# Audibility of all-pass components in binaural synthesis

Minnaar, Pauli; Christensen, Flemming; Møller, Henrik; Olesen, Søren Krarup; Plogsties, Jan

*Published in:*
Proceedings of 106th Audio Engineering Society Convention, May 8-11, 1999, Munich, Germany

*Publication date:*
1999

*Document Version*
Accepted author manuscript, peer reviewed version

**Audibility of All-Pass Components in Binaural Synthesis**

4911 (L5)

Pauli Minnaar, Flemming Christensen, Henrik Møller, Søren Krarup Olesen & Jan Plogsties

*Aalborg University, Aalborg, Denmark*

**Presented at the 106th Convention 1999 May 8–11 Munich, Germany**



**AN AUDIO ENGINEERING SOCIETY PREPRINT**

# Audibility of all-pass components in binaural synthesis

Pauli Minnaar, Flemming Christensen, Henrik Møller, Søren Krarup Olesen & Jan Plogsties

Acoustics Laboratory, Aalborg University
Fredrik Bajers Vej 7 B4, DK-9220 Aalborg Ø, Denmark
acoustics@kom.auc.dk

## Abstract

Transfer functions can generally be decomposed into a minimum phase component, an all-pass component and a pure delay. Earlier investigations have shown that head related transfer functions and headphone transfer functions have all-pass components, but if these do not have an audible impact on the sound it might be beneficial to implement binaural synthesis without them. A subjective experiment is described in which listeners are presented with signals that differ only in terms of a second order all-pass section. The results of the experiment are threshold curves dependent on the centre frequency and Q-factor of the all-pass section.

## 1 Introduction

Binaural technology has in recent years received a considerable amount of attention since it makes reproduction of a spatial auditory event possible through two audio channels only. Binaural signals are obtained in one of two ways: 1) binaural recording and 2) binaural synthesis. These binaural signals can then be played to a person through headphones. In binaural recording the aim of the sound reproduction system is to provide the same pressure fluctuations in front of the eardrums that would have been there in a listening situation. Then the same auditory event may result. Binaural synthesis aims to generate signals that will create the perception of a spatial auditory event that, unlike the case of binaural recording, may not have existed in real life. This is usually done by convolving anechoic signals with head related transfer functions (HRTFs).

Before playback both channels have to be corrected for the transmission through the reproduction chain, here referred to as the headphone transfer functions (PTFs). HRTFs and PTFs are measured on people having microphones inserted into their ears. The HRTFs and inverse PTFs are implemented as digital filters and it is the properties of these filters that are discussed here.

A transfer function can be broken up into three components 1) a minimum phase component, 2) an all-pass component and 3) a linear phase component as in the following equation.

$$\mathrm{h}(n) = \mathrm{h}(n)_{\text{minimum phase}} * \mathrm{h}(n)_{\text{all-pass}} * \mathrm{h}(n)_{\text{linear phase}} \tag{1}$$

The corresponding frequency response function is:

$$\mathrm{H}(f) = \mathrm{H}(f)_{\text{minimum phase}} \cdot \mathrm{H}(f)_{\text{all-pass}} \cdot \mathrm{H}(f)_{\text{linear phase}} \tag{2}$$

The minimum phase component has a magnitude exactly as the original transfer function and the smallest possible phase angle for the corresponding magnitude. The all-pass component has a magnitude of unity and a phase structure that can be decomposed into an integer number of second order sections (all-pass sections). Similarly the linear phase component has a unity magnitude and it corresponds to a pure delay. The question of interest here is what perceptual role the all-pass component plays in the transfer function and more specifically how the perception of the all-pass sections relates to those found in typical HRTFs and PTFs.

## 2 Theory

The transfer function for a single second order all-pass section is as follows:

$$\mathrm{H}(s) = \frac{s^2 - \frac{\omega_0}{Q} s + \omega_0^{\,2}}{s^2 + \frac{\omega_0}{Q} s + \omega_0^{\,2}} \tag{3}$$

$$\text{where } \omega_0 = 2\pi f_0 . \tag{4}$$

$Q$ is the Q-factor and $f_0$ the centre frequency of the all-pass section.

In order to clarify the influence that the variables $Q$ and $f_0$ has on the impulse response, four examples are given in Figure 1. The examples correspond to four extreme situations:

a) Low $f_0$, High $Q$ b) High $f_0$, High $Q$
c) Low $f_0$, Low $Q$ d) High $f_0$, Low $Q$

In case a) the low frequency ringing is clearly seen after the initial impulse. Since the Q-factor is high the ringing will continue beyond the time window shown here. The same is seen in b), but the frequency is clearly higher. Notice that in both cases a) and b) the impulse response starts with a single high amplitude impulse. The first values in the impulse response are not generally the highest for low Q-factors, though, as seen in cases c) and d). The Q-factor controls the decay time of the impulse response. When the Q-factor is low the impulse response dies away quickly and energy is delayed in the impulse response. The all-pass sections shown are all causal. A non-causal all-pass section can be obtained that has the same time structure as the causal all-pass but is a time reversed version thereof.

Since thresholds of audibility of all-pass sections can be related to group delay, equations are given here in terms of the centre frequency and Q-factor [1]. The group delay at a given frequency is defined as the negative derivative of the phase with respect to frequency. The group delay of an all-pass section when evaluated at its centre frequency ($f = f_0$) is given as:

$$\tau_g(f)\Big|_{f=f_0} = \frac{4Q}{2\pi f_0} \tag{5}$$

Furthermore an expression for the group delay evaluated at $f = 0$ is given as:

$$\tau_g(f)\Big|_{f=0} = \frac{2}{2\pi f_0 Q} \tag{6}$$

## 3 Previous work

As a starting point for the analysis of the audibility of all-pass components the effects of a single all-pass section (second order section) on the perception should be known. Most investigations on the sensitivity of the human auditory system to the phase structure of a signal aim at finding perceptual tolerances to phase changes in audio and sound reproduction systems. A review on the audibility of all-pass sections is given by Preis [2] for studies up to 1982. The studies found perceptual thresholds of relatively high Q-factor to depend on the test signals, the method of presentation and the training of the listeners. A contribution to this topic was made by Karjalainen et. al. [3] who investigated phase sensitivity of the hearing with regard to the application of loudspeaker equalisation.

The threshold of audibility of a second order all-pass section with a high Q-factor relates to the group delay, $\tau_g(f)$, at its centre frequency. The group delay values reported for mid range frequencies (approximately 500 Hz to 4 kHz) vary between 1-3 ms for the most critical test signals, which seem to be impulses and sharp transients. It is generally reported, that the perceptual threshold for other signals is higher. Deer et. al. [4] investigated the audibility of all-pass sections centred at 2 kHz presented diotically through headphones. The perceptual threshold for impulses is reached when the peak group delay values are approximately 2 ms.

## 4 Experimental methods

A psycho-acoustical experiment was done in order to determine thresholds of audibility for a single all-pass section. Eight untrained listeners (4 male and 4 female) participated that had normal hearing according to an audiometric test. Test signals were generated by convolving a signal with a single second order all-pass section. The original and treated signals were presented consecutively with a one second interval.

Signals were presented to a listener through Beyerdynamic DT990 headphones that were equalised for its minimum phase response. The task of the listener was to determine whether there was a difference between the two signals. The listener adjusted the Q-factor of the all-pass section by selecting a point on a scale drawn on an electronic tablet. The scale is shown in Figure 2. The range of Q-factors on the scale was 1.2 decades. Before the start of a session the listener was familiarised with the task.

Threshold curves were determined for 26 different conditions as shown in Table 1 below. A listener was presented with 2 repetitions of a condition. The order in which the conditions were presented was randomised for every listener.

| Presentation | Q-factor | Centre frequency | Direction of decay | Signal |
|---|---|---|---|---|
| Diotic | High $Q$ | 1,2,4,8,12 kHz | Causal | Clicks |
| Diotic | High $Q$ | 1,2,4,8,12 kHz | Non-causal | Clicks |
| Dichotic | High $Q$ | 4,8,12 kHz | Causal | Clicks |
| Dichotic | High $Q$ | 4,8,12 kHz | Non-causal | Clicks |
| Dichotic | Low $Q$ | 1,2,4,8,12 kHz | Causal | Noise |
| Dichotic | Low $Q$ | 1,2,4,8,12 kHz | Non-causal | Noise |

Table 1. Conditions used for presentation in the listening experiment.

High $Q$ thresholds are determined for diotic and dichotic presentation using a click (single impulse) and for low $Q$ dichotic presentation using a 0.7 second pink noise signal. The term diotic refers to a presentation of the same signals to both ears with an all-pass section applied to both signals. A dichotic presentation refers to the same signals in both ears but an all-pass is applied to only one signal. Notice that to determine the high $Q$ threshold for diotic presentation 1, 2, 4, 8 and 12 kHz centre frequencies were used whereas in the high $Q$, dichotic case 1 kHz and 2 kHz centre frequencies were omitted. Pilot experiments revealed that a click containing an all-pass section below 2 kHz was always audible for dichotic presentation.

Pilot experiments indicated that the lateralisation created by a low $Q$ all-pass section presented dichotically depends little on the signal. It is therefore reasonable to expect the low $Q$ threshold to be very similar if not the same for different signals. A pink noise signal was used here and 1 kHz and 2 kHz centre frequencies were included although ringing (above the high $Q$ threshold) was potentially audible at these frequencies. However the pink noise tend to mask the tone to the extent that the determination of a low $Q$ threshold could be made.

## 5 Results

Six different threshold curves were derived through the subjective experiment. The conditions used for every threshold is summarised in Table 1 and the corresponding threshold curves are given in Figure 3. The threshold at a centre frequency is the mean (on a logarithmic basis) of two attempts made by all listeners, i.e. the mean of 16 answers. Error bars on the mean indicate the standard deviation of the threshold.

The mean thresholds are all plotted in Figure 4. Thick lines refer to thresholds corresponding to causal all-pass sections whereas thin lines refer to those for non-causal all-pass sections. Curves with a positive slope correspond to the high $Q$ threshold whereas those with a negative slope are for the low $Q$ threshold.

The effects of an all-pass section are audible when the Q-factor is higher than the high $Q$ threshold line or lower than the low $Q$ threshold line. Notice that the threshold lines crossing each other at 2 kHz in Figure 4 do not correspond to the same method of presentation.

## 6 Discussion

In order to approach the question of the audibility of all-pass components in binaural synthesis we have started by inspecting a single second order section. More specifically thresholds were obtained for the psycho-acoustical effects that such a second order section will have when presented to a listener under diotic and dichotic conditions. The results of high $Q$ and low $Q$ all-pass sections are discussed here followed by an example of their application to inverse filtering of headphones for binaural sound reproduction and the implementation of HRTFs for binaural synthesis.

### *6.1 High Q all-pass*

A high $Q$ all-pass section typically 'rings' for a relatively long time and increasingly so as the Q-factor increases. As the Q-factor is reduced the ringing reduces until it becomes inaudible indicating a threshold. The thresholds for high $Q$ all-pass sections as shown in Figures 3 and 4 can be summarised as follows. A threshold exists for a click signal above which it can be distinguished from a click that contains an all-pass section with a high Q-factor. This threshold depends on the centre frequency and Q-factor. The application to either one or both reproduction channels does not influence the threshold very much since the ringing was detected equally well for the diotic and dichotic presentations. The threshold for a causal all-pass is slightly higher than for its non-causal counterpart which may be related to temporal (pre- or post-) masking of the click signal.

The group delay of an all-pass section when evaluated at its centre frequency ($f = f_0$) is given in equation (5). The high $Q$ curves in Figure 4 can very nearly be approximated by a straight line indicating a linear relationship between $f_0$ and $Q$ which in turn implies a constant group delay at the centre frequency. The constant group delay values for causal and non-causal all-pass sections are 1.5 ms and 1.2 ms here which correspond well with values of around 2 ms reported in Deer et. al. [4].

According to the literature [2], [3], [4] and pilot experiments the high $Q$ thresholds depend on the signal. The basis for the high $Q$ thresholds is the ringing which is well exposed by a transient signal and ultimately so by a single impulse (click). The ringing is heard as a tone before the impulse (for non-causal all-pass) or after the impulse (causal all-pass). The literature suggests that the threshold value for the group delay at the centre frequency is higher for other signals than for clicks. It must be clear that other signals such as speech, instrumental music, pink noise and white noise reveal the tone to a much lesser extent than a click. Therefore the high $Q$ threshold for most other signals is expected to be much higher than for a click. It is doubtful whether high $Q$ all-pass sections at high frequencies will be heard under normal listening circumstances. This holds for causal and non-causal all-pass sections as well as for diotic and dichotic presentations.

### 6.2 Low Q all-pass

A low $Q$ all-pass section is seen to 'displace' some energy to a point later in the impulse response. Generally, the lower the Q-factor the larger the shift in time and audibility becomes more likely. This is the case when signals are dichotically presented; the same signal in both ears but an all-pass section is applied to only one ear. Here the time difference between the two signals result in a lateralisation of the auditory event. When a low $Q$ all-pass section is applied to both ear signals (diotic listening) no lateralisation occurs since both channels are delayed equally much. Therefore a threshold does not exist for signals presented diotically.

Q-factors below the low $Q$ threshold lines (lines with negative slope) in Figures 3 and 4 will give detectably different lateralisations. Whether the all-pass is decaying in causal or non-causal time does not appear to effect the threshold much which should be intuitive since it simply yields lateral displacements in opposite directions.

An expression for the group delay evaluated at $f = 0$ is given in equation (6). The low $Q$ threshold as shown in Figure 4 is approximately linear giving a constant group delay for $f = 0$. The value for the group delay calculated from the threshold curve is approximately 30 μs. This means that if an all-pass section creates a group delay at $f = 0$ of more than 30 μs the lateral shift in the auditory event will be audible. In addition the fixed group delay calculated from the low $Q$ threshold curve determined in this experiment should be useful regardless of the signal.

### 6.3 Inverse filtering of headphones for binaural sound reproduction

A binaural signal should be played back at the point at which it was recorded (binaural recording) or at which the HRTF was measured (binaural synthesis). Therefore the transfer characteristics of the reproduction channel (here the PTF) is measured. The transmission of sound from the headphones to the microphone generally does not give a flat magnitude. As the effect that the channel has on the signal is considered undesirable it is usually compensated for by means of a digital filter.

An ideal compensation of the channel will result in a magnitude of unity and a phase of zero. This may require a non-causal filter which can be made causal by applying a delay. Whatever is not compensated (equalised) for by the inverse filter will be seen in the result after convolution. The high $Q$ threshold curves can be applied directly to determine the audibility of an all-pass section if not compensated for.

An example of five repeated measurements [5] made on a single person are given in Figure 5. Sennheiser HE60 electrostatic headphones were used to measure the PTF at the block entrance to the ear canal. Since the measurements are very repeatable we can be confident that the all-pass sections seen in the phase of the all-pass component are consistent. The corresponding centre frequencies and Q-factors of the all-pass sections are calculated and shown in Figure 6. On the same plot the high Q, diotic, causal threshold for a click signal is given. The Q-factor of the first all-pass section ($f_0$ slightly higher than 8000 Hz) is higher than the threshold line. Therefore if the inverse filter does not compensate for this all-pass section a ringing will be heard if a click signal is presented. Note that it is not the PTF or the inverse filter that rings but rather the result after convolution. Also we keep in mind that the high $Q$ threshold is expected to be higher when other signals are used.

In contrast to the given example measurements made on people wearing microphones and headphones [5], [6] show inconsistencies in the presence of all-pass sections. This implies that errors can be made when correcting the response. If all else is considered equal two scenarios can occur: 1) not correcting for an all-pass section that is present and 2) correcting for an all-pass section that is not present. Generally the error made in 1) will result in an impulse response that is causal whereas that in 2) is non-causal. Since the high $Q$ threshold for causal all-pass sections is higher than in the case of its non-causal equivalent it seems reasonable to suggest that it will be more safe not to equalise for an all-pass that is there than to equalise for an all-pass that is not there. However, this difference is not very pronounced.

### *6.4 Implementation of HRTFs for binaural synthesis*

The implementation of binaural synthesis may be simplified if minimum phase HRTFs can be used. An important question is whether it is possible to create a minimum phase HRTF in such a way that it cannot be distinguished from the original. An example of an HRTF measured on a person [7] on the contra-lateral side at an azimuth of 112.5 degrees and an elevation of –22.5 degrees is shown in Figure 7. The leading zeros (pure delay) are removed from the impulse response and the remaining part is decomposed into a minimum phase and an all-pass component.

In the phase of the all-pass component some all-pass sections are seen. In Figure 8 the corresponding Q-factors are shown as a function of its centre frequency as well as the causal low $Q$ threshold. Since at least one all-pass section has values for $f_0$ and $Q$ below the threshold line it is reasonable to expected in dichotic presentation that removing the all-pass component from the original HRTF will create an audible change. In other words if the shown HRTF (with pure delay replaced again) is compared to a minimum phase equivalent (with the same pure delay replaced) the difference may be audible. However, since the signals in the two ears are different the threshold may not apply directly.

## 7 Summary

The effects of an all-pass section (second order section) is audible under some circumstances. Six thresholds were determined by means of a psycho-acoustical listening experiment involving eight subjects wearing headphones. A threshold for high Q-factors was determined for diotic and dichotic presentation above which the audibility of the all-pass section will increase with its Q-factor. This threshold for a causal all-pass section is slightly higher than for its non-causal counterpart indicating that a non-causal all-pass section becomes audible at lower Q-factors. Whether the all-pass was applied diotically or dichotically does not influence the results markedly. On the other hand for low Q-factors thresholds were determined for dichotic presentation where the causality of the impulse response showed little influence.

## 8 Acknowledgement

This work was funded by the Danish National Centre for IT Research.

## 9 References

[1] M. E. Valkenburg: "Analog filter design", CBS College Publishing, 1982.

[2] D. Preis: "Phase distortion and phase equalisation in audio signal processing - A tutorial review", Journal of Audio Engineering Society, Vol.43, No. 5, 1982, pp. 774-794.

[3] M. Karjalainen, E. Piirilä, A. Järvinen, J. Huopaniemi: "Comparison of loudspeaker equalisation methods based on DSP techniques", Proceedings of 102nd Audio Engineering Society Convention, Munich, preprint 4437, 1997.

[4] J. A. Deer, P. J. Bloom, D. Preis: "Perception of phase distortion in all-pass filters" Journal of Audio Engineering Society, Vol.33, No. 10, 1985, pp. 782-786.

[5] Pauli Minnaar, Henrik Møller: "Inverse filtering of headphones for binaural sound reproduction", Presented at the 136th Meeting of Acoustical Society of America, October 12-16, 1998, Norfolk, Virginia, USA.

[6] Henrik Møller, Dorte Hammershøi, Clemen Boje Jensen, Michael Friis Sørensen: "Transfer characteristics of headphones measured on human ears", Journal of the Audio Engineering Society, Vol. 43, No. 4, April 1995, pp. 203-217.

[7] Henrik Møller, Michael Friis Sørensen, Dorte Hammershøi, Clemen Boje Jensen: "Head-related transfer functions of human subjects", Journal of the Audio Engineering Society, Vol. 43, No. 5, May 1995, pp. 300-321.

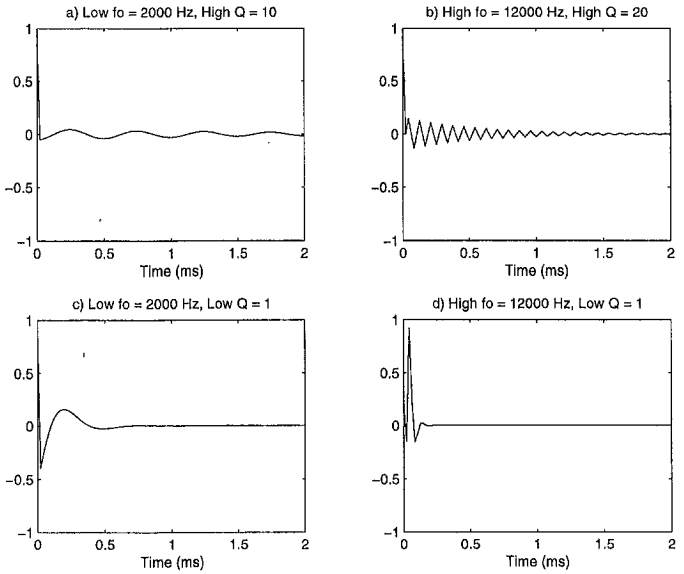


**Figure 1.** Impulse response functions for a single second order all-pass section. The all-pass is a function of the centre frequency, $f_0$, and Q-factor, $Q$.

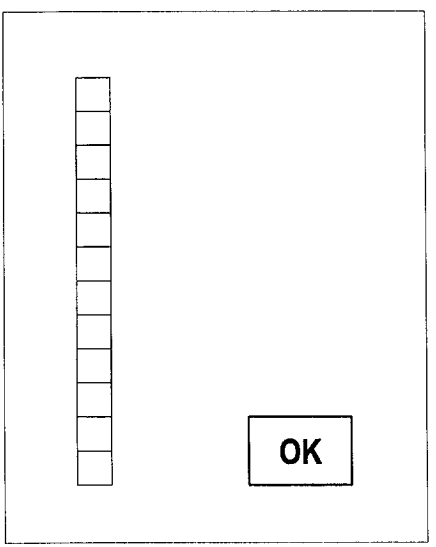


**Figure 2.** The interface used by listeners to determine the threshold. The condition at the top of the scale corresponded to an audible difference whereas the difference at the bottom was inaudible.

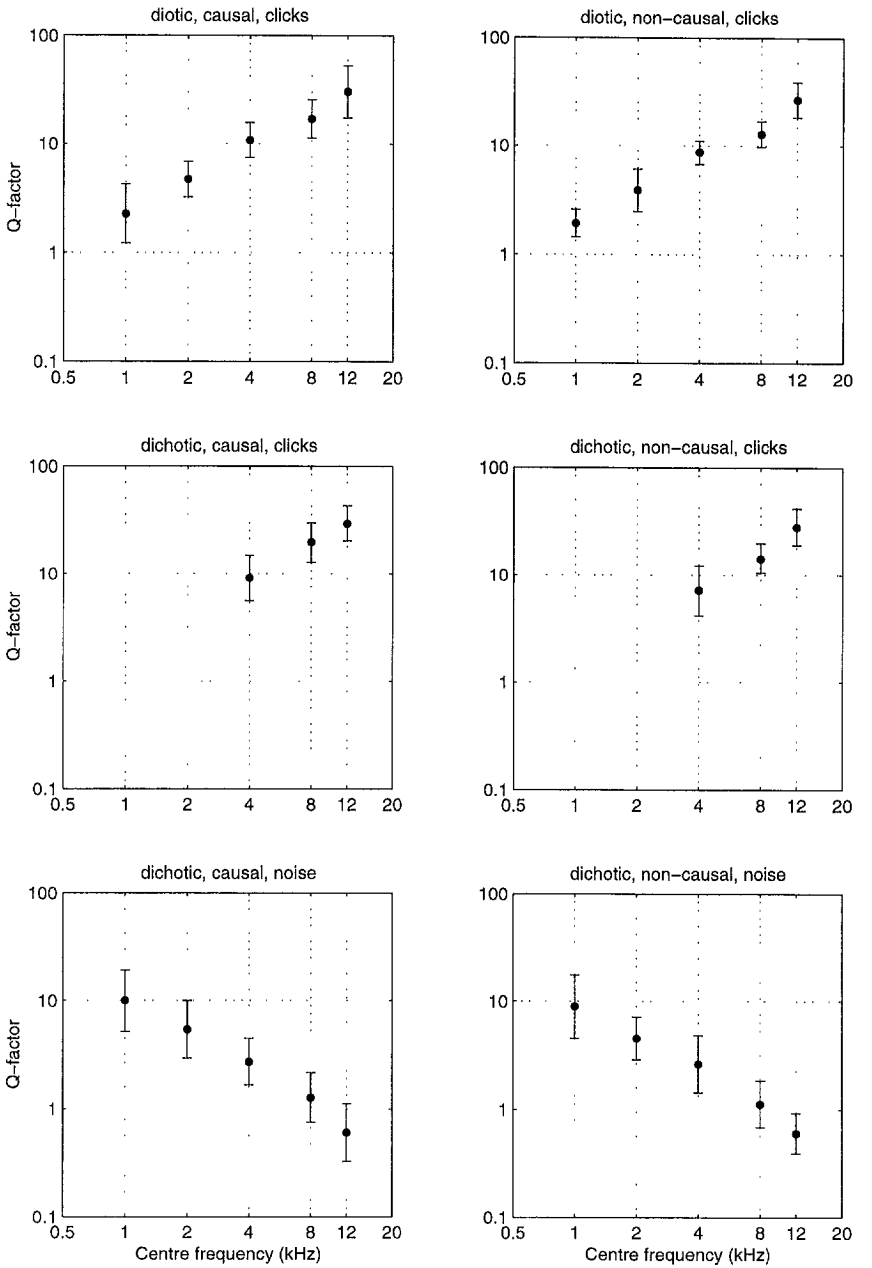


**Figure 3.** Mean and standard deviation (calculated on a logarithmic basis) for the threshold of audibility of all-pass sections under different conditions for 8 subjects.

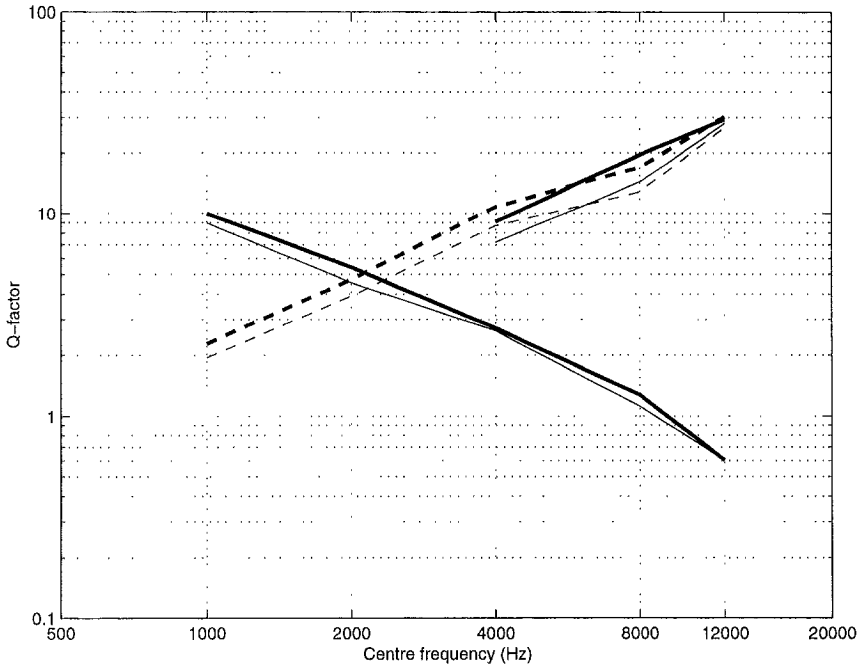


**Figure 4.** Thresholds of audibility of allpass components for 8 subjects. Solid lines refer to dichotic presentations whereas dashed lines refer to diotic presentations. Thick lines refer to causal and thin lines to non-causal all-pass sections. Curves with a positive slope correspond to a high $Q$ threshold whereas the curves with a negative slope represent a low $Q$ threshold.

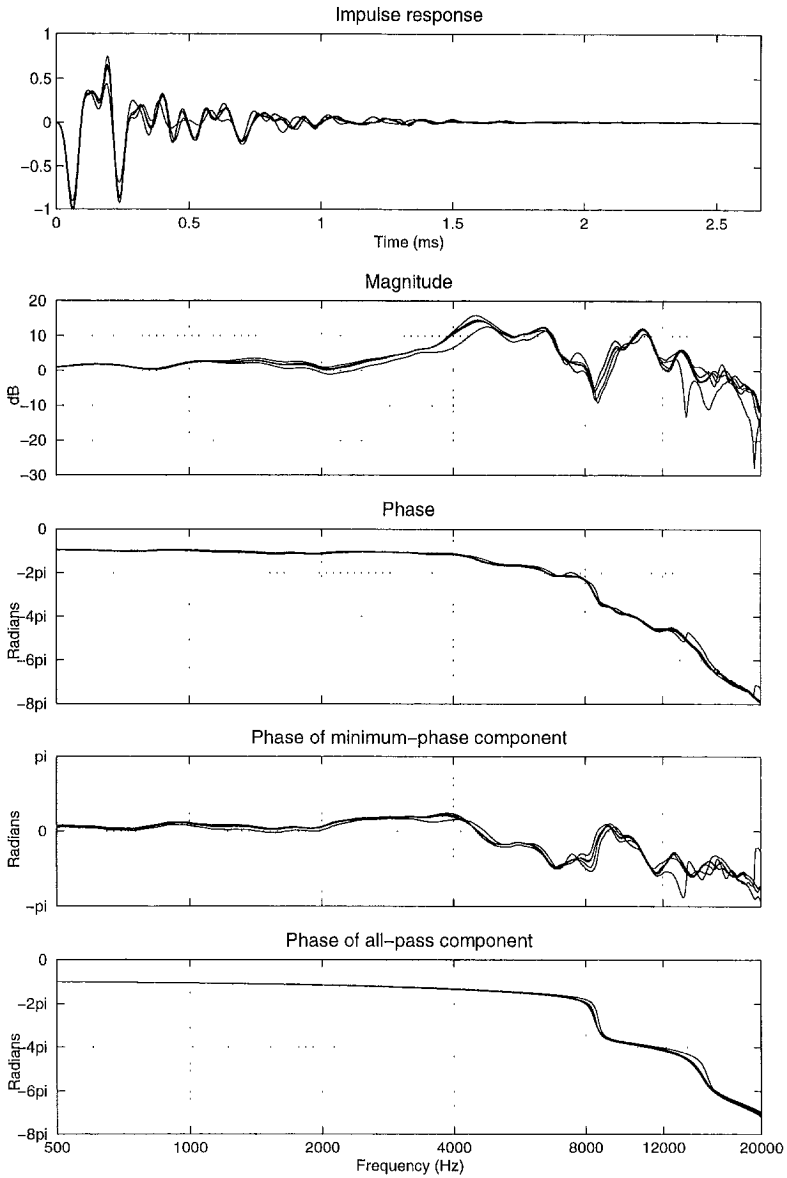


**Figure 5.** Five repeated measurements of a headphone transfer function (PTF) on one person with microphones at the blocked entrance to the ear canal [5].

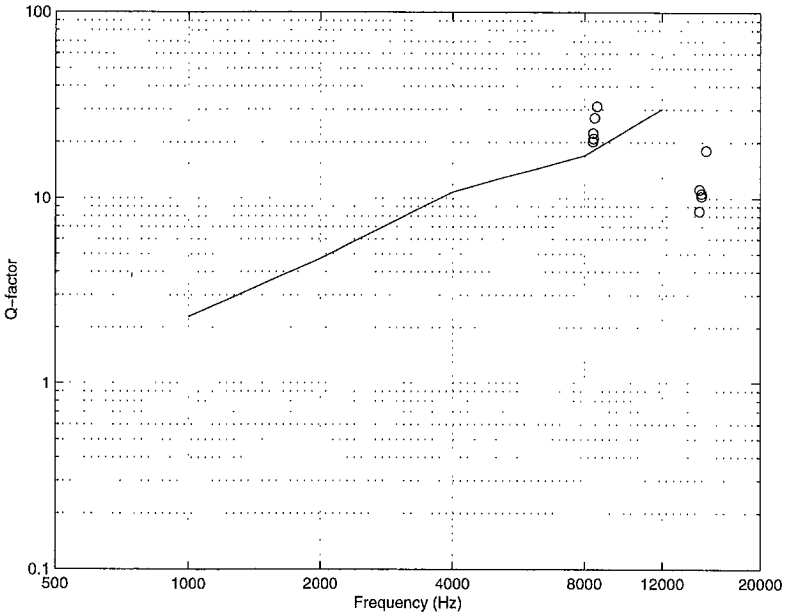


**Figure 6.** Centre frequencies and Q-factors for the five measurements shown in Figure 5 are related to the high $Q$, diotic, causal threshold for a click signal.

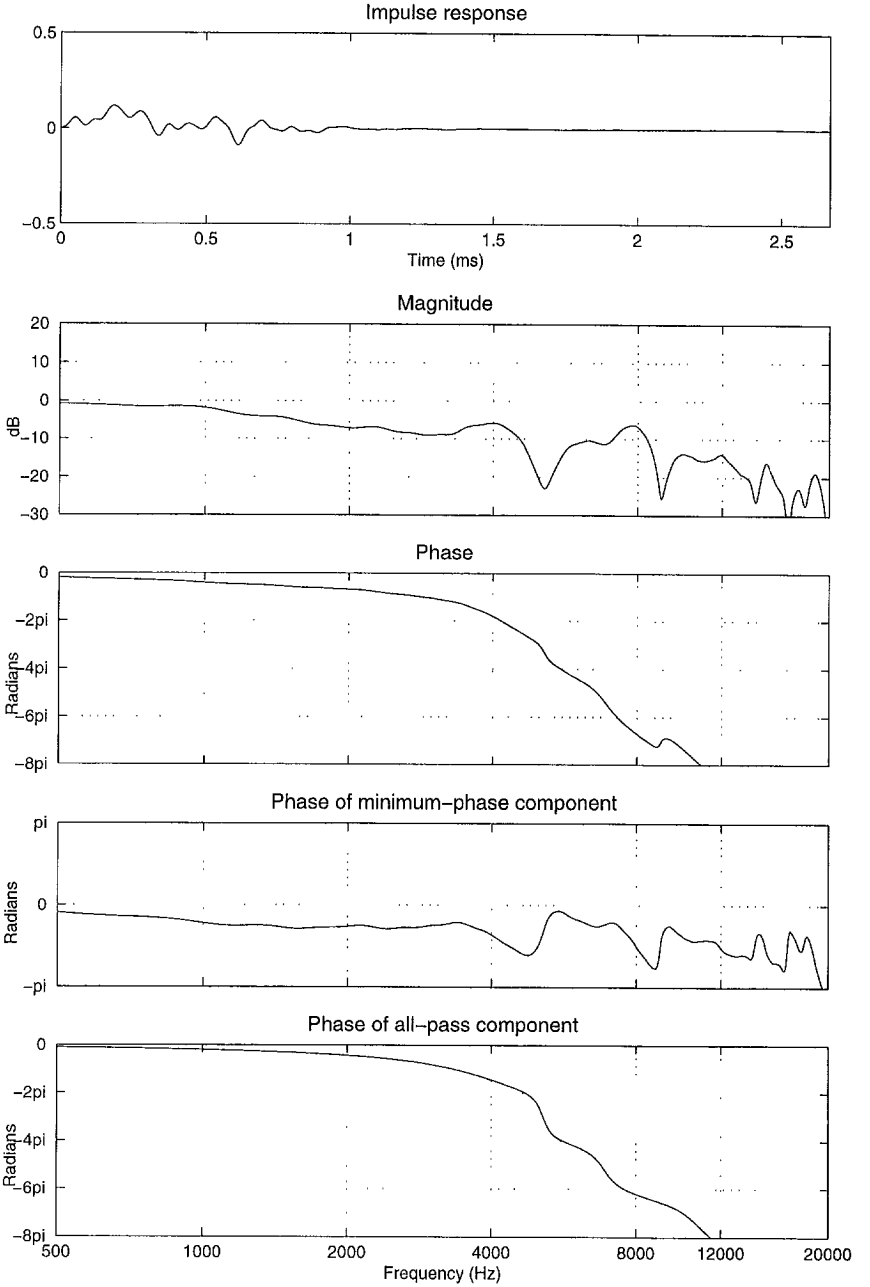


**Figure 7.** An HRTF on the contra-lateral side at 112.5 degrees azimuth and –22.5 degrees elevation measured at the blocked entrance to the ear canal of a person [7].

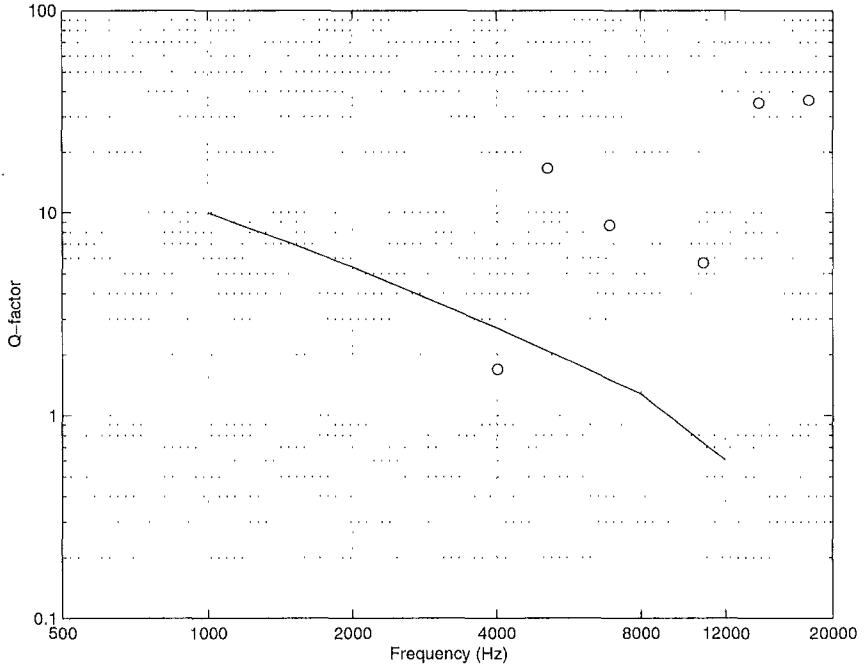


**Figure 8.** Centre frequencies and Q-factors for the measurement shown in Figure 7 are related to the low $Q$, dichotic, causal threshold.