informa
healthcare

Original Article

International Journal of Audiology 2010; 49: 634–644

Lars Bramsløw

Oticon A/S, Smørum, Denmark

Preferred signal path delay and high-pass cut-off in open fittings

Key Words

Hearing aids
Open fittings
Sound quality
Processing delay

Abbreviations

AFB: Anti-feedback
BTE: Behind the ear
BTL: Bradley-Terry-Luce
DSP: Digital signal processing
FFT: Fast Fourier transform
HA: Hearing aid
HI: Hearing impaired
HATS: Head and torso simulator
IG: Insertion gain
NH: Normal hearing
REAG: Real-ear aided gain
REIG: Real-ear insertion gain
REM: Real-ear measurement
REOG: Real-ear occluded gain
REUR: Real ear unaided response
RITE: Receiver in the ear
SII: Speech intelligibility index
VER: Vent external response

Abstract

The combination of delayed sound from a digital hearing aid with direct sound through an open or vented fitting can potentially degrade the sound quality due to audible changes in timbre and/or perception of echo. The present study was designed to test a number of delay and high-pass combinations under worst-case (i.e. most sensitive) conditions. Eighteen normal-hearing and 18 mildly hearing-impaired subjects performed the test in a paired comparison (A/B) task. The subjects were asked to select a preferred setting with respect to sound quality. The test was set in an anechoic chamber using recorded speech, environmental sounds, and own voice. Experimental hearing aids were fitted binaurally with open domes thus providing maximum ventilation. The preference data were processed using a statistical choice model that derives a ratio-scale. The analysis indicated that in these test conditions there was no change in sound quality when varying the delay in the range 5–10 ms and that there was a preference for 2000 Hz high-pass filtering in most conditions, regardless of the hearing losses tested.

Sumario

La combinación de sonido retrasado de un auxiliar auditivo con un sonido directo a través de una adaptación ventilada potencialmente puede degradar la calidad del sonido debido a cambios auditivos en timbre y/o la percepción de eco. El presente estudio fue designado para probar un número de retrasos y combinaciones pasa alto bajo las peores condiciones (i.e. las más sensibles). Dieciocho sujetos normales y 18 con hipoacusia superficial hicieron la prueba con comparación pareada (A/B). Se les pidió a los sujetos que escogieran su condición preferida con respecto a la calidad del sonido. La prueba se llevó a cabo en una cámara anecoica utilizando lenguaje grabado, sonidos ambientales y la propia voz. Se adaptaron auxiliares auditivos experimentales en forma binaural con cúpulas abiertas para máxima ventilación. Los datos preferidos fueron procesados utilizando un modelo de selección estadístico que deriva en una escala por ratios. El análisis indicó que en estas condiciones de prueba no hubo cambios en la calidad de sonido cuando se varió el retraso en el rango de 5–10 ms y que había una preferencia por la filtración pasa-alto de 2000Hz en la mayoría de condiciones, sin importar la hipoacusia en cuestión.

The processing delay in hearing aids is regarded as a critical design feature. Nevertheless there are not very precise limits for the delay. The general trend is to keep the delay as low as possible but this raises a number of problems in terms of e.g. filterbank resolution at low frequencies and efficiency of the anti-feedback (AFB) system. The present investigation will focus on the comb-filter effect that comes from the addition in the ear of direct sound through the vent and delayed sound through the hearing aid (HA), when the signals are of approximately the same level. For equal levels and a frequency-independent delay of ΔT there will be destructive interference ($-\infty$ dB) for all multiples of the corresponding frequencies $(2N+1)/(2\Delta T)$ and constructive interference (+ 6 dB) for frequencies $N/\Delta T$.

These effects were measured on a head and torso simulator (HATS) using experimental receiver-in-the-ear (RITE) instruments fitted with open domes. The instruments were set at 10 dB flat insertion gain (occluded situation) and the measurement signal was white noise presented through a loudspeaker. The resulting insertion gains curves are shown in Figure 1 for signal path delays 5, 7, and 10 ms. The low-frequency roll-off of the large ‘vent’ in the dome will cause the amplified (and delayed) path level to drop at low frequencies and the direct path level to drop at high frequencies. There will be a transition zone where the two contributions are similar in level and interference may occur. Depending on the gain in the instrument, this region may move up and down in frequency. In the example shown here, using 5 ms delay, the transition region is roughly 500–2000 Hz. If a high-pass filter is introduced in the signal path, the bandwidth of the transition zone can be made smaller such that the comb filter effect is diminished.

Perceptual effects of delay

The perceived effect of comb filtering is a change in timbre, often described as ‘barrel’, ‘sea shell’, ‘listening through a pipe’, or similar. This is caused by the spectral notches in the resulting spectrum as discussed earlier. Other effects can be ‘echo’, e.g. a direct temporal perception of the direct and delayed signals. It is the author’s experience through the years, that the timbre effects are the most notable of the potential perceptual effects for typical hearing aid processing delays. This was confirmed in the pilot phase of the present study: below 15 ms delay, the coloration is the most immediate and clear effect of the comb filter.

ISSN 1499-2027 print/ISSN 1708-8186 online
DOI: 10.3109/14992021003753482

Received:
February 18, 2009
Accepted:
March 6, 2010

Lars Bramsløw
Oticon A/S, Kongebakken 9, DK-2765 Smørum, Denmark.
E-mail: lab@oticon.dk

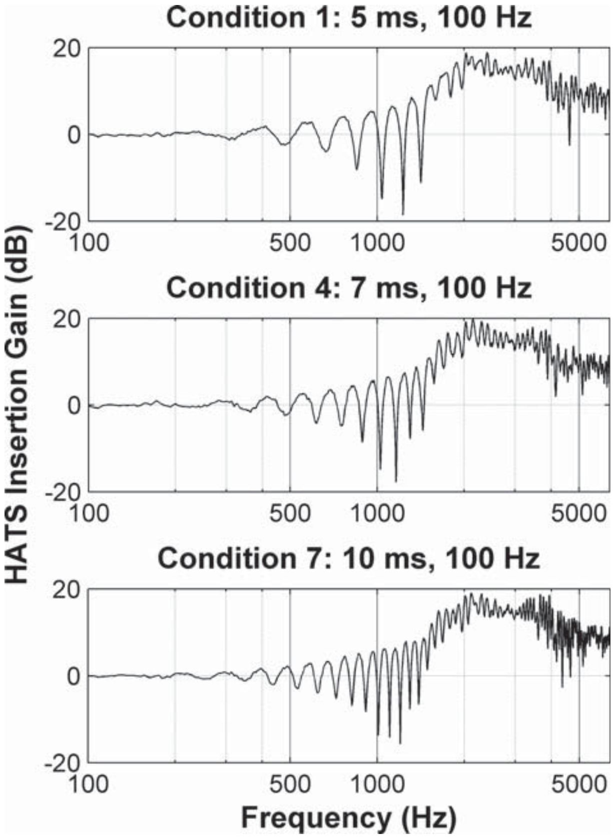

Figure 1. Comb filter effect measured on HATS. The instruments were set at 10 dB flat insertion gain (occluded) and the fitting to the ear canal was an open dome. The effect of increasing delay on the total output can be seen from the top and down.

Agnew and Thornton (2000) investigated noticeable and objectionable delays in an experimental hearing aid with adjustable delay. The study involved 18 normal-hearing test subjects and the experimental instrument was presumably set to 0 dB insertion gain independent of frequency. The instrument had one microphone and two receivers in a completely symmetric fitting. The receivers were placed in non-custom in-the-ear modules so there was a deliberate, individual leak providing a direct sound path. Own voice was used as the only stimulus and subjects were asked to adjust the delay to the point where it was noticeable and to the point where it was objectionable. Delays of 3–5 ms were noticeable in 76% of the trials, and longer than 10 ms delays were objectionable in 90% of the trials.

Groth and Søndergaard (2003) examined disturbance on a seven-point disturbance scale for binaural, symmetric, and non-occluding fittings. They used 2, 4, and 10 ms delay and 10 dB real-ear insertion gain (REIG) from 1000–6000 Hz. A 50-ms delay was used for training; the perceived artefact in this case would be an echo, unlike the test conditions which will be dominated by timbre changes. The test subjects were 10 normal-hearing listeners and 10 mildly hearing-impaired listeners. Stimuli were speech, music and own voice. Own voice was found to be the most critical signal of the three. They concluded that up to 10 ms delay is acceptable, based on an expected disturbance rating of maximum 3 on the seven-point scale. No test of noticeable delay was made. The authors claim that delay is even less critical in reverberant environments. In the study, there is not a clear distinction between the perceptual effects timbre and echo.

Müsch et al (2004) looked at monaural 'partially occluding' fittings. Six normal-hearing listeners judged naturalness in a magnitude estimation task. They also reported what acceptable naturalness was. Delay was 3, 6, 9, 12.4, 15, 30, and 60 ms, and REIG was 0, 5, 10, and 15 dB. REIG was flat above 500 Hz and lower for lower frequencies due to the open fitting. Speech was used as input, and no own voice was used. In a second experiment all signals were recorded from real ears, then normalized to the same REIG of 0 dB and played back over insert phones; this was done in an attempt to separate the delay effect from other effects of insertion gain (such as audibility of microphone noise and loudness imbalance). But it did not change the results; with 0 dB REIG up to 15 ms was acceptable, but with 15 dB REIG up to 6 ms was acceptable. As in Groth and Søndergaard (2003), noticeable delay was not tested.

Keidser et al (2007) studied the effects and interactions of low-frequency gain and DSP features, i.e. directional microphones and noise reduction. While not dealing with comb filter effects per se, the combined field and laboratory study does present some relevant results with respect to preferred low-frequency gain. The test compared NAL-NL1 prescriptions with modified gain at 250 Hz: 0 dB (vent dominated sound), 6 and 12 dB (HA dominated sound). These gain settings were tested in combination with features enabled and disabled. A large range of LF losses were tested from 12.5 to 70 dB HL at 500 Hz left-right average. The large majority of the subjects preferred the 0-dB gain condition in both lab and field test, even if it reduces the potential benefit from DIR and NR. There was no relation between preference and the amount of LF loss (500 Hz) or vent size. It was not investigated further what the physical and perceptual cause of the 0 dB preference could be, i.e. lack of comb filtering, less noise and distortion, or preferred frequency response. Localization as measured in a laboratory set-up was not affected by the 250 Hz gain amount. This suggests that the comb-filter effects can be reduced by decreasing LF gain—with no other negative side-effects.

Stone and colleagues have investigated tolerable delays in a series of five papers (Stone & Moore, 1999, 2002, 2003, 2005; Stone et al, 2008). The first four papers deal with closed fittings and thus little or no comb-filter effects. Most delay issues were related to echo and disturbance of own voice and delays up to 30 ms were found acceptable. The last paper (Stone et al, 2008) is an investigation on tolerable delay limits for open fittings. Many factors in the study were simulated: it used simulated vents, simulated instruments, and simulated hearing losses. Three effects were investigated in three experiments; delay, maximum gain, and gain slope (as function of frequency). Only normal-hearing test subjects were used, and the task was to rate the disturbance on a seven-point rating scale. For linear gain the disturbance increased with increasing delay and with decreasing gain slope. The authors conclude that the maximum tolerable delay is 5 ms for sloping frequency responses (insertion gain) and 10 ms for a flat frequency response. Regarding the signal path delay of the combined HA and open fitting, there is also a frequency-dependent delay: the delay in the low frequencies is determined by the vent and thus approximately 0 ms, whereas the high-frequency delay is determined by the delay of the hearing aid. There is a delay transition zone between these two values. This effect was also part of the Stone et al (2008) study, where it was studied independently of the amplitude response effects because it used a simulated setup. In the present study, a physical system is used and the amplitude (comb filter) and variable delay effects are inherently combined.

While the above-mentioned studies indicate limits for processing delay, they have various limitations that motivated the present study. Some studies have used own voice only, and some studies have used simulated systems which are convenient but perhaps not representative of the physical hearing aid and vent/dome system. The rating scales used may be less sensitive and subject to large inter-individual differences, thus a more sensitive test method was considered here. Most studies investigated objectionable or acceptable delay artifacts rather than noticeable delay, which presumably is more critical. From the literature as well as informal listening it appears that the objectionable delay is quite low and may set very rigid design constraints for a digital HA, but that the interaction with frequency response is not completely investigated. This interaction may be exploited to increase the processing delay and obtain other user benefits through more advanced signal processing.

Experiment

Purpose

The primary purpose of the test was to investigate the effects of comb filtering due to processing delay on sound quality. This should lead to more precise limits for setting the delay. A secondary purpose was to see if the perceived comb filter artifacts could be reduced by high-pass filtering in the hearing aid.

With the outset in a typical HA delay of approximately 5 ms, the two primary research questions were:

– What happens to the sound quality of the HA + vent system when the delay is increased above 5 ms?
– Can the sound quality in open fittings be improved if the comb-filter effect is reduced by reducing the low-frequency gain as in a high-pass cut-off?

It was determined in the pilot phase that the coloration (timbre) associated with the comb filtering is the predominant sound quality effect over echo and loudness changes, and thus timbre and sound quality. This is the focus of the present study. The overall goal was to achieve upper limits for signal path delay that are trustworthy and representative of critical use in critical environments. Therefore the test was designed to be as sensitive as possible.

Methodical issues

Test set-up

The test was carried out as a laboratory test only, because this was expected to be the most sensitive test of perceived comb-filter effects. The test was carried out in an anechoic chamber, considered to be the most sensitive room condition, without any confounding reflections (which also cause comb filtering).

The test setup consisted of:

– Experimental receiver-in-the-ear (RITE) hearing aids with adjustable frequency-independent signal path delay. Two instruments were used in a binaural, symmetrical fitting.
– PC and USB HiPro box to control the test instruments. The test was controlled by a Matlab application written for the purpose.
– PC and sound card (ECHO, Audiofire8), attenuator (HP 350D Attenuator Set), and an active loudspeaker (Genelec 8030A Bi-Amplified Monitoring System)
– Touch screen for test subject response.

The test subject was seated in a chair with the loudspeaker straight in front. Although other loudspeakers were mounted in the anechoic room, only the front speaker (0° azimuth and elevation) was used. The touch screen was mounted on the chair in such a way that it could be operated while reading from a book as done in the 'own voice' condition. The PC and monitor for the test leader were located outside the anechoic room. A talk-back clip microphone was attached to the test subject and connected to an intercom to provide possibility for communication at all times.

Test method

A number of test methods were considered, e.g. category rating as either absolute rating of each test condition or relative rating within sets of test conditions—see e.g. Bech & Zacharov (2006) for an overview. It was decided to use the simple A/B paired comparison with preference choice for primarily two reasons:

– Judged to be the most sensitive.
– Very easy task for the test subject.

The attribute used for preference choice was 'overall quality of reproduction'.

See Figure 2 for the test subject user interface (UI). The test subject listens to the stimulus in an infinite loop, and the two upper 'listen' A/B buttons will switch the hearing aid settings between the two value sets to be compared. Once the subject has decided on a preference, this is selected via the two lower 'prefer' A/B buttons. Pressing 'Next' will then stop the sound playback and proceed to the next paired comparison.

It was decided to perform the preference test as a two-alternative forced choice, meaning that there is not a 'no difference' reply option. The test subject has to indicate a preference even if no difference can be heard, and this random choice is then taken into account in the following statistical analysis of the results.

Material

Test subjects, normal hearing

The normal-hearing group was recruited amongst Oticon employees, who volunteered to participate in the study. It consisted of both expert listeners who were used to critical HA listening and

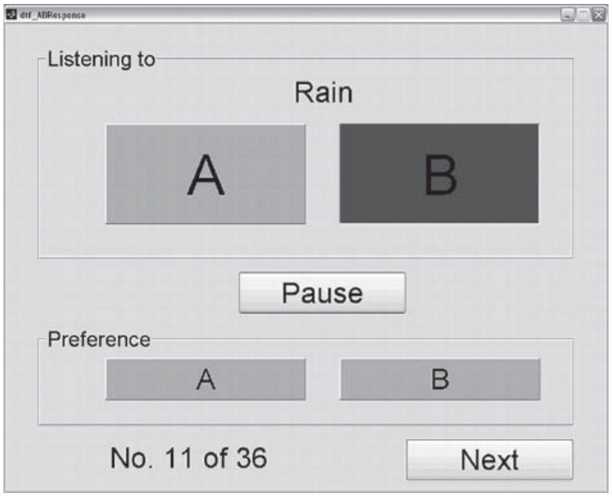

Figure 2. The test subject user interface, as shown on the touch screen.

non-experts. None of the test subjects knew the exact purpose and test settings used in the study. Danish as native language was not required because the client task only addressed quality and NOT speech intelligibility. All subjects were screened with an audiometer in a test booth and were only included in the test if the hearing loss was equal to or less than 20 dB HL at all audiometric frequencies from 125 to 8000 Hz. Eighteen subjects passed the normal-hearing screening and were included in the test.

Test subjects, hearing impaired

The 18 hearing-impaired test subjects were recruited from our internal database. All had hearing-aid experience.

The selected hearing losses were mild and moderate symmetrical sensorineural losses with near-normal low-frequency hearing (up to 30 dB HL) and up to severe high-frequency loss (up to 80 dB HL). This is within the typical fitting range of a modern open fitting hearing aid and the type of loss is expected to be the most sensitive to comb filtering because of the near-normal LF hearing. The audiograms are summarized in Figure 3.

Hearing instruments

The hearing instruments used for the test were binaurally fitted experimental digital BTE-RITE devices with adjustable frequency-independent signal path delay. They included a 16-channel filterbank for setting frequency response which in this case was used to produce the specified insertion gain (IG) and desired high-pass cut-off.

Settings

In order to keep the experimental conditions simple and easier to interpret, the instruments were programmed in a very simple setting:

– 10 dB flat insertion gain (when measured in a closed fitting) on HATS. This is similar to prescribed gain for a mild flat hearing loss and leaves fitting rationale out of the experiment for simplicity. The low gain also minimizes feedback issues. The same gain was used for NH and HI test subjects. For further discussion of insertion gain and high-pass filtering, see the ‘Physical effects’ section below.
– Same insertion gain (IG) for the all inputs level, i.e. linear gain and no compression.
– Noise reduction and other DSP features: OFF
– Anti-feedback system: OFF
– Directionality: Omnidirectional

This fixed setting was written to the EEPROM of the instruments. The delay and gain of both instruments as changed during the test was modified in working memory (RAM). The left and right instruments were always set identically.

Earpiece and venting

It was expected that the most open fittings would also provide the most comb filter effects because the low-pass cut-off for direct sound through the vent opening is the highest. Therefore it was decided to use Oticon ‘Open Dome’ for all test subjects, selected to appropriate ear canal size (6, 8, or 10 mm). The frequency response for incoming sound through the vent is also known as the vent external response (VER). The low-pass cut-off for the vent external response is approximately 2000 Hz for the Open Dome and the attenuation for high frequencies is less than 6 dB, as shown in the lower panel of Figure 4. The combined effect of delay hearing aid sound and direct vent sound is shown in the upper panel of Figure 4, which shows that the comb filter effects can extend up to 2000 Hz.

Signals

According to the test purpose, the test signals should be ‘critical’ in the sense that they should make the comb-filter effects as audible as possible. They should also be relevant to the test subject in the sense that they should be well-known and not annoying, e.g. traffic noise was not considered. The most effective signal is a pink or white noise signal, because there is energy at all frequencies making the comb-filter notches ‘visible’ in the long-term spectrum. However these are artificial signals. Transient sounds, such as clicks, have

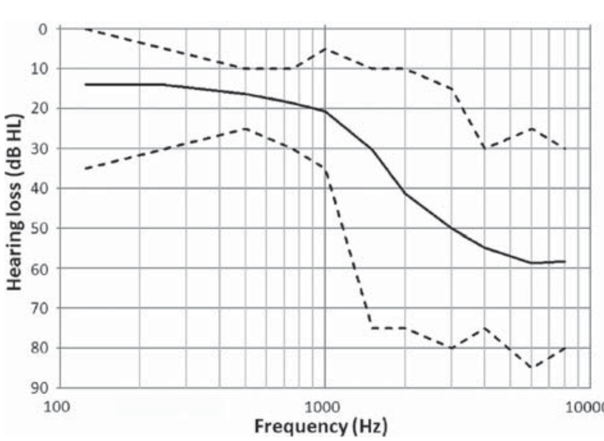

Figure 3. Summary of audiograms for the hearing-impaired listeners. The left and right sides have been combined. Minimum and maximum values are per frequency across all test subjects.

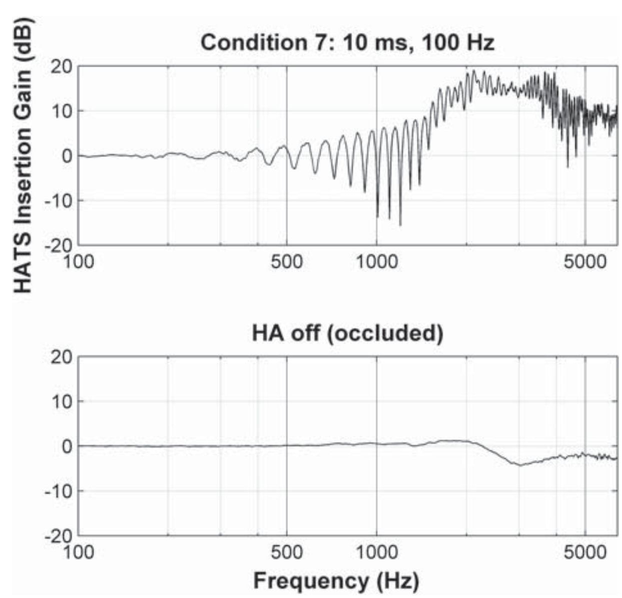

Figure 4. Occluded response (HA off – lower panel) compared to open fitting (upper panel). Notice the very little attenuation all the way up in frequency.

also been found useful, especially when the delay is so large that an echo or 'ringing' can be perceived.

After informal listening to a large number of signals, the following were selected:

- Waves on a beach. This signal is a natural broad-band random signal, similar to the ideal but artificial pink noise.
- Rain: soft rain on a porch. This signal has both random noise properties (background) and transient properties (single drops in the foreground).
- Speech: a male voice reading a text passage. A male was selected in order to have the lowest pitch and possibly making the dips more audible. Speech was not selected as a very critical signal, but simply because it is the most important signal for the HA user.
- Own voice: this has traditionally been regarded as a critical signal with regards to temporal aspects of hearing aids. The A-B test was performed while listening to own voice. The subject could either speak freely or a book was provided for reading aloud (Egner, 1953).

The three recorded signals were of 30–60 seconds duration, and played back in a loop. The signal presentation level was set at 63 dB SPL for speech, and the other signals were adjusted relatively in level to be judged natural by the author. An overview of the recorded signals and corresponding presentation levels is provided in Table 1.

Pilot test

The pilot listening test used a variety of signals and instrument settings (manipulated via a dedicated control panel). From theory it was expected that the following factors would affect the perceived comb filter effect:

- Delay
- Insertion gain
- High-pass cut-off

These factors were adjusted in a number of listening sessions with the author, a normal-hearing test subject, and two hearing-impaired test subjects. The following informal effects were noted:

- Varying the delay changes pitch but not the magnitude of the comb-filter coloration
- The comb-filter effect is clearly audible over a quite wide range of nominal insertion gains (IG), in the order of +5 – +15 dB, but the point of most audibility varies from subject to subject.
- The comb-filter effect is dramatically reduced by adding a high-pass cut-off. In this case the hearing aid insertion gain below the cut-off frequency was set at −40 dB, to effectively remove the hearing aid contribution completely.
- The changes made in delay and high-pass cut-off did not affect the perceived loudness—only the timbre.

Test design

Paired comparison

The task of the test subject was chosen to be a full paired comparison with preference choice. This test method is regarded as the most sensitive in order to detect and choose between small differences, see e.g. Bech & Zacharov (2006) for further discussion. Full paired comparison limits the number of test conditions to a small number, due to the square law increase of pairs with the number of conditions. With an estimated 30 seconds per comparison, the following time estimates can be made for a full paired comparison of nine conditions: $9 \times (9-1)/2 = 36$ comparisons $= 18$ minutes. This is a reasonable duration for one preference test. The study was thus designed for nine signal processing conditions repeated for each of the four test signals, with a short break between signals. The paired comparison was only done within each signal, so the different levels for different signals were never compared directly. Within pairs, the A/B assignment was randomized at run-time, i.e. when the each test session was initiated.

Test conditions

As mentioned previously, the important factors to consider when selecting test conditions are delay and high-pass cut-off. In this case it was important to test all combinations of delay and high-pass, so a full 3×3 factorial was the chosen design. The exact settings were:

- Delay: 5, 7, 10 ms: a typical HA delay plus 2 and 5 ms.
- High-pass cutoff: 100, 1250, 2200 Hz. These values were selected among the cut-offs available in the experimental HA.
- Insertion gain: fixed +10 dB flat on HATS with closed vent.

The resulting factorial design is summarized in Table 2.

Blocking

Each block in the experiment consisted of one test signal and a full comparison of all nine conditions, meaning 36 pairs. The sequence of pairs within a block was randomized once and then repeated across all test signals and test subjects. The test design is a balanced 'digram Latin square' (Hinkelmann & Kempthorne, 2008) which is used to balance out order effects in a group of four test subjects. A complete balanced design thus required an integer count of four test subjects, e.g. $4 \times 4 = 16$ test subjects.

Test protocol and execution

There was one visit including both training and testing, which lasted 1–2 hours including breaks.

Written consent

An information letter had been sent to the test subject prior to the visit. This letter also contained a written consent form which was signed by the test subject and handed over at the test. The experiment was approved by the regional research council ethical committee for Copenhagen as part of a larger study.

Instruction

Prior to the test a written instruction was presented to the test subject. The subject was instructed to select the setting with the

Table 1. Listing of the test signals and levels. The levels were measured at the position of the listener using a B&K free-field microphone and the B&K 3560 Pulse System. Own voice was not measured for each test subject.

Signal	L_{eq}
Waves	62.3 dB SPL
Rain	51.7 dB SPL
Speech	63.1 dB SPL
Own voice	Variable

best sound reproduction: '…This test is about selecting the setting that you think has the best sound. Do not judge whether speech is understandable or not, but focus on the sound quality'. The subject was encouraged to ask for further details. In some cases there was some doubt about the quality dimension (attribute) which was defined to be 'overall quality of the reproduction' and in some cases elaborated by the test leader as 'best reproduction', 'what you prefer', 'disregard speech intelligibility'. Some subjects would also use the term 'natural' even though this was not used directly in the instruction.

Training

There was one training block with 20 paired comparisons, which were selected to be the most different pairs in the full test and thus easier to make a preference choice on. In this phase of the experiment, the experimenter was in the test room (anechoic chamber) and in a constant dialogue with the test subject. After each paired comparison, the test subject was asked if there was a noticeable difference and if so to motivate the choice of either A or B. These subjective comments were written down.

In the training phase the NH group was able to detect audible differences on almost all pairs, and the HI group detected only slightly fewer different pairs. The high-pass conditions made the test subjects aware of the presence or absence of the comb filter artifact, described as for instance 'bee', 'airplane', 'echo', 'pipe', and 'noise'. The difference between delays was also detected and described as different tonal characters but there was often no clear preference for higher or lower delays. The listeners quickly became familiar with the test and the touch screen interface and performed the task easily.

One normal-hearing and one hearing-impaired subject experienced feedback at the default +10 dB IG, and consequently the gain was reduced to +5 dB. The response data were kept because both REM data and pilot listening had indicated that the comb filter effects were not very sensitive to the absolute gain within this range of values.

Data collection

Each test subject had to complete four test blocks, one for each of the four signals. The entire duration of a visit was two hours or less, and the test subjects quickly became familiar with the preference task. It was intended to complete the test for 16 test subjects, which would produce a balanced design. As a result more than 20 subjects were booked in each group, and the tests were eventually completed with 18 test subjects in each group, leaving the design slightly imbalanced with respect to the 4×4 Latin squares.

Physical effects

The listening test is used to determine the perceptual effects of delay and high-pass. In order to link this to the physical effects (= available cues), the in-situ gain of the nine test conditions was measured on a Brüel & Kjær 4128 Head and Torso Simulator (HATS). The HATS measurements were intended to produce average, standardized results for the physical effects. The HATS was seated in the test subject chair. All nine test conditions were measured on both L and R ears—and since they were identical only left ear is presented. Two additional conditions were measured: open ear (unaided = REUR), and occluded ear (HA off = REOG).

The Brüel & Kjær PULSE system was used for FFT power spectrum analysis. White noise at approximately 80 dB SPL was used for the measurement. The left ear power spectra were stored for the unaided and aided situation and used to calculate the insertion gain (IG) for the HATS by subtracting the above-mentioned REUR.

Table 2. Test condition numbers.

High-pass	*100 Hz*	*1250 Hz*	*2200 Hz*
Delay			
5 ms	1	2	3
7 ms	4	5	6
10 ms	7	8	9

Effect of delay

The effects of delay are clearly illustrated in the broadband (100 Hz) condition, i.e. conditions 1, 4, and 7. These are shown in Figure 1: For condition 1 (5 ms), the notches are clearly visible in the region 300–2000 Hz, and they are deepest in the region 750–1500 Hz. The interval between notches is 200 Hz corresponding to 1/5 ms. For increasing delay (top to bottom), the notches move closer as expected. For condition 7 (10 ms), the notches are spaced at approximately 100 Hz (= 1/(10 ms)) and visible from 300–2000 Hz.

Effect of high-pass

The high-pass characteristic is obtained by attenuating the low-frequency bands below cut-off to −40 dB. The high-pass effects on HATS signal are easily illustrated in Figure 5 by comparing the full bandwidth 100 Hz (condition 7, top panel) to the high-pass versions of the same (conditions 8 and 9, middle and bottom panels). It can be observed that the high-pass condition effectively removes all notches below cut-off. Similarly, the peaks below cut-off (up to 6 dB) are removed.

It should be noted that the most common high-frequency mild-moderate hearing loss will have near-normal hearing at low frequencies below 1000 Hz and thus often call for a target insertion gain smaller than the 10 dB tested in this study, but still close to 0 dB. Unlike the high-pass settings used in this study, the lower target will still cause comb filter artifacts because the direct and hearing aid contributions are roughly the same magnitudes. A slightly larger low-frequency loss of 20–30 dB would prescribe more than 0 dB target insertion gain and thus also cause comb filtering, because it is still suitable for open fittings.

Occluded ear

The effective insertion gain with the hearing instrument on and off is shown in Figure 4, top and bottom panel. Turning the HA off removes gain and comb-filter responses and leaves the occluded (vented) response: It can be seen that up to approximately 2000 Hz the open fitting has no attenuation, and for higher frequencies, the almost flat attenuation is 2–5 dB.

Effect of insertion gain

The effect of varying the HA gain was measured during the pilot tests on three individuals as rear-ear aided gain (REAG). Real-ear insertion gain (REIG) could not be calculated because the real-ear unaided response (REUR) was not measured on these test subjects.

An example of the REAG data is shown in Figure 6 for nominal HA insertion gains of 5, 10, and 15 dB. As seen in the plots, the effect of increasing IG is a downward shift in comb filtering. The combination of HA gain and the LF roll-off from the open dome vent moves the transition region down with increasing gain.

Perceptual preference effects

Data selection and processing

The data from all subjects were collected and analysed by means of dedicated Matlab functions. By counting the number of wins for each of the nine conditions and entering it into a 9×9 matrix, a preference matrix is built with an empty diagonal, and the winners in rows and losers in columns. Each cell can take value between 0 and 18 in the case of 18 test subjects. An example of a preference matrix from the present study is shown in Table 3. The preference matrix could be used directly as a non-parametric ranking of preference; however it is preferable to use a statistical model of the choice process and collapse the matrix into a score for each of the nine conditions, with confidence intervals based on the consistency of choices in the entire matrix. Such a model is the Bradley-Terry-Luce (BTL) model (Bradley, 1984), as used by e.g. Ellermeier et al (2004).

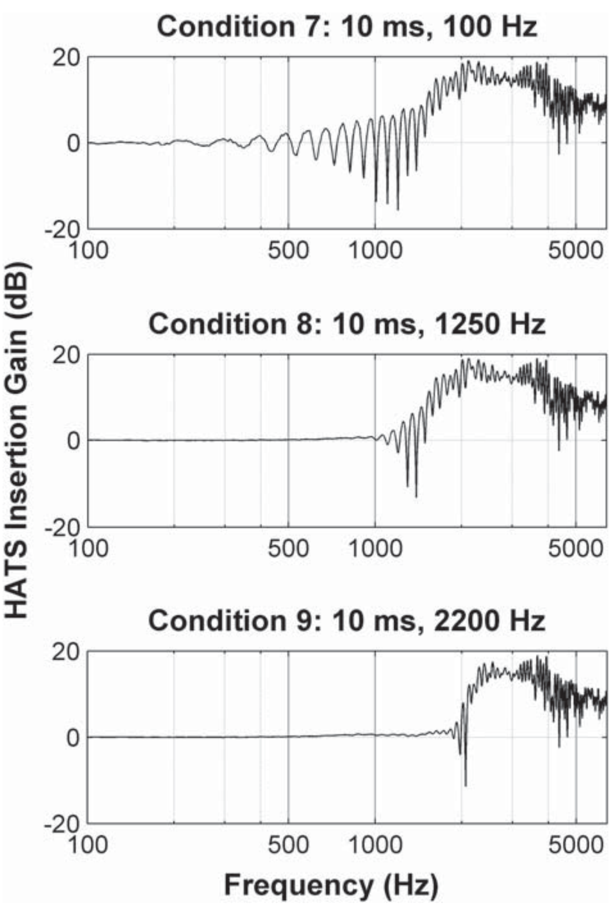

Figure 5. Comb filter effect measured on HATS. The instruments were set at 10 dB flat insertion gain (occluded), and the fitting to the ear canal was an open dome. The effect of increasing the high-pass cut-off on the total output can be seen from the top and down.

The BTL-model is available for Matlab, via the OptiPt.m function written by Wickelmaier & Schmid (2004). The preference matrix was used as input to the OptiPt function which can derive the preference score for each condition, thus nine scores are calculated. The function was slightly modified to be able to handle zeros in cells corresponding to 0% or 100% preference. This function takes the preference matrix as input.

The output of the BTL model is the BTL score which is a ratio scale, meaning that doubling and halving are equidistant, and the meaningful way to plot this is using a logarithmic scale. In the present study, the nine BTL scores—one per test condition— were normalized so that sum (BTL scores) = 1. This means that the scores can also be interpreted as a probability that a given condition will win in all of the paired comparisons.

In addition to the BTL score, the OptiPt function outputs other relevant parameters; two of these are used in the present study: A chi-square statistic used to test the validity of the model, and a covariance matrix, which can be used to calculate the 95% confidence intervals according to Wickelmaier & Schmid (2004). For more information on the models and their use and interpretation, see Wickelmaier & Schmid (2004) and Bradley (1984). Some good examples of application in acoustics are provided by Ellermeier et al (2004) and Choisel & Wickelmaier (2007).

Sample result from BTL analysis

The following is a walk through the results of the BTL analysis for the 'waves' signal and the NH group. Table 2 provides a convenient overview of the nine conditions and their numbering.

The preference matrix for this combination is shown as an example in Table 3. It is clear from the numbers that rows 3, 6, and 9 have very high scores. The resulting nine BTL scores for the 'waves' signal are shown in Figure 7, upper panel. There is a clear preference for all 2200 Hz conditions over the 100 and 1250 Hz conditions, which are rated very low (10^{-7} or lower!) and the latter two are not significantly different. Among the 2200 Hz conditions, the 5-ms delay is judged significantly lower than 7 and 10 ms. The dash-dotted line in this and the following plots indicate chance probability = 1/9 (0.11), and all BTL scores converge on this value if there are no significant effect.

It should be noted that in this particular case the chistat test shows a rejection of the null hypothesis: c*hisquare test:* $v = 53.3$, $df = 28$, $p <= 0.00270$. This means that the ordinary BTL model is inappropriate, probably due to the strong preferences, i.e. 18–0 for condition 3, 6, and 9 over several other conditions which also lead to numerical problems in the analysis. The alternative analysis in this case is a preference tree (Tversky & Sattah, 1979), in which the preference criteria between different pairs in not the same but based on different perceptual dimensions, e.g. that some choices are based on timbre while others are based on noise or artifacts. Preference trees can be created based on different hypotheses for the choice criteria and then validated using the chistat test. This is discussed in Zimmer et al (2004). One attempt did not show a very different outcome at this and was thus not pursued further. For all other signals and test subjects following this, the BTL model was valid, i.e. the null hypothesis was accepted.

Main effects

The main effects of delay and high-pass were calculated from the 9 BTL scores as follows:

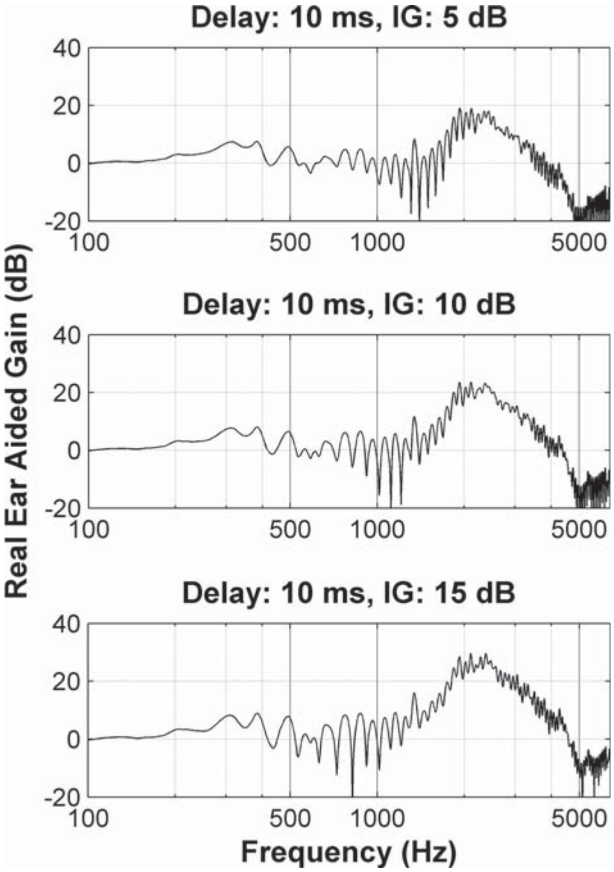

Figure 6. Example of real-ear aided gain recorded on an individual for different nominal insertion gains (5, 10, 15 dB). The effect of increasing gain is a downward shift in the frequency for the comb filter region.

– Mean BTL-score for delay were calculated by combining across all high-pass conditions and vice versa. The means were calculated in the log domain as appropriate for the ratio scale. This is equivalent to the geometric mean.

Table 3. Example of a preference matrix for the entire group (NH, 'waves'). Winners are indicated by rows, and losers by columns. For instance condition #4 has been preferred 14 times (bold) over condition #1 and likewise #1 has lost four times to #4 (bold). The diagonal is 0 since identical conditions in A and B are never compared.

Condition	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
1	0	2	0	**4**	0	0	14	3	0
2	16	0	0	17	7	0	17	3	0
3	18	18	0	18	18	8	17	18	8
4	**14**	1	0	0	3	0	10	2	0
5	18	11	0	15	0	0	16	7	0
6	18	18	10	18	18	0	18	18	9
7	4	1	1	8	2	0	0	1	0
8	15	15	0	16	11	0	17	0	0
9	18	18	10	18	18	9	18	18	0

– 95% confidence intervals were calculated by summing the variances across the other dimension, e.g. across high-pass for delay confidence intervals, and dividing this by the associated degrees of freedom (df = 6).

Delay

The main effects for delay are shown in Figure 8 for all four signals and both hearing loss groups. As indicated by the overlapping confidence intervals, there were no significant effects of delay for any signal or subject group.

A few tendencies can be noted: The plot for normal-hearing preference for 'own voice' (Figure 8, lower right) shows a trend towards less preference with increasing delay—especially 10 ms—but the differences are not significant.

The lack of significant preference for delay can be interpreted in two ways: Either the difference is not audible, or the difference is audible but there is nevertheless no clear preference. The perceptual effect of changing the delay is a shift of the perceived pitch of the coloration in the opposite direction, e.g. increased delay leads to a lower pitch. This change was clearly indicated by most subjects in the initial training session and thus the latter interpretation holds: Delay effects are audible and discernable, but they are equally poor in terms of sound quality.

High-pass cut-off

The main effects and associated confidence intervals for high-pass cut-off were calculated as previously described. The resulting BTL-scores are shown in Figure 9. For the normal-hearing listeners, there was a clear preference for increasing high-pass filtering, meaning

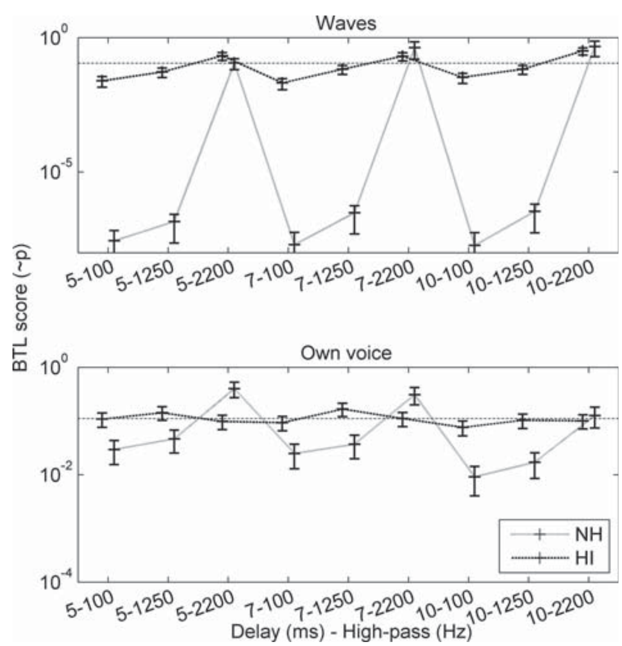

Figure 7. Preference scores for all nine conditions, NH listeners (grey) and HI listeners (dotted), signals 'waves' and 'own voice'. The effect of 'waves' is very strong, thus the different y-axis for this signal.

that 2200 Hz was preferred in all cases and 1250 in some cases over 100 Hz. For the hearing-impaired listeners the same result was seen for two signals. The relevant three examples are discussed here:

1) The first example is for normal-hearing listeners and the 'speech' signal (Figure 9, lower left). In this case all three high-pass conditions are significantly different. Thus the high-pass filter reduces the perceived comb filter effect and the more dips are removed the better for sound quality. It could be argued that NH test subjects prefer high-pass filtering because they have no need for the insertion gain. The same result is seen for the other signals.
2) The second example is for hearing-impaired listeners and the 'waves' signal (Figure 9, upper left). This example shows a similar situation, but for the hearing-impaired group. They do have the need for insertion gain, yet they still prefer the high-pass conditions with less gain and less comb filter artifacts.
3) The third example is for hearing-impaired listeners and the 'speech' signal (Figure 9, lower left). The HI 'speech' condition has no effect of high-pass which demonstrates that the most typical use case—speech—is not a critical signal for detecting comb filter effects for these listeners.

Interaction effects

Besides main effects a factorial experiment like the present one will also test interactions, i.e. preferences for specific combinations of the two main factors delay and high-pass cut-off. These can be determined by inspecting confidence intervals across all nine conditions. The BTL scores for the signals 'waves' and 'own voice' are shown in Figure 7. The other two signals 'rain' and 'speech' did not add further information and are therefore not shown.

The nine BTL scores for the 'waves' signal are shown in Figure 7, upper pane. Apart from the strong main effect of high-pass for both groups, there is a significant difference within the highly preferred 2200 Hz conditions, NH group: 5 ms is not preferred.

The nine BTL scores for the 'own voice' signal are shown in Figure 7, lower pane. For 'own voice' there is—as for the other signals—a clear and significant preference for 2200 Hz high-pass. Within these three high-pass conditions, the 10 ms delay is not preferred over 5 and 7 ms for the NH group. This is also the case for the 'rain' signal (not shown).

Speech intelligibility effects (simulated)

The primary purpose of a hearing aid is to improve speech intelligibility, especially in noise. The preferred hearing aid may not be the one that provides the best speech intelligibility for the user, so it is important to check that the goal of high sound quality does not sacrifice speech intelligibility.

In this study there was no subjective test of speech intelligibility, e.g. as a sentence test performed on all test subjects. Instead, the effect of delay and high-pass filtering on the speech intelligibility were assessed by means of a standardized, objective model: the speech intelligibility index (ANSI, 1997; Pavlovic et al, 2009). This method is applicable for linear systems in stationary noise, and provides an index in the range 0...1, predicting speech intelligibility score.

The speech stimulus used in the listening test was recorded on HATS left ear in 10 conditions: The 3×3 aided gain combinations of delay and high-pass plus the unaided (open ear) condition. This recorded signals were then analysed in 1/3 octaves to produce a 1/3 octave power spectrum which was fed to the SII calculation model (Pavlovic et al, 2009). The model assumed binaural listening, the

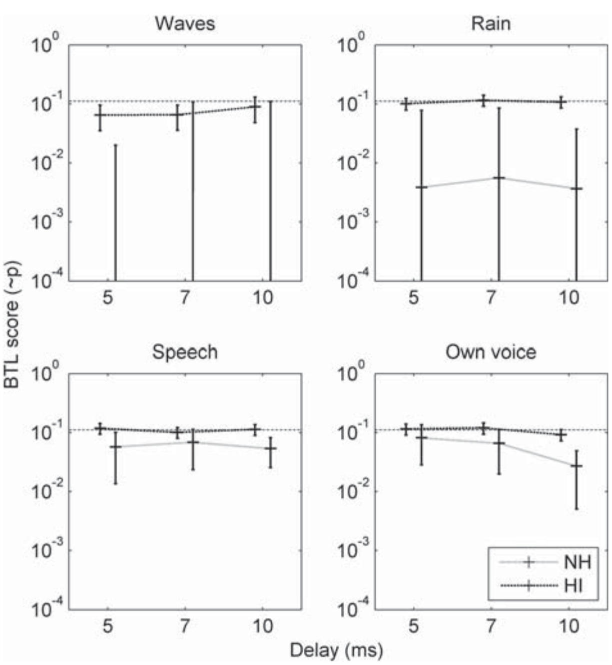

Figure 8. Main effect of delay for NH listeners (grey) and HI listeners (dotted). Geometric means are shown with associated 95% confidence intervals. There are no significant effects of delay.

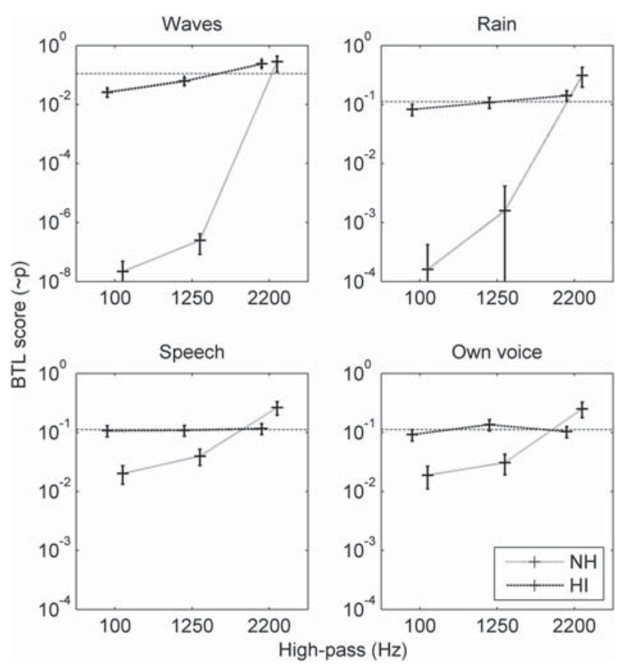

Figure 9. Main effect of high-pass cutoff for NH listeners (grey) and HI listeners (dotted). Geometric means are shown with associated 95% confidence intervals. Note different y-axis for 'waves'.

average audiogram of all test subjects (Figure 3), speech as 'average speech' material and a noise spectrum with the same spectral shape as the speech spectrum. The noise spectrum was shifted up and down to produce a range of signal-to-noise ratios: −5 dB, 0 dB, 10 dB, 20 dB.

For the four SNRs applied here, the following results were obtained across the nine test conditions:

– SNR = −5 dB: $0.238 < \text{SII} < 0.242$: 0% difference
– SNR = 0 dB: $0.384 < \text{SII} < 0.391$: 1% difference
– SNR = 5 dB: $0.669 < \text{SII} < 0.681$: 1% difference
– SNR = 10 dB: $0.835 < \text{SII} < 0.850$: 2% difference

It is clear from these results that there is no practical difference across the nine test conditions, so neither delay nor high-pass setting in the range used here affects predicted speech intelligibility.

As previously stated, the instruments used in the study had the noise reduction system disabled. Since most noise reduction system reduce gain in frequency bands for both noise and signal, based on the modulation in the band, the SNR in each band is not changed. Furthermore, with open fitting the effective gain reduction in the low frequencies is 0 dB, because of the direct sound bypassing the instrument. Thus, using traditional single-channel noise reduction will not change the above simulation results. If, on the other hand, a directional microphone is used, this will change the SNR in the affected frequency range. Thus, the noise reduction benefit obtained from a directional microphone will be reduced due to the high-pass filtering.

Discussion

From the BTL analysis and associated plots, the following observations were made:

Main effects, delay

– No significant effects of delay in either group (see Figure 8). The reason is not lack of audible difference but rather lack of preference; all delay artifacts are audible. The result is consistent with previous findings by Agnew & Thornton (2000) that delays above 3 ms are noticeable; and all delays in the present study are above this value and noticeable.

Main effects, high pass cut-off

– 2200 Hz is preferred over 1250 and 100 Hz in both groups and all conditions, except for HI 'speech' and HI 'own voice' (see Figure 9) where no preference is found. The latter two signals thus seem less sensitive to this type of artifact. There is nothing in the test results that points against using this highest cut-off for subjects in the audiogram range used here (Figure 3), even though they in principle need the amplification (for other purposes than sound quality). The above-mentioned speech intelligibility simulations confirm that there is no loss in speech intelligibility due to the proposed high-pass cutoff.
– 1250 Hz is preferred over 100 Hz in a few cases: NH 'waves' and 'speech' and HI 'waves'. 'Waves' is thus the most sensitive signal in this test. This can be explained by the random, broadband nature of the spectrum ensures components at all frequencies as required to detect the notches created by the comb-filtering process.

Interactions, delay × high-pass

Having established that delay changes have no systematic main effect on preference, and that 2200 Hz in general is the preferred high-pass cut-off, it is relevant to look at the effect of delay for the 2200 Hz cut-off:

– For 2200 Hz, the NH group prefers 5 and 7 ms over 10 ms for the 'rain' and 'own voice' signals, and prefers 7 and 10 ms over 5 ms for the 'waves' signal. The HI group does not have a preference for delay within the 2200 Hz conditions. For normal-hearing listeners this suggests that 10 ms delay can reduce sound quality for specific sensitive signals, although in some cases ('waves') in opposite direction. 7 ms is not critical at any time, for normal-hearing listeners.
– For hearing-impaired listeners there is no preference for any of the delays tested in the high-pass filtered condition.

In summary, the trend for the NH group and the HI group is the same, but the NH group is more sensitive.

No delay is better than the other; however 10 ms may be the upper limit regarding sound quality, as indicated in the 2200 Hz conditions for normal-hearing listeners. This is consistent with Groth and Søndergaard (2003) and Stone et al (2008) who both determined 10 ms as the upper limit for acceptable delay. Within the 2200 Hz high-pass conditions, there are specific and significant preferences for specific signals but they point in opposite directions for different signals. For a general listening program in a hearing aid it is thus not possible to recommend one delay over the other.

The lower limit for delay has not been found in this study, since even 5 ms caused audible artifacts in the 100 Hz broad band condition (REIG = 10 dB), which is comparable to Müsch et al (2004), who found 6 ms for 15 dB REIG and 15 ms for 0 dB REIG to be acceptable. Informal listening with a simulated hearing aid indicated approximately 2 ms to be the lower limit that was noticeable by a normal-hearing listener, quite similar to the 3–5 ms limit for found to be noticeable by Agnew and Thornton (2000). Below this limit there is no audible coloration from the comb-filter effect. The other quoted studies investigating the lower delay limit obtained higher results, because they looked at tolerable or acceptable delay.

The lower limit has other drawbacks in digital hearing aid applications because there is less time for beneficial signal processing and because anti-feedback control becomes increasingly difficult as the delay is reduced. Some digital hearing aids have frequency-dependent delay that is inversely proportional to frequency, with quite short delays at 2000–3000 Hz. Lower frequencies will have longer delays in order to provide enough gain shaping flexibility and therefore the comb-filter effects will end up being similar to those in a hearing aid with frequency-independent delay.

Regarding the upper limit of delay it may be even higher than 10 ms for hearing-impaired listeners but this has not been investigated. A hearing aid with 10 ms delay has good processing power in all aspects and there is typically no need to increase the delay further. The next limitation is expected to be 30 ms that cause speech-production problems (Stone & Moore, 2002).

The preference of high-pass filtering is important to keep in mind when designing the gain strategy in the fitting software. The low-frequency bands can be switched off to minimize comb filter effects. Recent results from NAL (Keidser et al, 2007) indicate that listeners prefer no increase in LF gain in vented fittings, confirming that it is preferred to leave these bands off so the direct sound will dominate. The present study was carried out for IG =

10 dB, which is lower gain than many typical fittings. However an increasing high-frequency gain would move the comb filter artifacts down in frequency if no high-pass filter was applied as shown in Figure 6. Consequently, the preferred filter cut-off will also eliminate the comb filter if they are moved down.

The effect on speech intelligibility was addressed in the previous section, and found not to be a complicating issue regarding both delay and high-pass.

In the case of compression—as in most commercial devices, the comb filter effect may be dynamic, appearing only when high input levels cause the compression to reduce insertion gain to near 0 dB. However it can be expected that the pitch of this dynamic phenomenon is equally bad for the any of the delays tested here and that clever control of gain can minimize this problem.

There was a good point in using environmental sounds as examples of sounds with little information (unlike speech) but still desired sounds. The 'rain' and 'wave' signals worked well for this purpose.

Conclusion

The present study investigated the perceived sound quality for hearing-aid processing delays in the range 5–10 ms and high-pass cutoffs in the range 100–2200 Hz. Eighteen normal-hearing (NH) and 18 hearing-impaired (HI) listeners were tested. The hearing instruments were fitted using open domes and the standard insertion gain was set to 10 dB linear for the closed vent situation. The instruments had no other signal processing, e.g. directional microphones, noise reduction, and digital feedback cancellation were switched off.

The main conclusion is that regarding sound quality there is no preference for delay in the range 5–10 ms. This is the case for both listener groups. The comb filter artifact is clearly audible and changes coloration pitch as a function of delay—however there is no preference for one coloration pitch over another.

The other dimension in the investigation is the effect of the three high-pass conditions in the range 100–2200 Hz. The NH subjects clearly preferred the 2200 Hz condition over the lower cut-offs, and the HI subjects showed the same tendency although only for the natural sounds. The result is especially interesting for the HI group who seem to prefer no gain—with respect to sound quality—in a frequency range where most fitting rationales would prescribe more than 0 dB gain for the hearing losses used here.

Simulations of speech intelligibility for a range of relevant signal-to-noise ratios showed no significant changes across all combinations of delay and high-pass, so there is no penalty in applying the above findings.

Looking at the preferred 2200 Hz high-pass alone, there is a small, but significant degradation of sound quality for 10 ms *for the NH group only and specific sensitive signals*. However, when considering all signals and hearing losses tested, a signal path delay up to 10 ms can be allowed without harming the sound quality.

Acknowledgements

The author wishes to thank Oticon colleagues for review and help on test design, Peter Djørup for hardware/software support, Niels Henrik Pontoppidan for hearing-aid firmware, and Jette Nissen for assistance with test subject scheduling. Furthermore thanks to all the test subjects, who volunteered for the listening test. Parts of this study was also presented as poster at IHCON 2008 'International Hearing Aid Conference', Lake Tahoe, USA, and at the Annual Meeting of the Danish Acoustical Society, March 2009.

Declaration of Interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Agnew J. & Thornton J.M. 2000. Just noticeable and objectionable group delays in digital hearing aids. *J Am Acad Audiol*, 11, 330–336.

ANSI 1997. Methods for Calculation of the Speech Intelligibility Index. *ANSI S3.5-1997*.

Bech S. & Zacharov N. 2006. *Perceptual Audio Evaluation: Theory, Method and Application*, Wiley.

Bradley R.A. 1984. Paired comparisons: Some basic procedures and examples. In: P.R. Krishnaiah & P.K. Sen (eds.), *Handbook of Statistics, Volume 4*. Amsterdam: Elsevier, 299–326.

Choisel & Wickelmaier. 2007. Evaluation of multichannel reproduced sound: Scaling auditory attributes underlying listener preference, *J Acoust Soc Am*, 121, 388–400.

Egner T.1953. *Klatremus og de andre dyr i Hakkebakkeskoven* (In Danish), Gyldendal.

Ellermeier W., Made M. & Danie P. 2004. Scaling the unpleasantness of sounds according to the BTL model: Ratio-scale representation and psychoacoustical analysis. *Acustica*, 90, 101–107.

Groth J. & Søndergaard M.B. 2003. Disturbance caused by varying propagation delay in non-occluding hearing aid fittings. *Int J Audiol*, 43, 594–599.

Hinkelmann K. & Kempthorne O. 2008. *Design and Analysis of Experiments, Volume 1, Introduction to Experimental Design*. Wiley.

Keidser G., Carter L., Chalupper J. & Dillon H. 2007. Effect of low-frequency gain and venting effects on the benefit derived from directionality and noise reduction in hearing aids. *Int J Audiol*, 46, 554–568.

Müsch H., Nooraei N., Edwards B. & Pavlovic C. 2004. Permissible delay in a partially occluding monaurally fit hearing aid. Poster presented at IHCON 2004, LakeTahoe, USA.

Pavlovic C., Hou J. & Müsch H. 2009. SII: Speech Intelligibility Index. http://www.sii.to

Stone M.A. & Moore B.C.J. 1999. Tolerable hearing aid delays. I. Estimation of limits imposed by the auditory path alone using simulated hearing losses. *Ear Hear*, 20, 182–192.

Stone M.A. & Moore B.C.J. 2002. Tolerable hearing aid delays. II. Estimation of limits imposed during speech production. *Ear Hear*, 23, 325–338.

Stone M.A. & Moore B.C.J. 2003. Tolerable hearing aid delays. III. Effects on speech production and perception of across-frequency variation in delay. *Ear Hear*, 24, 175–183.

Stone M.A. & Moore B.C.J. 2005. Tolerable hearing-aid delays IV. Effects on subjective disturbance during speech production by hearing-impaired subjects. *Ear Hear*, 26, 225–235.

Stone M.A., Moore B.C.J., Meisenbacher K. & Derleth R.P. 2008. Tolerable hearing aid delays. V. Estimation of limits for open canal fittings. *Ear Hear*, 29, 601–617.

Tversky A. & Sattah S. 1979. Preference trees. *Psychological Review*, 86, 542–573.

Wickelmaier F. & Schmid C. 2004. A Matlab function to estimate choice model parameters from paired-comparison data. *Behavior Research Methods, Instruments & Computers*, 36, 29–40.

Zimmer K., Ellermeier W. & Schmid C. 2004. Using probabilistic choice models to investigate auditory unpleasantness. *Acustica*, 90, 1019–1028.