Taylor & Francis
Taylor & Francis Group

International Journal of Audiology

ISSN: 1499-2027 (Print) 1708-8186 (Online) Journal homepage: http://www.tandfonline.com/loi/iija20

Tolerable delay for speech production and perception: effects of hearing ability and experience with hearing aids

Tobias Goehring, Josie L. Chapman, Stefan Bleeck & Jessica J. M. Monaghan*

To cite this article: Tobias Goehring, Josie L. Chapman, Stefan Bleeck & Jessica J. M. Monaghan* (2017): Tolerable delay for speech production and perception: effects of hearing ability and experience with hearing aids, International Journal of Audiology, DOI: 10.1080/14992027.2017.1367848

To link to this article: http://dx.doi.org/10.1080/14992027.2017.1367848

Published online: 24 Aug 2017.

Taylor & Francis
Taylor & Francis Group

International Journal of Audiology 2017; Early Online: 1–8

International Journal of Audiology

Check for updates

Original Article

Tolerable delay for speech production and perception: effects of hearing ability and experience with hearing aids

Tobias Goehring, Josie L. Chapman, Stefan Bleeck, and Jessica J. M. Monaghan*

Institute of Sound and Vibration Research, University of Southampton, Southampton, UK

The British Society of Audiology

The International Society of Audiology

NAS NORDIC AUDIOLOGICAL SOCIETY

Abstract

Objective: Processing delay is one of the important factors that limit the development of novel algorithms for hearing devices. In this study, both normal-hearing listeners and listeners with hearing loss were tested for their tolerance of processing delay up to 50 ms using a real-time setup for own-voice and external-voice conditions based on linear processing to avoid confounding effects of time-dependent gain. *Design*: Participants rated their perceived subjective annoyance for each condition on a 7-point Likert scale. *Study sample*: Twenty normal-hearing participants and twenty participants with a range of mild to moderate hearing losses. *Results*: Delay tolerance was significantly greater for the participants with hearing loss in two out of three voice conditions. The average slopes of annoyance ratings were negatively correlated with the degree of hearing loss across participants. A small trend of higher tolerance of delay by experienced users of hearing aids in comparison to new users was not significant. *Conclusion*: The increased tolerance of processing delay for speech production and perception with hearing loss and reduced sensitivity to changes in delay with stronger hearing loss may be beneficial for novel algorithms for hearing devices but the setup used in this study differed from commercial hearing aids.

Key Words: Speech perception, psychoacoustics/hearing science, hearing aids, hearing aid satisfaction

Introduction

State-of-the-art hearing devices such as hearing aids employ digital signal processing to improve perceptual aspects of audio signals for the user. When an audio signal is processed by a hearing aid, the air-conducted signal produced by the device is delayed relative to the direct signal that enters the ear canal. The addition or interaction of these two signals may be perceived by hearing aid users and cause annoyance. The amount of delay depends on the device; in contrast to analogue hearing aids that allow for negligible throughput delays between microphone and receiver (under 1 ms; Dillon et al. 2003), digital technology introduces delays in the range of several milliseconds (typically 2–10 ms). As processing delay increases, and as the levels of the direct and processed sound become more equal at the eardrum, sound quality deteriorates and user annoyance ratings increase (Stone and Moore 1999). In order to avoid this problem, hearing devices are designed to operate with low delay (typically <10 ms, as suggested by Groth and Sondergaard 2004; Stone et al. 2008; Bramslow 2010). However, this requirement of low processing delay places a strong restriction on the development of novel algorithms for digital hearing devices. If larger delays can, in fact, be tolerated by typical users, this would allow for the use of more sophisticated signal processing techniques and the integration of information over wireless connections (e.g. for binaural algorithms or integration of smartphones), for which a much greater delay is required.

The perceptual effects of processing delay depend on its magnitude (Stone and Moore 1999). With longer processing delays, the desynchronisation of audio–visual information (>80 ms; Summerfield 1992) and auditory-proprioceptive feedback (>50–200 ms; Stuart et al. 2002), perception of distinct echo sounds (>50 ms; Litovsky et al. 1999) or altered speech production rates (>30 ms; Stone and Moore 2002; Stuart et al. 2002) have been reported. With shorter delays, a comb-filtering effect may dominate the disruption of speech perception. The comb-filtering effect results from superposition of the direct and delayed sounds. The superposition introduces spectral ripples spaced at the inverse of the delay and leads to a spectral colouration affecting the timbre of the sound that is perceived as unnatural or annoying. A natural situation in which comb filtering is produced is when a speaker is listening to their own voice. Here, the sound is transmitted via bone conduction

*Now at The Australian Hearing Hub, Sydney, Australia
Correspondence: Tobias Goehring, Institute of Sound and Vibration Research, University of Southampton, UK. Email: goehring.tobias@gmail.com

(Received 10 October 2016; revised 3 August 2017; accepted 7 August 2017)
ISSN 1499-2027 print/ISSN 1708-8186 online
DOI: 10.1080/14992027.2017.1367848

and via air conduction to the cochlea. The speech reaches the cochlea via these two paths with different delays (about 0.7 ms difference; Stromsta 1962) and the two signals sum to result in a comb-filtering effect. Long-term acclimatisation to the spectral colouration produced during own-voice listening is thought to be one of the reasons why short delays below about 4 ms are not perceived as bothersome or even noticeable by human listeners (Agnew and Thornton 2000; Groth and Sondergaard 2004, Zakis et al. 2012). However, for longer delays in the range of 10–50 ms, the increased spectral colouration due to the comb-filtering effect and the starting perception of echoes leads to a disturbing percept for the listeners (Agnew and Thornton 2000; Stone and Moore 1999, 2002, 2005).

Stone and Moore performed a series of experiments to investigate subjective tolerance of processing delay. In one study (Stone and Moore 1999), normal-hearing (NH) participants listened via headphones to recorded stimuli that were obtained by mixing above-ear and in-ear recordings of a person while talking, and processed with hearing aid and hearing loss simulations. Four different hearing losses from mild to moderately severe were simulated to account for the effects of threshold elevation and loudness recruitment. Results showed monotonically increasing disturbance with increasing delay. They found the higher the degree of simulated hearing loss, the greater the tolerance to delay (tolerable up to 40 ms for moderately severe losses). They concluded that the delay should not exceed 20 ms for mild to moderate hearing losses. Agnew and Thornton (2000) also tested delay tolerance for expert NH participants listening to their own voice while reading without using hearing loss simulations. Participants controlled the level of the delayed sound and the amount of delay while being able to compare the delayed condition to the undelayed condition at any point. They found delays about 14 ms to be objectionable. However, their experimental design prevented short-term acclimatisation effects that were later reported to increase tolerance to delay (Stone and Moore 2005) and they reported that the use of expert listeners can be seen as a worst case in terms of delay tolerance. Stone and Moore (2002) provided NH participants with behind-the-ear devices using real-time processing that used 0 dB insertion gain (at 65 dB SPL) and asked them to rate their disturbance during speech production (reading from a book) for a range of delays (from 7 to 43 ms) without hearing loss simulations. Again, they found tolerance limits of about 15–20 ms depending on the acoustic environment and when defining a rating of 3 out of 7 as the tolerance limit (with 1 labelled as ‘not at all disturbing’, 4 labelled as ‘disturbing’ and 7 labelled as ‘highly disturbing’). No effects on speech production were reported up to delays of about 30 ms.

A similar experiment was performed using participants with hearing loss (HL) fitted with behind-the-ear devices with real-time processing and using a realistic gain prescription to achieve the same overall loudness as would be perceived by NH participants (Stone and Moore 2005). The delay tolerance limits were about 14–30 ms, with a non-monotonic effect of low frequency hearing loss on delay tolerance. The participants were partly unblinded and trained to hear the effects of changing delay settings. The signal processing implemented a four-channel, fast-acting, wide-dynamic range compression as commonly used in hearing aids. The very low compression threshold of 32 dB SPL and attack and release times of around 9 ms introduced level- and time-dependent variations to the processed signal, but not to the direct signal and thus may have interacted with the perceptual effect of delay. The time-varying dynamics of the processed signal and the direct signal were further increased by the use of a non-linear gain prescription with level compression. Although this processing is similar to what would be used in commercial hearing aids, it is unclear to what degree the subjective ratings of delay tolerance were affected by the introduction of non-linear processing. The authors suggest that this processing may have led to the non-monotonic effect of low frequency hearing loss on delay tolerance and emphasise that the use of other prescription methods and compression parameters may lead to a different pattern of results. Furthermore, perception of external speech was not assessed in this study.

In another study, Stone et al. (2008) investigated delay tolerance for NH participants with open-canal fittings and several types of gain prescription for the delayed signal. Participants were listening to and asked to rate above-ear and in-ear recordings of a person reading that were mixed and processed with a hearing loss and hearing aid simulation similar to the methodology followed by the experiment in Stone and Moore (1999). In this study, participants were much more sensitive and ratings reached tolerance thresholds at delays of just 5–6 ms. However, also in this study, the authors argued that the use of dynamic range compression in combination with a hearing loss simulation for loudness recruitment introduced disturbing dynamic side effects that could have interacted with the disturbing effects of delay and made the interpretation of results difficult. Several studies reported no significant effects on delay tolerance or preference for delays up to 10 ms for HL participants using own-voice and external sounds (Groth and Sondergaard 2004; Bramslow 2010; Zakis et al. 2012) and support the common choice of hearing aid manufacturers to restrict processing delay to values below 10 ms. However, in these studies, the greatest tested delay was 10 ms or less, providing no information about tolerance of longer delay.

As a whole, the literature has been taken to support an upper limit of 10 ms for processing delay. Current hearing devices are constrained to use processing delays below this limit, restricting the development of more complex signal processing methods such as acoustic feedback and noise cancellation algorithms (Dillon et al. 2003). Recommendations in the literature were predominantly based on data obtained using NH participants (Stone and Moore 1999, 2002; Stone et al. (2008) Agnew and Thornton 2000) and there are indications that HL participants might tolerate longer delays (Stone and Moore 1999; Groth and Sondergaard 2004; Stone and Moore 2005; Bramslow 2010). This would be beneficial for the development of algorithms that try to improve speech understanding in noisy environments or the suppression of acoustic feedback for people with hearing loss (Löllmann and Vary 2009; Chen et al. 2016; Monaghan et al. 2017). Previous studies using HL participants with delays above 10 ms have mainly focussed on measuring delay tolerance to own-voice stimuli (Stone and Moore 2005) or tested across-frequency delay (Stone and Moore 2003). State-of-the-art hearing devices can actively reduce the occlusion effect by using active feedback cancellation systems (Mejia et al. 2008; Borges et al. 2013) or by detecting when the user speaks to decrease the amplification in lower frequencies for improved own-voice perception (US patents US9094766 B2; US8477973 B2). This means that delay tolerance to external-voice stimuli may become more relevant for future hearing devices. In order to increase the validity of subjective ratings, the use of real-time processing is preferred over pre-recorded stimuli because this provides a more natural perception of own-voice and external-voice sounds by maintaining proprioceptive and audio–visual cues.

We aimed to address some of the limitations of previous studies by testing both NH and HL participants with the same real-time processing setup for both own- and external-voice stimuli over the most relevant delay range of 10–50 ms. As previously proposed by Bramslow (2010), we kept the experimental setup simple and did not include any non-linear processing methods (such as dynamic range compression and active noise cancellation) to avoid confounding effects to the perception of delay as reported by Stone and Moore (1999, 2005, Stone et al. (2008)). A downside of this approach is that such a simple processing setup is not a realistic and accurate simulation for commercial hearing aids that employ a multitude of adaptive and non-linear processing techniques to improve sound perception. However, the main motivation of this study was to compare the tolerance of processing delay between NH and HL groups in a way that would be informative to different types of hearing devices (hearing aids, hearables, etc.) and not the evaluation of a specific setup. Consequently, we chose a linear fitting gain in contrast to the non-linear prescription gain commonly used in hearing aids. However, the choice of the fitting rationale may not have a strong effect on delay tolerance as reported by Stone and Moore (2002) who found no significant difference in disturbance ratings between linear and non-linear fittings for own-voice perception during speech production with NH participants. Several studies have chosen to use a frequency-independent gain for HL participants regardless of their degree of hearing loss (Groth and Sondergaard, 2004; Bramslow, 2010). We aimed to compare NH and HL groups and, therefore, provided frequency-dependent gain to compensate for elevated hearing thresholds in the HL group. In the external-voice condition, two level ratios of 0 and 20 dB between the delayed and direct sounds were investigated to account for the worst-case scenario of comb filtering and a less strong case that may occur in practice with a hearing device, respectively. We investigated whether HL participants show increased tolerance to output delay for own-voice and external-voice conditions in comparison with NH participants when tested under the same conditions in a real-time processing setup. Furthermore, we assessed the dependence of delay tolerance on the degree of hearing loss and the level difference between the simulated ‘‘direct’’ and ‘‘delayed’’ sound paths reaching the eardrum. We also aimed to investigate whether there is a difference in tolerable delay between new and experienced users of hearing aids, due to long-term acclimatisation effects to delayed sounds that may occur with the regular use of hearing aids.

Method

Participants

Twenty NH participants aged 18–45 years (12 females, eight males, an average age of 24 years) and 20 HL participants aged 45–81 years (12 females, eight males, an average age of 67 years) were recruited at the University of Southampton (Southampton, UK) and at the Audiology department of the Royal Berkshire Hospital (Reading, UK). Participants were not paid but reimbursement of travel expenses was offered. The normal-hearing group had hearing levels not exceeding 20 dB at octave intervals between 500 and 8000 Hz. The hearing loss group was tested without hearing aids (HA) and had hearing levels higher than 20 dB but not exceeding 70 dB at octave intervals between 500 and 8000 Hz. Of the 20 HL participants, 10 were regular users of hearing aids and 10 were new to the use of hearing aids (out of whom seven received their HA on the day of testing and three decided against using a HA). Hearing thresholds were confirmed via an audiogram measured by the experimenter before the start of the experiment or were provided by the clinics for the regular users, if a recent measurement was available that was no older than 2 years and that was used for the fitting of their hearing aids. Within the HL group, there was a significant correlation between age and pure-tone averages (PTA, with a Pearson correlation of $r = 0.52$, $p = 0.019$). Table 1 shows average pure-tone thresholds at 500, 1000 and 2000 Hz and further information for the HL group.

Equipment and stimuli

For the NH group, the experiments were performed in a quiet but not sound-proof, carpeted meeting room (5 m $\times$ 5 m $\times$ 2.5 m,

Table 1. HL group participant information and HL subgroup allocation according to PTA.

Participant	*Gender*	*Age*	*HA usage*	*Mould type*	*PTA*	*HL subgroup*
HL1	Male	71	Experienced	Closed	44	mid
HL2	Female	66	New	Open	40	mid
HL3	Female	77	Experienced	Closed	64	high
HL4	Male	69	Experienced	Open	54	high
HL5	Male	68	Experienced	Closed	58	high
HL6	Female	77	New	Open	34	low
HL7	Female	76	New	no HA	46	mid
HL8	Female	68	New	Open	42	mid
HL9	Male	50	New	Open	29	low
HL10	Female	83	Experienced	Open	46	mid
HL11	Male	76	New	Open	50	high
HL12	Female	75	New	Open	39	mid
HL13	Female	81	New	Open	41	mid
HL14	Female	79	Experienced	Open	53	high
HL15	Female	53	Experienced	Open	29	low
HL16	Female	75	Experienced	Open	65	high
HL17	Male	48	New	no HA	29	low
HL18	Male	52	New	no HA	24	low
HL19	Female	51	Experienced	Open	38	mid
HL20	Male	45	Experienced	Closed	51	high

RT60 = 0.3 s) at the University of Southampton audiology clinics. For the HL group, the experiments took place in a similar room (3 m × 3 m × 2.5 m) at the Royal Berkshire Hospital. The equipment was based on a digital signal processing board (Analog Devices ADSP-BF537-EZLITE) that delivers a minimum throughput delay of 1.5 ms (measured with RME Fireface UC). Stimuli were picked up with a dynamic microphone (JHS GS67), amplified with a pre-amplifier (ART Dual pre USB) and further processed with the DSP board before presentation to the participant via circumaural, closed headphones (Sennheiser HD380pro). The DSP board performed analogue-to-digital conversion (48 kHz, 16 bit) and split the signal to produce the two simulated sound paths: the direct air-conducted sound path and the delayed sound path of a hearing device.

Three different stimulus conditions were generated in real time to simulate three scenarios: own-voice perception (OwnV), external-voice perception with no level difference between the delayed and direct sound (Ext0dB) and external-voice perception with a level difference of 20 dB between the delayed and direct sound (Ext20dB). Accordingly, the direct sound path was muted (OwnV), kept at the input level (Ext0dB) or attenuated by 20 dB (Ext20dB) (Hétu and Quoc, 1992). In the OwnV condition, the participants spoke into the microphone and listened to their own voice, which they heard directly via bone conduction and via the simulated hearing device delayed by D through the headphones. In the two external voice conditions, the experimenter spoke into the microphone and the participant listened to the stimuli through the headphones. The delay D was set to 10, 20, 30, 40 or 50 ms. The direct sound signal was then added to the delayed signal. In the case of the HL participants, a hearing loss dependent gain was applied as final step to compensate for their hearing thresholds via a linear half-gain rule according to the audiogram of each participant using a 10-Band graphic equaliser (MXR M108). It should be noted that the level-ratio between the direct and delayed sound paths was the same for both NH and HL participants. The fitting gain for the HL group was applied after the mixing of the two sound paths in order to compensate for the hearing thresholds, and not to simulate the fitting gain of a hearing aid that would only affect the delayed sound path. The rationale behind this choice was to directly compare the annoyance of delay between NH and HL groups, not confounded by the level-ratio between the signals.

The complete setup was calibrated with a sound level metre (B&K2260) and artificial ear (B&K4153) to give an average sound level of 65 dB(A) when talking at a normal conversational level into the microphone. Participants were asked if the overall loudness was comfortable to them and were allowed to adjust the volume. All participants confirmed that the loudness was comfortable to them and thus no changes to the presentation level were made.

Experimental procedure

At the beginning of the experiment, two practice trials with ratings were performed for the own-voice (OwnV) and one external-voice (Ext0dB) condition. The practice trials used the minimal possible delay produced by the DSP board of 1.5 ms to allow the participant to acclimatise to the sound and usage of the setup. A total of 15 conditions comprising three different stimuli (OwnV, Ext0dB, Ext20dB) and five different delay settings (10, 20, 30, 40 and 50 ms) were presented. The presentation order of these conditions was randomised for each participant using a Latin square. The participants were asked to rate each condition on a 7-point Likert rating scale for the perceived subjective annoyance (with 1 labelled as 'not at all annoying' to 7 labelled as 'very annoying', and no other labels used).

In line with Stone and Moore (2005), we used a popular narrative book ('Harry Potter and the Philosopher's Stone' by J. K. Rowling) and let the respective speaker (participant or experimenter) read an arbitrarily selected passage of approximately one-minute length for each condition. The participants were instructed to read in a conversational manner not emphasising or raising their voice for special parts of the text. The total testing time was around 30 min and a short break was offered half way through the experiment. All experiments were approved by the local ethics board (ref ID 8978) and the NHS research ethics committee (REC reference number 8978).

Results

The group mean annoyance ratings for all conditions (5 delays per stimulus) are shown in Figure 1 for the NH and HL groups. The ratings for the training condition are shown in the figures but were not included in the statistical analysis. A difference in delay tolerance between the two groups occurred for the own-voice condition for delays above 20 ms and for the external voice Ext0dB condition over the whole range of delays. The external voice Ext20dB condition led to smaller differences between the two groups and lower ratings overall than for the other conditions. A repeated measures ANOVA (within-subject factors of delay and voice condition and between-subject factor of hearing ability) was performed including both NH and HL groups. There were significant main effects of delay [$F(4,152) = 85.875$, $p < 0.001$], voice condition [$F(2,76) = 13.228$, $p < 0.001$] and hearing ability [$F(1,38) = 4.619$, $p = 0.038$]. Post-hoc tests revealed statistically significant differences between groups for the delay of 30 ms (marginal, $p = 0.068$) and 50 ms in the OwnV condition ($p = 0.043$) and for all delays from 10 to 50 ms in the Ext0dB condition ($p < 0.05$). There was no significant effect of gender when included as the third factor in the repeated measures ANOVA. There was a significant interaction between voice condition and hearing ability [$F(2,76) = 6.503$, $p = 0.002$] that can be explained with the noticeable effect of hearing ability on delay tolerance for the Ext0dB condition but not for the Ext20dB condition. No other interactions were significant.

To assess the effect of degree of hearing loss on delay tolerance, the HL group was split based on their hearing thresholds into low PTA (<35 dB HL, $n = 5$, 3 males, an average age of 56 years), mid PTA (35 < PTA <50 dB HL, $n = 8$, 1 males, an average age of 71 years) and high PTA (>=50 dB HL, $n = 7$, 4 males, an average age of 70 years) subgroups. Group mean annoyance ratings for the three subgroups are shown in Figure 2. Consistent with differences between the HL and NH groups, a pattern of larger tolerable delay with greater hearing loss occurred in the ratings for OwnV and Ext20dB conditions. The difference between the HL subgroups was most evident for the Ext20dB condition for delays above 20 ms. The difference also occurred for the Ext0dB condition for delays of 30 and 50 ms, but the differences between groups were smaller than for the other two conditions.

A repeated measures ANOVA (within-subject factors of delay and voice condition and between-subjects factor of hearing loss subgroup) was performed for the HL group only. There was a significant main effect of delay [$F(4,68) = 31.337$, $p < 0.001$], but no effect of voice condition or hearing loss subgroup. All two-way

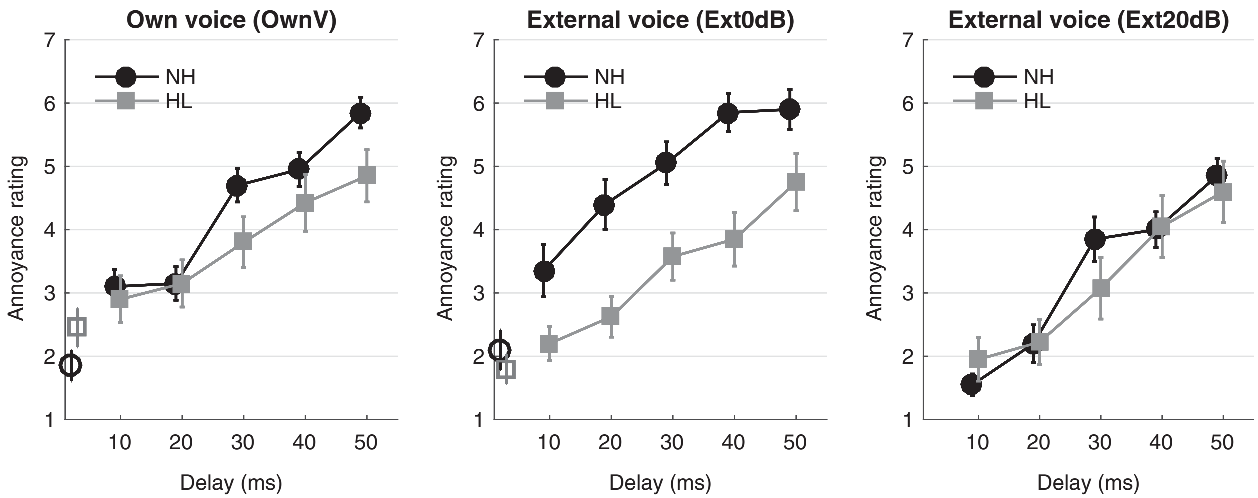

Figure 1. Group mean annoyance ratings for the normal-hearing ($n = 20$) and hearing loss ($n = 20$) groups in conditions (from left to right): own voice, external voice without level difference and external voice with level difference. Error bars show the standard error of the mean. Points plotted at 1.5 ms show average practice trial ratings (OwnV and Ext0dB only).

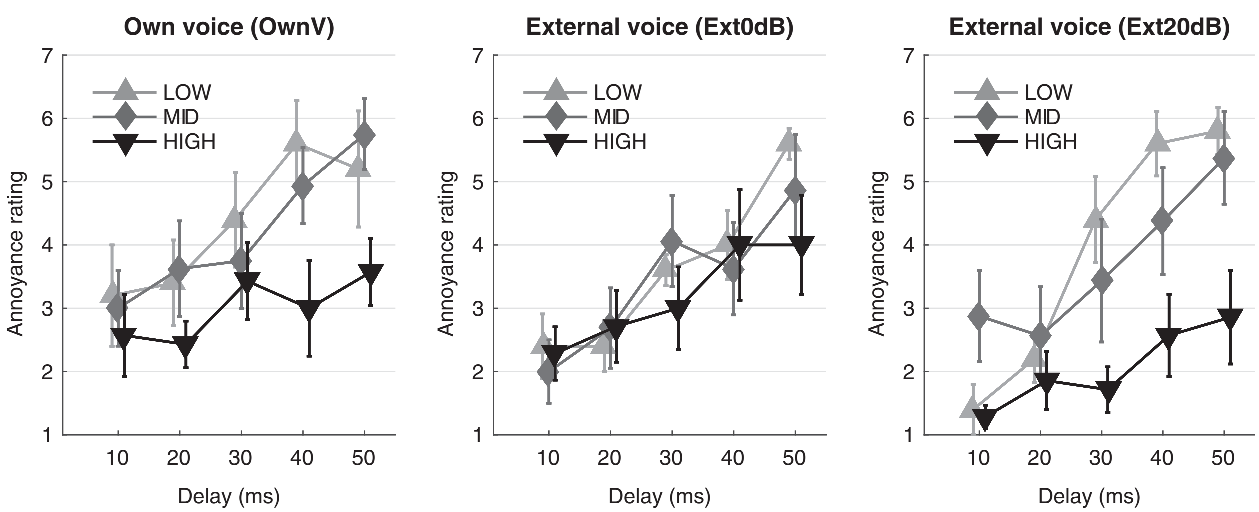

Figure 2. Annoyance ratings for the three hearing loss subgroups (low ($n = 5$), mid ($n = 8$), high ($n = 7$)). Otherwise as Figure 1.

interactions were non-significant but there was a marginally significant three-way interaction among delay, voice and hearing-loss subgroup [$F(16,136) = 1.716$, $p = 0.051$]. This interaction can be interpreted as the differences in the effect of hearing loss subgroup on delay tolerance between the three voice conditions. Furthermore, we analysed the relationship between PTA and rating scores within the HL group. The rate of change of annoyance with delay was estimated as the slope value of a linear fit to the five delay ratings for each condition and then averaged across all three conditions (OwnV, Ext0dB and Ext20dB) for each participant. Across participants, there was no significant correlation between PTA and average rating scores but there was a significant negative correlation between PTA and the average slope of ratings (Pearson correlation, $r = -0.51$, $p = 0.022$). Figure 3 shows the average slope values plotted against the PTA for each participant. Average slopes decreased from 0.9 for the low PTA group to 0.7 for the mid PTA group and down to 0.4 for the high PTA group. The decrease in average slopes with higher PTA indicates a reduced sensitivity to changes in delay with stronger hearing loss and is in line with the trend in mean scores between HL subgroups.

Group mean scores for experienced and new hearing aid users are shown in Figure 4. There was a difference between groups for the OwnV and Ext20dB conditions, for which the experienced users gave lower ratings by about 0.5–1 units for all delays (OwnV) or delays above 20 ms (Ext20dB). Ratings of the new hearing aid users in the Ext0dB condition did not monotonically increase with delay and were even lower than for experienced users at delays of 20 and 40 ms. A repeated measures ANOVA (within-subject factors of delay and voice condition and between-subject factor of hearing aid experience) was performed for the HL group only. There was a significant effect of delay [$F(4,72) = 27.453$, $p < 0.001$], but no significant effects were found for voice condition and hearing aid experience and there were no significant interactions.

Discussion

Our data demonstrate that participants with a broad range of hearing losses have significantly greater delay tolerance than normal-hearing participants. Generally, the results are consistent with previous work, especially with Stone and Moore (2005). However, we add to this work by comparing the annoyance ratings for normal-hearing and hearing loss groups using own-voice and external-voice stimuli and a larger range of delays, up to 50 ms. We used real-time processing to allow for a more natural perception than using recorded stimuli and restricted the setup to linear fitting and processing paradigms in an attempt to avoid non-linear side-effects that may interact with the perception of delay. In addition to the perception of delay during speech production (own voice condition, OwnV), the external voice condition with no level difference (Ext0dB) represents a worst-case scenario where the two sound paths (direct and delayed) have equal level over the whole frequency range. This means that the colouration by comb filtering is maximal in the Ext0dB condition. In contrast, the external voice condition with level difference (Ext20dB) was intended to simulate the attenuation of the direct sound for example by an earmould. The age difference between the NH and HL group may have confounded the effect of hearing loss and compromises the interpretation of the results. Furthermore, the wide range of hearing losses for the HL group (from 24 up to 65 dB PTA) makes it difficult to formulate quantitative limits for the tolerance of processing delay.

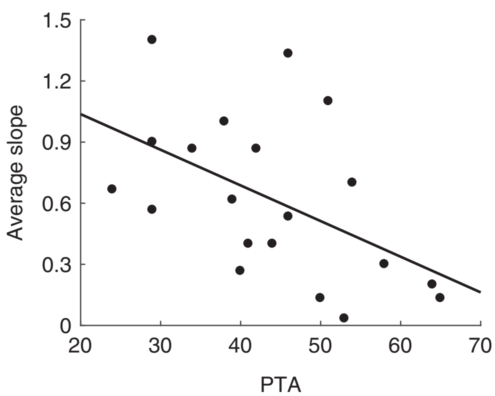

Figure 3. Participants' pure-tone averages plotted against the average slopes of annoyance ratings (averaged across voice conditions OwnV, Ext0dB and Ext20dB for each participant) for the group with hearing loss ($n = 20$). The line shows a linear fit to the data points.

In order to compare our results with previous findings, we apply the same threshold as Stone and Moore who chose a rating of 3 to quantify tolerance limits. NH listeners gave annoyance ratings above 3 for all delays in the OwnV and Ext0dB conditions and for a delay of 30 ms for the Ext20dB condition. This suggests that for NH listeners the delay should not exceed 10 ms, which is in line with previous results. The HL group gave ratings above 3 for delays of 20 ms in the OwnV condition, 30 ms in the Ext0dB condition and 30 ms in the Ext20dB condition. This indicates that the HL group tolerates longer delays than the NH group of about 10–20 ms and is in line with results from Stone and Moore for listeners with stronger hearing loss. However, the practice trials in our study showed ratings about 2 when a delay of 1.5 ms was applied, that would normally not be perceived as bothersome or even noticeable (Agnew and Thornton 2000; Groth and Sondergaard 2004). This could suggest that a threshold of 3 might be too strict for the setup or scale used in this study, or that there was some bias annoyance due to the setup used here, which differs from previous studies.

Overall, the annoyance ratings for the HL group were significantly lower than for the NH group and the tolerable delay limit can be estimated to be about 10–30 ms for the average listener with hearing loss, with a lower limit for the own-voice condition than for the external-voice condition. This is consistent with results reported by Stone and Moore (2005). The HL group gave ratings with average slopes that were negatively correlated with PTA. This indicates that with higher hearing loss the sensitivity to changes in delay is reduced. Especially, the high PTA group showed shallower rating profiles in comparison with the low and mid PTA groups (Figure 2). The third group (high) did not pass the threshold limit of

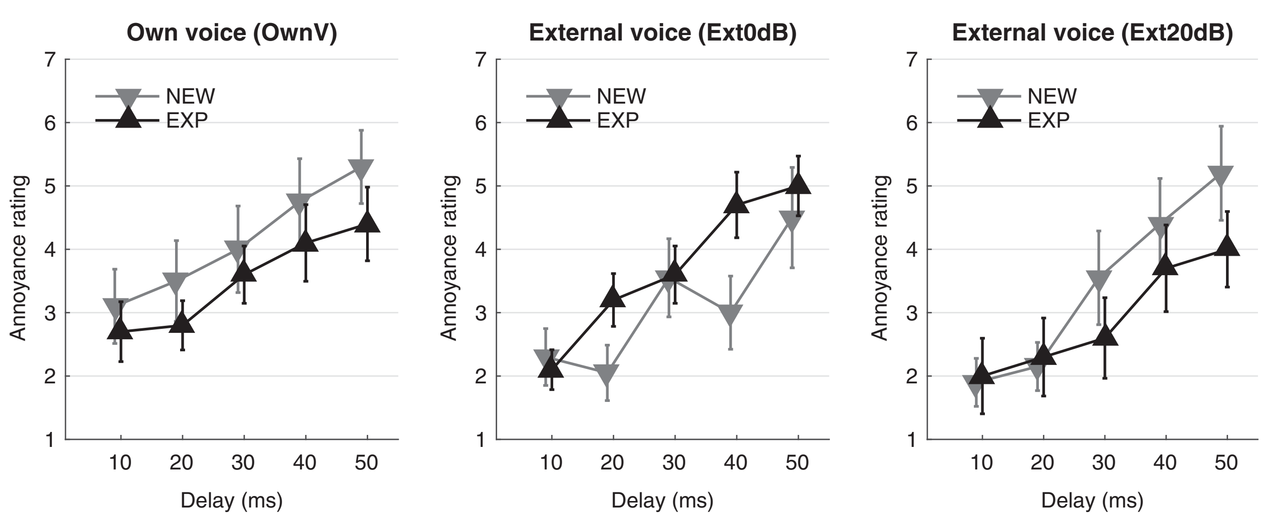

Figure 4. As Figure 1 but for the experienced ($n = 10$) and new ($n = 10$) hearing aid users.

3 for the Ext20dB condition for any of the delays. This reflects comments by participants from the third group that they could not hear any difference between the conditions. Thus, also within the HL group, we find a trend towards higher tolerance to delay with higher degree of hearing loss as reported by Stone and Moore (2005).

Annoyance ratings were higher for the new, than for the experienced, hearing aid users in two out of three conditions (OwnV, Ext20dB), although this trend was not significant overall. These stimuli represent in a simplified manner the scenarios of own-speech production and listening to external speech with a hearing aid. Potential long-term acclimatisation, as suggested by Stone and Moore (2005), could explain this result. In contrast, the Ext0dB condition represents an extreme-case scenario that would normally not occur in a hearing aid (equal level of direct and delayed signals over the whole frequency range). Consequently, the potential long-term acclimatisation effect to delay might not apply to the Ext0dB condition. Although the effect of long-term acclimatisation seems likely to account for the higher tolerance in the OwnV and Ext0dB conditions and the two groups were of same average age (67.1 versus 66.9 years), the experienced users had higher PTA values (50.2 versus 37.4 dB) than the new users. It is possible that both differences between groups (experience and higher degree of hearing loss) contributed to the increased tolerance to delay shown by the experienced user group. However, this effect did not reach the statistical significance level and has to be treated carefully. Future work that aims to address this question should test a larger sample and try to match PTAs between new and experienced participants. The acclimatisation to longer delays could alternatively be investigated by providing participants with custom hearing aids that use longer processing delays to see if participants are able to acclimatise to specific delays over longer time courses.

There are several differences between the technical setup and signal processing implemented in this study (e.g. the use of headphones and linear processing) and commercial hearing aids with earmoulds that make use of a large variety of combinations of non-linear processing techniques. This reduces the potential of the findings in this study for predicting tolerable delay for users of a specific type of commercial hearing aid. Given the results presented in this study and earlier results of Stone and Moore (1999), the level ratio between the direct and delayed signals scales the magnitude of the comb-filtering effect and thus directly influences the tolerable delay. This is supported by the finding of lower annoyance ratings for the Ext20dB than for to the Ext0dB condition for both NH and HL groups. Different types of earmoulds and amplification settings of commercial hearing aids will alter the level ratio and thus change the tolerance to delay. However, we considered this by choosing conservative mixing ratios between the direct and delayed sounds to represent worst-case scenarios in terms of comb filtering. Interestingly, the strongest and highly significant difference between NH and HL groups occurred in the most extreme case of comb filtering for the Ext0dB condition. In practice, it is likely that users of hearing aids encounter lower magnitudes of comb filtering due to larger level ratios between the direct and delayed sound paths. It should be noted that this study made use of linear processing in an attempt to isolate the perceptual effect of delay by avoiding confounds with dynamic level changes between the direct and delayed sounds. While this approach may give a better estimation of the annoyance caused by processing delay alone, extrapolation of the findings to commercial hearing aids and other hearing devices that make use of non-linear processing is limited by the simplified processing used. Furthermore, stimuli were presented via circumaural, closed headphones and generated using a simulation of the summation of direct and delayed sounds that may have led to a different perception of speech than with state-of-the-art hearing aids with earmoulds.

Conclusion

The tolerable processing delay is one of the factors that limit the development of novel and more complex algorithms for hearing devices such as hearing aids. This study found a significantly greater tolerance of processing delay for listeners with hearing loss than for normal-hearing listeners for own-voice and external-voice conditions when tested with the same setup. Accordingly there was a trend of increased tolerable delay with higher degree of hearing loss within the hearing loss group, but this effect did not reach statistical significance. However, there was a significant negative correlation across participants between the average slopes of ratings and their PTA, with shallower slopes for higher PTA. This relationship indicated that there was a decrease in sensitivity to changes in delay with stronger hearing loss.

Quantitatively, results from this study indicate that delays up to 20 ms are not expected to exceed tolerance limits for the average listener with hearing loss. Higher degrees of hearing loss may allow an increase in delay up to 20–30 ms without causing excessive annoyance. These findings are in line with previous recommendations for listeners with hearing loss and extend the testing conditions to the scenario of external-voice stimuli with linearly processed signals. It should be noted that the testing setup used in this study differed from commercial hearing aids in several aspects and, therefore, limits the extrapolation of the findings.

Experienced hearing aid users showed a slightly increased tolerance to delay in two out of three voice conditions, possibly due to potential long-term acclimatisation. However, this effect was not statistically significant and could have been confounded by a greater degree of hearing loss in the experienced group. The potential beneficial effects of acclimatisation on delay tolerance remain an interesting question for future studies.

Acknowledgements

The authors would like to express our gratitude to all participants who took part in this study. The authors would also like to thank Prof. Brian C. J. Moore for his helpful comments on an earlier draft of this paper and four anonymous reviewers for their constructive feedback. The work leading to this deliverable and the results described therein has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/under REA Grant agreement no. PITN-GA-2012-317521 and by EPSRC grant EP/K020501/1.

Declaration of interest: No potential conflict of interest was reported by the authors.

References

Agnew, J., and J. M. Thornton. 2000. ''Just Noticeable and Objectionable Group Delays in Digital Hearing Aids.'' *Journal of the American Academy of Audiology* 11: 330–336.

Borges, R. C., M. H. Costa, J. A. Cordioli, and L. F. Assuiti. 2013. ''An Adaptive Occlusion Canceller for Hearing Aids.'' *Paper presented at* *the Proceedings of the IEEE 21st European Signal Processing Conference (EUSIPCO)*, 1–5.

Bramslow, L. 2010. ''Preferred Signal Path Delay and High-Pass Cut-Off in Open Fittings.'' *International Journal of Audiology* 49: 634–644. doi:10.3109/14992021003753482.

Chen, J., Y. Wang, S. E. Yoho, D. Wang, and E. W. Healy. 2016. ''Large-Scale Training to Increase Speech Intelligibility for Hearing-Impaired Listeners in Novel Noises.'' *The Journal of the Acoustical Society of America* 139: 2604–2612. doi:10.1121/1.4948445.

Dillon, H., G. Keidser, A. O'Brien, and H. Silberstein. 2003. ''Sound Quality Comparisons of Advanced Hearing Aids.'' *The Hearing Journal* 56: 30–32. doi:10.1097/01.HJ.0000293908.50552.34.

Groth, J., and M. B. Sondergaard. 2004. ''Disturbance Caused by Varying Propagation Delay in Non-Occluding Hearing Aid Fittings.'' *International Journal of Audiology* 43: 594–599.

Hétu, R., and H. T. Quoc. 1992. ''Can an Inactivated Hearing Aid Act as a Hearing Protector?'' *Canadian Acoustics* 20: 35–36.

Litovsky, R. Y., H. S. Colburn, W. A. Yost, and S. J. Guzman. 1999. ''The Precedence Effect.'' *The Journal of the Acoustical Society of America* 106: 1633–1654.

Löllmann, H. W., and P. Vary. 2009. ''Low Delay Noise Reduction and Dereverberation for Hearing Aids.'' *EURASIP Journal on Advances in Signal Processing* 2009: 437807. doi:10.1155/2009/437807.

Mejia, J., H. Dillon, and M. Fisher. 2008. ''Active Cancellation Of Occlusion: An Electronic Vent for Hearing Aids and Hearing Protectors.'' *The Journal of the Acoustical Society of America* 124: 235–240. doi:10.1121/1.2908279.

Monaghan, J. J. M., T. Goehring, X. Yang, F. Bolner, S. Wang, M. C. Wright, and S. Bleeck. 2017. ''Auditory Inspired Machine Learning Techniques can Improve Speech Intelligibility and Quality for Hearing-Impaired Listeners.'' *The Journal of the Acoustical Society of America* 141: 1985–1998. doi:10.1121/1.4977197.

Stone, M. A., and B. C. J. Moore. 1999. ''Tolerable Hearing Aid Delays. I. Estimation of Limits Imposed by the Auditory Path Alone Using Simulated Hearing Losses.'' *Ear and Hearing* 20: 182–192.

Stone, M. A., and B. C. J. Moore. 2002. ''Tolerable Hearing Aid Delays. II. Estimation of Limits Imposed During Speech Production.'' *Ear and Hearing* 23: 325–338.

Stone, M. A., and B. C. J. Moore. 2003. ''Tolerable Hearing Aid Delays. III. Effects on Speech Production and Perception of Across-Frequency Variation in Delay.'' *Ear and Hearing* 24: 175–183. doi:10.1097/01.AUD.0000058106.68049.9C.

Stone, M. A., and B. C. J. Moore. 2005. ''Tolerable Hearing-Aid Delays: IV. Effects on Subjective Disturbance During Speech Production by Hearing-Impaired Subjects.'' *Ear and Hearing* 26: 225–235.

Stone, M. A., B. C. J. Moore, K. Meisenbacher, and R. P. Derleth. 2008. ''Tolerable Hearing aid Delays. V. Estimation of Limits for Open Canal Fittings.'' *Ear and Hearing* 29: 601–617.

Stromsta, C.1962. ''Delays Associated with Certain Sidetone Pathways.'' *The Journal of the Acoustical Society of America* 34: 392–396. doi:10.1121/1.1918139.

Stuart, A., J. Kalinowski, M. P. Rastatter, and K. Lynch. 2002. ''Effect of Delayed Auditory Feedback on Normal Speakers at Two Speech Rates.'' *The Journal of the Acoustical Society of America* 111: 2237–2241.

Summerfield, Q. 1992. ''Lipreading and Audio-Visual Speech Perception.'' *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 335: 71–78. doi:10.1098/rstb.1992.0009.

Zakis, J. A., B. Fulton, and B. R. Steele. 2012. ''Preferred Delay and Phase-Frequency Response of Open-Canal Hearing Aids with Music at Low Insertion Gain.'' *International Journal of Audiology* 51: 906–913. doi:10.3109/14992027.2012.701020.