Evaluation of a multi-channel algorithm for reducing transient sounds

Journal:	*International Journal of Audiology*
Manuscript ID	TIJA-2017-10-0328.R1
Manuscript Type:	Original Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Keshavarzi, Mahmoud; University of Cambridge, Experimental Psychology Baer, Thomas; University of Cambridge, Department of Experimental Psychology Moore, Brian; University of Cambridge, Dept of Experimental Psychology
Keywords:	Hearing Aids, Noise, Psychoacoustics/Hearing Science, Speech Perception

Evaluation of a multi-channel algorithm for reducing transient sounds

Mahmoud Keshavarzi, Thomas Baer & Brian C.J. Moore

Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

Correspondence: Mahmoud Kesharvarsi, Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK.

E-mail: mahmoud.keshavarzi.ir@ieee.org

Abbreviations

AGC	Automatic gain control
FFT	Fast Fourier Transform
MCTR	Multi-channel transient reduction
RMS	Root-mean square

Key words: Hearing aid; transient noise reduction; multi-channel analysis; acoustic annoyance, preference judgment

Evaluation of a multi-channel algorithm for reducing transient sounds

Mahmoud Keshavarzi, Thomas Baer and Brian C.J. Moore

Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

First revision, January 2018.

Correspondence: Mahmoud Kesharvarsi, Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK.
E-mail: mahmoud.keshavarzi.ir@ieee.org

Key words: Hearing aid; transient noise reduction; multi-channel analysis; acoustic annoyance, preference judgment

Abbreviations

AGC	Automatic gain control
FFT	Fast Fourier Transform
MCTR	Multi-channel transient reduction
RMS	Root-mean square

Abstract

Objective: The objective was to evaluate and select appropriate parameters for a multi-channel transient reduction (MCTR) algorithm for detecting and attenuating transient sounds in speech. *Design*: In each trial, the same sentence was played twice. A transient sound was presented in both sentences, but its level varied across the two depending on whether or not it had been processed by the MCTR and on the "strength" of the processing. The participant indicated their preference for which one was better and by how much in terms of the balance between the annoyance produced by the transient and the audibility of the transient (they were told that the transient should still be audible). *Study sample*: Twenty English-speaking participants were tested, ten with normal hearing and ten with mild-to-moderate hearing impairment. Frequency-dependent linear amplification was provided for the latter. *Results*: The results for both participant groups indicated that sounds processed using the MCTR were preferred over the unprocessed sounds. For the hearing-impaired participants, the medium and strong settings of the MCTR were preferred over the weak setting. *Conclusions:* The medium and strong settings of the MCTR reduced the annoyance produced by the transients while maintaining their audibility.

Key words: Hearing aid; transient noise reduction; multi-channel analysis; acoustic annoyance, preference judgment

Introduction

Despite great advances in digital noise reduction systems and automatic gain control (AGC) systems, users of cochlear implants and hearing aids still have problems related to speech intelligibility and discomfort and/or annoyance in the presence of environmental noises. Transient sounds such as a door slamming, a hammer hitting a nail, or a knife hitting a plate can be especially problematic, since such sounds often have a short-term level that is well above the long-term average level in a given acoustic situation, and since users of cochlear implants and hearing aids often have a very small range of levels between the detection threshold and the level at which sounds become uncomfortably loud (Zeng & Shannon, 1999; Moore, 2007).

According to Dyballa et al. (2015), transient sounds have three main characteristics: a rapid onset (sometimes with a rise time less than 1 ms), a rapid decline (over tens of ms), and a short overall duration (usually less than a few hundred ms). In addition to causing annoyance or discomfort, an intense transient may cause the AGC system in a cochlear implant or hearing aid to decrease the gain, with the result that speech sounds following shortly after the transient may be barely, if at all, audible (Moore et al., 1991).

All hearing aids and cochlear implants incorporate some form of amplitude compression or AGC to “squeeze” the large range of sound levels encountered in everyday life into the small dynamic range of the user. AGC systems in hearing aids usually filter the incoming signal into several frequency “channels” and apply the AGC independently in each channel. The AGC in each channel is characterized by an attack time, a measure of the time taken to reduce the gain when the sound level suddenly increases, and the release time, a measure of the time taken for the gain to increase when the sound level suddenly decreases (ANSI, 2003). The attack time is usually chosen to be reasonably small, typically in the range 5-50 ms, so that when the input sound level suddenly increases the gain is rapidly reduced, thereby protecting the user from possible discomfort. However, even an attack time as small as 5 ms may be too long to provide adequate protection from intense transients (Korhonen et al., 2013). For example, Keidser et al. (2009) reported that users of hearing aids complained about transient sounds causing loudness discomfort, and Moore and Füllgrabe (2010) reported complaints about the loudness of transient sounds for users of hearing aids fitted using the CAM2 method (Moore et al., 2010). The dual time-constant AGC system (Moore & Glasberg, 1988; Moore et al., 1991; Stone et al., 1999; Boyle et al., 2009) was designed to reduce the gain rapidly in response to a transient sound, but to restore the gain to the value that applied before the transient after cessation of the transient. However, even this system may not react sufficiently quickly to provide adequate protection from transient sounds.

One method of providing protection from intense transients is peak clipping or fast-acting limiting. This is incorporated in most hearing aids, but it has the problem that it introduces distortion and reduces sound quality and speech intelligibility (Stelmachowicz et al., 1999; Tan & Moore, 2008). Furthermore, peak clipping does not operate for transient sounds whose level does not reach the threshold for clipping or limiting.

Several hearing aid manufacturers have developed transient or impulse sound reduction systems to protect the hearing aid user from discomfort and/or annoyance. The objective of these systems is to selectively attenuate transient sounds, so that they remain audible, but are not uncomfortable or annoying (Luo, 2009; Launer et al., 2016). Such systems mostly operate on the broadband signal, and any short-term gain reduction is applied to the entire signal. However, transient sounds can vary markedly in their spectral content. Some transient sounds, such as keys jingling, are dominated by high-frequency components, with little energy at low frequencies. An overall reduction in gain produced in response to such a signal would result in a brief reduction in level of any low-frequency components that were present, such as vowel sounds in speech, giving the misleading impression that the low-frequency sounds were interrupted. Conversely, some transient sounds, such as a book being abruptly closed, have most of their energy at low frequencies. An overall reduction in gain produced in response to such a sound would result in a brief reduction in level of any high-frequency components that were present, such as fricatives in speech. This might, for example, make a sound like a sustained /s/ be perceived as /st/.

The current study describes and evaluates the benefits of a newly developed multi-

97 channel transient reduction (MCTR) algorithm. The algorithm is intended to provide a brief
98 reduction in gain only for frequency regions in which the transient sound has significant energy,
99 thereby avoiding disturbing perceptual effects in other frequency regions. The gain reduction is
100 designed to be progressive: weak transients are not attenuated at all, moderately intense
101 transients are attenuated by a medium amount, and intense transients are attenuated considerably.
102 It is intended that, for applications in hearing aids and cochlear implants, the algorithm would be
103 applied as a side chain or in parallel with the main multi-channel AGC system. For example, the
104 main AGC system could be slow-acting, keeping the long-term average level in each channel
105 within a certain range, with the transient reduction system providing protection from transient
106 sounds.

107 Methods of reducing intense transient sounds with some similarities to our method were
108 described in a patent (Schneider et al., 2010). A "pattern analysis" approach was used, including
109 the use of multiple frequency channels. The intended application of the patent was the prevention
110 of "acoustic shock" for users of headphones and headsets. Acoustic shock refers to effects of
111 very intense sounds that may occur unintentionally as a result of equipment malfunctioning. The
112 effects include temporary or permanent hearing loss, tinnitus, and hyperacusis (McFerran &
113 Baguley, 2007). The aim of the methods described in the patent was to reduce the level of the
114 intense sound (which was not necessarily a brief transient) to a predetermined safe level. This
115 contrasts with our MCTR algorithm, for which the goal was to selectively attenuate transient
116 sounds so that they remained audible but were not uncomfortable or annoying. We have not
117 found any published evaluations of the methods described by Schneider et al. (2010).

118 A frequency-selective method for attenuating transients was described by Hirszhorn et al.
119 (2012). The method was based on estimating the power spectral density of the transient and
120 using that information to selectively filter the sound so as to attenuate the transient. However,
121 this method was not intended for application in hearing aids or cochlear implants, and it involved
122 delays between 40 and 250 ms, which would be unacceptably long for use in hearing aids (Stone
123 & Moore, 1999; Stone et al., 2008).

A four-channel transient reduction system for cochlear implant users was evaluated by Dyballa et al. (2016). Details of the system, such as the method used for frequency analysis, were not described. The authors stated that "Transient detection in each band was carried out as in the original single-band algorithm", but no further details were provided, so it is difficult to assess the similarity between their system and our MCTR system. They did not report any attempt to optimize the parameters of the processing. The results of their evaluation with experienced cochlear-implant users showed that the transient-reduction system improved reception thresholds for speech in cafeteria noise and office noise and gave higher comfort and clarity ratings for speech in cafeteria noise.

For the MCTR algorithm used here, pilot experiments with normal-hearing and hearing-impaired participants indicated that the algorithm did not have any influence on the intelligibility of the speech on which the transients were superimposed, as was found for the broadband transient reduction system described by Korhonen et al. (2013). Therefore, the focus of this study was on the annoyance produced by the transient sounds, as determined in a paired-comparisons task. We reasoned that the MCTR algorithm should reduce the level of the transients to prevent them from being too loud or annoying, but the level reduction should not be so great that the transients became unnatural or difficult to hear. Therefore, the instructions to the participants emphasized that their judgments should be based on the balance between the loudness/annoyance produced by the transients and their audibility/naturalness.

Method

Transient reduction algorithm

The sampling rate used in the MCTR algorithm was 22,050 Hz, which allows processing of the whole frequency range covered by conventional hearing aids (up to about 10,000 Hz). The signal was segmented into frames with a duration of approximately 1 ms (22 samples) and there was a 12-samples overlap between successive frames. The signal in each frame was windowed using a Tukey window shape defined by:

$$\begin{cases} w(n) = 0.5\left(1-\cos\left(\dfrac{2\pi n}{12}\right)\right) & 0 \le n \le 5 \\ w(n) = 1 & 6 \le n \le 15 \\ w(n) = 0.5\left(1-\cos\left(\dfrac{2\pi(n-9)}{12}\right)\right) & 16 \le n \le 21 \end{cases} \quad (1)$$

where n is the sample number. This window shape was chosen because a concatenation method rather than an overlap-add method (Allen, 1977) was used to reconstruct the signal, as described below. The specific window used, with 10 samples at maximum amplitude, gave a good compromise between spectral resolution and temporal fidelity (Harris, 1978).

Each windowed frame was zero padded on either side to give 32 samples and the Fast Fourier Transform (FFT) was used to obtain a frequency-domain representation of the frame with 16 bins. Bins were grouped so as to form five frequency channels. The number of bins in frequency channels 1 to 5 was 1, 1, 2, 3, and 9, respectively.

The MCTR algorithm detects transient sounds by comparing the short-term magnitude (amplitude) in channel i and frame j, M_{ij}, to a running estimate of the root-mean-square (RMS) magnitude in that channel at the time of frame j, RMS_{ij}. If the ratio of these two exceeds a criterion value (different for each channel) then a transient is deemed to be present. We used the following criteria for detecting a transient in the ith channel of the jth frame:

$$M_{ij}/RMS_{ij} > \delta_i \qquad i = 1, \ldots, 5 \quad (2)$$

where the constants δ_i were chosen in such a way that the MCTR correctly detected frames including transients, while not responding to short-term peaks in the speech. The values of δ_i were 12, 21, 12, 8, and 7 for channels 1 to 5, respectively. With these values, the detection of transients was perfect for the sentences and transient levels used in our experiment (see below for details). In other words, all transients were detected, and there were no false detections in parts of the sentences where no transient was present.

The running RMS magnitude of the ith frequency channel for the jth frame, normalized by the number of bins in that channel, was calculated as:

$$RMS_{i,j} = \frac{\sum_{n=j-m}^{j-1} \sqrt{\frac{\sum_{l_i} \left|FFT_n(k_i)\right|^2}{l_i}}}{m} \tag{3}$$

where $FFT_n(k_i)$ is the kth FFT bin within channel i for frame n, l_i represents the number of FFT bins within channel i (so that Σ_{li} represents summation over all bins within channel i), and m represents the number of frames contributing to the calculation of the running RMS magnitude for frame j. The appropriate value of m depends on several factors, such as sampling rate, frame length and the overlap between successive frames. It should not be so small that the moving RMS value is affected by brief pauses in the speech. The value used in the MCTR algorithm was 1500, corresponding to 0.68 s. If a transient was detected in a given frame, the running RMS magnitude was not updated using samples from that frame. This was done to prevent the running estimate of the RMS magnitude being influenced by the superimposed transient. When a transient was detected, the value of m was kept at 1500 by not dropping the earliest samples.

When a transient was detected in the jth frame, the magnitude for the ith channel of that frame was attenuated by an amount, C_{ij}, whose value in dB was defined by:

$$C_{ij}(R_{ij},\alpha) = \begin{cases} \alpha R_{ij} & R_{ij} > 0 \quad i = 1,2,\ldots,5 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where parameter α is a positive real number and R_{ij} is $20\log_{10}(M_{ij}/RMS_{ij})$. Thus, when the ratio M_{ij}/RMS_{ij} was ≤1, no attenuation was applied. When the ratio was above 1, the attenuation increased progressively as the ratio increased. Figure 1a shows the attenuation as a function of R_{ij} for three values of α, 0.267, 0.467, and 0.933. Figure 1b shows the resulting output levels. For the middle value shown here, when R_{ij} was 20 dB, corresponding to a magnitude ratio of 10, C_{ij} was 9.3 dB and the resulting value of M_{ij}/RMS_{ij}, converted to dB, was 10.7 dB. When R_{ij} was 29.5 dB, corresponding to a magnitude ratio of 30, C_{ij} was13.8 dB and the resulting value of M_{ij}/RMS_{ij}, converted to dB, was 15.7 dB. Thus, after application of the attenuation, the output

199 level of the transient increased as the input level increased, to give some impression of the
200 magnitude of the transient, but the increase in output level was more gradual than the increase in
201 input level. For frames in which a transient was detected and attenuated, the output signal for that
202 frame was obtained by performing an inverse FFT of the modified spectral magnitudes. If no
203 transient was detected, the untransformed input frame was used. The final output was produced
204 by concatenating the central 10 samples (the flat portion of the window) from each frame.

205 The procedure of using the untransformed frame when no transient was detected had the
206 advantage that numerical errors produced by the FFT/IFFT processing were avoided when no
207 transient was detected. The windowing and FFT/IFFT transformations had unity (0 dB) gain
208 when C_{ij} = 0 dB.

209

210 *Participants*

211 Twenty English-speaking participants were tested. Audiometric thresholds were measured for
212 audiometric frequencies from 0.25 to 8 kHz for all participants, using a Grason-Stadler GSI-61
213 audiometer. Ten of the participants had normal hearing, with all audiometric thresholds ≤20 dB
214 HL in both ears, and ten had hearing loss, with audiometric thresholds over the range 0.5 to 4
215 kHz not greater than 75 dB HL. The hearing threshold was 40 dB HL or more for at least one
216 frequency over that range.

217

218 *Sound signals*

219 To evaluate the effects of the MCTR algorithm, we investigated participants' preferences for
220 different amounts of attenuation of the transient, produced by varying parameter α. Nine types of
221 transient sounds were used, as described in Table 1. Eight out of these sounds were the same as
222 used by Korhonen et al. (2013) and the remaining one was obtained from the ROOMSIM sounds
223 (Campbell et al., 2008). Transients were presented in nine different sentences, and each sentence
224 included only one transient sound. The combination of transient and sentence varied across
225 participants. The RMS input level (before frequency-dependent amplification for the hearing-

impaired participants) of the sentences (excluding the transients) was 60 dB SPL. For each type of transient sound there were four conditions, based on the amount of attenuation applied by the MCTR algorithm. The first condition was a baseline condition using transients whose peak levels (measured in a 10-ms interval) relative to the RMS level of the speech are specified in Table 1; no MCTR was applied. These peak levels were chosen so that the transients were perceived as loud and somewhat unpleasant, but not excessively so, based on pilot experiments. The baseline condition is referred to as condition “none” (no attenuation). The second, third, and fourth conditions used the signals from condition none, but processed using the MCTR algorithm with $\alpha = 0.267$, 0.467 and 0.933. These conditions are referred to according to the strength of the attenuation as weak, medium, and strong, respectively. Accordingly, there were 36 stimuli: 9 types of transient sounds × 4 processing conditions (no processing and processing with three values of α). The sounds for the hearing-impaired subjects were given linear frequency-dependent amplification according to the "Cambridge formula" (Moore & Glasberg, 1998). This was done using a finite impulse response filter created using Matlab.

Figure 2 illustrates the operation of the MCTR algorithm using T6 (A metal can filled with metal bolts, shaken once). The panels show the waveform of the speech+transient for conditions none, weak, medium, and strong. It can be seen that the amplitude of the transient decreases progressively as the strength of the MCTR increases, while the speech waveform occurring before and after the transient is not affected by the MCTR.

Procedure

The participants were seated in a quiet room and wore Sennheiser HD580 headphones connected to the sound card of a computer (24 bit resolution, sampling rate = 22050 Hz). For each transient sound, six paired comparisons were performed: condition none versus condition weak, condition none versus condition medium, condition none versus condition strong, condition weak versus condition medium, condition weak versus condition strong, and condition medium versus condition strong. The procedure was the same as described by Moore and Sek (2013). The two sounds to be compared were presented in succession with a 200-ms silent interval between them. The possible orders were used equally often and the order was randomized across trials. The instructions to the participant, which appeared on the computer screen, were as follows: "On each trial you will hear the same sentence twice in succession. A transient background sound (e.g. the sound of glasses clinking) has been added to each sentence. The background sound should be clearly audible and it should sound natural, but it should not be too loud or too annoying and it should not interfere with your perception of the sentence. Please decide whether you prefer the sound in the first interval or the sound in the second interval, and by how much, by using the mouse to position the slider on the screen. Your judgment should be based on the balance between the audibility/naturalness of the transient sound and its loudness/annoyance. For example, if the transient sound is barely audible or does not sound natural in the first interval and is clearly audible and natural but not too loud or annoying in the second interval, you should indicate a preference for interval 2. On the other hand, if the sound is clearly audible and natural in both intervals, but is comfortably loud in interval 1 and louder or more annoying in interval 2, you should indicate a preference for interval 1."

On each trial, each pair of sounds was presented only once. Participants responded using a mouse to select the position of a slider on the screen along a continuum labeled "1 much better", "1 moderately better", "1 slightly better", "equal", "2 slightly better", "2 moderately better", and "2 much better", where 1 refers to the first sound and 2 refers to the second sound. Choices were not restricted to the labeled points; any point along the slider could be chosen. Within a given session (block of trials), each of the six pairs of conditions was presented in both orders for each of the nine transient sounds, so there were 108 trials in a session.

Preference scores for each participant and each comparison were computed in the following way. The extreme positions of the slider were arbitrarily assigned values of −3 and +3. Regardless of the order of presentation in a given trial (condition X first or condition Y first), if X was preferred the slider position was coded as a negative number and if Y was preferred the slider position was coded as a positive number. For example, if the order on a given trial was Y

280 first and X second, and the participant set the slider position midway between "2 slightly better"
281 and "2 moderately better", the score for that trial was assigned a value of –1.5. The overall score
282 for a given comparison and a given transient type was obtained by averaging the scores for the
283 two orders for that comparison and transient type for each participant. Scores were then averaged
284 across participants, but separately for the normal-hearing and hearing-impaired participants.
285 Preference scores therefore were constrained to fall in the range –3 to +3.

286

287 **Results**

288 *Preferences for normal-hearing participants*

289 Figure 3 shows mean preference scores for the normal-hearing participants for each transient and
290 each comparison. In what follows, two-tailed *t*-tests were used to assess whether the mean
291 preference scores across transient types differed significantly from zero for each comparison.
292 Outcomes of the *t*-tests are indicated in parentheses. Given that six *t*-tests were being performed
293 for each participant group, the significance level was set to 0.05/6 = 0.008. Significant *t* values
294 are indicated by *. For the none vs weak comparison (panel a), the preference scores all indicated
295 a small preference for condition weak, with a mean of 0.52 ($t = 8.83$, $p = 0.000021$*). For the
296 none vs medium comparison (panel b), the preference scores all indicated a preference for
297 condition medium, with a mean of 1.20 ($t = 9.38$, $p = 0.000013$*). The preferences were stronger
298 for some transients (T6 and T7) than for others (T8 and T9) and were stronger than for
299 comparison none vs weak. For the none vs strong comparison (panel c), the preference scores all
300 indicated a preference for condition strong, with a mean of 0.64 ($t = 5.50$, $p = 0.00057$*).
301 However, the strengths of the preferences were very small for some transients (e.g. T3 and T8)
302 and were smaller overall than for comparison none vs. medium.

303 For the medium vs weak comparison (panel d), the preference scores all indicated a small
304 preference for condition medium, with a mean of 0.43 ($t = 9.85$, $p = 0.000009$*). For the strong
305 vs weak comparison (panel e), the preference scores were small and variable in sign, with a mean
306 of 0.10, indicating no clear overall preference for one condition over the other ($t = 1.14$, p =

0.2857). For the medium vs strong comparison (panel f), the preference scores all indicated a preference for condition medium, but the size of the preference was small, with a mean of −0.29 (t = 5.02, p = 0.0010*).

Overall, the results for the normal-hearing participants indicate that the stimuli processed using the MCTR algorithm were preferred over the unprocessed stimuli, and that the medium attenuation setting was slightly preferred over the weak attenuation setting and the strong attenuation setting.

Preferences for hearing-impaired participants

Figure 4 shows the mean preference scores for the hearing-impaired participants. For the none vs weak comparison (panel a), the preference scores all indicated a small preference for condition weak, with a mean of 0.44 (t = 11.50, p = 0.000002*). For the none vs medium comparison (panel b), the preference scores all indicated a preference for condition medium, with a mean of 0.92 (t = 15.58, p = 0.0000002*). The preferences were stronger than for comparison none vs weak. For the none vs strong comparison (panel c), the preference scores all indicated a preference for condition strong, with a mean of 0.73 (t = 7.27, p = 0.00008*). However, the strengths of the preferences were very small for some transients (e.g. T8 and T9).

For the medium vs weak comparison (panel d), the preference scores all indicated a small preference for condition medium, with a mean of 0.21 (t = 5.95, p = 0.00034*). For the strong vs weak comparison (panel e), the preference scores were small and variable in sign, with a mean of 0.21 (t = 3.03, p = 0.016). For the medium vs strong comparison (panel f), the preference scores were close to zero, with a mean of −0.11 (t = 3.24, p = 0.012), which was not significant after allowing for multiple comparisons.

Overall, the results for the hearing-impaired participants indicate that the stimuli processed using the MCTR algorithm were preferred over the unprocessed stimuli, and that the medium and strong attenuation settings were preferred over the weak attenuation setting.

334 **Discussion**

335 The results for both participant groups indicated that stimuli processed using the MCTR
336 algorithm were preferred over the unprocessed stimuli. The normal-hearing participants showed
337 a small preference for the medium setting of the MCTR algorithm relative to both the weak and
338 strong settings, while the hearing-impaired participants tended to prefer the medium and strong
339 settings relative to the weak setting, but showed no clear preference when comparing the
340 medium and strong settings. The difference between the two groups probably reflects the effects
341 of loudness recruitment for the hearing-impaired participants, which, given the frequency-
342 dependent linear amplification provided for them, probably led to the transients being louder and
343 more annoying than for the normal-hearing participants. Hence, the hearing-impaired
344 participants preferred slightly greater attenuation of the transients.

345 Informal questioning indicated that both the normal-hearing and hearing-impaired
346 participants could hear the transient sounds in all conditions, although they were sometimes
347 heard as being weak for condition strong. However, the subjective quality of the transients was
348 sometimes reported to be somewhat changed, especially for condition strong. This may have
349 partly been caused by waveform discontinuities that could occur as a consequence of the
350 concatenation procedure used in the MCTR, although the participants did not reported hearing
351 any clicks superimposed on the transients. In practice, a value of α between 0.467 and 0.933,
352 perhaps $\alpha = 0.66$, would seem to be suitable for use with hearing-impaired participants. This
353 would be sufficient to reduce the loudness and annoyance of the transients while maintaining the
354 audibility and sound quality of the transients.

355 For the transients and sentences used in our experiment, the transient detection part of the
356 MCTR algorithm worked perfectly. However, it did not always work perfectly with transients
357 whose level was somewhat lower than used here. In pilot work it was found that, except for
358 channel 5, transient detection (especially for transients whose peak levels in the original time-
359 domain speech signal were less than 15 dB above the RMS level) was more reliable (i.e., there
360 were fewer false positives and fewer misses) when it was required that the criterion be met for

two adjacent channels (and the values of δ_i were adjusted). Transient detection based on more channels could be used in further work.

Generally, the strengths of the preferences were weak, rarely exceeding 1 scale unit, on a scale where a score of −3 or +3 would indicate a perfectly consistent and strong preference for one condition over the other. The small preference scores probably reflect four (not mutually exclusive) factors: (1) Participants were not completely consistent in their judgments. Since the maximum absolute value of the score on a single trial was 3, any variability leads to a mean score above –3 and below 3; (2) Participants are usually reluctant to use the extremes of a rating scale (Poulton, 1979; Moore & Tan, 2003). Hence, scores of –3 or 3 were very rare; (3) The preferences reflected a balance between the annoyance produced by the transients and the audibility of the transients; (4) Some of the weaker transients may not have been very annoying, for some participants leaving little room for improvement to be produced by the MCTR.

The MCTR algorithm used here differs from most transient-reduction algorithms described in the literature (with the exception of Dyballa et al., 2016), in that transients are detected and attenuated in a frequency-selective manner. Thus, attenuation of transients dominated by high frequencies did not affect the gain applied to low frequencies, and vice versa. This was intended to avoid disturbing effects of the transient reduction on the perception of speech components falling in frequency regions remote from the dominant frequencies in the transient. Although we did not evaluate the effects of the MCTR on speech quality or intelligibility, participants reported that both the quality and the intelligibility of the speech were high and did not vary across conditions. Hence, the MCTR algorithm appears to be successful in reducing the loudness and annoyance of transient sounds without affecting the quality and subjective intelligibility of the speech on which the transients are imposed.

Unlike some transient reduction systems (Hirszhorn et al., 2012), the MCTR algorithm has a very low inherent delay of about 1 ms, owing to the use of short frames and a concatenation method rather than an overlap-add method. This delay is well within the range that is acceptable for hearing aid applications (Stone & Moore, 1999; Stone et al., 2008).

Summary and conclusions

We evaluated a multi-channel transient reduction (MCTR) algorithm for detecting and attenuating transient sounds added to speech. In contrast to most previous transient-reduction algorithms, the transients were detected and attenuated in a frequency-selective manner. The MCTR was evaluated using different “strengths” of the transient reduction, using ten participants with normal hearing and ten with mild-to-moderate hearing impairment. Frequency-dependent linear amplification was provided for the latter. The results for both participant groups indicated that sounds processed using the MCTR were preferred over the unprocessed sounds. For the normal-hearing participants, the medium setting of the MCTR was preferred over the weak and strong settings. For the hearing-impaired participants, the medium and strong settings of the MCTR were preferred over the weak setting. The medium and strong settings of the MCTR reduced the annoyance produced by the transients while maintaining their audibility and without any obvious effects on speech quality or subjective speech intelligibility.

Acknowledgments

This work was supported by the Engineering and Physical Sciences Research Council (UK, grant number RG78536). We thank Aleksander Sęk for helping with the computer software for performing paired comparisons. We thank three reviewers for helpful comments on an earlier version of this paper.

Declaration of interests: The authors have no conflicts of interest to report.

References

Allen, J.B. 1977. "Short Term Spectral Analysis, Synthesis and Modification by Discrete Fourier Transform." *IEEE Transactions on Acoustics, Speech and Signal Processing*, 25, 235-238.

ANSI 2003. *ANSI S3.22-2003, Specification of Hearing Aid Characteristics*. New York: American National Standards Institute.

Boyle, P.J., Büchner, A., Stone, M.A., Lenarz, T. & Moore, B.C.J. 2009. "Comparison of Dual-Time-Constant and Fast-Acting Automatic Gain Control (Agc) Systems in Cochlear Implants." *International Journal of Audiology*, 48, 211-221.

Campbell, D.R., Palomaki, K.J. & Brown, G.J. 2008. Roomsim, a Matlab Simulation of Shoebox Room Acoustics for Use in Teaching and Research. Retrieved from http://media.paisley.ac.uk/¡«campbell/Roomsim.

Dyballa, K.H., Hehrmann, P., Hamacher, V., Lenarz, T. & Buechner, A. 2016. "Transient Noise Reduction in Cochlear Implant Users: A Multi-Band Approach." *Audiology Research*, 6, 154.

Dyballa, K.H., Hehrmann, P., Hamacher, V., Nogueira, W., Lenarz, T. et al 2015. "Evaluation of a Transient Noise Reduction Algorithm in Cochlear Implant Users." *Audiology Research*, 5, 116.

Harris, F.J. 1978. "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform." *Proceedings of the IEEE*, 66, 51-83.

Hirszhorn, A., Dov, D., Talmon, R. & Cohen, I. 2012. Transient Interference Suppression in Speech Signals Based on the OM-LSA Algorithm. *International Workshop on Acoustic Signal Enhancement*, Aachen, Germany.

Keidser, G., Convery, E., Kiessling, J. & Bentler, R. 2009. "Is the Hearing Instrument to Blame When Things Get Really Noisy?" *Hearing Review*, 16, 12-19.

Korhonen, P., Kuk, F., Lau, C., Keenan, D., Schumacher, J. et al 2013. "Effects of a Transient Noise Reduction Algorithm on Speech Understanding, Subjective Preference, and Preferred Gain." *Journal of the American Academy of Audiology*, 24, 845-858.

Launer, S., Zakis, J. & Moore, B.C.J. 2016. Hearing Aid Signal Processing. *In*: G.R. Popelka, B.C.J. Moore, A.N. Popper & R.R. Fay (eds.) *Hearing Aids*, New York: Springer, pp. 93-130.

Luo, H. 2009. "Acoustic Shock Detection and Control." *Canadian Acoustics*, 37, 96-97.

McFerran, D.J. & Baguley, D.M. 2007. "Acoustic Shock." *Journal of Laryngology and Otology*, 121, 301-305.

Moore, B.C.J. 2007. *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues,* *2nd Ed.* Chichester: Wiley.

Moore, B.C.J. & Füllgrabe, C. 2010. "Evaluation of the CAMEQ2-HF Method for Fitting Hearing Aids with Multi-Channel Amplitude Compression." *Ear and Hearing*, 31, 657-666.

Moore, B.C.J. & Glasberg, B.R. 1988. "A Comparison of Four Methods of Implementing Automatic Gain Control (Agc) in Hearing Aids." *British Journal of Audiology*, 22, 93-104.

Moore, B.C.J. & Glasberg, B.R. 1998. "Use of a Loudness Model for Hearing Aid Fitting. I. Linear Hearing Aids." *British Journal of Audiology*, 32, 317-335.

Moore, B.C.J., Glasberg, B.R. & Stone, M.A. 1991. "Optimization of a Slow-Acting Automatic Gain Control System for Use in Hearing Aids." *British Journal of Audiology*, 25, 171-182.

Moore, B.C.J., Glasberg, B.R. & Stone, M.A. 2010. "Development of a New Method for Deriving Initial Fittings for Hearing Aids with Multi-Channel Compression: CAMEQ2-HF." *International Journal of Audiology*, 49, 216-227.

Moore, B.C.J. & Sek, A. 2013. "Comparison of the Cam2 and NAL-NL2 Hearing-Aid Fitting Methods." *Ear and Hearing*, 34, 83-95.

Moore, B.C.J. & Tan, C.T. 2003. "Perceived Naturalness of Spectrally Distorted Speech and Music." *Journal of the Acoustical Society of America*, 114, 408-419.

Poulton, E.C. 1979. "Models for the Biases in Judging Sensory Magnitude." *Psychological Bulletin*, 86, 777-803.

Schneider, T., Brennan, R.L., Hermann, D. & Soltani, T. 2010. Method and System for Acoustic Shock Protection. USA Patent, US7672462 B2.

Stelmachowicz, P.G., Lewis, D.E., Hoover, B. & Keefe, D.H. 1999. "Subjective Effects of Peak Clipping and Compression Limiting in Normal and Hearing-Impaired Children and Adults." *Journal of the Acoustical Society of America*, 105, 412-422.

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 and Hearing*, 20, 182-192.

Stone, M.A., Moore, B.C.J., Alcántara, J.I. & Glasberg, B.R. 1999. "Comparison of Different Forms of Compression Using Wearable Digital Hearing Aids." *Journal of the Acoustical Society of America*, 106, 3603-3619.

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 and Hearing*, 29, 601-617.

Tan, C.T. & Moore, B.C.J. 2008. "Perception of Nonlinear Distortion by Hearing-Impaired People." *International Journal of Audiology*, 47, 246-256.

Zeng, F.G. & Shannon, R.V. 1999. "Psychophysical Laws Revealed by Electric Hearing." *Neuroreport*, 10, 1931-1935.

482 **Table 1.** Characteristics of the transient sounds and the peak levels used for condition none
483 (see text for details), measured over a 10-ms time interval and expressed relative to the RMS
484 level of the speech.

485

Transient number	Description	Rise time (ms)	10-ms peak level
T1	A concrete block hit with a metal hammer	1	20 dB
T2	Two water glasses tapped together	3	20 dB
T3	A glass jar filled with glass marbles, shaken once	4	20 dB
T4	A metal object struck with a metal hammer	<1	16 dB
T5	A set of keys dropped on a wooden table	<7	16 dB
T6	A metal can filled with metal bolts, shaken once	7	18 dB
T7	Two metal rails hit together	1	20 dB
T8	A plastic ball-point pen being clicked	<1	18 dB
T9	A metal spoon being swirled in a porcelain cup	4	16 dB

486

487

Figure captions

Figure 1. Panel a (top) shows the attention C_{ij} (in dB) plotted as a function of the ratio M_{ij}/RMS_{ij} (in dB) for three values of constant α, 0.267 (right-pointing triangles), 0.467 (circles), and 0.933 (crosses). For values of the ratio below 0 dB, no attenuation was applied. Panel b (bottom) shows the resulting output level as a function of input level.

Figure 2. Illustration of the operation of the MCTR algorithm for conditions none (no transient reduction), weak, medium, and strong. The waveform of the speech+transient is shown for each condition.

Figure 3. Preference scores for each transient and each comparison for the normal-hearing participants. Each panel shows results for a different comparison, as indicated in the key.

Figure 4. As Figure 3 but for the hearing-impaired participants.

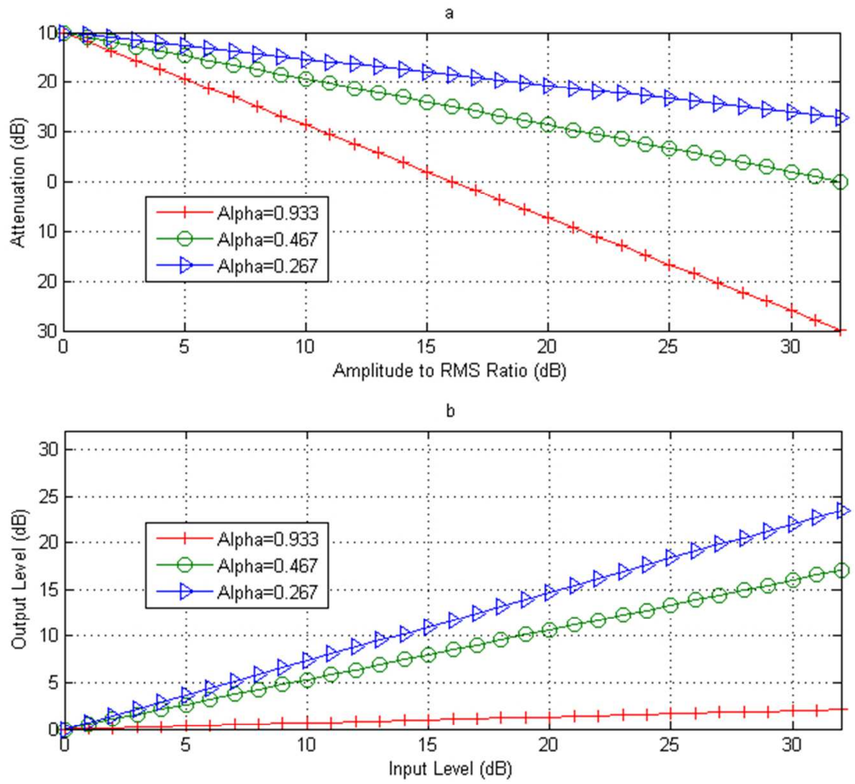

Figure 1. Panel a (top) shows the attention C_{ij} (in dB) plotted as a function of the ratio M_{ij}/RMS_{ij} (in dB) for three values of constant a, 0.267 (right-pointing triangles), 0.467 (circles), and 0.933 (crosses). For values of the ratio below 0 dB, no attenuation was applied. Panel b (bottom) shows the resulting output level as a function of input level.

190x161mm (96 x 96 DPI)

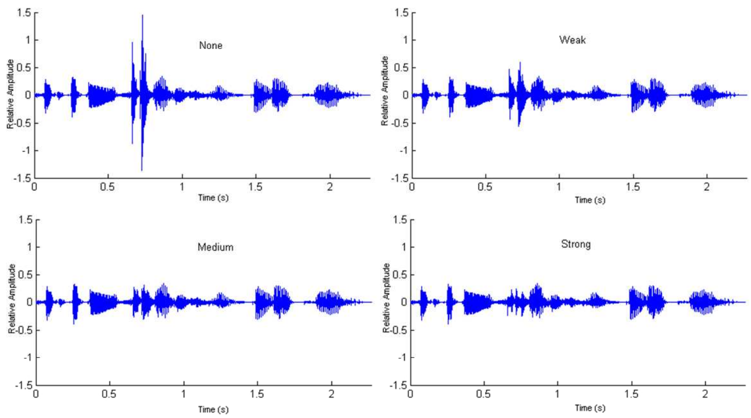

Figure 2. Illustration of the operation of the MCTR algorithm for conditions none (no transient reduction), weak, medium, and strong. The waveform of the speech+transient is shown for each condition.

338x174mm (96 x 96 DPI)

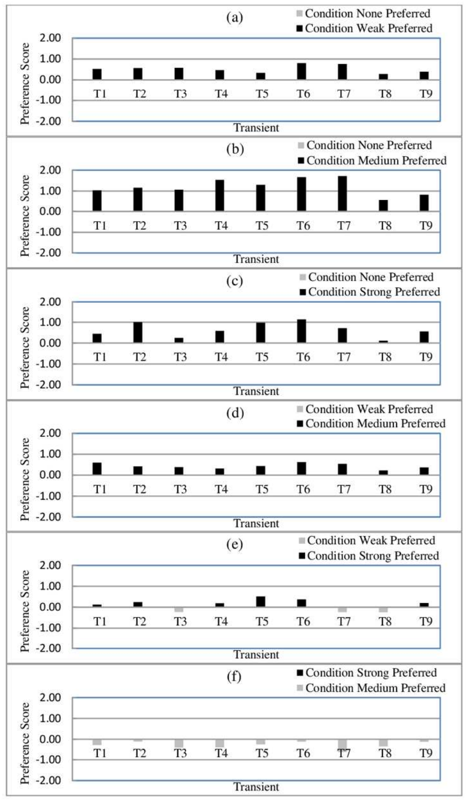

Figure 3. Preference scores for each transient and each comparison for the normal-hearing participants. Each panel shows results for a different comparison, as indicated in the key.

215x279mm (200 x 200 DPI)

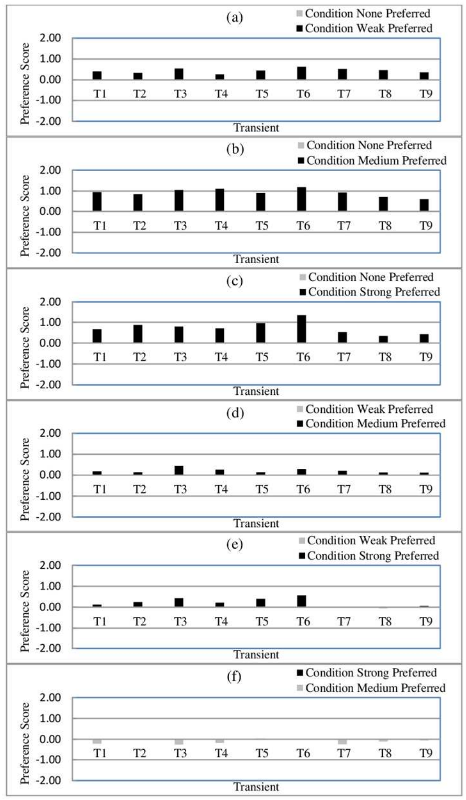

Figure 4. As Figure 3 but for the hearing-impaired participants.

215x279mm (200 x 200 DPI)