Audio Engineering Society

Convention Paper 5550

Presented at the 112th Convention
2002 May 10–13 Munich, Germany

Binaural Modeling of Multiple Sound Source Perception: Coloration of Wideband Sound

Kazuho Ono[1], Ville Pulkki[2], and Matti Karjalainen[2]

[1] *NHK Science and Technical Research Laboratories,*
1-10-11 Kinuta Setagaya-ku, Tokyo 157-8510, JAPAN
ono.k-gs@nhk.or.jp

[2] *Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing,*
P. O. Box 3000, FIN-02150 HUT, FINLAND
Ville.Pulkki@hut.fi and Matti.Karjalainen@hut.fi

ABSTRACT

Binaural modeling of coloration perceived due to multiple coherent sources is studied under the condition that sounds arrive at a listener successively within a certain time delay. Our former work showed that a binaural model of timbre perception describes the coloration systematically and almost independently of the directional perception, based on listening tests using 1-Bark-width bandlimited noise and pulse trains. The present work is an extension of the study to ensure our modeling, including listening experiments with broadband noise and pulse trains. The results show in general that the binaural model still has a good agreement with listening test results for broadband signals, but in some cases clear deviation from bandlimited noise case especially at low frequencies is also found.

INTRODUCTION

This work is continuation of a study in which binaural modeling of multiple source perception is investigated [1]. The former work proposed a new methodology of coloration experiments, in which the sound signals were recorded at the entrances of ear canals simultaneously with listening tests. After the tests the responses of listeners were explained by analyzing the recorded ear canal signals with a binaural auditory model. Listening tests were conducted based on the method of adjustments, using 1-Bark-width bandlimited noise and pulse trains with less than 4 ms delay between loudspeakers. The model-based analysis also suggested that the coloration of a virtual source produced by two real sound sources could be modeled systematically and almost independently of the directional perception, related to binaural loudness as a function of critical band.

The present work extends the previous study to ensure and refine the model. First, listening experiments were done with other signals than in the former work, mainly with wideband signals. The listening test results from wideband signals were compared with those from bandlimited signals, showing that both cases behave similarly in general. However, in some specific cases they were clearly different, especially where prominent coloration existed in the bandlimited noise case at low frequencies. The methodology of recording the ear canal signal simultaneously with listening tests, proposed in the former work, was also adopted in the present work.

Secondly, based on the listening test results from the present and the former work, the validity of the model is discussed. The model calculates the Composite Loudness Level (CLL) which is a simple summation of each ear's loudness and uses it as a parameter describing binaural loudness. If this model is perfect for binaural loudness, CLL for a virtual and a single real source should be identical when listeners adjusted both loudnesses to be equal. In this case, CLL difference (defined here as CLL for single source minus CLL for virtual source) should be zero.

The model-based analysis of the signals recorded during the listening tests, in which listeners adjusted the loudness of a virtual source to be equal to that of a single real source, showed that the CLL difference was 0 ± 2 phon in most cases. The analysis also indicated that when the CLL difference showed a clear deviation from zero phon level, the deviation typically means that the coloration of virtual source was less prominent than the timbre model suggests. One example of these cases was a small hump of the CLL difference at 1.7 kHz in the stereophonic case with no delay between loudspeakers. Another example was a large hump of CLL difference in the low frequency range in the 2 ms delay case, which existed only in the wideband signal case.

1. PERCEPTION OF MULTIPLE SOUND SOURCE

1.1 Directional perception

Spatial and directional hearing have been studied intensively; for overviews, see for example [2] or [3]. The duplex theory of sound localization states that the two main cues of sound source localization are the *interaural time difference* and the *interaural level difference* which are caused respectively by the wave propagation time difference (primarily below 1.5 kHz) and the shadowing effect by the head (primarily above 1.5 kHz). The auditory system decodes the cues in a frequency-dependent manner. The cues resolve in which cone of confusion the sound source lies. A cone of confusion can be approximated by a cone having an axis of symmetry along a line passing through the listener's ears and the apex in center point on the line between ears.

In anechoic conditions the arrival times of sound from any direction to the ear canals is within 1 ms. In a real room, however, the time delays of reflections from walls usually exceed 1 ms and then the precedence effect should be taken into account to evaluate the perceptual quality. There exist many experimental studies on the precedence effect [3]. The perceived direction is reported to depend generally on the time delays of primary and secondary sound, but there are many other factors which affect the behavior of the precedence effect. Transient components of sound signals play an important role. A stable precedence effect requires the bandwidth of sound signal to be wide enough [4]. In recent years, the adaptive aspects and high-level processes in the precedence effect have gained increased interest.

1.2 Timbre perception

Timbre is defined by American Standards Association [5] as "that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar". Timbre is the attribute in which we perceive loudness-frequency characteristics, in other words, subjective amplitude-frequency characteristics. Since these are expected to change owing to the difference in arrival time at the entrance of ear canals in multiple source perception, it is interesting to study how these physical changes in sound will affect the perception of timbre.

In summing localization in two-channel stereophony, the idea of binaural suppression [3] has been studied to explain our experience that the coloration of a phantom source is not as prominently perceived as it is expected physically from the comb filter effect. Theile proposes an "Association Model" [6] stating that when our perception of localization is clearer, our perception of timbre is suppressed and is not sensitive to the change of amplitude-frequency characteristics, based on the listening tests using dummy head recordings. This model indeed explains why we do not perceive coloration of a virtual source in 2-channel stereophony so well, but it has not yet been verified in relation to the ear canal signal, and there has been opponents to this model [7].

In the present work, timbre perception of a virtual source created by a multiple source and its relation to the ear canal signals are of interest.

1.3 Binaural Loudness

The sound field reproduced by a multiple source is usually asymmetric to human head unless both the loudspeaker setup and the listening conditions are symmetric. It is hence necessary to consider binaural loudness in the binaural modeling of coloration. Binaural loudness has been studied mainly with stationary sounds. It is reported to depend on binaural level difference [8], and the phase difference [9]. In the present work, binaural loudness under the precedence effect condition is of interest.

2. BINAURAL AUDITORY MODEL

2.1 Composite Loudness Level (CLL)

In the present study, the perceived timbre is modeled as a binaural loudness level spectrum over frequency channels. A schematic diagram for the binaural model of neural decoding is presented in Fig. 1. It takes as input the sound signal arriving to the ear canals and computes the decoded frequency-dependent loudness level spectra. It models the middle ear, the cochlea, and the auditory nerve. The model is explained in more detail in [10] and [11].

The middle ear, cochlea, and auditory nerve models have been implemented based on the HUTear 2.0 software package [12]. The middle ear is modeled using a filter that approximates a response function derived from the minimum audible pressure curve [13]. The cochlear filtering of the inner ear is modeled using a 42-band gammatone filter bank [14]. Center frequencies of the filter bank follow the ERB (equivalent rectangular bandwidth) scale [15]. The outputs from the auditory models are interpreted and compared with the listening test results.

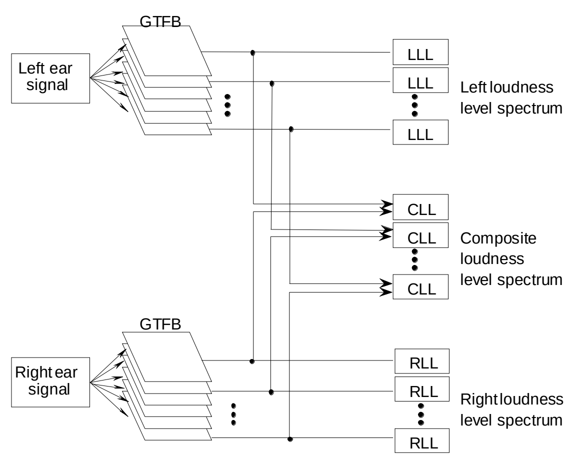

Fig. 1 Modeling of right-ear, left-ear, and composite loudness level spectra (RLL, LLL and CLL spectra). LL denotes loudness level.

The loudness of each frequency band in each ear is computed using Zwicker's formulae [16]. Due to simplicity, this model is used instead of the more thorough model proposed in [17]. From the left-ear loudness level spectrum (LLL) and right-ear loudness level spectrum (RLL), composite loudness level spectrum (CLL) is computed by summing the loudnesses of each ear at each frequency band.

2.2 Experimental results using narrowband signals [1]

In the former work, a binaural auditory model which calculates CLL as binaural loudness was proposed with the hypothesis that loudness perception can be modeled separately from directional perception. Listening tests were conducted to ensure this modeling and it was found that CLL describes binaural loudness well for narrowband noise and pulse trains approximately below 3 kHz where the head shadowing effect due to HRTFs is not prominent and the interaural difference of sound pressure is not high.

3. LISTENING TESTS

3.1 Listening test setups

In the present work, listening tests with wideband signals were conducted instead of narrowband signals that were used in the former work. All tests were conducted in an anechoic room with two loudspeakers reproducing the same sound signal with a controlled time delay from one loudspeaker to another, which was same as in the former work. The subjects adjusted the test signal in subbands just like using a graphic equalizer, band by band and iteratively, until it sounded maximally similar in timbre with the reference sound.

A methodology proposed in the former work, based on simultaneous recording of sound signals at the entrances of the ear canals, was also adopted. These recorded ear canal signals were used as input to auditory modeling. This methodology perfectly enables comparison between subjective listening test results and corresponding outputs of the auditory model, in which way the individuality of HRTFs and small variations of listening position are effectively taken into account. The details of listening test setups are discussed below.

3.1.1 Coloration evaluation

In the tests of this study the subjects adjusted the loudness level of each band of a wideband virtual source created by two real sources by using level control to match as well as possible in timbre (and level) with a single real reference source.

The reference single source was presented first, followed by the multiple source virtual sound, as shown in Fig. 3, and this was repeated until the timbre match was done. The sound signal used for reference and that composing the virtual source were identical except for a delay between loudspeakers. The sound signals were wideband noises for Listening test 1, 1a, and 2, and wideband pulse trains for Listening test 3. The bandwidth of sound signal was 6 Barks for all test signals, which was found as an upper limit of bandwidth to get stable adjustment results from experienced listeners (See also 3.2). To get enough bandwidth, several 6-Bark bandwidth signals that had adjacent frequency bands but no overlap were used, instead of one sound signal having wider than a 6-Bark bandwidth.

A preliminary experiment also showed that for non-experienced listeners even 6 Bark bands were too wide to be adjusted to get consistent results. Hence they did not attend the listening tests in the present work. In general these experiments were found very demanding, but the subjects used in the final study reported that it was always possible to find a timbre match that was satisfactory.

The magnitude spectrum characteristics of the reference signal were always kept flat over the 6-Bark bandwidth. The magnitude spectrum of the test signal that composed the virtual source was adjusted so that each time the subject pressed one key in a column of vertically adjacent four keys of a computer keyboard, for example, '1','Q','A' or 'Z', the computer changed the level of the corresponding component within a 1 Bark bandwidth. Among the four keys, two keys (in this example 'Q'/'A') were allocated for the change of +/- 1dB, and other two keys ('1'/'Z') were allocated for +/- 3 dB adjustment. The four-key columns were aligned horizontally adjacent so that the lowest Bark band corresponded to the keys '1','Q','A', and 'Z', and the highest band to the keys '6','Y','H', and 'N'. The sequence was repeated until the subject pressed 'enter' key to denote that he had selected the best matching level for each band. The reason for using a computer keyboard instead of a slider board of a graphic equalizer was to minimize visual bias in the results.

The loudspeakers were in front of the listener at azimuth directions θ_1, θ_2 in the *horizontal* plane, as shown in Fig. 2. The sound pressure levels from both loudspeakers, generating the virtual source, were equal. The listener was requested to keep his/her head immobile. As a result a set of gain adjustment coefficients as a function of frequency channel were obtained. If a virtual source were not colored, the listeners' judgments would be constant. Colorations of virtual sources are shown as higher or lower adjustment values depending on frequency channel, corresponding to dips or peaks in virtual source coloration spectra.

3.1.2 Simultaneous recording of ear signals

In the former work we proposed a new methodology for evaluating the binaural modeling of multiple source perception. It is based on recording the signals entering the ear canals simultaneously with the listening test, which makes it possible to analyze these signals that correspond to the listening condition and to be free from the problem of individuality of HRTFs or that of minor variation in listening position. The ear canal signals were recorded using small electret microphones set close to the entrance of open ear canals of the subjects. The details are shown in [1].

The recorded sound was windowed by a rectangular window for each reference single source and test virtual source sound respectively to pick out the corresponding section for both the reference source and the virtual source. The windowed sections were used as the inputs of the auditory model.

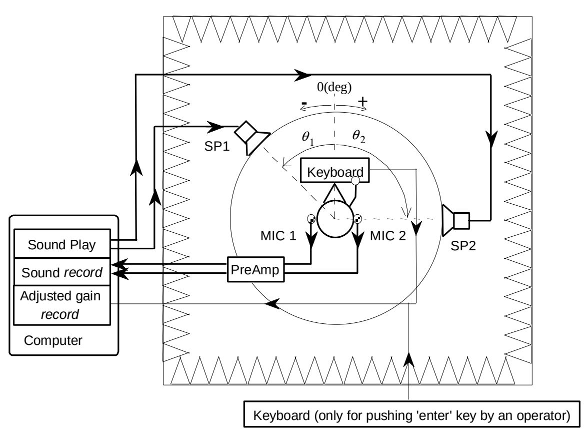

Fig. 2: Experimental setup for listening tests in an anechoic chamber at HUT Acoustics Laboratory.

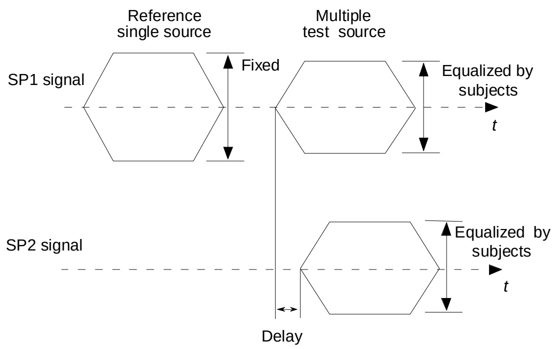

Fig. 3: Sequence of sound presentation in timbre adjustment experiments. The sequence was repeated until the subject had found best match between the reference source and the multiple source test sound. The duration of each sound was 700 ms (400ms only for Listening test 3) and the pauses between them were approximately 200 ms.

In the auditory model all outputs from the auditory filter bank should in principle be used to compute the total loudness, but in this study, in order to get a higher signal to noise ratio, only the outputs of filter banks whose center frequencies were within the bandwidth of test signals were used to calculate loudness values. Although the absolute values of total loudness calculated by this ‘equivalent’ filtering is not exactly that of the input signal from microphones, the values are expected to be valid as long as relative values, for example reference vs. test, are concerned.

3.2 Listening test 1

A listening test for wideband noise was conducted first. This experiment corresponds to the experiment 2 in the former work [1], with the same experimental conditions except that the sound signal used was wideband noise, instead of narrowband noise in the previous study.

Two loudspeakers were used in two different setups, which are the stereophonic setup (θ_1=–30°, θ_2=+30°) and a front-side setup (θ_1=0°, θ_2=+90°). The test was conducted in an anechoic chamber with two GENELEC 2029A loudspeakers located at 2.2 meters distance from the listener for stereophonic setup and 1.6 meters distance for front-side setup. The SPL of the single reference source was 62 dB, measured at the listening position. The test attendees were 3 males with normal hearing and they repeated each condition twice. All of them were experienced listeners.

Delay values used were 0 ms and 2 ms for the stereophonic setup, and 2 ms for the front-side setup. The delay was added to the right hand loudspeaker (located at +30°) for stereophonic setup and the right hand side loudspeaker (located at +90°) for front-side setup. Frequency bands used covered 12 Barks, being 2…13 Barks for the stereophonic setup with 0 ms delay and 1…12 Barks for the stereophonic and front-side setups with 2 ms delay. The 12 Bark band was covered with two 6-Bark bandwidth signals, as described in 3.1.1. These conditions were chosen to cover the frequency range where prominent coloration was found in the narrowband signals of the former work, which cases were around 1.7 kHz in the stereophonic setup without delay and around 200 Hz in both setups with 2 ms delay.

Through a preliminary experiment that was conducted with experienced listeners as well as non-experienced ones it was found that the task to ‘equalize’ the timbre of a multiple source was harder than the task to adjust the loudness of a narrowband signal. Experienced listeners always reported that they were successful in adjusting, and results were consistent although they spent a relatively long time (approx. 10 min. on average) and the deviations were not very small. Non-expert listeners found the task so hard that their adjusted values sometimes diverged to an extreme value, for example, different by more than 20 phon from the averaged data of expert listeners. Hence it was not possible to simply compare the data from expert listeners and non-expert listeners, and further experiments by non-expert listeners were not conducted.

The noise signals used were noises whose bandwidth equaled to 6 Bark. They were generated by inverse FFT from a frequency domain signal whose amplitude-frequency characteristics were flat (for reference single source) or followed the gain data adjusted by listeners (for test virtual source), and whose phase characteristics were randomized. The reference source was presented from the left loudspeaker (located at –30°) for stereophonic setup and from the front loudspeaker (located at 0°) for front-side setup.

The perceived directions of the virtual sources were expected to be different from the reference single source, but the subjects were instructed to concentrate on loudness evaluation only. They were not required to answer the perceived direction of the sounds but just to comment the perceived direction voluntarily.

Fig. 4 shows the calculated results of CLL (Composite Loudness Level) difference in thick line, plotted with average values of adjusted gain in thin line, for the case of stereophonic setup with 0 ms delay, which corresponds to the coloration of conventional stereophonic playback. The CLL difference stands for reference single source *minus* test virtual source CLL difference. If this were zero in all cases, it would be a perfect parameter to describe binaural loudness and timbre perception.

It should be noticed that this figure shows merged results for 12-Bark bandwidth obtained from two experimental data each of which using 6-Bark bandwidth sound signals. It should also be commented that the CLL calculation was based on filterbanks using ERB channels, while the sound signals used and the way of adjustment was based on Bark channels.

From Fig. 4 it is easily seen that the adjusted gain has a hump near 1.7 kHz. This hump existed in the results of narrowband signal cases in the former work [1], and also in the work where the coloration of an amplitude-panned virtual source was studied [18].

On the other hand, CLL difference values are stable within 0 ± 2 phon. There is a small hump around 1.7 kHz. These tendencies in CLL difference were also found in the results from narrowband signal cases including the small hump.

Fig. 5 and Fig. 6 show the calculated results of CLL difference for the cases of stereophonic setup and front-side setup with 2 ms delay, respectively. (The adjusted gain is plotted again by a thinner line.) In these cases, CLL shows a prominent hump around 200 Hz. This is clearly different from the results obtained in narrowband case, where CLL difference was stable around zero.

3.3 Listening test 1a

The Listening test 1 for wideband noise, as described above, showed clear difference from the narrowband noise case. Since it was not possible to conduct the experiment for whole 12-Bark bandwidth at one time due to the difficulty of adjustment, experiments were divided into two bands, lower band and higher band, as mentioned before. All result plots in Listening test 1 were hence ‘joined’ plots of two bands.

In this test, the subjects reported that adjustments were much easier at the lowest and highest bands, i.e., at band edges. One possible reason for this would be that at band edges less masking effect is expected because one of the two adjacent band component does not exist. Listening test 1a was conducted just to ensure if this 'end' effect affects the results or not. The test was conducted only for the stereophonic setup and the front-side setup with 2 ms delay (Fig. 5 and Fig. 6) under the same conditions as Listening test 1, except for a frequency range from 3 Barks to 8 Barks. This range was selected in order to have the low-end frequency roughly equal to the peak in CLL difference (see Fig. 5 and Fig. 6) at 3 Bark.

Fig. 7 and Fig. 8 illustrate the results. It is easily seen that the CLL difference was similar to the results in Fig. 5 and Fig. 6 within the corresponding frequency range, and a peak in CLL difference at 3 Barks was clearly found even if it was located at the low end of the band.

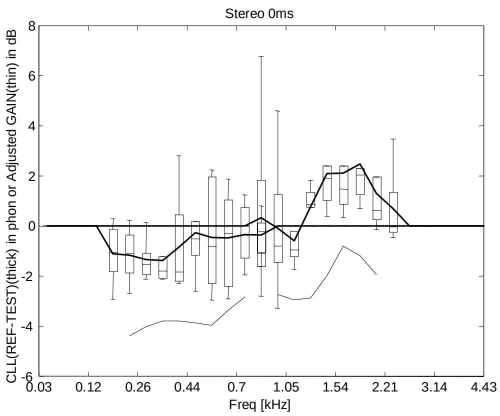

Fig. 4. Composite loudness level (CLL) difference for wideband noise. Stereophonic setup is used, no delay between channels. Adjusted gain is plotted by thin line (in dB).

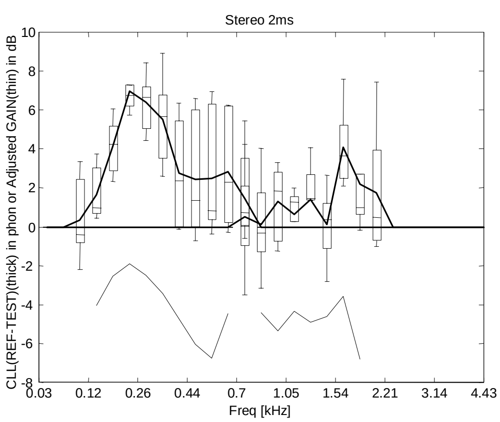

Fig. 5. CLL difference for wideband noise. Stereophonic setup is used, 2 ms delay between channels. Adjusted gain is plotted by thin line (in dB).

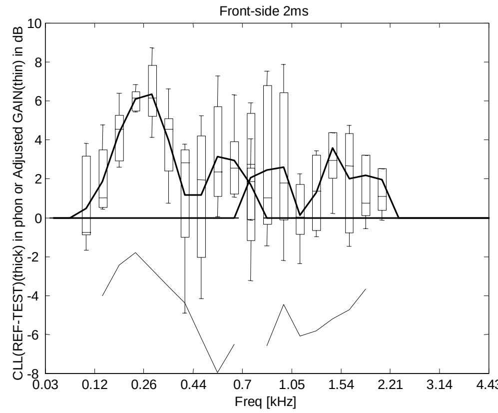

Fig. 6 CLL difference for wideband noise. Front-side setup is used, 2 ms delay between channels. Adjusted gain is plotted by thin line (in dB).

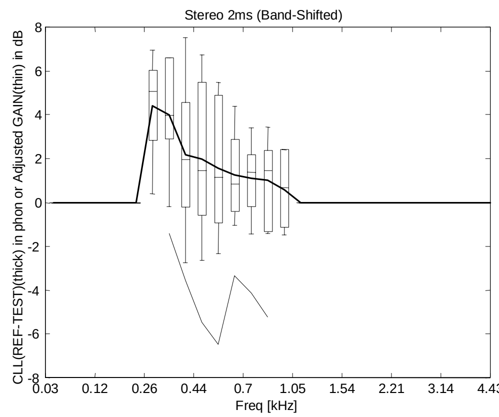

Fig. 7. CLL difference for wideband noise. Stereophonic setup is used, 2 ms delay between channels. Adjusted gain is plotted by thin line (in dB).

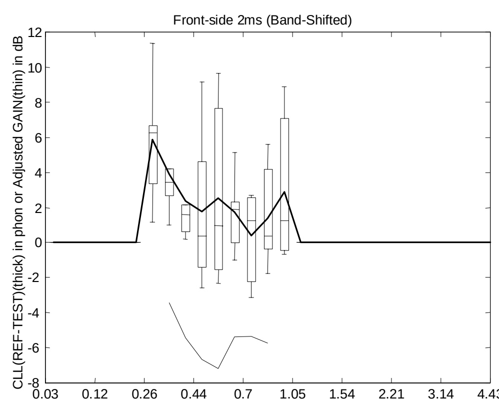

Fig. 8. CLL difference for wideband noise. Front-side setup is used, 2 ms delay between channels. Adjusted gain is plotted by thin line (in dB).

3.4 Listening test 2

In Listening test 1, where the conventional stereophonic case with no delay was studied, the CLL difference had a hump (positive deviation from zero) by 2 phon around 1.7 kHz. This hump is not so prominent, but it is remarkable that this hump existed also with the stereophonic setup with no delay in the results of the narrowband case of the former work. These humps mean that CLL for the reference single source was larger than CLL for the test multiple source, even though subjects adjusted so that both sources were perceived equally loud. This is, in other words, a hump in adjusted gain is not so prominent than the timbre model based on CLL suggests.

There were several possible reasons for this hump. Listening test 2 deals with one of them, which is the difference of the sound direction between the reference and the test. In all experiments for the stereophonic setup above, the reference single source was the left loudspeaker (located at -30°) which was one of the two single sources composing the virtual source, while the direction of the virtual source was from center. This test focused on that difference in direction through the two following listening tests.

The first one was to find out if this hump occurs between the single source and the virtual source when both sources are in the same direction. This experiment was conducted by placing the reference single source in center, which was expected to be the same direction as the virtual source (phantom center image).

Fig. 9 depicts the calculated results of CLL difference, plotted with adjusted gain by thin line. There again exists a hump around 1.7 kHz, similar to the results when the reference single source was located at the left loudspeaker. This implies that the hump is not due to the direction of the reference source.

The other experiment was to find whether this hump occurs when listeners compared two single sources from different directions, which are, left and center. In this experiment the reference single source was located at –30 degrees and the test single source was located at 0 degrees.

Fig. 10 depicts the calculated results of CLL difference, plotted with adjusted gain again in thin line. In this case the hump around 1.7 kHz was relatively smaller, which implies that the hump in CLL difference is not due to the direction of sound source.

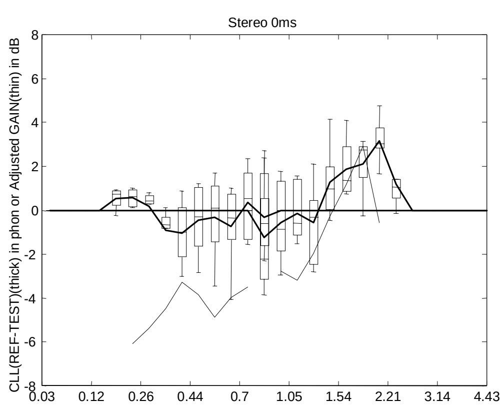

Fig. 9. CLL difference for wideband noise. Stereophonic setup is used, no delay between channels. Adjusted gain is plotted by thin line (in dB). Reference single source was in center. (From left in Fig. 4)

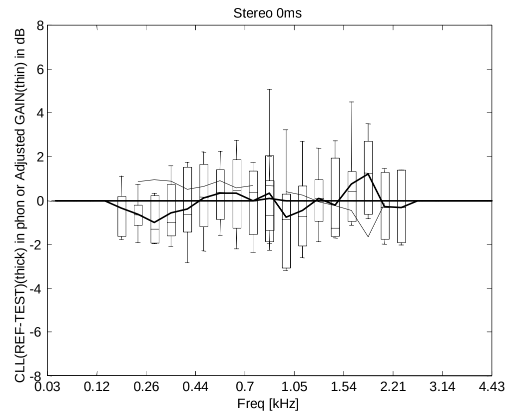

Fig. 10. CLL difference calculated using binaural auditory model. Adjusted gain is plotted by thin line. Comparison between single source at –30 degrees (reference) and 0 degrees (test).

3.5 Listening test 3

Listening tests 1, 1a, and 2 dealt with wideband noise, while this test deals with wideband pulse train. The pulse train was obtained as the minimum-phase component of 6-Bark wideband noise whose duration was cut to 100 ms and by repeating this four times. This retains the magnitude spectrum practically same as with the noise stimulus. Other conditions were the same as in Listening test 1, except that frequency ranges used were limited to where coloration was prominent in the narrowband signal case, that is, the higher band used in stereophonic setup with no delay (8-13 Bark) and lower band in both stereophonic and front-side setups with 2 ms delay (1-6 Bark).

Fig. 11 shows the calculated results of the CLL difference, plotted also with adjusted gain, for stereophonic setup with no delay. It shows that the CLL difference was negative, different from all other results with the same setup.

Fig. 12 and Fig. 13 show the calculated results of CLL difference, including adjusted gain, for the stereophonic and the front-side setup, respectively. Each figure has a clear hump in CLL difference at 3 Bark, which is similar to the results obtained in the wideband noise case, and dissimilar to the results with the narrowband signal case.

4. Discussion

In the listening tests of this study the subjects adjusted the timbre of a virtual source to match with a real source as a function of frequency. The adjusted gains were not constant with frequency, which means that the virtual sources without adjustment are colored in some degree.

In this study, Composite Loudness Level (CLL) difference, which is a calculated difference of CLL for a reference single source minus CLL for a test virtual source, was used to explain the test results. It was found that CLL is in general a good parameter for binaural modeling for wideband signals as well as for narrowband signals. There were two cases in which CLL showed clear variation from zero: a 2 dB hump around 1.7 kHz in the stereophonic setup with no delay, and a great hump nearly 10 dB at 3 Bark (around 200 Hz) in both the stereophonic and the front-side setup with 2ms delay.

In these cases the model could not explain the listening test results. In similar narrow-band tests there were less prominent discrepancies. This suggests that introducing a broader bandwidth signal changes the timbre perception. In this point there are some possible ways to try to explain these phenomena. At least some of the problematic cases occur at frequency bands where it is known that the binaural properties of sound are highly unnatural [10], [18]. The introduction of such changes in a broadband material may affect perceived timbre. Possible temporal changes of timbre during stimulus signals were not monitored in this study, which could also explain at least some features in the obtained data. Exploring these issues is left to future projects.

5. Conclusion

Binaural modeling of virtual source coloration was studied under precedence effect conditions, based on a methodology of using recorded ear canal signals, which enables complete matching between the listening test conditions and binaural modeling. It was found that the composite loudness level (CLL) obtained by summing up the loudnesses of each ear is in general a good measure for binaural loudness and timbre for wideband signals, although in some cases there are clear deviations of subjective test results from model-based simulations.

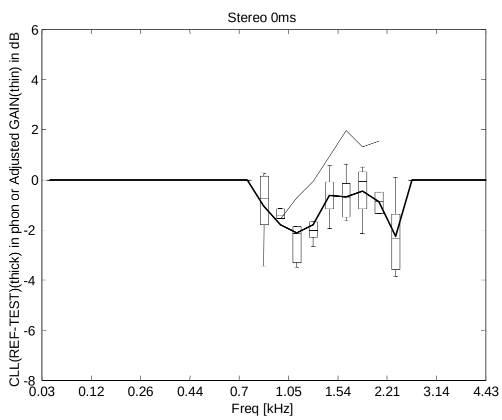

Fig. 11. CLL difference by sideband pulse train. Stereophonic setup is used, no delay between channels. Plotted with adjusted gain in thin line.

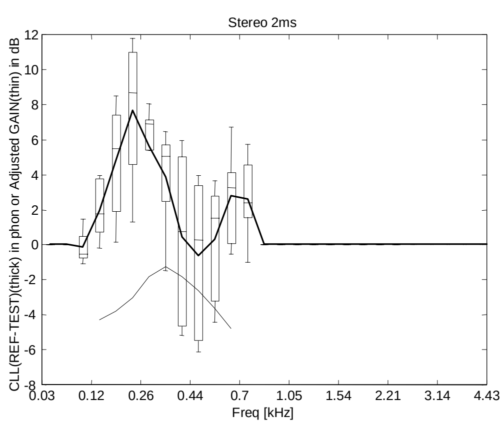

Fig. 12. CLL difference by wideband pulse train. Stereophonic setup is used, 2 ms delay between channels. Plotted with adjusted gain in thin line.

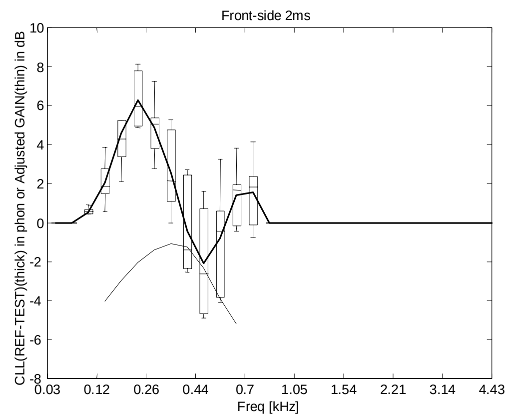

Fig. 13. CLL difference calculated using binaural auditory model. Front-side setup is used, 2 ms delay between channels. Plotted with adjusted gain in thin line.

References

[1] K. Ono, V. Pulkki and M. Karjalainen, Binaural Modeling of Multiple Sound Source Perception: Methodology and Coloration Experiments, In *Proceedings of the 111th Convention of the Audio Engineering Society*, New York, December. 2001, Preprint 5446.

[2] R. H. Gilkey and T. R. Anderson, editors. *Binaural and Spatial Hearing in Real and Virtual Environments*, Lawrence Erlbaum Assoc., 1997.

[3] J. Blauert. *Spatial Hearing*, Revised edition, The MIT Press, 1997, Cambridge, MA, USA.

[4] J. Blauert and W. Cobben, Some Consideration of Binaural Cross Correlation Analysis, *Acoustica*, 39:96-104, 1978.

[5] American Standards Association. *Acoustical Terminology SI*, American Standards Association, New York, 1960.

[6] G. Theile, Über die Lokalisation im Überlagerten Schallfeld, *Doctoral thesis, Technical University of Berlin*, 1980.

[7] H.-J. Platte and K Genuit, Ein Beitrag Zum Verständnis der Summenlokalisation, *DAGA '80*, VDE-Verlag, pp. 595-598.

[8] E. Zwicker and U. T. Zwicker. Dependence of binaural loudness summation on interaural level differences, spectral distribution, and temporal distribution. *J. Acoust. Soc. Am.*, 89(2):756-764, February 1991.

[9] B. E. Mulligan, L. S. Goodman, D. P. Gleisner, M. L. Faupel. Steps in loudness summation. *J. Acoust. Soc. Am.*, 77(3):1141-1154, March 1985.

[10] V. Pulkki, M. Karjalainen and J. Huopaniemi. Analyzing Virtual Sound Source Attributes Using a Binaural Auditory Model. *J. Audio Eng. Soc*., 47(4):203-217, April 1999.

[11] V. Pulkki and M. Karjalainen. Localization of Amplitude-Panned Virtual Sources I: Stereophonic panning, *J. Audio Eng. Soc.*, 49(9):739-752, September 2001.

[12] A. Härmä and K. Palomäki. HUTear -- a free Matlab toolbox for modeling of auditory system, 96-99,*Proc. Matlab DSP Conference*,Comsol Ltd., Espoo, Finland, Nov. 1999. http://www.acoustics.hut.fi/software/HUTear/

[13] B. C. J. Moore. *An introduction to the psychology of hearing*, Academic Press, San Diego, fourth edition, 1997

[14] R. Patterson, K. Robinson, J. Holdsworth, D. Mckeown, C. Zhang and M. H. Allerhand. Complex sounds and auditory images. In L. Demany, Y. Cazals and K. Horner, editors, *Auditory Physiology and Perception*, pages 429-446. Pergamon, Oxford, 1992.

[15] B. C. J. Moore, R. W. Peters and B. R. Glasberg. Auditory filter shapes at low center frequencies. *J. Acoust. Soc. Am.*, 88(1):132-140, July 1990.

[16] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*, Springer-Verlag, Heidelberg, Germany, 1990.
[17] B. C. J. Moore. A Model for the Prediction of Thresholds, Loudness, and Partial Loudness, *J. Audio Eng. Soc.*, 45(4):224-240, 1997.
[18] V. Pulkki and M. Karjalainen. Coloration of Amplitude-Panned virtual Sources, In *Proceedings of the 110th Convention of the Audio Engineering Society*, Amsterdam, May. 2001, Preprint 5402.