# Detection mechanisms for processing delays in simulated vented hearing devices

Cite as: JASA Express Lett. **1**, 014402 (2021); https://doi.org/10.1121/10.0003064
Submitted: 04 September 2020 . Accepted: 30 November 2020 . Published Online: 12 January 2021

Florian Denk, Kristin Ohlmann, and Birger Kollmeier

JASA EXPRESS LETTERS ARTICLE asa.scitation.org/journal/jel

# Detection mechanisms for processing delays in simulated vented hearing devices

**Florian Denk,**[a] **Kristin Ohlmann, and Birger Kollmeier**
*Medizinische Physik and Cluster of Excellence Hearing4all, Universität Oldenburg, 26111 Oldenburg, Germany*
*florian.denk@uol.de, cynthia.kristin.ohlmann@uol.de, birger.kollmeier@uol.de*

**Abstract:** Processing delays are a disturbing factor in hearing devices, especially with vented or open fits. While the disturbance due to delays is well characterized, neither have the perception thresholds of delays been systematically assessed, nor are the perceptual detection mechanisms clear. This study presents experiments determining the delay detection thresholds in simulated linear vented hearing devices in normal-hearing listeners, where spectral effects of delays were either compensated or not. Furthermore, the psychometric function for the detection of delays was determined for an example condition and linked to model predictions, showing that delay detection can be well predicted from spectral artefacts. 

[Editor: Chatterjee Monita] https://doi.org/10.1121/10.0003064

**Received:** 4 September 2020 **Accepted:** 30 November 2020 **Published Online:** 12 January 2021

## 1. Introduction

Digital hearing devices like hearing aids or hearables allow for flexible advanced signal processing techniques like adaptive noise reduction or dynamic compression tailored to the needs of a user. However, the benefit of digital processing comes at the cost of a processing delay that is typically within several milliseconds.[1] Effects of processing delays have been reported to be disturbing, produce a hollow sound, and can have negative effects on the ability to localize sounds.[2–5] A processing delay is most critical when the device is vented, i.e., a leakage component that directly enters the ear canal may interfere with the delayed output of the hearing device. Since most modern hearing aids and many hearables feature a vented, loose, or open fit to improve the wearing comfort, the disturbing effects of processing delays are one critical issue that limit the performance and acceptance of current hearing devices.

It is well established that increasing delays cause increasing disturbances.[2,3] Agnew and Thornton[4] reported on just noticeable delays in closed-fit hearing aids of 2–6 ms with a direct switching paradigm and trained subjects, and other studies confirmed that delays in this range are perceptible and slightly annoying.[2,3] However, to our best knowledge, the thresholds for the detection of processing delays have not been assessed systematically. In consequence, it is not clear what the exact perceptual mechanism for the detection of processing delays looks like. While a delay of several milliseconds is too short to be perceived as an echo, the interference of leakage and hearing device output leads to spectral ripples, which can provide a prominent spectral cue. Stone *et al.*[2] included a reduction (but no complete elimination) of delay-generated spectral ripples by means of a post-filter and found only a minor effect on the annoyance of the delay. The effect of complete elimination of spectral cues on the delay detection threshold would give important insights on the perceptual detection mechanisms, although such a compensation of spectral ripples is arguably not applicable in real hearing devices.

We here present a psychoacoustic study that systematically assesses the detection thresholds and psychometric functions for processing delays in simulated linear hearing devices, with the aim to pinpoint the underlying perceptual mechanisms. To this end, a simple but representative simulation of hearing devices similar to previous studies based on a pair of low- and high-pass filters was used.[2,5] To test the importance of spectral cues for the detection of delays, spectral ripples caused by the delay were either compensated by a high-order post-filter or not. Detection thresholds for processing delays were determined in normal-hearing listeners using an adaptive up-down procedure for different simulated linear hearing devices, where acoustic parameters representing the vent size and insertion gain of the hearing device were varied. For one example hearing device configuration, the psychometric function around and above the threshold was also determined. Furthermore, the detection of delay artefacts was modeled by regarding the spectrotemporal resolution of the human auditory system as represented by a Gammatone filterbank.[6] This study thus provides insights into how humans

[a]Present address: German Institute of Hearing Aids, 23562 Lübeck, Germany. Author to whom correspondence should be addressed, ORCID: 0000-0003-3490-123X.

detect the processing delay of hearing devices. This knowledge could be exploited for the reduction of the perceivable artefacts of delays and thus improve the performance and sound quality of future hearing devices.

## 2. Methods

### *2.1 Simulation of hearing devices and stimuli*

Listening through linear vented hearing devices was simulated by convolving stimuli with an impulse response that was constructed as shown in Fig. 1. A vent causes both the leakage of low-frequency sounds into the ear canal and acoustic high-pass filtering of the hearing device output, where both cut-off frequencies depend on the vent size. Hearing device transmission was thus simulated by superimposing a low-pass component that represents the leakage, and a high-pass component including a gain and variable delay that represents the hearing device output. High- and low-pass filters were implemented as second-order Butterworth filters. The high-pass filter was phase-matched to the low-pass filter such that the addition was fully constructive. An additional compensation filter (see below) applied to the sum of low- and high-pass component assured an either coarsely or exactly flat frequency response of the simulated device. Finally, playback from headphones resulted in an overall frequency response that matched the open ear. Calculations and sound presentation were conducted with a sampling rate of 192 kHz to circumvent the necessity to implement fractional sample delays. This simple simulation of hearing devices is similar to previous studies[2,5] and represents the *relative* contributions of leakage and hearing device output at the eardrum of a user, while the overall frequency response is appropriate for normal-hearing listeners. To test the importance of spectral effects of delays, two different compensation filter designs were included:

- Peak compensation: Inverse of the magnitude response for a delay of zero (see "Intermediate" and upper "Transfer function" panels of Fig. 1, light gray lines), independent of the applied delay. This compensates for a peak that is generated around $f_c$ with 0 delay due to constructive superposition of leakage and hearing device output, as well as for the high-frequency amplification $G_{HP}$. As a result, for all conditions a flat response is achieved at 0 delay, while higher delays produce spectral notches but no peaks within the spectral interaction region. In real hearing devices, very similar spectral ripple artefacts are created.
- Ripple compensation: Inverse of the magnitude response for a given delay (see "Intermediate" and lower "Transfer function" panels of Fig. 1, dark gray lines). This results in a flat frequency response despite the presence of the processing delay. The necessary compensation filter has a long impulse response and results in an addition of delayed components in the impulse response, somewhat similar to reverberation (see Fig. 1, bottom right).

Both compensation filters were implemented as minimum-phase FIR filters with the target magnitude response. Due to the chosen implementation, the compensation filter order was slightly, but for the compensation accuracy insignificantly, dependent on the delay [$80\,004 + 768 \cdot$ delay (ms)]. This corresponds to impulse response lengths between 415 and 450 ms and 2.2–2.4 Hz frequency resolution in the range of obtained delay thresholds.

Different characteristics of the hearing device were implemented by changing the center frequency of the overlap between low- and high-pass component $f_c$ (350, 700, and 1400 Hz), the overlap width $\delta_f$ (difference between $-3$ dB cut-off

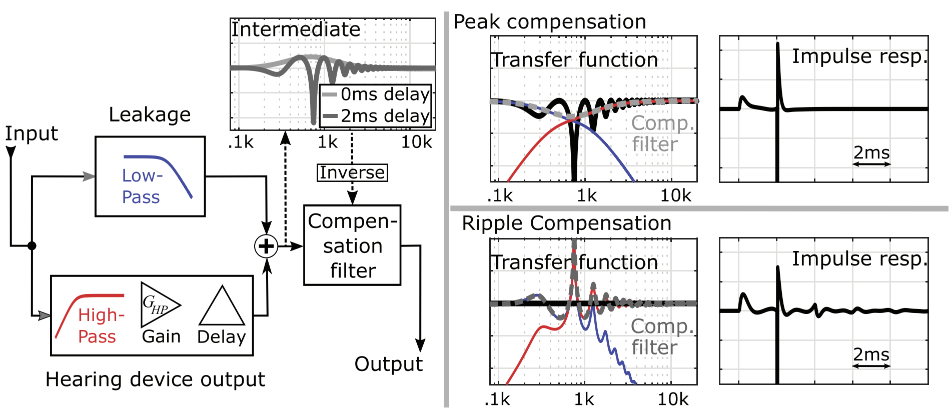


Fig. 1. Left: Flow chart of hearing device simulation. Low- and high-pass filters simulate the leakage and delayed hearing device output, respectively. A compensation filter is applied to the sum of low- and high-pass filter output, creating the output signal. The compensation filter has the inverse magnitude response of the intermediate response for either 0 delay (peak compensation) or the given delay (ripple compensation, here shown for 2 ms). Right: Examples of the resulting transfer functions and impulse responses with peak (top) and ripple compensation (bottom) and 2 ms delay, $f_c = 700$ Hz overlap center frequency, $\delta_f = 4$ ERBn overlap width, and $G_{HP} = 0$ dB high-pass gain. Black lines denote the output, dashed gray lines the compensation filters, and red and blue lines the parts of the output made up by the hearing device output and leakage, respectively.

frequencies, set to 1, 4, and 7 equivalent rectangular bandwidth numbers (ERBn) of auditory filters in normal-hearing subjects), and the gain of the high-pass component $G_{HP}$ (0, 10 dB). These parameters cover the range of vent sizes from loose-fit headphones or an earmould with an approximately 1 mm hole ($f_c = 350$ Hz) via vents with 1–2 mm diameter ($f_c = 700$ Hz) to open-fit hearing aids ($f_c = 1400$ Hz),[7] as well as different relative levels between leakage and hearing device output associated with insertion gains appropriate for near-to-normal-hearing users. Note that the compensation filter is seen as another experimental parameter.

Three different stimuli were considered: pink noise, a speech snippet (spoken name "Anna," male voice, from OLLO database[8]) and a click sound (snapping finger[9]) The stimuli had a length of 700 ms, including 400 ms of silence that included output of the compensation filters, with a pause of 500 ms between intervals. They were presented at 62 dB SPL equivalent to the free field on Sennheiser HD650 headphones while the subjects were seated in an auditory booth.

### 2.2 Experiment I: Detection thresholds

In the first experiment, thresholds for the detection of delays were determined using an adaptive 1-up 2-down 3-interval 1-alternative forced choice procedure (yielding the 70.1% detection point).[10] For 36 parameter combinations as shown in Sec. 3.1, the delay was utilized as the experimental variable and the delay perception threshold determined with respect to a reference stimulus with identical processing but the delay set to 0. The order of conditions was randomized, and no interleaving was included. The start value of the delay was chosen individually for each condition, such that the delay was clearly audible but not too far away from a sample threshold pre-determined by one author who did not participate in the experiment. The delay was reduced in steps of 1/2, $1/2^{2.5}$, and $1/2^4$ of the difference between start value and sample threshold on each upper reversal, and the mean of four reversals in the measurement phase was taken as the threshold. Ten normal-hearing subjects participated (five male, five female, aged 20–34 yr, hearing thresholds better than 20 dB HL for frequencies $\leq$ 8 kHz, including one author).

### 2.3 Experiment II: Psychometric functions

For one example condition ($f_c = 700$ Hz, $\delta_f = 1$ ERBn, $G_{HP} = 0$ dB) and both compensation approaches, the psychometric function of the delay detection was determined in a second experiment. A 3-interval 1-alternative forced choice paradigm was used with the constant stimuli method, where each 20 repetitions of 15 different delays around and above the threshold determined in experiment I were presented. Furthermore, the thresholds with the adaptive procedure were determined in the same way as above. Six subjects (four males, two females, aged 21–38 yr, hearing thresholds better than 20 dB HL for frequencies $\leq$ 8 kHz, independent of experiment I, including one different author) participated in the second experiment.

### 2.4 Auditory modeling

The perception of delays was additionally analysed through a simple auditory model consisting of a linear Gammatone filterbank with 1-ERBn wide auditory filters covering the full audible range, as implemented by Hohmann.[6] To avoid spectral sampling artefacts, the spacing of auditory filters was set to 0.1 ERBn. An auditory difference between the reference and a test stimulus was determined by subtracting the average power outputs (in dB) of the auditory filterbank. From this auditory difference spectrum, the maximum absolute value across all auditory filters $\Delta_{Aud,Max}$ was computed. Taking the root-mean-square average across all auditory filters yielded very comparable results but less differentiated values; thus, only the maximum is considered here.

## 3. Results

### 3.1 Delay perception thresholds

Figure 2 shows the obtained delay detection thresholds, summarizing the results with both compensation filters and all hearing device configurations made up by combinations of $f_c$, $\delta_f$, and $G_{HP}$. Conditions have been sorted along the $x$ axis by the highest frequency for which the level difference between low- and high-pass component is $\leq$ 3 dB, referred to as highest interaction frequency $f_h$. Note that in the conditions with $G_{HP} = 10$ dB, $f_h$ lies below the nominal interaction region because of the gain on the high-pass component. Each color represents one stimulus, filled markers indicate results with peak compensation, and results with ripple compensation are denoted by hollow markers.

Generally, the thresholds depend strongly on the condition and lie in a range between 0.1 and 8 ms. The arrangement in Fig. 2 reveals an approximate $1/f_h$ dependence of most perception thresholds. The $1/f_h$ dependence is stronger for peak compensation, where most delay perception thresholds with speech and pink noise can be approximated reasonably well as $0.35/f_h$ (see dotted line). Notches at a specific frequency are generally created by delays of $\left(n+\frac{1}{2}\right)/f$ (see dashed lines). Many thresholds obtained with ripple compensation lie at delays that fulfill this dependence for the appropriate highest interaction frequency $f_h$ (see especially speech stimulus). For peak compensation, the thresholds for the stimuli pink noise and speech are quite consistent, besides some exceptions that are discussed in Sec. 4, and with slightly lower thresholds with the pink noise. The thresholds with the click stimulus are generally higher and have no clear $1/f_h$

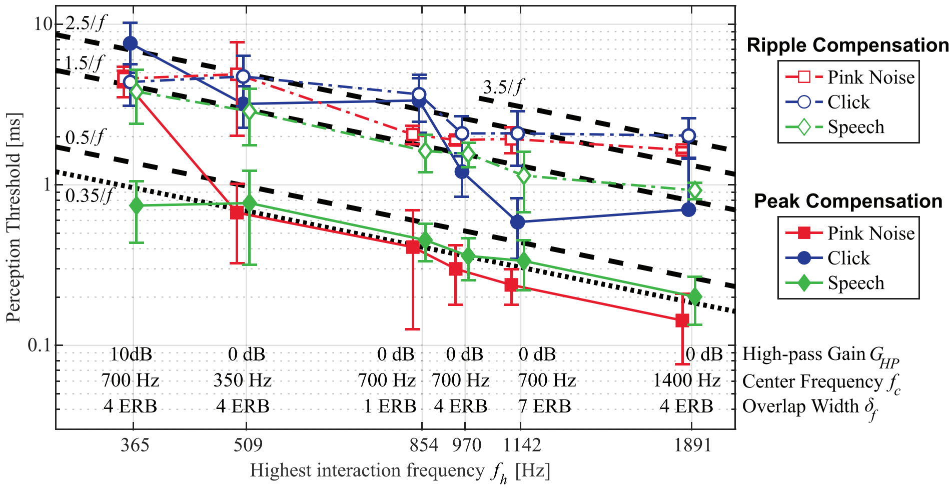


Fig. 2. Determined delay perception thresholds for all assessed hearing device parameters (see bottom of panel), compensation filters (filled/hollow markers), and stimuli (indicated by color). Each combination of $f_c$, $\delta_f$, and $G_{HP}$ is characterized and positioned on the $x$ axis by the highest interaction frequency $f_h$, i.e., the highest frequency where the level difference between low- and high-pass component is $\leq 3$ dB. Symbols show mean $\pm$ 1 standard deviation across subjects. Black lines denote $n/f$ functions for reference.

dependence. For ripple compensation, the lowest thresholds tend to be seen with the speech stimulus, and the result for the click stimulus is much more consistent with the other two stimuli.

### 3.2 Psychometric functions

The top panels of Fig. 3 show the psychometric functions, i.e., percentage of detections of a delay as denoted on the $x$ axis, for a hearing device configuration with $f_c = 700$ Hz, $\delta_f = 1$ ERBn, and $G_{HP} = 0$ dB with a pink noise stimulus. The left and right panels show the results for peak and ripple compensation, respectively. The psychometric function looks very different for both compensation approaches but shows a non-monotonic behavior in both cases. For peak compensation, it increases quickly for delays below the average threshold for this condition in experiment I (0.4 ms, vertical line), reaches near 100% just above 0.5 ms, and then slightly decreases again down to 87% around 1.5 ms. At delays larger than 1 ms, the variation between subjects is larger than in the ascending area. The subjects participating in experiment II reached a 70.1% detection, as expected from the threshold in the adaptive procedure, at lower delays than the average of experiment

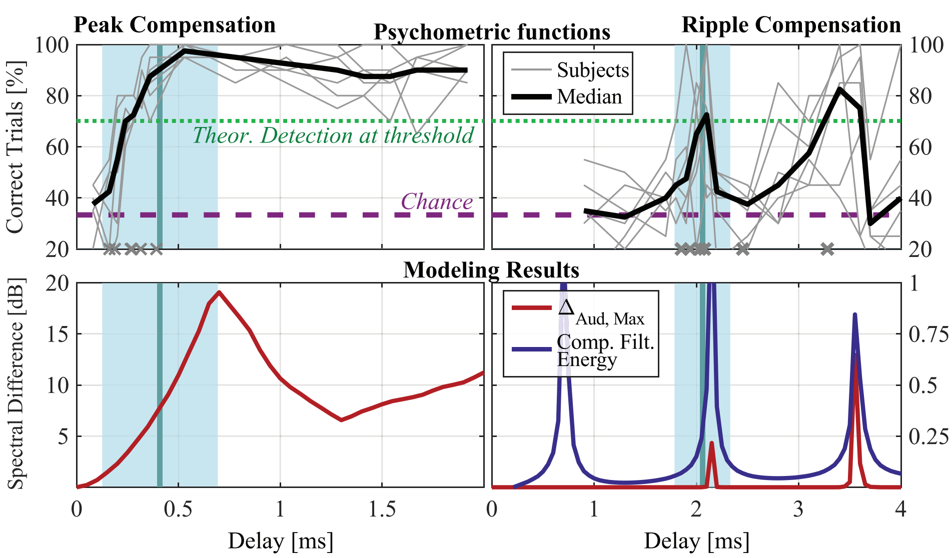


Fig. 3. Top: Individual and median psychometric functions (lines) and thresholds with the adaptive procedure (crosses on $x$ axis) from experiment II. Bottom: Maximum auditory spectral difference between reference and test stimuli depending on delay; for ripple compensation (right), the broadband energy of the compensation filter is also given. The vertical light blue line and shade denote the average threshold and standard deviation for this condition from experiment I. Pink noise stimulus, $f_c = 700$ Hz, $\delta_f = 1$ ERBn, $G_{HP} = 0$ dB.

I. This is, however, consistent with their individual thresholds determined within experiment II (see gray crosses on $x$ axis, top left panel of Fig. 3). For ripple compensation, the psychometric function shows distinct peaks around delays of 2.14 and 3.55 ms, which corresponds well to $1.5/f_c$ and $2.5/f_c$ for this condition (cf. Sec. 3.1). Between the peaks, the detection rate approaches the chance level of 1/3. Here, the thresholds expected from the psychometric function are well in line with both the average threshold from experiment I and the thresholds determined in experiment II using the adaptive procedure. While for most subjects, they lie at the lowest peak of the psychometric function, in one subject the adaptive procedure yielded a threshold at the second peak (~3.3 ms).

### *3.3 Auditory modeling*

The bottom panels of Fig. 3 show the maximum of the auditory difference spectrum $\Delta_{Aud,Max}$ depending on delay. Please note the different $y$-scale between the left and right panels showing results for peak and ripple compensation, respectively. For ripple compensation, the broadband energy of the compensation filter (in dB) is also given. The auditory difference curves are very consistent with the psychometric functions and resemble their non-monotonic behavior.

For peak compensation, the $\Delta_{Aud,Max}$ increases with delay up to approximately 0.7 ms, where the spectral deviation reaches 19 dB. For larger delays, it decreases again until a minimum of 6.5 dB is reached around 1.3 ms, after which the spectral difference increases again. The locations of maxima and minima roughly correspond to the characteristic points of the psychometric function; the spectral difference for the average delay threshold from experiment I for the assessed condition is 7.6 dB. For ripple compensation, the spectral difference shows distinct narrow peaks around the same delays as the psychometric function. However, the maximum auditory spectral difference is below 0.25 dB at the first peak and around 0.75 dB at the second peak, i.e., much lower than for peak compensation around threshold. At these peaks (and also at 0.71 ms or $0.5/f_c$), the energy of the compensation filter also shows very distinct peaks, which we verified to originate mainly from a gain at $f_c$ (e.g., 40 dB at 2.1 ms). Thus, at delays around $\left(N+\frac{1}{2}\right)/f_c$, corresponding to multiples of the inverse of the frequency where the most critical superposition of low-and high-pass component occurs, the ripple compensation could not completely suppress spectral effects of delays. Since no gain limitation was imposed on the compensation filter, the limiting factor was probably the discrete Fourier transform (DFT) frequency resolution of approximately 2.3 Hz.

## 4. Discussion

The obtained delay detection thresholds depend significantly on the acoustic parameters of the simulated hearing devices and the compensation filter and lie in a range between 0.1 and 8 ms. Previous studies reported higher typical thresholds, presumably because annoyance instead of detection was tested,[2,3] and non-vented fits were used when detection was tested in a comparable paradigm.[4] It should be stressed that the present thresholds are not necessarily the values at which delays become prominent or annoying in real hearing devices. The reported thresholds rather constitute a lower boundary for the detection of highly controlled artefacts created by delays and thus allow an inference on the perceptual mechanisms involved.

The presented detection thresholds with peak compensation, which resembles real hearing devices reasonably well, can be consistently explained by a spectral detection mechanism at least for pink noise and speech as a stimulus. The superposition of leakage and delayed hearing device output creates spectral notches that move towards lower frequencies with increasing delay (dashed lines in Fig. 2). However, these notches can only occur within the interaction region of low-pass and high-pass components, since a large level difference between both components would avoid considerable interference effects. The perception thresholds therefore occur at delays where the notches approach the upper end of the interaction region as characterized by $f_h$, such that a noticeable spectral difference is created. The auditory model assessment showed that the thresholds are well described by a difference of 5–8 dB in the most critical auditory filter band, also for other conditions not explicitly shown in Fig. 3. These values are in line with previous notch detection experiments.[11,12] Spectral differences are more easily detected in speech or pink noise as compared to the short click stimulus, for which much higher, and inconsistent, thresholds were observed. Omitting the peak compensation filter would not influence this interpretation, since it affects the test and reference spectrum (with 0 delay) equally.

If the delay gets increased beyond threshold, the primary notch at $f = 0.5/\text{delay}$ may move beyond the lower frequency limit of significant interaction between low- and high-pass component, resulting in a decreasing spectral difference that may get close to the spectral difference at threshold. This is seen both for the psychometric functions and the model calculations shown in Fig. 3 (left panels). Another increase of spectral difference with increasing delay is then seen if the secondary notch at $f = 1.5/\text{delay}$ moves into the interaction region. The non-monotonicity of the perception cue means that the measured threshold can be sensitive to the start value and step size. This issue indeed shows up in the present data in the form of outliers, e.g., the threshold with pink noise in the leftmost condition of Fig. 2. This threshold is about a factor of 3 higher than the appropriate threshold with the speech stimulus but coincides well with the delay creating a secondary notch at the upper end of the interaction region (delay $= 1.5/f_h$). Since the start value of the adaptive procedure was apparently chosen too high, the experiment consistently converged at a local minimum of the spectral difference. It should be noted that due to the high-pass gain in this condition, the significant interaction region is particularly narrow, since it lies about an octave below the high-pass cut-off frequency. This result verifies that appropriate filtering of the hearing device output to minimize the interaction region with the leakage can reduce the perceptual effects of delays.[2,13,14] With narrow interaction regions as in the example above it could be possible to tailor (potentially *increase*) the delay such that no notches are created inside the interaction region. Effectively, the spectral artefacts due to delays are then shifted into frequency regions where they have no effect due to level differences between the leakage and the hearing device output.

If the spectral effects of hearing device delays were compensated through a post-filter in the ripple compensation condition, detection thresholds rose by a factor of 3–5 as compared to the peak compensation condition (cf. Fig. 2). However, a similar dependence on $f_h$ was noted, and the thresholds lie at delays that create a particularly deep notch. A comparison of the psychometric function and residual spectral cues (Fig. 3) shows that detection occurred at selected delays where a small (<1 dB) residual spectral difference is noted after the compensation. The detection cues for ripple compensation are thus residual spectral cues or, more likely, compensation artefacts due to high amplification at single frequencies (cf. compensation filter energy in Fig. 3) that can create reverberation-like effects, rather than a temporal "echo" detection of the delay. Similar artefacts are expected to be much more prominent in practical applications, e.g., through high amplification of microphone noise. Nevertheless, the results with ripple compensation highlight that perceptual effects of processing delays at or near threshold are driven by spectral ripple artefacts.

## 5. Conclusion

1. Processing delays in vented hearing devices, simulated here by a delayed and high-pass hearing device output and the low-pass leakage component in superposition, appear to be detected primarily by spectral cues, specifically spectral notches in the interaction region between both signal paths. Just detectable delays created a spectral difference of approximately 5–8 dB at the most critical auditory filter band.
2. Delay artefacts may in some cases be reduced by *increasing* the delay, since the notches may be moved out of the spectral interaction region.
3. Compensation of spectral ripples by means of high-order filters led to higher delay detection thresholds in the present experiment but is probably not applicable in practice due to computational effort, high amplification of microphone noise, and potential instabilities.
4. The consistency between psychoacoustic measurements and a simple auditory spectral detection model provides evidence that such models can be employed to optimize the perceived quality of vented hearing devices.

## Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project ID 352015383, SFB 1330 *HAPPAA* A4.

## References and links

[1] B. Kollmeier and J. Kiessling, "Functionality of hearing aids: State-of-the-art and future model-based solutions," Int. J Audiol. **57**(sup3), S3–S28 (2018).
[2] M. A. Stone, B. C. J. Moore, K. Meisenbacher, and R. P. Derleth, "Tolerable hearing aid delays. V. Estimation of limits for open canal fittings," Ear Hear. **29**(4), 601–617 (2008).
[3] J. Groth and M. B. Søndergaard, "Disturbance caused by varying propagation delay in non-occluding hearing aid fittings," Int. J Audiol. **43**(10), 594–599 (2004).
[4] J. Agnew and J. M. Thornton, "Just noticeable and objectionable group delays in digital hearing aids," J. Am. Acad. Audiol. **11**(6), 330–336 (2000).
[5] F. Denk, S. D. Ewert, and B. Kollmeier, "On the limitations of sound localization with hearing devices," J. Acoust. Soc. Am. **146**(3), 1732–1744 (2019).
[6] V. Hohmann, "Frequency analysis and synthesis using a Gammatone filterbank," Acta Acust. united Acust. **88**(3), 433–442 (2002).
[7] F. Kuk, D. Keenan, and C. Lau, "Comparison of vent effects between a solid earmold and a hollow earmold," J. Am. Acad. Audiol. **20**, 480–491 (2009).
[8] B. T. Meyer, T. Jürgens, T. Wesker, T. Brand, and B. Kollmeier, "Human phoneme recognition depending on speech-intrinsic variability," J. Acoust. Soc. Am. **128**(5), 3126–3141 (2010).
[9] M. Fix, "Finger snapping sound" (2012), http://soundbible.com/1978-Finger-Snapping.html (Last viewed 9/30/2019).
[10] S. Ewert, "AFC—A modular framework for running psychoacoustic experiments and computational perception models," in *Proceedings of the International Conference on Acoustics AIA-DAGA*, Merano, Italy (March 18–21, 2013), pp. 1326–1329.
[11] B. C. J. Moore, S. Oldfield, and G. J. Dooley, "Detection and discrimination of spectral peaks and notches at 1 and 8 kHz," J. Acoust. Soc. Am. **85**(2), 820–836 (1989).
[12] B. Caswell-Midwinter and W. M. Whitmer, "Discrimination of gain increments in speech," Trends Hear. **23**, 2331216519886684 (2019).
[13] L. Bramsløw, "Preferred signal path delay and high-pass cut-off in open fittings," Int. J. Audiol. **49**(9), 634–644 (2010).
[14] F. Denk, H. Schepker, S. Doclo, and B. Kollmeier, "Equalization filter design for achieving acoustic transparency in a semi-open fit hearing device," *Proceedings of the 13th ITG Conference on Speech Communication*, Oldenburg, Germany (October 10–12, 2018), pp. 226–230.