lateral directions, which implies that this phenomenon is not perceived clearly.

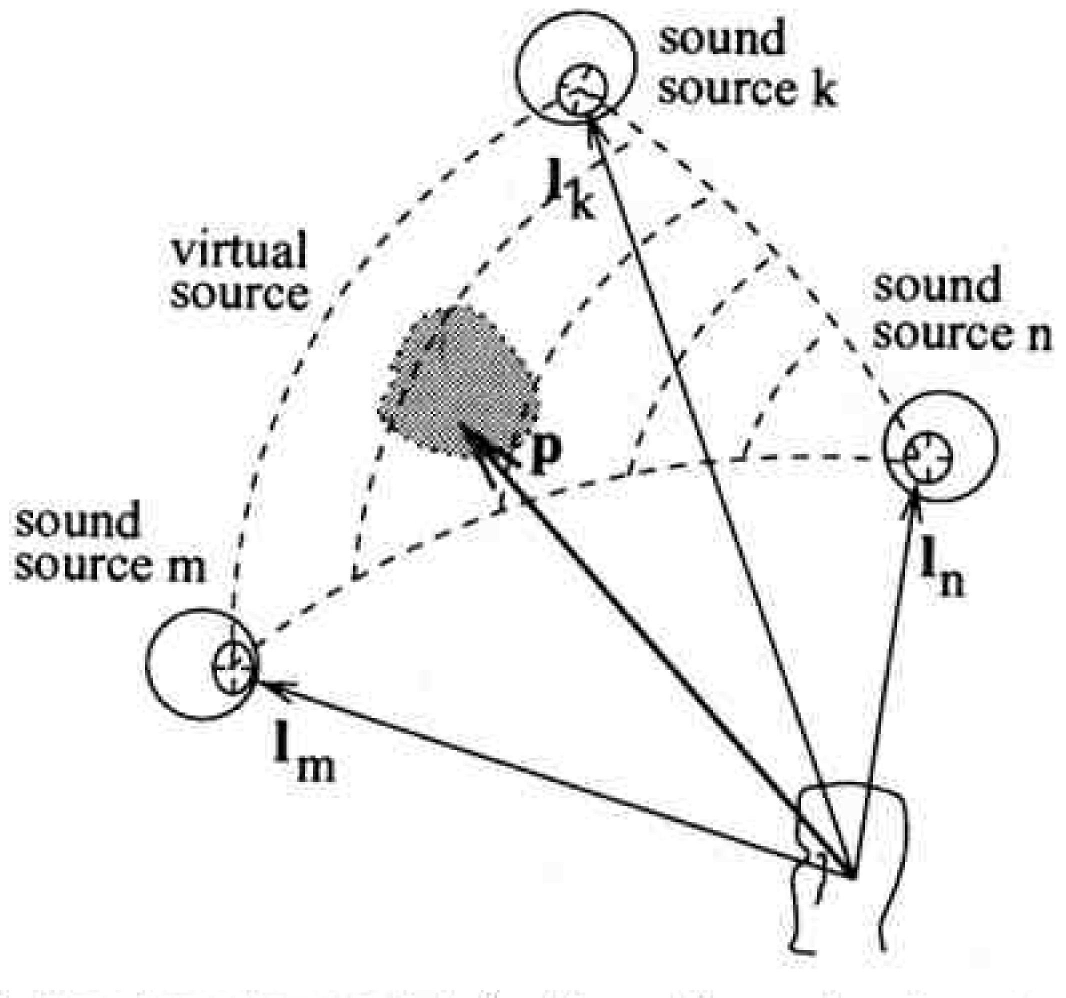


Fig. 3. Loudspeaker triplet for three-dimensional vector base amplitude panning (VBAP).

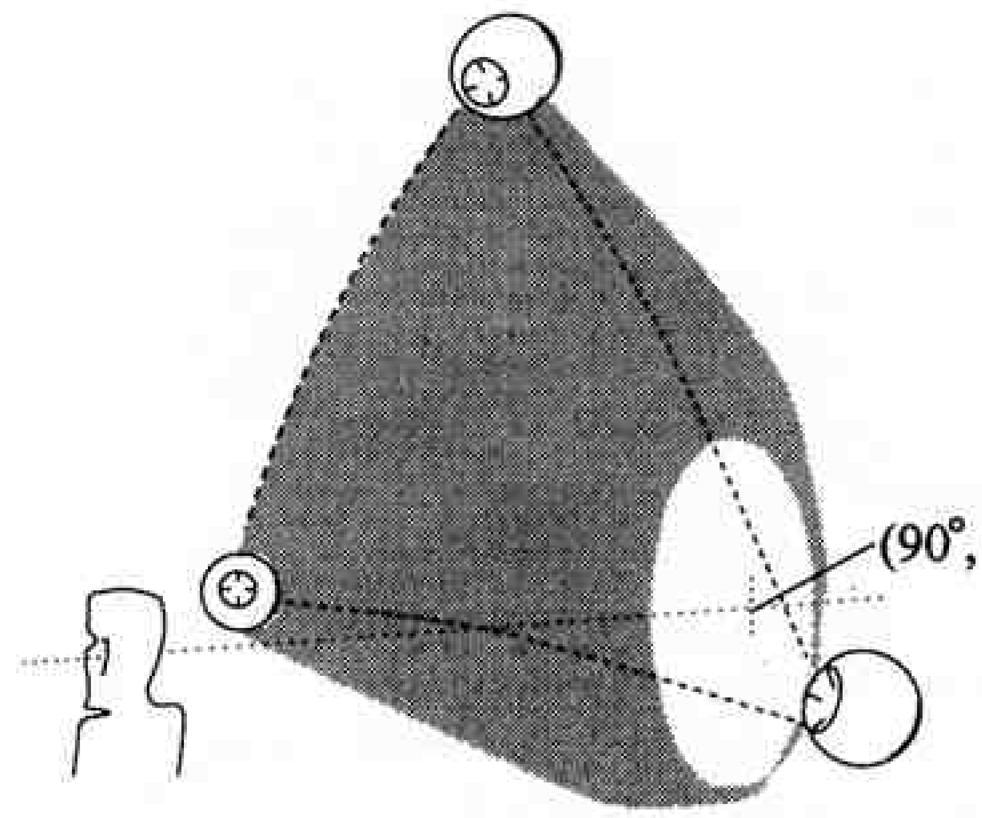


Fig. 4. Triangle that includes all panning directions defined by a loudspeaker triplet (dashed lines). Virtual source cannot be positioned to circle around direction (90°, 0°). Directions where virtual sources may occur are shaded.

2 MODELING THE DIRECTION PERCEPTION OF VIRTUAL SOURCES

To simulate the perception of virtual sources, a binaural auditory model is used to calculate localization cues for signals arriving at the ear canals. The θ_{cc} directions that the main localization cues suggest are calculated with the binaural auditory model presented in a companion paper [3].

The model of auditory localization used in this study consists of the following parts:

- Simulation of ear canal signals using HRTFs
- A binaural model of neural decoding of directional cues
- A model of high-level perceptual processing for the θ_{cc} direction decoding based on a database search.

The model estimates the ITD angle (ITDA) and the ILD angle (ILDA) as functions of the equivalent rectangular bandwidth (ERB) pitch scale. The ITDAs and the ILDAs present the θ_{cc} directions where a real source should be placed to produce the virtual source ITD or ILD. As an addendum to the model in the companion paper, monaural loudness level spectra are also simulated. The loudness levels of each ERB band are calculated using the formulas of Zwicker and Fastl [23]. The spectra include features that do not carry directional information. These features are estimated by a spectrum that is formed by averaging loudness level spectra in 15 directions evenly around the listener. This spectrum is subtracted from a measured loudness level spectrum, forming a directional loudness level (DLL) spectrum that is an auditory analog to the directional transfer functions used in some studies [24].

3 EXPERIMENTAL METHODS

The methodology is described briefly, since the techniques used in the listening tests are very similar to those used in the companion paper [3].

3.1 Listening Tests

In the listening test used in this study the subject matched the perceived direction of a virtual source as accurately as possible with the perceived direction of a reference real source (a loudspeaker). When a virtual source adjustment was finished with one real source, the next adjustment was conducted with another real source. The subject adjusted the panning direction using (θ_{cc}, ϕ_{cc}) coordinates. The virtual source was produced with two or three loudspeakers, for which the gain factors were computed and controlled with VBAP. As a test result, a set of panning angles was obtained that produced the perception of a virtual source in the same direction as the reference real source. The initial panning angle was selected randomly.

The tests were conducted in a large anechoic chamber. Subjects were sitting in a chair with a lightweight head rest. They were asked to hold their heads still during listening. The chair was adjusted to align the listener's ears to the same horizontal plane with a loudspeaker at 0° of ϕ_{cc}. The loudspeakers were visible, but the subjects were asked to keep their eyes closed during the adjustment process. The test signals were broad-band or octave-band pink noise.

In the test the reference real source was presented first, followed by a presentation of the virtual source. The sequence was repeated after a short pause. The temporal envelope of the stimuli is presented in Fig. 5. Some details of the tests are given in Table 1.

3.2 Analyzing the Listening Test Results

We may interpret the subjects' performance by simulating the auditory cues produced by the virtual sources. The ITD, ILD, and DLL spectra are simulated using the auditory model presented in Section 2 for virtual and reference real sources. Since the decoding mechanisms for spectral cues are not well known, the analysis can be conducted only in a qualitative fashion. The spectra of real and virtual sources are compared and discussed.

Due to large variations between individual DLL spectra, personal HRTFs of each subject had to be used to simulate the spectra that appeared in a subject's ears during the listening test. Measured HRTFs were available for eight subjects, which were measured with the same system that was used in the listening tests. The analysis was completed after all of the listening tests had been conducted. The subjects thus could not make their decisions match expectations based on the simulations.

Table 1. Listening test details.

Loudspeakers	Spherical: Ø 160 mm; Driver: Audax AT080M0
Loudspeaker distance	2m
Subjects	Ages 19–31; hearing threshold ≤20 dB
Familiarization	One virtual source adjustment
Pink noise	1.4-s sample of broad-band pink noise
Octave-band noise	Slope attenuation −25 dB/octave; 10 center frequencies according to ERB scale
Loudness	60 dB(A) ±4 dB

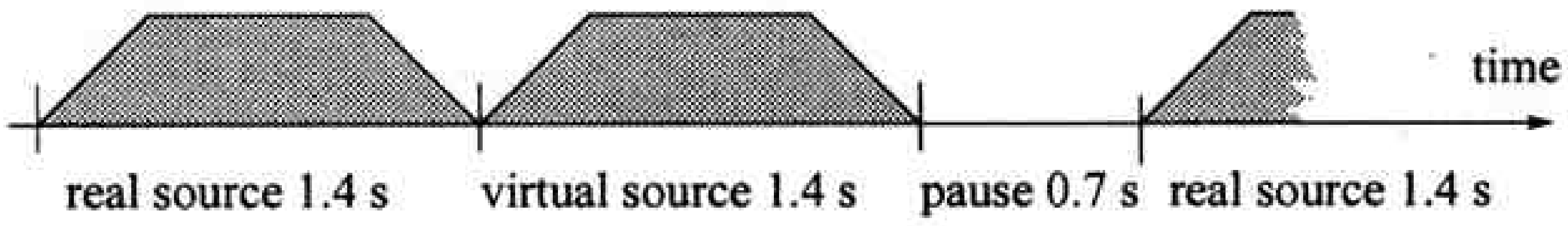


Fig. 5. Temporal envelope of test stimuli.

4 EXPERIMENT I: PANNING IN THE MEDIAN PLANE

The localization of virtual sources was investigated with amplitude panning within a loudspeaker pair in the median plane. Loudspeakers for virtual source production were in the ϕ_{cc} directions of $-15°$ and 30°, and loudspeakers that were used as reference real sources were in the ϕ_{cc} directions of 0° and 15°, as shown in Fig. 6.

4.1 Experiences of Subjects

The virtual source was perceived to be spread differently by different subjects. Some subjects reported that they could adjust the direction of the virtual source quite consistently, and that the spread of the virtual source was relatively small. Some subjects reported that they perceived the virtual source to be very diffuse. Subject VV reported that the virtual source was spread between $-15°$ and 30° of the ϕ_{cc} direction, but still he could adjust the direction of the "center of gravity" of the virtual source.

Some subjects also reported that the spread of virtual sources was dependent on the panning angle. In these cases the virtual source was more spread out near one of the reference real sources, which made the adjustment process more difficult for that source. The timbre of the virtual source was reported to be different from the reference real source.

Subject TL reported that the virtual source was localized mostly toward the directions of the loudspeakers that generated the virtual sources. He localized the virtual sources to directions between the sources, with very few panning angles. Some other subjects reported a similar behavior of the virtual source. However, they perceived the directions between the sources with a greater number of panning angles.

Two test candidates could not conduct the test because they localized the virtual sources inside their heads. They were replaced by other subjects.

4.2 Pink Noise Listening Test

To investigate the dependence between amplitude panning laws and perceived virtual source direction with broad-band signals, pink noise was used in the first test. There were 14 subjects, and they completed the adjustment four times for both reference real sources.

4.2.1 Results

The data were tested to fulfill the assumptions for an analysis of variance (ANOVA) of normal distribution of data and residuals. The ANOVA test was run to find the effects of the following variables: *subject*, *source* (0° or 15° of ϕ_{cc}), and *repetition*. Also, all two-way interactions were tested. The ANOVA test results are shown in Table 2.

The variable *repetition* did not affect the results significantly; neither was its interaction with *subject* or *source* significant. This implies that learning was not a significant factor in the judgments.

The effect of the variable *source* was significant. The results are presented in Fig. 7 as a box plot for each reference real source. The responses have greater values for the reference real source at 15° than for the reference real source at 0°. The listeners thus had to adjust the panning angle to larger values to perceive the virtual source in higher ϕ_{cc} directions. This suggests that the perceived ϕ_{cc} direction varies monotonically with the ϕ_{cc} panning angle.

The median responses correspond fairly well to the

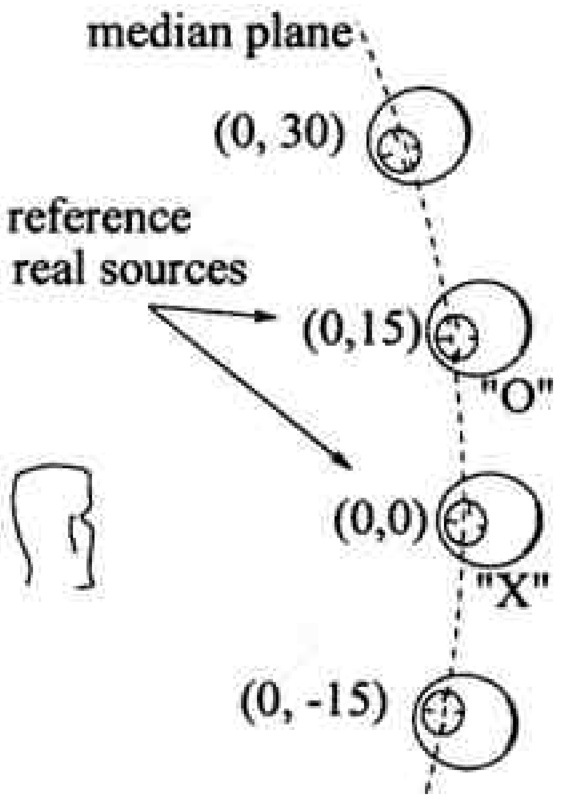


Fig. 6. Setup used in median plane listening tests. Directions expressed as (θ_{cc}, ϕ_{cc}) pairs.

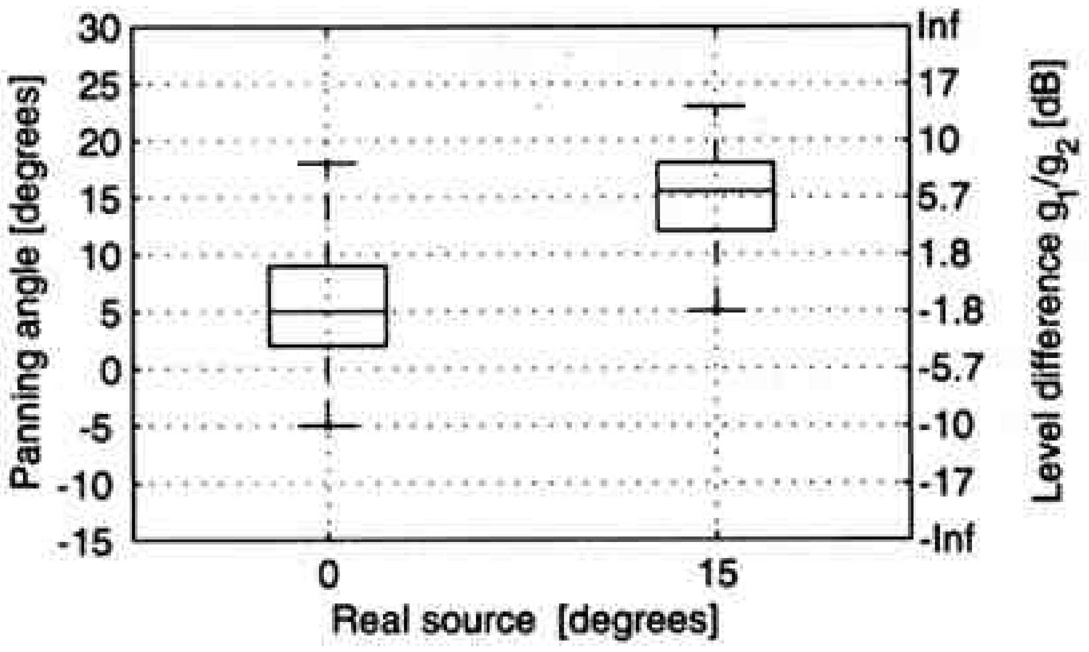


Fig. 7. Box plot for results of median plane listening test with pink noise. Panning angles judged best for reference real sources at 0° or 15° ϕ_{cc} directions. 14 listeners; adjustments conducted four times to each real source; loudspeaker span 45°; setup as in Fig. 6.

Table 2. ANOVA results of data from vertical pink noise tests.

Source	Sum of Squares	df	Mean Square	F	σ
Subject	1252	13	96.3	20.1	<0.001
Source	2283	1	2283	477.7	<0.001
Repetition	5.52	3	1.84	0.385	0.765
Subj * source	578.9	13	44.5	9.3	<0.001
Subj * repetition	291	39	7.46	1.56	0.087
Source * repetition	17.4	3	5.81	1.22	0.318

reference real source directions. However, the data ranges are quite wide, approximately 20° of the panning angle with each real reference source. In the companion paper it was reported that in a similar test, using stereophonic listening, the range of the horizontal angles was only 3°. This suggests that the panning angle predicts the θ_{cc} direction much more accurately than the ϕ_{cc} direction, at least in these setups.

The effects of the variable *subject* and of the interaction *subject * source* were significant. These features suggest that each subject needs an individual panning angle to perceive the virtual source in the ϕ_{cc} direction of each reference real source. Based on this result it is clear that the perception of the ϕ_{cc} direction of an amplitude-panned virtual source is individual, which yields that a single law that would predict the virtual source direction correctly for all individuals cannot be formed.

The responses of individual subjects are now discussed. Fig. 8 shows that subjects other than NP, LS, and TL panned the virtual sources in a relatively consistent way. Here to be consistent means that the virtual sources matched to the 0° reference real source have a lower ϕ_{cc} panning angle than the virtual sources matched to the 15° reference real source. Also, the panning angle distributions of each individual to each reference real source are relatively narrow. This suggests that the perception of the virtual source is similar each time.

Subjects NP, LS, and TL panned virtual sources corresponding to different reference real sources to almost the same panning angle values. This suggests that these subjects perceived the virtual source to jump from $-15°$ to 30° of ϕ_{cc} with very few panning steps. This is in agreement with TL's subjective experiences.

4.2.2 Analysis

The goal of this analysis is to find the reasons why subjects favored certain panning angles to perceive virtual sources in the same directions as reference real sources. The analysis was performed with the responses of subjects HJ, LS, MR, RV, TL, VP, and VV, since their HRTFs were available. The spectral contents of the signals were monitored with DLL spectra, which are shown in Fig. 9 for the real and virtual sources for both ears of subject RV and for the right ear of subjects TL and VV. Each virtual source was simulated with the median panning angle of four adjustments to one reference real source. Not all simulation results are shown since the figures would require several pages and would not yield any new information.

The correspondence between virtual and real source DLL spectra is studied. The dominant theory of the perception of the ϕ_{cc} is based on decoding the location of the peaks and notches in frequency [12]. The virtual source DLL spectrum should therefore have peaks or notches at the same frequencies with the corresponding reference real source spectra. However, when the virtual source spectra are compared with reference real sources, there seems to be no correspondence between peaks or notches for any of the subjects analyzed. It seems that summing localization has removed the narrow notches and peaks from the DLL spectra and created distortions to them. The performances of subjects who perceived the virtual sources consistently can therefore not be explained with this theory of ϕ_{cc} direction localization.

The DLL spectra of virtual sources are different for each ear, as can be seen in subject RV's plots. This might be at least a partial reason why there existed prominent spreading in virtual sources. Subject TL did not consistently hear the virtual sources. His real source spectra seem to behave differently with ϕ_{cc} when compared to other subjects' spectra. This may be a reason why his virtual source perceptions were in general different from those of other subjects.

4.3 Octave-Band Noise Listening Tests

To explore the panning performance with stimuli of a narrower band, a test similar to the one in the previous section was run with octave-band noise at 10 center frequencies selected according to the ERB scale. There

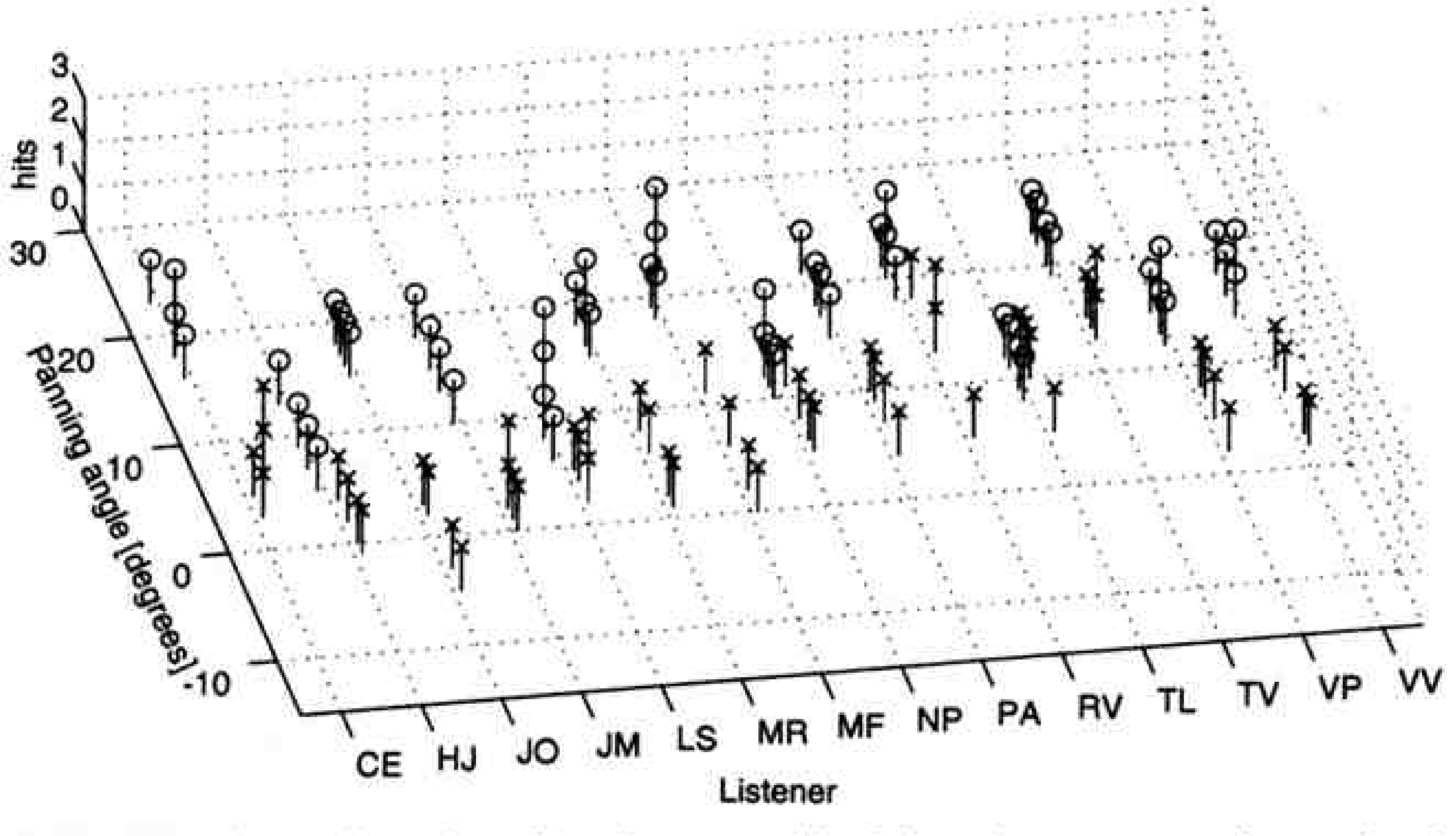


Fig. 8. Panning angles judged best in median plane listening test with pink noise. × —panning angles for reference real source at a 0° ϕ_{cc} direction; ○—panning angles for reference real source at a 15° ϕ_{cc} direction. 14 listeners; adjustments conducted four times to each real source; loudspeaker span 45°; setup as in Fig. 6.

were ten subjects, and the subjects completed the adjustment process twice for both reference real sources with each stimulus. In preliminary trials for the listening test it was found that at some frequency bands the localization of a virtual source could not be adjusted at all. A "not possible" button was added, and participants were asked to press it when they were unable to complete the virtual source adjustment.

4.3.1 Test Results

The percentage of how often subjects pressed the "not possible" button is plotted in Fig. 10. Subjects pressed it frequently at low frequencies, and for signals with center frequencies above 4 kHz they did not use it at all. This coincides with earlier results, in which it was found that high-frequency components have to be presented to perceive the sound source elevation [4, p. 102].

The panning angles judged best for each frequency band are shown in Fig. 11. If the subject had pressed the "not possible" button, his or her judgment was replaced with a random value in the data analysis. This is consistent with the fact that in listening tests the subjects were advised to leave the virtual source in a random panning angle value if a judgment could not be made.

In the results it can be seen that at low frequencies the panning angle distribution has a broad data range, in some cases spanning all the defined panning angle values. For stimuli with center frequencies above 3 kHz the data ranges become narrower but are still quite wide, and the medians deviate considerably from the corresponding reference real source directions.

Investigations are carried out to determine which stimuli were localized consistently. As in the previous test, it is assumed that the subject perceived the virtual source consistently if he or she adjusted the panning angles to lower values for the 0° real source than for the 15° real source. To find out at which frequency bands consistent perception occurred, a Wilcoxon signed rank test was run between the panning angles corresponding to the upper and lower reference real sources at each frequency band of each subject. The null hypothesis was that the direction of the reference real source did not affect the center of the panning angle distribution in one frequency band. The test results are shown in Table 3.

It was found that the effect of the direction of the reference real source was not significant at frequency bands below 3000 Hz. Above this the effect is significant. This, together with the frequency of pressing the "not possible" button, suggests that the subjects' responses at frequency bands above 3 kHz are based on adjusting the direction of the virtual source to approximately the same direction as the reference real source. At lower frequency bands it can be assumed that amplitude panning did not create a stable virtual source between loudspeakers, and the subjects were guessing or judging the adjustment to be "not possible." Thus amplitude panning does not affect the perceived ϕ_{cc} direction if high frequencies are not present.

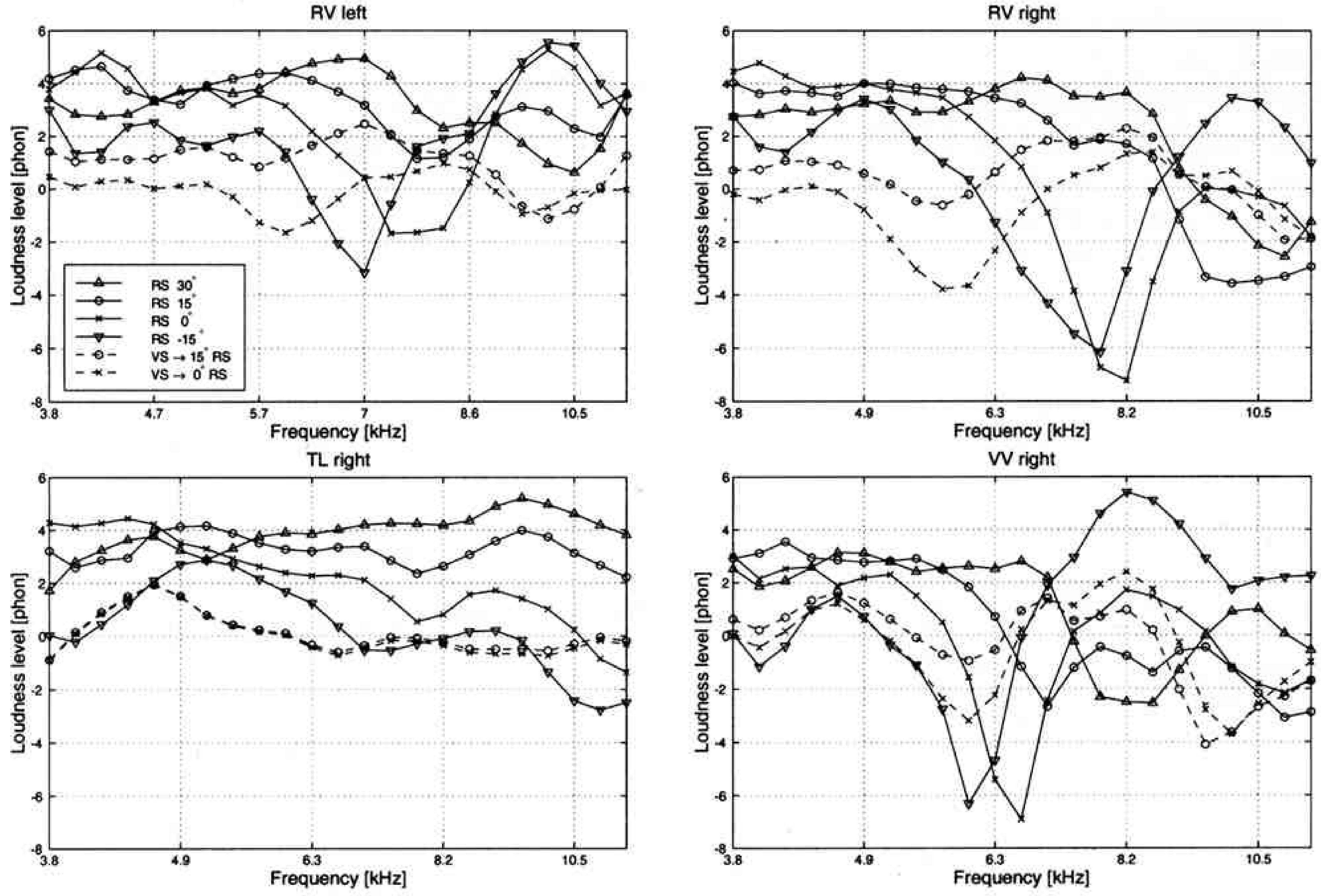


Fig. 9. Monaural directional loudness level spectra for 3 subjects. Median plane listening test with pink noise. RS—reference real source; VS ––– virtual source that has been matched to real source at 0° or 15° of ϕ_{cc}.

4.3.2 Analysis

For data analysis with the auditory model, HRTFs were available for subjects RV, HJ, LS, TL, and VP. Model-based simulations were run for these five individuals. When the DLL spectra of real and virtual sources were investigated, similar phenomena as in the previous test were found, that is, the peak or notch frequencies did not agree for real and virtual sources. The subjects' performances therefore cannot be explained by peak or notch frequency cues.

However, since the subjects perceived the virtual sources consistently at frequency bands with a center frequency above 3 kHz, there must be some information in the spectral content of the virtual sources that affected the perception of the ϕ_{cc} direction. A spectral characteristic should exist in octave band signals with center frequencies equal to or above 3.9 kHz, and should not exist with center frequencies below 3.9 kHz. Subject HJ's left-ear virtual source DLL spectra, when panned from $-15°$ to 30°, are shown in Fig. 12 for octave bands with center frequencies above or equal to 1.7 kHz.

A spectral characteristic that meets these requirements can be found: the magnitude near 6 kHz increases with the panning angle. This characteristic is illustrated in Fig. 12 by double lines. For octave-band noises with center frequencies below 3.9 kHz, this characteristic is more than 20 phons ($\approx$20 dB) below the highest level of the stimulus and can be assumed not to be significant. It can be assumed that this feature produced the changing perception of the ϕ_{cc} direction. The analysis results for other subjects were similar.

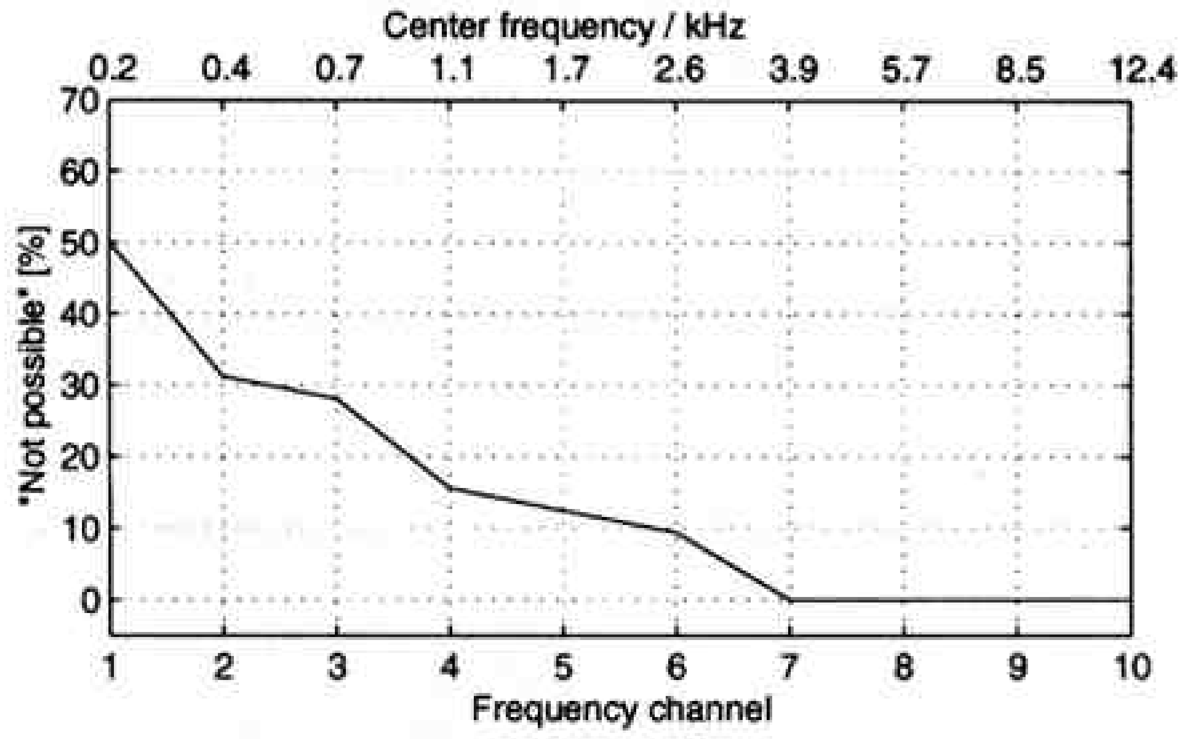


Fig. 10. Median plane listening test with octave-band noise; percentage of "not possible" judgments. Reference real sources at 0° or 15° ϕ_{cc} directions. 10 listeners; adjustments conducted twice to each real source with each frequency channel; loudspeaker span 45°; setup as in Fig. 6.

5 HYPOTHESIZING A CUE FOR PERCEPTION OF THE ϕ_{cc} DIRECTION

It was shown that the spectral peak or notch cues for decoding the ϕ_{cc} direction are not sufficient to explain the test results discussed. Different cues must also exist in ϕ_{cc} decoding. A hypothesis of a new type of cue is proposed here, based on our results and on studies of the modes of the pinna [10]. The loudness level L at a pinna mode frequency f_m, denoted by $L(f_m)$, varies with the ϕ_{cc} direction of a sound source. If the loudness level at a mode frequency $L_0(f_m)$ that would appear if the pinna modes did not exist is also known, then the ϕ_{cc} direction can be decoded.

Table 3. Octave-band median plane test*

Frequency Channel	Z	p
1	-0.26	0.79
2	-1.08	0.28
3	-0.50	0.61
4	-0.40	0.69
5	-0.88	0.38
6	-1.44	0.15
7	-3.18	0.001
8	-3.456	0.001
9	-3.48	0.001
10	-3.81	<0.001

* Wilcoxon signed-rank test results of test if centers of panning angle distributions corresponding to lower and upper real sources were different. Test was run on data at each frequency channel separately.

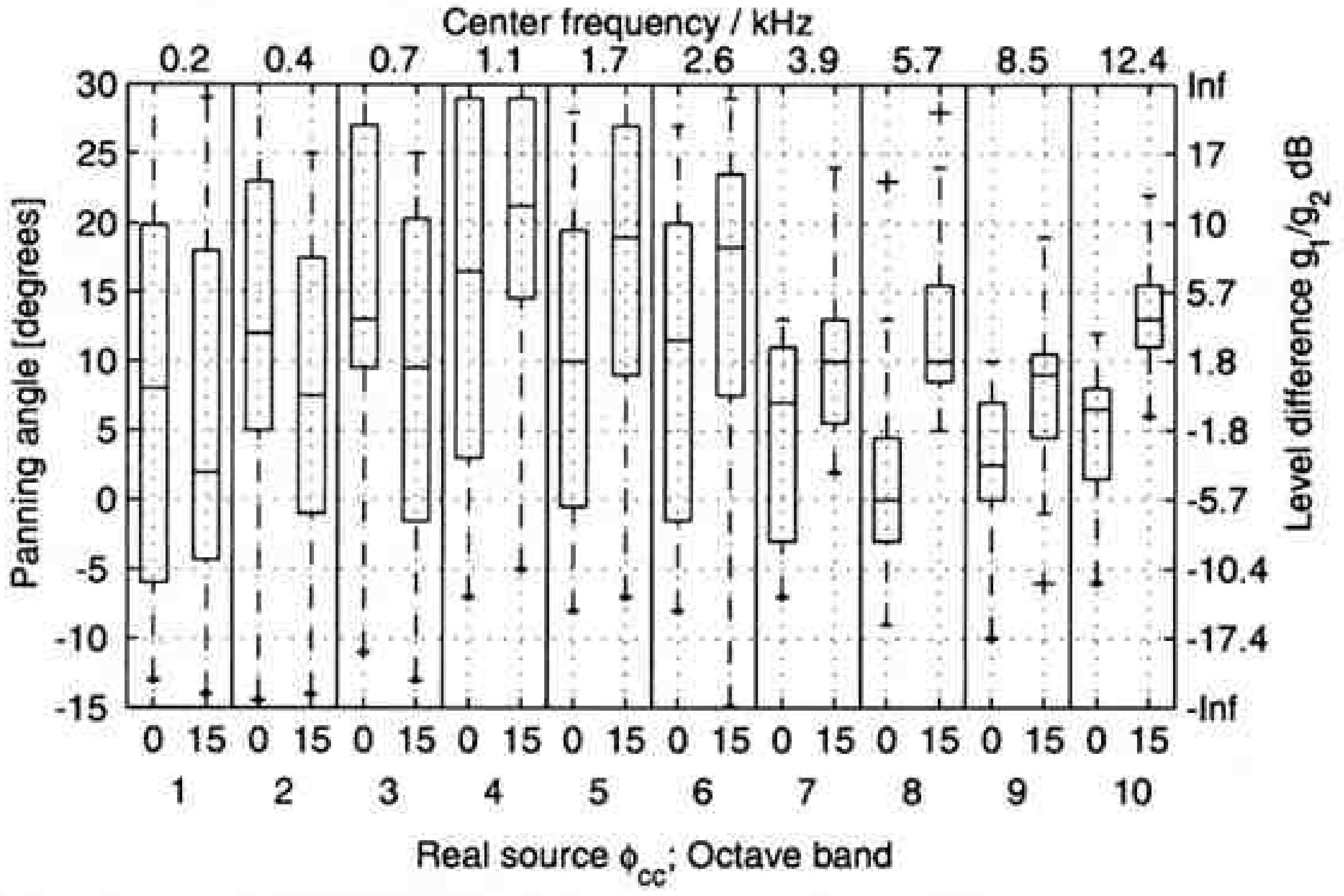


Fig. 11. Median plane listening test with octave-band noise. Panning angles judged best for each frequency channel and each reference source. Each graph subdivision denotes a frequency channel (1–10). Data for each reference source (0° or 15° of ϕ_{cc}) shown in each block. 10 listeners; adjustments conducted twice to each real source with each frequency channel; loudspeaker span 45°; setup as in Fig. 6.

Humans do not know L_0 a priori. However, if the spectrum is relatively flat near a mode frequency, the value of L at frequencies around the mode frequency can be used to estimate $L_0(f_m)$. The estimate is denoted here by $L_0^*(f_m)$. If the sound signal is known, the estimation can be completed with better accuracy. The ϕ_{cc} direction cue $q(f_m)$ can be formulated as

$$q(f_m) = \frac{L(f_m)}{L_0^*(f_m)} \tag{7}$$

which may be computed at several mode frequencies.

The behavior of $q(f_m)$ when a narrow-band signal is presented is of interest. When the frequency of a narrow-band signal is near f_m of a mode that has a maximum with elevated ϕ_{cc} values, the $q(f_m)$ of that mode would suggest elevated ϕ_{cc} directions, since the value of $L(f_m)$ is then amplified, and $L_0^*(f_m)$ is not. This agrees at least qualitatively with narrow-band localization tests conducted earlier [12], [13]. Also, when peaks or notches are introduced to a broad sound spectrum, the values of different $q(f_m)$ vary, which coincides with results of Bloom [15] and Watkins [16].

The directional cue presented explains at least partially and qualitatively the results achieved in this study. However, based on the listening test results presented in this paper the hypothesis cannot be tested further. The colored spectrum of the virtual sources and the different virtual source spectra in each ear of the subject generate too many degrees of freedom. Tests specially designed for this hypothesis should be conducted, such as listening tests in which the spectral magnitude of stimuli would be changed at the modal frequencies of the pinna.

6 EXPERIMENTAL II: PANNING IN THREE DIMENSIONS

In tests with three-dimensional panning the localization of virtual sources was studied in two different loudspeaker triplets (see Fig. 13). Eight subjects judged virtual sources for the same reference real sources four times. Pink noise was used as the stimulus. The results are analyzed with respect to the θ_{cc} perception. The perception of the ϕ_{cc} direction is not simulated, since it was found in the previous sections that decoding mechanisms are not known well enough.

6.1 Triplet 1

Triplet 1 consisted of a stereophonic setup declined to $-17°$ of ϕ_{cc} and one elevated loudspeaker in the median plane, as shown in Fig. 13(a). The loudspeakers used as reference real sources were located in the median plane at ϕ_{cc} angles of 0° and 15°. The test results are shown as a boxpot in Fig. 14. The θ_{cc} panning angles have been judged quite accurately for 0°. The ϕ_{cc} panning angles have wide data ranges. However, their median values correspond roughly to reference real source ϕ_{cc} values.

The source of the wide data ranges of ϕ_{cc} panning angles is now studied. The best matched panning angles for four subjects and both reference real sources are presented in Fig. 15. The virtual source ϕ_{cc} panning angles have been judged individually and quite similarly in different repetitions. For example, to match to a real source at 15° of ϕ_{cc}, subjects HJ and PA panned the

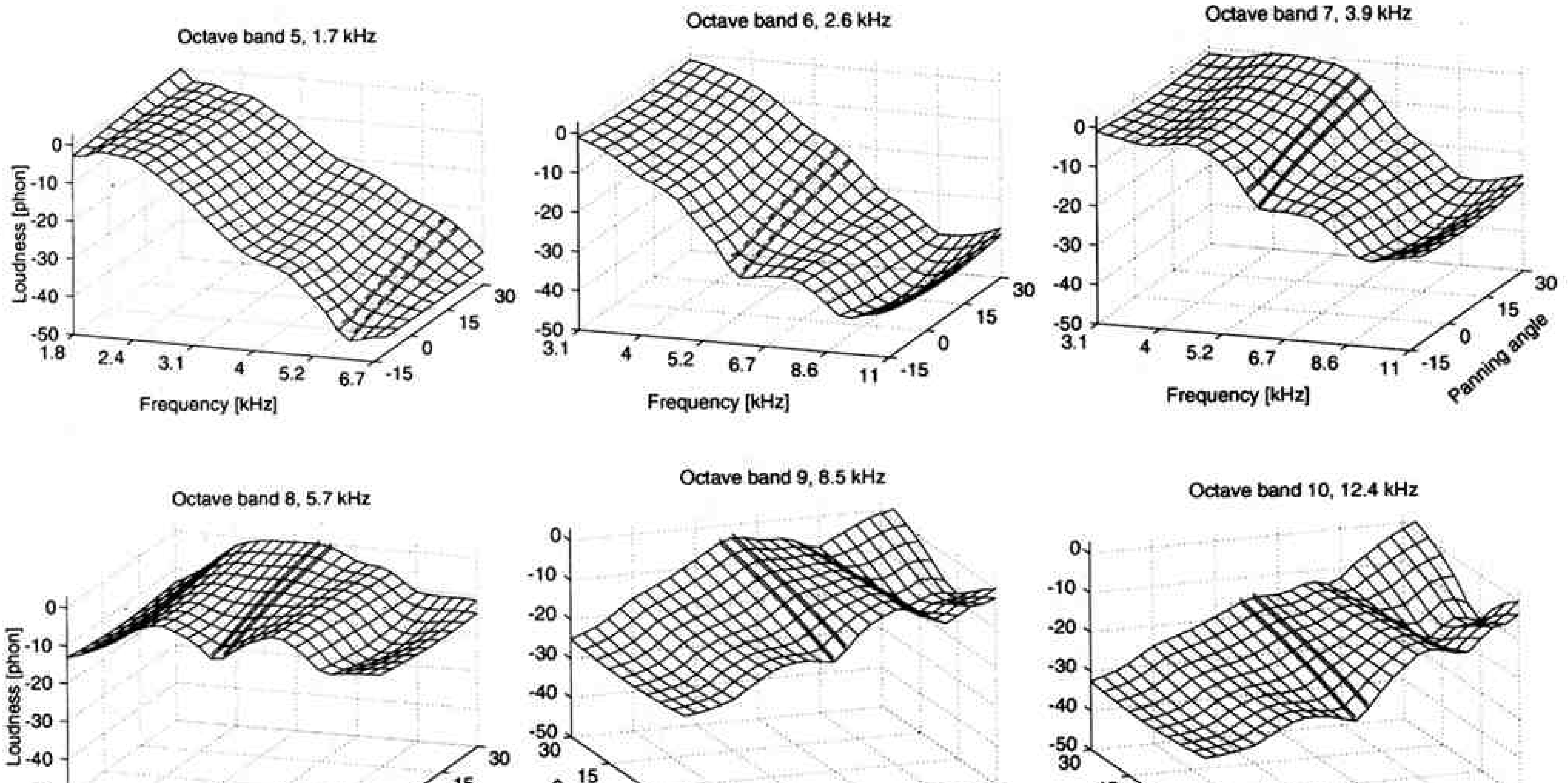


Fig. 12. Directional loudness level spectra of virtual sources at panning angles of $-15°$ to 30° of ϕ_{cc}. Octave band-passed pink noise bands 5–10. Left ear, subject HJ in median plane listening test. Highest magnitude peak of spectra is scaled to 0 phons. A spectral characteristic that may have affected perception of the ϕ_{cc} direction is shown with dashed double lines for octave bands where subjects did not perceive virtual sources consistently (bands 5 and 6), and with solid lines when they perceived them consistently (bands 7–10).

virtual sources to almost 30° of ϕ_{cc}, whereas subject VV favored ϕ_{cc} panning angles closer to 0°. Subject NP's panning angles are very near the reference real source directions. The results of the four subjects that are not shown are similar to those presented.

The ITD and ILD values are zero for reference real sources in the median plane. Due to the symmetrical setup, the same cue values should be obtained when the θ_{cc} panning angle is 0°. In the results the θ_{cc} panning angle deviated slightly from this predicted value, which may be caused by some nonideal factor in the listening setup.

6.2 Triplet 2

Triplet 2 was formed by two loudspeakers in the median plane at ϕ_{cc} values of $-15°$ and 30° and one loudspeaker to the left of the median plane, as shown in Fig. 13(b). The loudspeakers used as reference real sources were mounted at horizontal plane θ_{cc} directions of $-20°$ and $-10°$.

The best matched panning angles for each reference real source are presented in Fig. 16 for four subjects. As in previous tests, the subjects have judged the θ_{cc} panning angles values consistently, and the ϕ_{cc} panning angle values individually but quite consistently for each subject. The results for the other four subjects were similar to the results shown.

In Fig. 17 the results are presented as box plots. There exists a tendency that the virtual sources adjusted to the $-10°$ reference source θ_{cc} direction have higher ϕ_{cc} panning angles when compared with panning angle values corresponding to the $-20°$ reference source. The reason for this phenomenon is not known.

The θ_{cc} panning angles are biased further away from the median plane when compared with the corresponding reference source directions. This implies that the perceived direction of the virtual source is biased toward the median plane from the panning direction. The ITDA and ILDA cues were simulated for subjects VP and TL, since HRTFs were available only for them. The ITDA values match accurately with the reference real source near 700 Hz (see Fig. 18). It seems that the subjects have adjusted the virtual source ITD around this band to agree with the reference real source ITD. This is similar to the performance found in tests with broadband noise reported in a companion paper [3]. The ILD

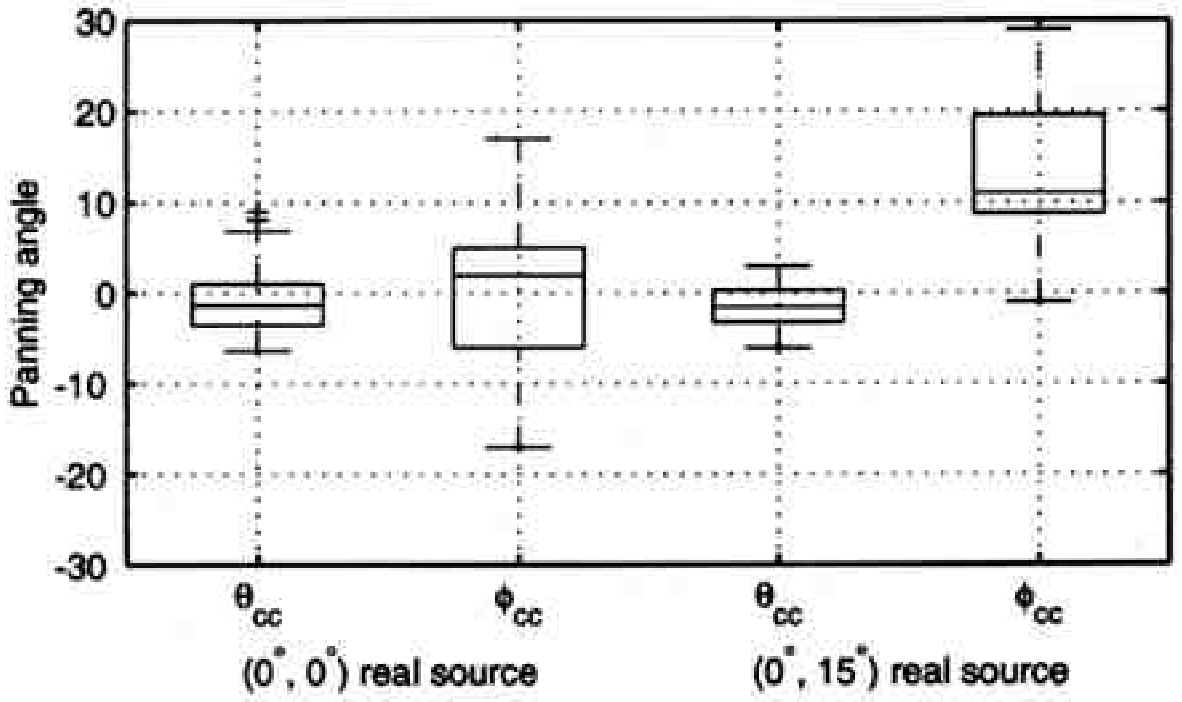


Fig. 14. Panning angles judged best in triplet 1 for reference real sources in directions (0°, 0°) and (0°, 15°). Pink noise; 8 listeners; adjustments conducted four times to each real source; loudspeaker setup as in Fig. 13(a).

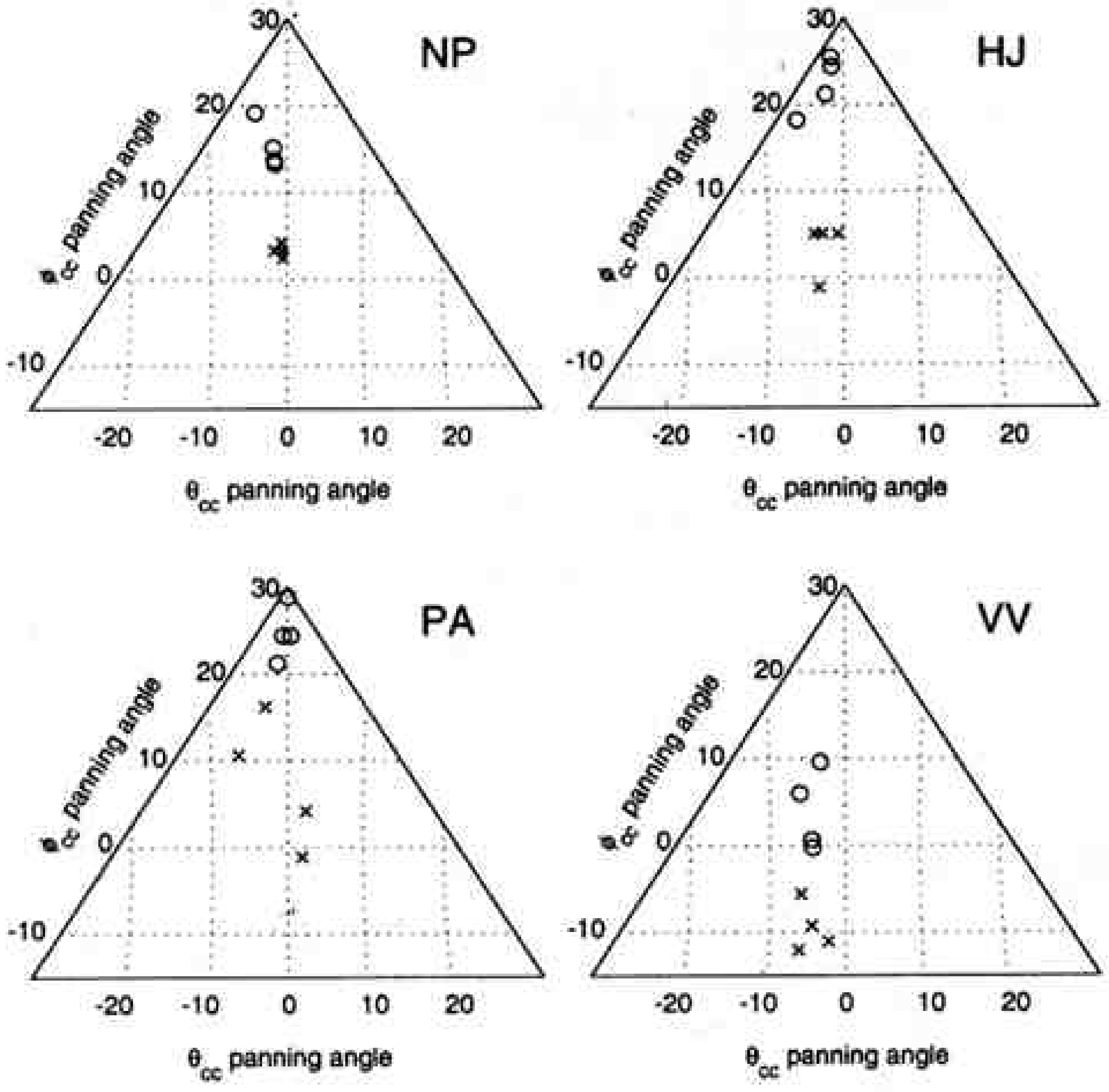


Fig. 15. Panning directions matched best for triplet 1 with reference real sources. × —directions (0°, 0°); ○—directions (0°, 15°). Pink noise; 8 listeners (4 shown), adjustments conducted four times to each real source; loudspeaker setup as in Fig. 13(a).

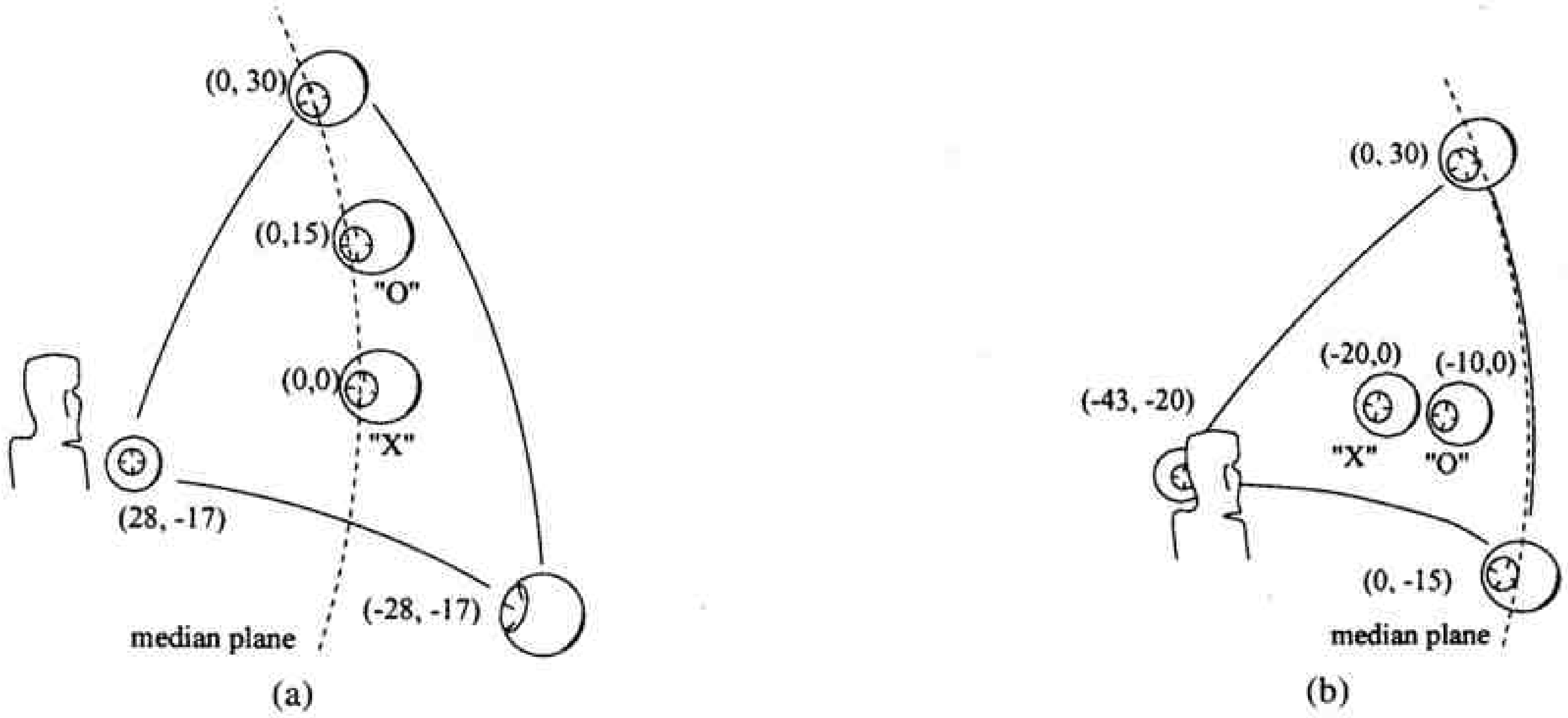


Fig. 13. (a) Triplet 1. (b) Triplet 2. Directions expressed as (θ_{cc}, ϕ_{cc}) pairs.

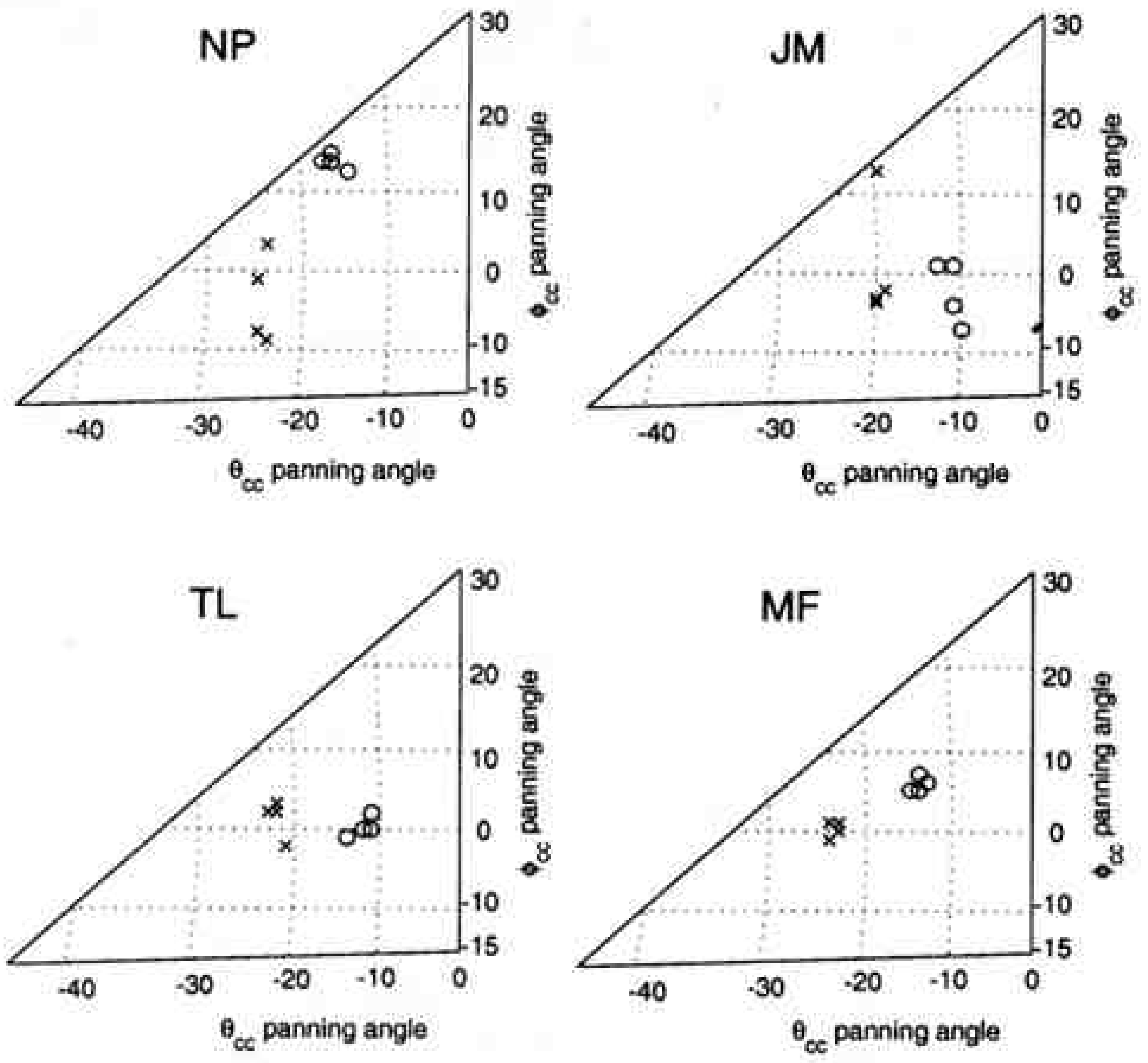


Fig. 16. Panning directions matched best in three dimensional panning tests for triplet 2 with a −20° reference real source at (−20°, 0°) (×) and a −10° reference real source at (−10°, 0°) (○). Pink noise; 8 listeners (4 shown); adjustments conducted four times to each real source; loudspeaker setup as in Fig. 13(b).

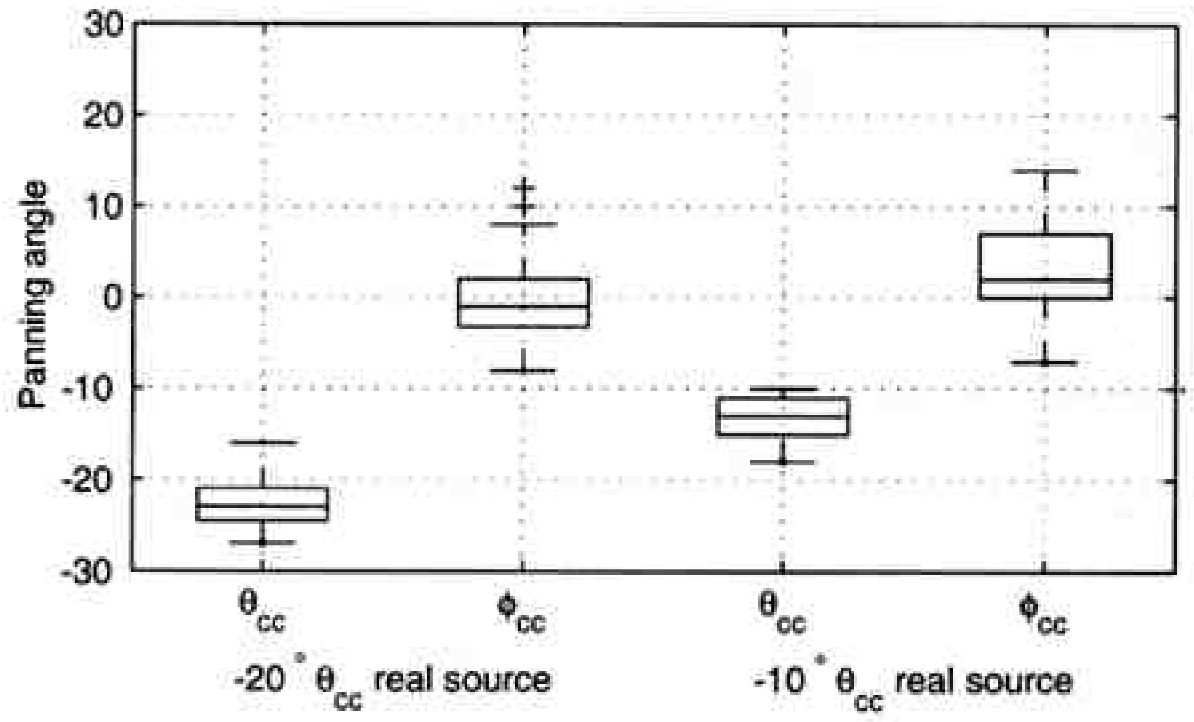


Fig. 17. Panning angles judged best in triplet 2 for reference real sources in directions (−20°, 0°) and (−10°, 0°). Pink noise; 8 listeners; adjustments conducted four times to each real source; loudspeaker setup as in Fig. 13(b).

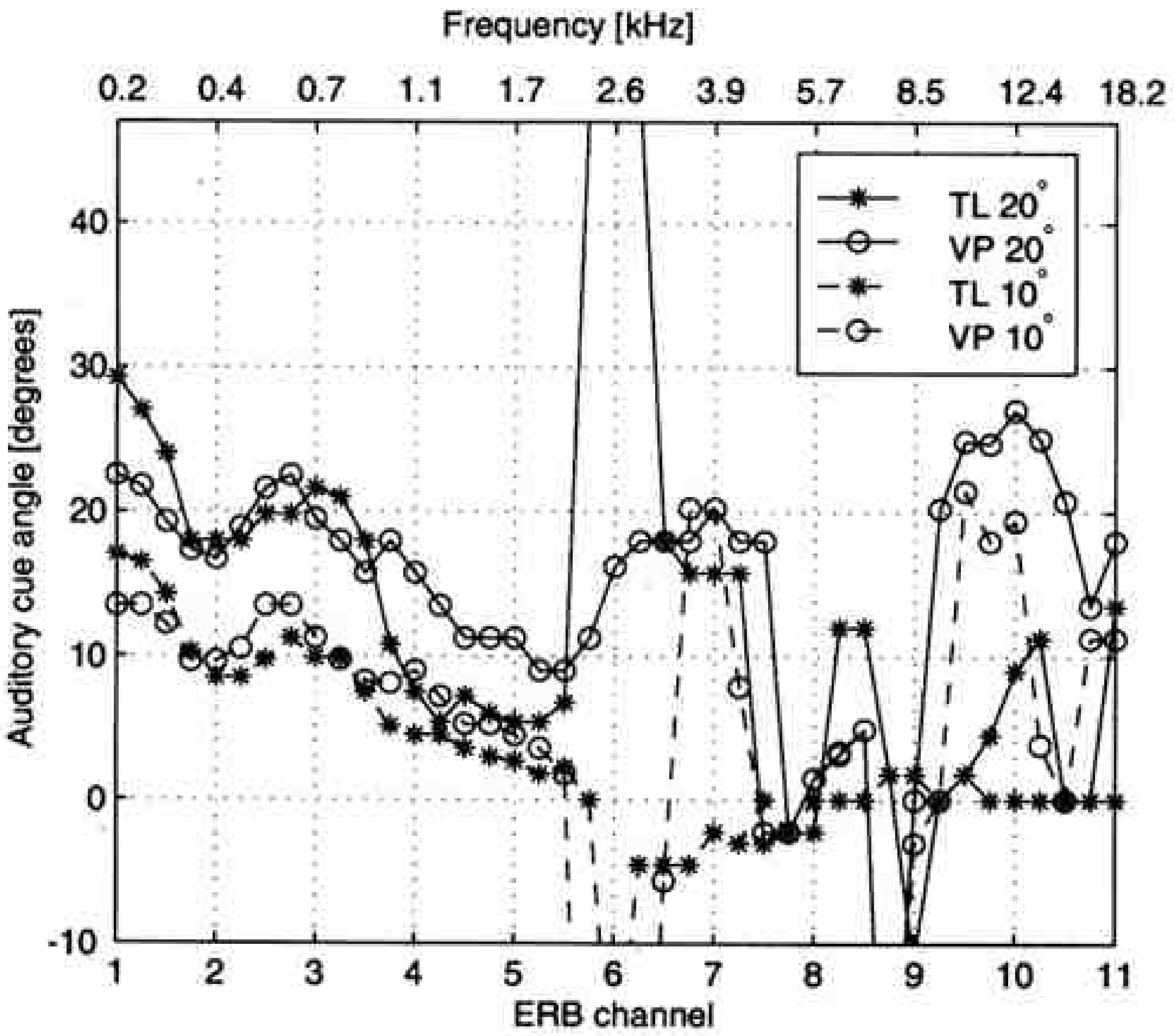


Fig. 18. Triplet 2, ITDAs of virtual sources that subjects TL and VP matched best for reference real sources with pink noise.

cues and the high-frequency ITD cues are unstable. ILD cues are not shown.

6.3 Discussion of Triplet Tests

The results suggest that virtual sources can be positioned inside a loudspeaker triplet at least in the present setups. The subjects adjusted the θ_{cc} panning angles of virtual sources consistently with each other and the ϕ_{cc} panning angles very individually. This seems to be a superposition of earlier test results: the θ_{cc} direction is perceived similarly as it is perceived with a loudspeaker pair in the horizontal plane, and ϕ_{cc} is perceived as with a loudspeaker pair in the median plane. Since the loudspeaker triplet listening tests were conducted with only two different setups, generalizations of the results are searched for within the next simulations described next.

7 SIMULATED LOCALIZATION OF BROAD-BAND VIRTUAL SOURCES

To investigate the perception of the θ_{cc} direction with different setups, model-based simulations with three triplets and two pairs were conducted. The loudspeaker sets and panning directions inside each set are shown in Fig. 19. The simulations were conducted with sets in four directions with respect to the listener. The centroids of the sets were set to θ_{cc} directions of 0°, 30°, 60°, and 90°.

The signals appearing at the subject's ears were simulated with measured KEMAR HRTFs [25]. The ITDAs and ILDAs were simulated with the auditory model described in Section 2. The deviations between the panning angles and the resulting ITDAs and ILDAs were calculated in all cases. The ITDA deviations below 1000 Hz and the ILDA deviations at all frequencies are shown in Fig. 20. The ITDA deviations above 1000 Hz are not presented since they behave mostly erroneously and in an unstable manner.

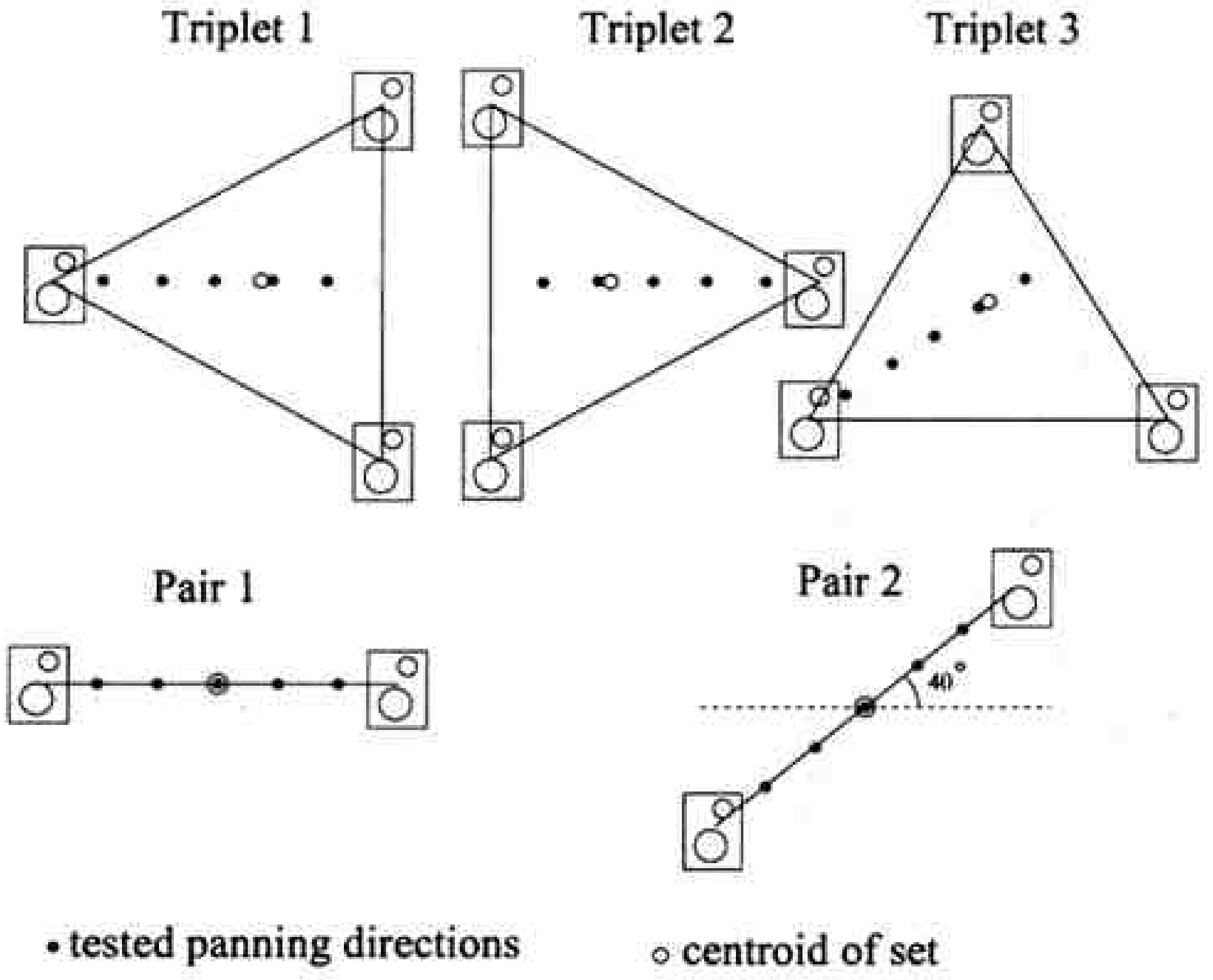


Fig. 19. Loudspeaker sets used in simulation of virtual source θ_{cc} direction perception. Each solid line between two loudspeakers spans a 60° angle from the listener's viewpoint.

The ITDA simulations are discussed. When the centroid of a panning set is at 0° of θ_{cc}, the ITDA deviations are small regardless of the loudspeaker set. This could be assumed for pairs based on results in the companion paper, but it is a new result for triplets. The spread of triplet deviations is slightly larger, though. However, this suggests that the panning angle describes reasonably accurately the perception of the virtual source θ_{cc} direction also in triplet panning.

When the centroid of a set is at 90° of θ_{cc}, the loudspeakers of that set are all at 60° of θ_{cc}, that is, they are thus in the same cone of confusion. Theoretically it was proposed by Eq. (6) that the θ_{cc} directions of the virtual sources be between the θ_{cc} values of the loudspeakers. Thus the perceived θ_{cc} direction of the virtual sources should in these cases have a value near 60°, which should create a prominent negative deviation. Indeed, the average deviation is on the order of $-25°$ and follows this proposition. This seems to be valid with both pairs and triplets.

With a set centroid direction of 0°, the negative bias phenomenon does not exist. When the set centroid direction increases, the negative bias increases as well. This explains the listening test results of Theile and Plenge [26], where it was found that virtual sources are localized nearer the median plane when a loudspeaker pair is moved toward the side of the subject. According to this simulation the same phenomenon is also present with triplets, which explains the source of bias of the panning angle in the listening test with triplet 2 (Section 6.2).

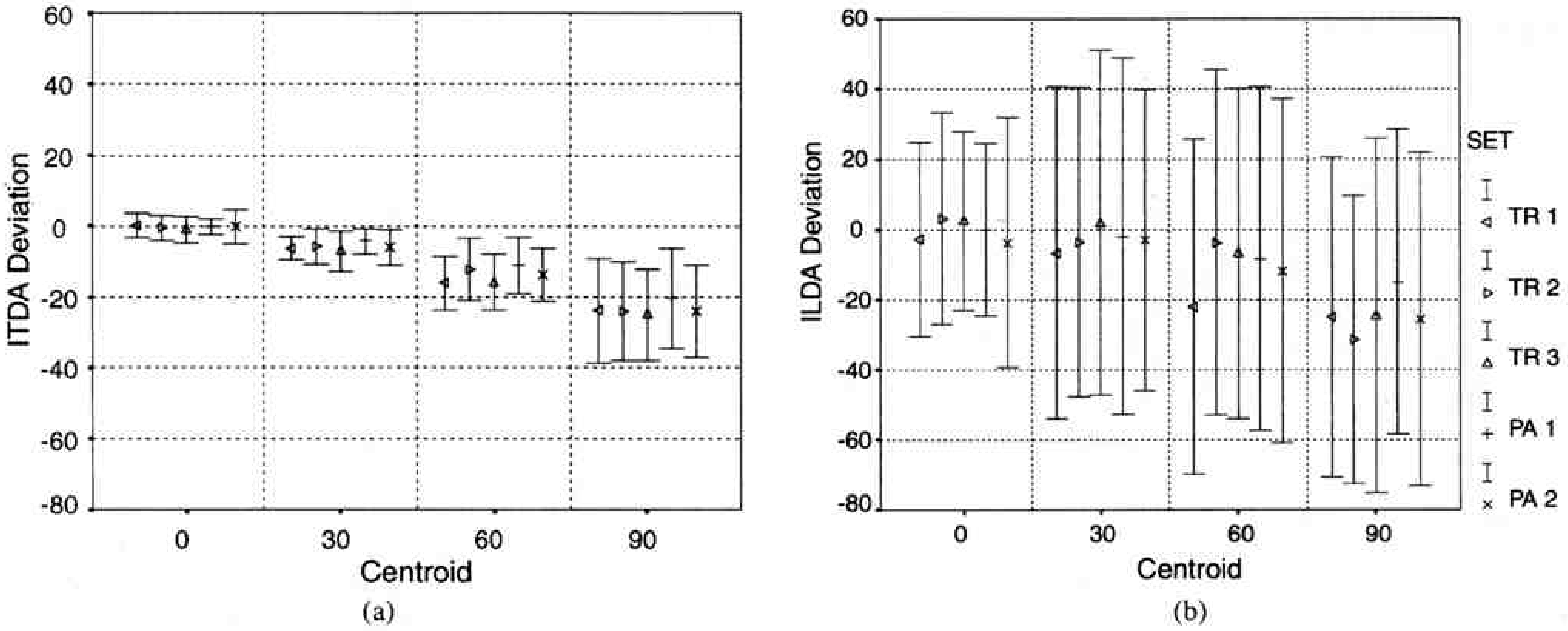
The spread of ITDA deviations increases also with the set centroid θ_{cc} direction. To explore this, the ITDA deviations were plotted with each loudspeaker set direction with respect to frequency and panning angle. The plots (not shown here) implied that the deviation is not related to frequency, but the effect of the panning angle is prominent. This suggests that the ITD cues of virtual sources are quite consistent, but are biased from the panning angle.

The ILDA deviations also have more negative values with increasing direction of the set centroid θ_{cc} direction. However, the ILDA deviations are spread out widely for all centroid directions. When the source of the spread was explored, it was found that the ILDA deviation is dependent on both frequency and panning angle. This suggests that the ILDA cues of amplitude-panned virtual sources are not in general consistent. Only when the loudspeaker set is symmetric with the median plane, do the ILDA values have roughly correct values at high frequencies. However, when interpreting ILDA simulations, the fact that ILD behaves nonmonotonously with real sources at large θ_{cc} absolute values [4] has to be taken into account. The ILD does not provide relevant cues for real source localization at such directions with all frequency bands. The inconsistent values of ILDAs may thus be neglected in part in virtual source localization by humans.

In the listening tests conducted in this paper it was found that the perception of the virtual source ϕ_{cc} direction is highly individual in the median plane. Since the localization is based on similar mechanisms also in other cones of confusion, it can be assumed that the perception of the amplitude-panned virtual source ϕ_{cc} direction is individual in all directions.

8 CONCLUSIONS

The main goal of this and the companion paper [3] was to investigate the perception of the direction of amplitude-panned virtual sources in different setups, and how well panning laws describe the perception. The results achieved in earlier studies and in these two papers are now summarized.



(a) (b)

Fig. 20. θ_{cc} direction simulation with different loudspeaker sets. Sets are shown in Fig. 19. TR1–TR3 denote different triangles, with data points shown as small triangles with corresponding shapes. PA1 and PA2 denote pairs 1 and 2, with data points shown as crosses with different orientations. (a) Deviations between θ_{cc} panning angle and ITDAs below 1000 Hz (12 bands). (b) Deviations between θ_{cc} panning angle and ILDAs (42 bands). x axis—direction of a loudspeaker set centroid. A mean value is used as a measure of the centroid, and standard deviation is used as a measure of spread.

The panning direction describes quite accurately the perceived virtual source θ_{cc} direction with loudspeaker pairs and also with loudspeaker triplets when the center of a pair or a triplet is near the median plane. The low-frequency ITDA and, roughly, the high-frequency ILDA coincide with the θ_{cc} panning direction.

If a pair or a triplet is to the left or right of a subject, the virtual source is located closer to the median plane than the θ_{cc} panning angle predicts, mostly due to low-frequency ITD behavior. If the direction (90°, 0°) is inside a triplet or in a pair, there exist θ_{cc} directions where the virtual source cannot be positioned with any panning angle value. ILD cues are generally distorted heavily with pairs or triplets in lateral directions.

The ϕ_{cc} directions of amplitude-panned virtual sources are perceived individually. Only if the loudspeakers share the same ϕ_{cc} direction is the virtual source localized mostly to that ϕ_{cc} direction. In other cases ϕ_{cc} direction perception cannot be predicted accurately. However, in our listening tests most listeners could find ϕ_{cc} panning angles that created a virtual source direction perception to a reference real source. The ϕ_{cc} direction perception was also shown to be monotonic with the ϕ_{cc} panning angle.

When the spectral contents of the best matched virtual sources were investigated in median plane listening tests, it was found that the peaks or notches of the virtual source spectra did not correspond to the peaks or notches of the corresponding reference real source spectra. This suggests that the localization cues based on decoding the frequencies of spectral notches or peaks are not sufficient to explain all phenomena of directional hearing in the median plane. A new cue type was hypothesized based on the decoding of the narrow-band loudness at pinna mode frequencies. This loudness is compared to the loudness that would appear without the pinna mode effect, which is estimated from the loudness around the mode frequency.

9 ACKNOWLEDGMENT

The author's work has been supported by the Graduate School in Electronics, Telecommunications and Automation (GETA) of the Academy of Finland and by Tekniikan Edistämissäätiö. The author would like to thank Prof. Matti Karjalainen for discussions and comments on this paper, two anonymous reviewers for suggestions and corrections, all listening test participants for their contributions, Toomas Altosaar for his help in improving the language, and Bill Gardner and Klaus Riederer for HRTF data.

10 REFERENCES

[1] L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen, "Creating Interactive Virtual Acoustic Environments," *J. Audio Eng. Soc.*, vol. 47, pp. 675–705 (1999 Sept.).

[2] H. Fouad, J. A. Ballas, and D. Brock, "An Extensible Toolkit for Creating Virtual Sonic Environments," in *Proc. Int. Conf. on Auditory Display 2000* (Atlanta, GA, 2000).

[3] V. Pulkki and M. Karjalainen, "Localization of Amplitude-Panned Virtual Sources, I: Stereophonic Panning," *J. Audio Eng. Soc.*, vol. 49, pp. 739–752 (this issue).

[4] J. Blauert, *Spatial Hearing*, rev. ed. (MIT Press, Cambridge, MA, 1997).

[5] R. H. Gilkey and T. R. Anderson, Eds., *Binaural and Spatial Hearing in Real and Virtual Environments* (Lawrence Erlbaum Assoc., Mahwah, NJ, 1997).

[6] R. O. Duda, "Elevation Dependence of the Interaural Transfer Function," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R. H. Gilkey and T. R. Anderson, Eds. (Lawrence Erlbaum Assoc., Mahwah, NJ, 1997), pp. 49–75).

[7] M. B. Gardner and R. S. Gardner, "Problem of Localization in the Median Plane: Effect of Pinnae Cavity Occlusion," *J. Acoust. Soc. Am.*, vol. 53, pp. 400–408 (1973).

[8] K. Iida, M. Yairi, and M. Morimoto, "Role of Pinna Cavities in Median Plane Localization," *J. Acoust. Soc. Am.*, vol. 103, p. 2844 (1998 May).

[9] E. A. G. Shaw and R. Teranishi, "Sound Pressure Generated in an External-Ear Replica and Real Human Ears by a Nearby Sound Source," *J. Acoust. Soc. Am.*, vol. 44, pp. 240–249 (1968).

[10] E. A. G. Shaw, "Acoustical Features of the Human External Ear," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R. H. Gilkey and T. R. Anderson, Eds. (Lawrence Erlbaum Assoc., Mahwah, NJ, 1997), pp. 25–47.

[11] V. R. Algazi and C. Avendano, "Low-Frequency ILD Elevation Cues," *J. Acoust. Soc. Am.*, vol. 106, p. 2237 (1999 Oct.).

[12] J. Blauert, "Sound Localization in the Median Plane," *Acustica*, vol. 22, pp. 205–213 (1969/1970).

[13] J. C. Middlebrooks, "Narrow-Band Sound Localization Related to External Ear Acoustics," *J. Acoust. Soc. Am.*, vol. 92, pp. 2607–2624 (1992 Nov.).

[14] J. A. Janko, T. R. Anderson, and R. H. Gilkey, "Using Neural Networks to Evaluate the Viability of Monaural and Interaural Cues for Sound Localization," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R. H. Gilkey and T. R. Anderson, Eds. (Lawrence Erlbaum Assoc., Mahwah, NJ, 1997), pp. 557–570.

[15] P. J. Bloom, "Creating Source Elevation Illusions by Spectral Manipulation," *J. Audio Eng. Soc.*, vol. 25, pp. 560–565 (1977 Sept.).

[16] A. J. Watkins, "Psychoacoustical Aspects of Synthesized Vertical Local Cues," *J. Acoust. Soc. Am.*, vol. 63, pp. 1152–1165 (1978).

[17] P. Zakarauskas and M. S. Cynader, "A Computational Theory of Spectral Cue Localization," *J. Acoust. Soc. Am.*, vol. 94, pp. 1323–1331 (1993 Sept.).

[18] J. Backman and M. Karjalainen, "Modelling of Human Directional and Spatial Hearing Using Neural Networks," in *Proc. ICASSP-93* (Minneapolis, MN, 1993), pp. I-125–I-128.

[19] H. A. M. Clark, G. F. Dutton, and P. B. Vanderlyn, "The 'Stereosonic' Recording and Reproducing System," *J. Audio Eng. Soc.*, vol. 6, no. 2, p. 102 (1958), reprinted in *Stereophonic Techniques* (Audio Engineering Society, New York, 1986).

[20] B. B. Bauer, "Phasor Analysis of Some Stereophonic Phenomena," *J. Acoust. Soc. Am.*, vol. 33, pp. 1536–1539 (1961 Nov.).

[21] J. C. Bennett, K. Barker, and F. O. Edeko, "A New Approach to the Assessment of Stereophonic Sound System Performance," *J. Audio Eng. Soc.*, vol. 33, pp. 314–321 (1985 May).

[22] V. Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning," *J. Audio Eng. Soc.*, vol. 45, pp. 456–466 (1997 June).

[23] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models* (Springer, Heidelberg, Germany, 1990).

[24] F. L. Wightman and D. J. Kistler, "Factors Affecting the Relative Salience of Sound Localization Cues," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R. H. Gilkey and T. R. Anderson, Eds. (Lawrence Erlbaum Assoc., Mahwah, NJ, 1997).

[25] B. Gardner and K. Martin, "HRTF Measurements of a KEMAR Dummy-Head Microphone," Tech. Rep. 280, MIT Media Lab Perceptual Computing, Cambridge, MA (1994).

[26] G. Theile and G. Plenge, "Localization of Lateral Phantom Sources," *J. Audio Eng. Soc. (Project Notes/Engineering Briefs)*, vol. 25, pp. 196–200 (1977 Apr.).

The biography for Ville Pulkki appears on page 751 of this issue.