

EFFECT OF DIGITAL NOISE REDUCTION AND DIRECTIONALITY ALGORITHMS IN HEARING AIDS ON TEMPORAL ENVELOPE DISTORTION AND SPEECH RECOGNITION

Contributions:

A Study design/planning
B Data collection/entry
C Data analysis/statistics
D Data interpretation
E Preparation of manuscript
F Literature analysis/search
G Funds collection

Geetha Chinnaraj^{A,C-E,G}, Hemanth Narayan Shetty^{A,D-E},
Vinodhini Palaniyappan^{B-C,E-F}

Department of Audiology, All India Institute of Speech and Hearing, India

Corresponding author: Geetha Chinnaraj; Department of Audiology, All India Institute of Speech and Hearing, Manasagangothri, 570006, Mysore, India; email: geethamysore.cs@gmail.com

Abstract

Background: Digital signal processing algorithms tend to alter temporal cues in the speech signal and yet individuals with hearing impairment rely strongly on these cues for speech perception. Hence, there is a need to assess the effect of these algorithms on temporal cues and speech perception. The present study aimed to quantify, in a wide dynamic range compression hearing aid, the individual and combined effects of digital noise reduction and directionality algorithms on temporal cues, syllable recognition, and sentence recognition. Temporal cues were quantified by the envelope difference index (EDI).

Material and methods: The study included 20 individuals (in the age range of 21 to 44 years) with mild to moderate sensorineural hearing loss. Sentence recognition, syllable recognition, and EDI were obtained at different levels in four different aided conditions – digital noise reduction algorithm only, directionality only, both digital noise reduction and directionality on, and both algorithms off. Sentences were presented in noise at +5 dB signal-to-noise ratio from 180° azimuth.

Results: Compared to independent activation of the algorithms, the combined algorithms significantly improved speech recognition scores at all presentation levels. The temporal changes induced by the algorithms were only mild, even though EDI was the highest when all the algorithms and the directional microphone were activated.

Conclusions: The noise reduction algorithms and compression induce temporal changes; however speech recognition improved when the algorithms were activated, presumably due to countervailing psychophysical factors.

Key words: envelope difference index • sentence recognition • temporal envelope • directionality • DNR

WPLYW ALGORYTMÓW CYFROWEJ REDUKCJI SZUMÓW I KIERUNKOWOŚCI W APARATACH SŁUCHOWYCH NA ZNIEKSZTAŁCENIE OBWIEDNI CZASOWEJ I ROZPOZNAWANIE MOWY

Streszczenie

Wprowadzenie: Algorytmy cyfrowego przetwarzania sygnału mogą zmieniać struktury czasowe w sygnale mowy, przy czym osoby z niedosłuchem w znacznym stopniu polegają na tych strukturach w percepcji mowy. Stąd wynika potrzeba oceny wpływu tych algorytmów na struktury czasowe i percepcję mowy. Celem obecnego badania było oszacowanie – dla aparatu słuchowego z kompresją szerokiego zakresu dynamiki – indywidualnego i łącznego wpływu algorytmów cyfrowej redukcji szumów i kierunkowości na struktury czasowe, rozpoznawanie sylab i zdań. Wskazówki czasowe były szacowane na podstawie wskaźnika różnicy obwiedni (ang. *envelope difference index*, EDI).

Material i metody: W badaniu uczestniczyło 20 osób (w wieku od 21 do 44 lat) z lekkim do umiarkowanego niedosłuchem odbiorczym. Uzyskano wyniki rozpoznawania zdań, rozpoznawania sylab oraz EDI dla różnych poziomów w czterech różnych sytuacjach wspomaganego słyszenia: tylko algorytm cyfrowej redukcji szumów, tylko kierunkowość, zarówno redukcja szumów, jak i kierunkowość, obydwa algorytmy włączone. Zdania były prezentowane w szumie przy stosunku sygnału do szumu +5 dB z azymutu 180°.

Wyniki: Połączone algorytmy – w porównaniu do niezależnego włączenia algorytmów – w znacznym stopniu poprawiły wyniki rozpoznawania mowy na wszystkich poziomach prezentacji. Zmiany struktury czasowej wywołane przez algorytmy były niewielkie, chociaż wskaźnik EDI był najwyższy, kiedy były włączone wszystkie algorytmy i mikrofon kierunkowy.

Wnioski: Algorytmy redukcji szumów i kompresji wywołują zmiany struktury czasowej, jednak rozpoznawanie mowy poprawiało się przy włączonych algorytmach, przypuszczalnie w związku z wyrównawczymi czynnikami psychofizycznymi.

Słowa kluczowe: wskaźnik różnicy obwiedni • rozpoznawanie zdań • obwiednia czasowa • kierunkowość • DNR

Background

Difficulty in speech recognition is a common problem in individuals with hearing impairment. Sensorineural hearing impairment affects speech recognition more in noise than in quiet [1]. In the natural environment, when noise accompanies speech it alters the envelope of the speech signal. Along with natural changes in the speech envelope due to noise, digital signal processing algorithms in hearing aids tend to further alter the characteristics of speech. Commonly used algorithms are wide dynamic range compression (WDRC), directionality, and digital noise reduction (DNR) [2]. Compared to linear amplification, WDRC alters the temporal envelope of speech [3,4]. It has also been reported that when DNR is used to improve comfort in the presence of noise it alters the temporal envelope of the incoming speech signal [5].

Temporal envelope effects on speech recognition

Several studies have examined the effects of the temporal envelope on speech recognition. Reports have shown that changes in the signal envelope result in poorer perception [6–9]. Tachibana and colleagues [10] reported that the contribution of temporal envelope cues was more important than spectral cues in some speech recognition tasks. The contribution of temporal envelope cues on speech recognition increased with an increase in the linguistic structure of the material. Li and colleagues [11] assessed the contribution of temporal and spectral envelope cues on consonant and vowel recognition and compared individuals with normal hearing sensitivity to those with hearing impairment. Their results revealed that the role of the temporal envelope is greater for recognition of consonants while spectral cues contribute more to recognition of vowels.

However, individuals with hearing impairment are assumed to rely more on the temporal envelope of the speech signal as they have broadened auditory filters, resulting in poor spectral resolution. Hence, in the present study the temporal envelope and its effect on speech recognition was measured in individuals with hearing impairment. In most of the above studies the temporal envelope was quantified using the envelope difference index (EDI), which is an objective way of computing the difference in the temporal envelope of two signals, and hence, this index was also used here.

Envelope Difference Index

The EDI was originally developed by Fortune and colleagues [12]. EDI ranges from 0 to 1 and quantifies the changes in the envelope of a processed stimulus compared to the unprocessed stimulus. A value of 0 means identical envelopes while 1 indicates completely different envelopes. The present study also used EDI to quantify the temporal changes since there is substantial evidence that EDI correlates well with speech recognition [4,13,14]. However, most of the studies on EDI have been carried out to evaluate the effect of compression on the temporal envelope. Jenstad and Souza [13,14] reported that an increase in EDI resulted in a decrease in speech recognition scores (SRSs) for release times of 12, 100, and 800 ms in a WDRC hearing

aid at 3:1 compression ratio. Souza et al. [4] studied the effect of compression using combinations of compression ratios (1:1, 2:1, 4:1, 12:1) and release times (12 and 800 ms) on the processing of vowel–consonant–vowel (VCV) stimuli. The higher the value of EDI, the greater was the voicing error and manner error for plosives and fricatives. In the case of fricatives, compression resulted in distortion of the speech signal. The above-mentioned studies have mostly focused on WDRC, but not other algorithms such as DNR and directionality.

DNR and directionality algorithms also tend to alter the temporal cues in a speech signal. Noise reduction strategies focus on envelope differences between the speech and noise across different frequency bands. Most hearing aids use modulation detectors to gauge the modulation depth of the signal and to reduce gain in the channels where the modulation depth is smaller [15]. This process of cutting down noise is supposed to deepen the temporal modulation of the signal [16] helping individuals with hearing impairment who tend to rely more on temporal cues.

Further, in real-world applications, many digital signal processing (DSP) algorithms may work simultaneously, depending on the environment; however, most of the above studies have assessed the independent effects of the DSP algorithms. Nevertheless, Walden et al. [17] have assessed the speech recognition benefits of these algorithms in elderly hearing aid users both in isolation and when combined. They found that directionality did improve speech recognition compared to the omnidirectional mode. However, DNR did not add any significant benefit. Nordrum et al. [18] also evaluated directionality and DNR algorithms in adult listeners with a hearing aid and found similar results.

A study by Geetha and Manjula [19] has quantified the acoustic changes induced by compression, DNR, and directionality – in isolation and in combination – using EDI and measured the quality ratings. The authors recorded the output of the hearing aid in a 2 cc coupler and processed it to obtain an EDI. They also obtained quality ratings from normal hearing individuals based on the recorded outputs. They found that the temporal changes observed in the output of the hearing aid did not depend on the number of algorithms activated. The results of the quality analysis revealed no significant difference in clarity rating across the aided conditions. However, there was only a single measure of EDI in the 2cc coupler, and so a statistical correlation could not be obtained between the EDI and the quality rating. Furthermore, the study included individuals with normal hearing sensitivity and hence generalization of the results is questionable. In addition, the study did not include any perceptual measures. Vinodhini [20] studied the combined effect of WDRC, DNR, and directionality and reported a moderate-level positive correlation between EDI and speech perception for sentences. The effects of noise reduction strategies were not measured independently.

Hence, in the present study an attempt is made to answer the following research questions: 1. Does the temporal envelope (as quantified by EDI) and perceptual scores change with activation of DNR and directionality algorithms alone

and together? 2. Does the EDI correlate with the perceptual scores under these conditions?

Effect of different linguistic structures

It is known that a hearing aid's behavior for sentences can be different from that of syllables. Most studies on EDI have used syllables and have found a good correlation between EDI and speech recognition. However, contrary to previous studies, Vinodhini [20] found no correlation between EDI and aided sentence recognition. This implies that the temporal envelope is the most important cue for syllables, but for sentence recognition redundancy cues also play a significant role [4]. Nevertheless, Jenstad and Souza [14] reported that the EDI obtained for syllables and sentences are comparable for similar hearing aid compression conditions. Jenstad and Souza [15] found a relationship between an increased EDI and a reduction in sentence recognition rate in quiet. Similar effects in noise were reported by Walaszek [21].

The effect of the length of the stimuli (syllables vs. sentences) on EDI is unclear, and hence in the present study two kinds of target stimuli – sentences and VCV stimuli – were included. For VCV stimuli, consonants combined with the vowel /i/ in the initial and final positions were used. A second main aim was to gauge the importance of the type of stimulus (its length) to assess the role of temporal envelope distortion on the perception of different type of stimuli in adult individuals with hearing impairment.

Material and methods

Participants

The present study included a total of 20 individuals (12 right ears and 8 left ears; 7 females, 13 males) who had a mild to moderate degree of sensorineural post-lingual hearing loss. The configuration was considered flat if the threshold difference between two adjacent octave frequencies was not more than 10 dB HL in the frequency range 0.25 to 8 kHz [22]. All the participants had a flat or gradual slope hearing loss. The participants' ages ranged from 21 to 44 years (mean = 36.8; SD = 6.8). The speech perception ability of each participant was assessed by obtaining word recognition scores. The unaided word recognition scores in each ear were not less than 70% for all the cases. All the participants had 'A' or 'As' type tympanograms and had acoustic reflexes. All had no prior hearing aid experience and were native speakers of the Kannada language. None of them had any history of psychological or neurological issues.

Hearing aid fitting

The participants were fitted with a 16-channel hearing aid connected to a personal computer with NOAH-3 software using a Noah link interface. The hearing aid was programmed using NAL-NL1 formula and acclimatization level was set to 2. NAL-NL1 was used as it is the most studied and accepted method for fitting hearing aids and because it was the latest set of NAL equations available in our clinic at the time of the study. The gain was further optimized until all the sounds in the Ling's six sounds test

were identified at normal conversational level. A routine hearing aid evaluation was carried out by obtaining word recognition scores for 25 words at 40 dB HL.

Test conditions

Different aided conditions for which the hearing aid was programmed were: 1) DNR only; 2) directionality only; 3) both DNR and directionality on; and 4) both DNR and directionality off. The presentation levels were 55, 65, and 80 dB SPL. In all the aided conditions, WDRC was enabled.

Testing was also done in the unaided condition. Tristate noise reduction and multi-band adaptive directionality were used in the hearing aid. The compression knee point was 50 dB SPL and the compression ratio was 3:1. The compression system used an attack time of 8 ms and release time of 350 ms. The testing was done using vowel-consonant-vowel (VCV) stimuli and sentence recognition scores (SRSs).

VCV recognition and SRS

Recorded sentences developed by Geetha et al. [23] were used to obtain aided recognition scores for sentences and recorded VCV stimuli were used to obtain VCV perceptual scores. The sentence test consisted of 25 equivalent phonemically balanced sentence lists. Each list had 10 sentences and 40 keywords. The VCV stimuli contained 21 consonants (/k/, /g/, /tʃ/, /t/, /d/, /ŋ/, /t/, /d/, /n/, /p/, /b/, /m/, /j/, /r/, /l/, /v/, /ʃ/, /s/, /h/, /l/, and /dʒ/), which frequently occur in the Kannada language [24]. These consonants were paired with low short central vowel /i/ in the initial and final position. Vowel /i/ was used as it has high-frequency energy extending up to 6 kHz and hence contains a longer transition [25]. The consonants were spoken by three female native speakers of Kannada. The recordings were done by placing the microphone 10 cm from the mouth of the speaker [26] and using Adobe Audition version 3.0. The stimuli were digitized using a 32-bit processor with a sampling frequency of 44.1 kHz. In addition, the subjective naturalness of the stimuli was checked: 10 normal-hearing individuals were asked to rate the naturalness on a 3-point rating scale and the speaker rated to have the most natural utterance was selected. Recording was done in a sound-treated room.

Presentation of the stimuli was at 55, 65, and 80 dB SPL (measured at the listening position, before hearing aid processing), and sent through a Lynx aurora signal router to Genelec 8020B loudspeakers (Thomann GmbH, Burgebrach, Germany) mounted on an Iso-Pod vibration insulating stand placed at 0° azimuth, and speech-shaped noise (noise spectrally matched to speech and gated on and off with the stimuli) was presented through a loudspeaker placed at 180° azimuth. A pilot study was carried out with five participants to track the signal-to-noise ratio (SNR) at which 50% scores were obtained, hereafter called SNR-50. An SNR-50 was evaluated on the sentence material. The results showed that a range of +4 to +7 dB SNR was required to obtain an SNR-50. Hence, the experiment was carried out at +5 dB SNR. Listeners were fitted using soft moulds with no venting and were asked to repeat the

words in the sentences. The responses were noted down on a response sheet. The SRS was calculated based on the total number of key words repeated correctly for each list. The maximum number of key words in each list was 40. The same procedure was done in unaided and all aided conditions. The VCV stimuli were presented in quiet at three different input levels using the same set up. The number of correctly repeated VCV stimuli was computed. The entire testing was done in a sound-treated double room.

Computation of EDI

The hearing aid was fitted using routine hearing aid fitting protocols (details in ‘hearing aid fitting’ section). The hearing aid programmed for each individual was fitted to a KEMAR head and torso simulator (Knowles Electronics Manikin for Acoustic Research, Model No. 45BB, manufactured by G.R.A.S. Sound and Vibration, mounted with an RA0045 ear simulator and ½-inch microphone) using soft moulds with no venting, and the output was recorded using a sound level meter in all the conditions mentioned in the ‘test conditions’ section. The recorded stimuli were edited using Adobe Audition. The aided recordings were compared to the unaided conditions with EDI software [12] to obtain an EDI score.

The method of EDI developed by Fortune et al. (1994) was adopted to determine the extent to which the hearing aid altered the natural temporal characteristic of the sentence in each experimental condition. The EDI was calculated using a custom script in Matlab (version R2009b, MathWorks, Inc., USA). The envelope of both signals was extracted separately by a full-wave rectification process followed by low-pass filtering at 50 Hz using a 6th-order Butterworth filter. Filtering was done in both forward and backward directions. Both envelopes were scaled to a common reference point by dividing them by the respective mean amplitude of the signal. Then the processed and unprocessed stimuli were cross-correlated and time-aligned before EDI measurement to avoid errors arising from temporal misalignment [27]. The EDI was computed using the Matlab code for each of the experimental conditions using the following equation.

$$EDI = \left(\sum_{n=1}^N |Env1_n - Env2_n| \right) / 2N$$

where Env1 is the envelope of the unprocessed signal (stimuli recorded in the KEMAR stimulus using the same set-up and procedure as used in the recording of the processed stimuli), Env2 is the envelope of signal 2 (stimuli recorded in the KEMAR stimulus with hearing aid on using the same set up and procedure as used in the recording of unprocessed stimuli), and N is the number of samples in each signal.

Results

The EDI obtained for sentences and EDI obtained for VCV stimuli were compared across different aided conditions. Correlation analysis was done between EDI and the subjective perceptual scores. The results are presented below.

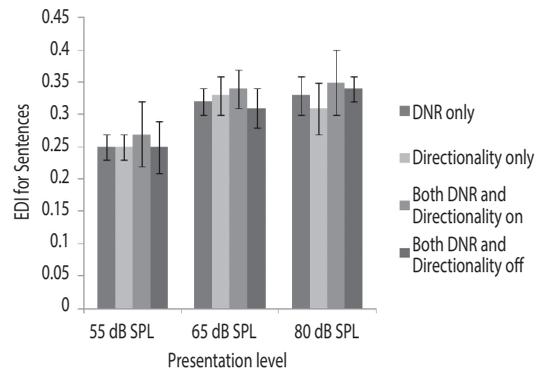


Figure 1. Envelope Difference Index (EDI) measured for sentences is compared across different aided conditions at 55, 65, and 85 dB SPL presentation levels. DNR = digital noise reduction algorithm

EDI for sentences and VCV stimuli

The EDI obtained for sentences are shown in Figure 1. The EDI for sentences ranged from 0.25 to 0.35. A greater EDI score indicates a greater change in the temporal envelope of the processed signal compared to the unprocessed signal.

A Shapiro–Wilk test of normality was carried out, and the results showed that the EDI data did not follow a normal distribution. Hence, a Friedman test was used to compare EDI across different aided conditions at each presentation level. The result of Friedman’s test showed that the EDI was significantly different across aided conditions at 65 dB SPL ($\chi^2(3) = 22.061, p < 0.001$) and at 80 dB SPL ($\chi^2(3) = 9.091, p < 0.001$) and was not significantly different at 55 dB SPL. Hence, a pairwise analysis was carried out using the Wilcoxon signed rank test across different aided conditions at 65 and 80 dB SPL.

Results of the Wilcoxon signed rank test comparing different conditions at 65 dB SPL revealed that the ‘Directionality only’ condition resulted in a significantly greater EDI compared to the ‘DNR only’ condition ($|Z| = -1.918, p < 0.05, r = -0.428$) and the reverse results were seen at 80 dB SPL – that is, the Directionality condition resulted in a significantly smaller EDI compared to the ‘DNR only’ condition ($|Z| = -2.580, p < 0.01, r = -0.576$). At 65 dB SPL, the ‘Both DNR and directionality on’ condition resulted in greater EDI compared to the ‘DNR only’ condition ($|Z| = -2.933, p < 0.01, r = -0.655$) and ‘Both DNR and directionality off’ condition ($|Z| = -2.765, p < 0.01, r = 0.618$). In addition, the ‘Directionality only’ condition resulted in greater EDI compared to the ‘Both DNR and directionality off’ condition ($|Z| = -3.175, p < 0.001, r = -0.709$). In comparison, at 80 dB SPL, the ‘Both DNR and directionality on’ ($|Z| = -2.839, p < 0.001, r = -0.634$) and ‘Both DNR and directionality off’ ($|Z| = -2.724, p < 0.01, r = -0.609$) conditions resulted in greater EDI compared to the ‘Directionality only’ condition.

The EDI obtained for VCV stimuli across the conditions are given in Figure 2. The EDI obtained for 21 VCV stimuli were averaged for each aided condition and computed for analysis. The EDI ranged from 0.140 to 0.220 for VCV stimuli.

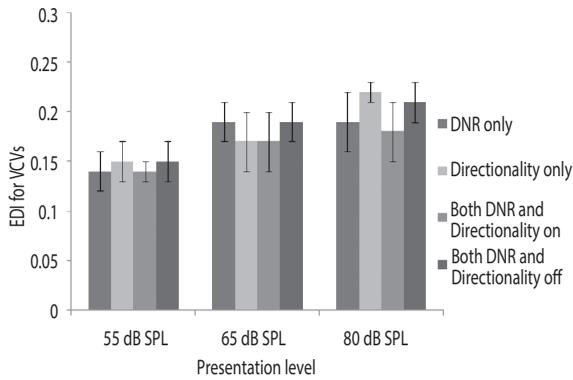


Figure 2. Envelope Difference Index (EDI) measured for VCV stimuli is compared across different aided conditions at 55, 65, and 85 dB SPL presentation levels. DNR = digital noise reduction algorithm

The EDI for VCV stimuli was also statistically compared across different aided conditions using Friedman’s test. The results showed that the EDI obtained was significantly different across aided conditions at 65 dB SPL ($\chi^2(3) = 14.589, p < 0.01$) and 80 dB SPL ($\chi^2(3) = 10.706, p < 0.05$) presentation levels. Further, pairwise analysis was carried out with the Wilcoxon signed rank test. At 65 dB SPL, EDI for VCV stimuli in the ‘DNR only’ condition was significantly greater than the ‘Directionality only’ ($|Z| = -2.879, p < 0.01, r = -0.643$) and ‘Both DNR and directionality on’ conditions ($|Z| = -2.162, p < 0.05, r = -0.483$). The ‘Both off’ condition yielded a significantly greater EDI compared to the ‘Directionality only’ ($|Z| = -2.844, p < 0.01, r = -0.635$) and ‘Both DNR and directionality on’ conditions ($|Z| = -1.984, p < 0.05, r = -0.443$). At 80 dB SPL, the ‘Directionality only’ condition yielded significantly greater EDI compared to the ‘DNR only’

($|Z| = -2.419, p < 0.05, r = -0.540$) and ‘Both DNR and directionality on’ conditions ($|Z| = -2.916, p < 0.01, r = -0.652$). The ‘Both DNR and directionality off’ condition resulted in significantly greater EDI than the ‘Both DNR and directionality on’ condition ($|Z| = -3.181, p < 0.001, r = -0.711$). There was no difference among any other conditions.

Analysis of SRS perceptual scores

SRS and VCV stimuli scores obtained across different conditions are given in Table 1. The results of a Shapiro–Wilk test showed that the data did not follow a normal distribution. Hence, Friedman’s test and Wilcoxon signed rank tests were carried out to compare SRS across different aided conditions.

The results of Friedman’s test showed a significant change in SRS across different aided conditions at 55 dB SPL ($\chi^2(4) = 69.048, p < 0.001$), at 65 dB SPL ($\chi^2(4) = 52.818, p < 0.001$), and at 80 dB SPL ($\chi^2(4) = 57.244, p < 0.001$). The results of the Wilcoxon signed rank test along with the effect size are given in Table 2.

The results also revealed that, among the aided conditions, combined activation of the algorithms resulted, at all presentation levels, in significantly higher SRSs compared to independent activation of the algorithms. Further, the SRS obtained from the ‘DNR only’ condition was significantly greater than the ‘Directionality only’ condition. SRSs obtained for the ‘Both DNR and Directionality on’ condition yielded a greater SRS compared to the ‘DNR only’ and ‘Directionality only’ conditions at all presentation levels. All the above differences had a large effect size.

The results of Friedman’s test showed that the scores for VCV stimuli syllables at 55 dB SPL ($\chi^2(4) = 48.764, p < 0.001$),

Table 1. Mean, median, and SD of SRS (number of correctly identified key words out of 40) obtained across aided conditions and presentation levels ($n = 20$)

Aided condition	Presentation level (SPL)	SRS			VCV stimulus recognition		
		Mean	Median	SD	Mean	Median	SD
Unaided	55 dB	04.55	03.00	5.06	01.15	1.00	1.03
	65 dB	14.50	12.50	5.95	01.90	2.00	1.54
	80 dB	23.00	24.50	5.69	04.75	4.50	2.08
DNR only	55 dB	15.95	16.50	3.69	04.97	5.00	1.55
	65 dB	32.35	33.00	3.57	11.67	12.00	1.24
	80 dB	33.90	35.00	3.55	12.25	12.50	2.01
Directionality only	55 dB	13.40	13.00	3.33	03.07	3.00	1.61
	65 dB	30.40	30.50	2.82	10.90	10.50	1.47
	80 dB	30.80	31.00	1.74	12.12	12.00	1.76
Both DNR and Directionality on	55 dB	20.40	21.00	3.25	05.65	5.00	2.40
	65 dB	33.10	32.50	2.63	12.02	12.50	2.09
	80 dB	34.50	35.00	2.70	12.17	12.00	1.83
Both DNR and Directionality off	55 dB	14.95	14.00	2.56	04.22	4.00	1.53
	65 dB	29.30	30.00	3.01	11.10	10.50	2.08
	80 dB	31.00	31.00	1.62	11.10	11.00	2.07

Note: Maximum possible score for SRS = 40; maximum possible score for VCV stimulus = 21; DNR = digital noise reduction algorithm

Table 2. Results of Wilcoxon signed rank test for comparison of SRS

	Conditions	Z	p	Effect size (r)
55 dB SPL	DNR only vs. Directionality only	-3.194	0.001**	-0.714
	DNR only vs. Both DNR and Directionality on	-3.947	0.000***	-0.882
	DNR only vs. Both DNR and Directionality off	-1.370	0.171	-
	Directionality only vs. Both DNR and Directionality on	-3.948	0.000***	-0.882
	Directionality only vs. Both DNR and Directionality off	-2.145	0.032*	-0.479
	Both DNR and Directionality on vs. Both DNR and Directionality off	-3.943	0.000***	-0.881
65 dB SPL	DNR only vs. Directionality only	-2.002	0.045*	-0.447
	DNR only vs. Both DNR and Directionality on	-0.986	0.324	-0.220
	DNR only vs. Both DNR and Directionality off	-3.561	0.000***	-0.796
	Directionality only vs. Both DNR and Directionality on	-3.026	0.002**	-0.676
	Directionality only vs. Both DNR and Directionality off	-0.965	0.334	-
	Both DNR and Directionality on vs. Both DNR and Directionality off	-3.414	0.001**	-0.763
80 dB SPL	DNR only vs. Directionality only	-3.266	0.001**	-0.730
	DNR only vs. Both DNR and Directionality on	-0.646	0.519	-
	DNR only vs. Both DNR and Directionality off	-2.738	0.006**	-0.612
	Directionality only vs. Both DNR and Directionality on	-3.595	0.000***	-0.803
	Directionality only vs. Both DNR and Directionality off	-0.418	0.676	-
	Both DNR and Directionality on vs. Both DNR and Directionality off	-3.619	0.000***	-0.809

Note: * $p < 0.05$; ** $p < 0.01$; DNR = digital noise reduction algorithm

Table 3. Comparison of VCV stimulus scores obtained across aided conditions using Wilcoxon signed rank test

Presentation level	Conditions	Z	p	Effect size (r)
55 dB SPL	DNR only vs. Directionality only	-3.280	0.001**	-0.733
	DNR only vs. Both DNR and Directionality on	-1.092	0.275	-
	DNR only vs. Both DNR and Directionality off	-1.761	0.078	-
	Directionality only vs. Both DNR and Directionality on	-3.100	0.002**	-0.693
	Directionality only vs. Both DNR and Directionality off	-3.195	0.001**	-0.714
	Both DNR and Directionality on vs. Both DNR and Directionality off	-1.734	0.083	-
65 dB SPL	DNR only vs. Directionality only	-2.131	0.033*	-0.516
	DNR only vs. Both DNR and Directionality on	-0.883	0.377	-
	DNR only vs. Both DNR and Directionality off	-1.454	0.146	-
	Directionality only vs. Both DNR and Directionality on	-1.730	0.084	-
	Directionality only vs. Both DNR and Directionality off	-0.339	0.734	-
	Both DNR and Directionality on vs. Both DNR and Directionality off	-1.763	0.078	-
80 dB SPL	DNR only vs. Directionality only	-0.749	0.454	-
	DNR only vs. Both DNR and Directionality on	-0.357	0.721	-
	DNR only vs. Both DNR and Directionality off	-2.099	0.036	-
	Directionality only vs. Both DNR and Directionality on	-0.546	0.585	-
	Directionality only vs. Both DNR and Directionality off	-2.346	0.019*	-0.524
	Both on vs. Both DNR and Directionality off	-2.477	0.013*	-0.553

Note: * $p < 0.05$; ** $p < 0.01$; DNR = digital noise reduction algorithm

65 dB SPL ($\chi^2(4) = 47.000$, $p < 0.001$), and 80 dB SPL ($\chi^2(4) = 44.699$, $p < 0.001$) presentation levels were significantly different between the aided conditions. Further

analysis was done using a Wilcoxon signed rank test. The results of this are given in Table 3.

Table 4. Comparison of EDI obtained across stimuli using Wilcoxon signed rank test

Presentation level	Aided condition	Z	Significance	Effect size (r)
55 dB SPL	DNR only	-3.922	<0.001***	-0.876
	Directionality only	-3.890	<0.001***	-0.869
	Both DNR and Directionality on	-3.885	<0.001***	-0.868
	Both DNR and Directionality off	-3.923	<0.001***	-0.877
65 dB SPL	DNR only	-3.922	<0.001***	-0.876
	Directionality only	-3.924	<0.001***	-0.877
	Both DNR and Directionality on	-3.922	<0.001***	-0.876
	Both DNR and Directionality off	-3.933	<0.001***	-0.879
80 dB SPL	DNR only	-3.925	<0.001***	-0.877
	Directionality only	-3.739	<0.001***	-0.836
	Both DNR and Directionality on	-3.923	<0.001***	-0.877
	Both DNR and Directionality off	-3.923	<0.001***	-0.877

Note: *** $p < 0.001$; EDI = Envelope Detection Index; DNR = digital noise reduction algorithm

Table 5. Comparison of speech perception scores obtained across stimuli using Wilcoxon signed rank test

Presentation level	Aided condition	Z	Significance	Effect size (r)
55 dB SPL	Unaided	-1.334	0.182	-
	DNR only	-3.608	0.000***	-0.806
	Directionality only	-3.628	0.000***	-0.811
	Both DNR and Directionality on	-3.683	0.000***	-0.823
	Both DNR and Directionality off	-3.926	0.000***	-0.877
65 dB SPL	Unaided	-3.926	0.000***	-0.877
	DNR only	-3.854	0.000***	-0.861
	Directionality only	-3.930	0.000***	-0.878
	Both DNR and Directionality on	-3.888	0.000***	-0.869
	Both DNR and Directionality off	-3.683	0.000***	-0.823
80 dB SPL	Unaided	-3.923	0.000***	-0.877
	DNR only	-3.928	0.000***	-0.878
	Directionality only	-3.785	0.000***	-0.846
	Both DNR and Directionality on	-3.832	0.000***	-0.856
	Both DNR and Directionality off	-3.926	0.000***	-0.877

Note: *** $p < 0.001$; DNR = digital noise reduction algorithm

The results revealed that recognition scores for VCV stimuli in the 'Directionality only' condition were significantly lower compared to the 'DNR only', 'Both DNR and Directionality on', and 'Both DNR and Directionality off' conditions at 55 dB SPL. Even at 65 dB SPL, the scores for VCV stimuli were significantly smaller for the 'Directionality only' condition compared to the 'DNR only' condition. At 80 dB SPL, the 'Both DNR and Directionality off' condition resulted in significantly smaller scores for VCV stimuli compared to the 'Directionality only' and 'Both DNR and Directionality on' conditions. There was no significant difference among any other conditions at any presentation level.

Correlation between EDI and SRS

Spearman's correlation was used to find the association between EDI and SRS and VCV stimuli scores. The results revealed that there was no statistically significant correlation between EDI and SRS ($p > 0.05$) or EDI and VCV stimuli scores ($p > 0.05$) at all the aided conditions and presentation levels.

Effect of type of stimulus on EDI

The EDI obtained for sentences and with VCV syllables were compared using a Wilcoxon signed rank test. The results (Table 4) revealed that the scores obtained for sentences were significantly greater than the EDI obtained for VCV syllables in all the conditions. This result indicates that, irrespective of the aided condition, changes in the temporal

envelope of a longer duration stimulus yielded a greater EDI compared to a shorter duration stimulus.

Effect of type of stimulus on SRS

The SRSs obtained for sentences and VCV stimuli were compared using a Wilcoxon signed rank test. The results, given in Table 5, revealed that the scores obtained for sentence recognition were significantly greater than the scores obtained for VCV stimuli, except for the unaided condition at 55 dB SPL. This result indicates that, irrespective of the aided condition, perception of a longer duration stimulus is better than with a shorter duration one.

Discussion

The results have shown that combined activation of the algorithms significantly improved SRSs compared to independent activation of the algorithms. The algorithms gave rise to mild temporal changes, and the conditions where all the algorithms and directionality were activated gave the highest EDI. These results are discussed in more detail below.

Effect of algorithms on EDI for sentences

At lower presentation levels, the EDI remained statistically unchanged irrespective of whether one or more of these algorithms was activated, even though activation of multiple algorithms resulted in slightly higher EDI (as seen in Figure 1). A similar finding has been reported in the literature [18]. The reason for this could be that, although the compression knee point was set at 50 dB SPL, valleys in the speech envelope are almost 12 dB lower than 55 dB SPL, a region where compression is not active. Hence, the net change in the EDI is smaller when all the algorithms are activated.

At 65 dB SPL, activation of both algorithms (as well as the Directionality algorithm alone) resulted in greater EDI compared to the DNR only condition and deactivation of both algorithms. Hence, it is clear that activation of both Directionality and DNR increased temporal envelope distortion in the incoming signal when combined with WDRC. This could be because of the way that noise reduction strategies work – they increase the temporal envelope modulation depth in the process of cutting down noise [16], which results in higher EDI. However, these findings go against the results of the study by Geetha and Manjula [19]. The difference might be due to a difference in stimulus presentation: in the current study, the sources of speech and noise were spatially separate (meaning that, as the literature demonstrates, directionality could work more effectively [28]), whereas in the earlier study the speech and noise emanated from the same source.

At 80 dB SPL, ‘DNR only’, ‘Both on’, and ‘Both off’ conditions resulted in greater EDI compared to the ‘Directionality only’ condition. At this higher presentation level, compression would have a greater effect on the temporal envelope of the signal [29]. Compression together with directionality will then result in a reduced EDI. This shows that activation of Directionality at higher levels reduces the effect of WDRC on EDI. The possibility of a difference in the long-term speech spectrum across these different conditions can be ruled out as the gain–frequency profile of the signal was

matched across different aided conditions. It is also clear that the EDI induced by DNR is greater than that from Directionality. Hence, it is clear that noise reduction strategies reduce the effect of WDRC on the temporal envelope. Among the noise reduction strategies, Directionality reduced the effect of WDRC more than DNR. A similar nullifying effect of WDRC by DNR and Directionality has been reported earlier [20,30]. Finally, linear conditions were not included, as the main aim here was to compare different algorithms including WDRC. The frequency shaping resulting from using a linear equation would be different from that resulting from a non-linear prescriptive equation, and this might result in a different EDI value than that obtained in the current study. This is one of the limitations of the current study and might be considered in future studies.

Effect of algorithms on SRS

At all presentation levels, simultaneous activation of all the algorithms resulted in better sentence (or equivalent) recognition scores compared to when either the ‘DNR only’ or ‘Directionality only’ algorithm was activated. The results at 65 and 80 dB SPL showed that the ‘DNR only’ condition resulted in better SRS scores compared to ‘Directionality only’ or when both algorithms were deactivated. In addition, activation of both algorithms resulted in significantly greater SRSs compared to ‘Directionality only’ and ‘Both off’ conditions. Hence, it is clear that, at mid and high presentation levels, activation of DNR helps in speech perception in the presence of noise. These results are not in agreement with the results of Nordrum et al. [18] who did not find any improvement with the DNR algorithm in their group data. However, analysis of individual data in their study did reveal that 50% of individuals showed greater speech perception scores when the DNR was activated along with directionality. The reason for this was attributed to the masking effect of directional microphones, which resulted in only a small improvement in speech perception, whereas the addition of DNR reduces this masking effect, so that there is a significant improvement in speech perception scores when both algorithms are activated.

Correlation between SRS and EDI for sentences

There was no significant correlation between SRS and EDI in any of the conditions. Earlier studies linking EDI to speech intelligibility focused mostly on the effects of WDRC in isolation and recognition of VCs, VCVs, and sentences in quiet (across different compression ratios and compression time-constants (e.g. [13,14,31,32])). Generally, it has been shown that as the EDI increases, the recognition rate decreases, so that there is a strong negative correlation [13,31]. Vinodhini [20] studied the combined effect of WDRC, DNR, and directionality and reported a moderate-level positive correlation between EDI and speech perception, which was restricted to the non-linear condition at 55 dB SPL only. The reason for no significant correlation between EDI and SRS could be because of the type of stimuli used and the age of the participants. The sentences have high contextual cues, even in the presence of noise, which might have resulted in higher SRSs even though the EDI was higher (indicating greater temporal distortion). The effect of temporal distortion tends to be higher for older listeners, as listening to sentences in noise

requires higher listening effort and cognitive resources [33]. Adult listeners were the participants in the current study and hence, although the temporal distortions were higher when all the algorithms were activated, higher SRSs were obtained. In addition, directionality improved the signal-to-noise ratio significantly, so that the negative effects of temporal distortion measured by EDI are offset by reductions in the noise level.

Although the EDI effectively quantifies temporal distortion, it might not be a good predictor of sentence recognition. Other factors, some of which were discussed above, tend to outweigh distortion effects. This statement can be supported by the results of earlier studies [13,14,31–33].

Effect of algorithms on EDI for VCV stimuli

As seen with EDI for sentences, EDI for VCV stimuli also exhibited differences at each presentation level. At a low presentation level, there was no significant change in EDI across different conditions, so the reason put forward for no significant change in EDI for sentences could again hold good here. At 65 dB SPL, DNR resulted in higher EDI. That is, DNR alone resulted in increased distortion of the temporal envelope, and the addition of directionality led to a decrease in changes to the temporal envelope. This might be viewed as the ability of directionality algorithms to compensate for changes created by DNR.

Effect of algorithms on VCV perception

At low and mid presentation levels, activation of both directionality and DNR resulted in better VCV recognition than when both were deactivated. ‘DNR only’ also resulted in equivalent VCV recognition scores when compared to the ‘Both on’ condition. At 80 dB SPL, activation of directionality resulted in the best VCV perception, followed by addition of DNR to Directionality.

Nordrum et al. [18] reported similar results on sentence perception whereby the addition of DNR resulted in a decline in SRS. This could have been due to filtering of information by DNR in channels where noise is present at high levels, resulting in poorer recognition of VCV stimuli.

Another reason could relate to the difference in how the WDRC algorithm functions. At low presentation levels, there will be amplification of the signal, and at high presentation levels there would be compression of the signal. In general, vowels have more energy and consonants less. Hence, Dillon [2] and Souza et al. [4] have reported that the action of WDRC varies within VCV stimuli. Individuals also rely on a level difference between the consonant and vowel to identify a few consonants [22,34], especially those affecting the place of articulation [35–37]. Hence, this could have been a reason for variations in the results.

Correlation of VCV perception and EDI

The scores obtained for VCV perception were correlated with EDI obtained for VCV stimuli. There was no significant correlation observed across any of the conditions, similar to the situation observed for sentences. It

has been reported that, in early studies using VCV stimuli, EDI has a good correlation with speech perception [13,14]. Although the type of stimuli used was the same, the variations in the results between the current study and the earlier studies could be due to the type of algorithms activated. In the current study, WDRC was kept constant, and the effects of noise reduction algorithms alone were studied. However, in the earlier studies, the algorithm was restricted to WDRC alone.

In earlier studies, outputs of the hearing aid have been recorded using a 2 cc coupler [19], ear simulator [21], and real ear measurement [20]. The use of real ear measurement was considered a good strategy and was argued to give more realistic outcomes. In the current study, the programmed hearing aid of each individual was fitted on the KEMAR, and the output of the hearing aid was recorded. Hence, the acoustics of the ear canal was included, although this remained constant across the recordings. Hence, this could have been an additional factor that contributed to variations in the results.

Therefore, it is clear that even when the speech segment is audible and other cues like spectral cues and amplitude modulation are available, speech perception might remain unaltered even with variations in the temporal envelope of the signal [39].

Effect of type of stimulus on SRS

The perceptual scores obtained for sentences were significantly greater than the scores obtained for VCV stimuli at all the tested conditions (except at 55 dB SPL for the unaided condition). This could be because the noise reduction algorithm tries to segregate the speech segment from the background noise, and in this attempt some essential parts of speech segments might be filtered out [5,40]. Hence, more parts of speech would have been altered in a longer stimulus than in a shorter one. As the duration of the stimuli increases, the syntactic and semantic cues also increase. This, in turn, facilitates the perception of speech at a complex level. The internal redundancy of the stimuli could be one of the reasons for better perception of sentences. The segmentals and the supra-segmentals could have acted as another cue for better perception of sentences.

Effect of type of stimulus on EDI

The EDI values obtained for sentences were significantly greater than those obtained for VCV stimuli at most of the aided conditions and presentation levels. The reason could be inherent fluctuations within the speech stimuli. As the duration of stimuli increases, the variations in terms of the temporal envelope also increase. This could have led to an increased EDI for sentences. Since the WDRC algorithm amplifies low-level sounds and compresses high-level ones, there is more fluctuation in gain with sentences than with VCV stimuli [2]. Hence, the EDI for sentences is higher than with VCV stimuli.

Conclusions

The results of the present study show that combined activation of the directionality and noise reduction algorithms,

along with WDRC, significantly improves SRS compared to independent activation of the algorithms – this was true at all presentation levels in our individuals with hearing impairment. Although the temporal changes induced by these algorithms are greater than when they are deactivated, the effect of other factors such as type of stimulus used, age, and compression settings tend to outweigh the effects of temporal distortion. The presentation level of speech and the length of the stimuli also affect the working of these algorithms. The above results are restricted to the age group tested in the current study; similar studies are needed in the future for different age groups and types of hearing loss.

References

1. Souza P. Speech perception and hearing aids. In: Popelka GR, Moore BCJ, Fay RR, Popper AN, Hearing Aids. Switzerland: Springer International, 2016; p. 151–80.
2. Dillon H. Binaural and bilateral considerations in hearing aid fitting. In: Hearing Aids. New York, NY: Thieme, 2001; p. 370–403.
3. Arpita V, Manjula P. Effects of compression release time in hearing aid on acoustic and behavioral measures of speech. Articles based on dissertations done at AIISH, 2012; 10: 20–31.
4. Souza PE, Hoover H, Gallun F. Application of envelope difference index to spectrally sparse speech. *J Speech Lang Hear Res*, 2012; 50: 824–37.
5. Levitt H. Noise reduction in hearing aids: a review. *J Rehabil Res Dev*, 2001; 38: 7–19.
6. Levitt H, Bakke M, Kates J. Signal processing for hearing impairment. *Scand Audiol Suppl*, 1993; 38: 7–19.
7. Boymans M, Dreschler WA. Field trials using a digital hearing aid with active noise reduction and dual-microphone directionality. *Audiol*, 2000; 39: 260–8.
8. Alcántara JI, Moore BCJ, Kuhnelt V, Launer S. Evaluation of noise reduction system in a commercial digital hearing aid. *Inter J Audiol*, 2003; 42: 34–42.
9. Kuk F, Korhonen P, Slugocki C. Preserving the temporal envelope in hearing aid processed sounds. *Hear Review*, 2018; 25(10): 40–4.
10. Tachibana RO, Sasaki Y, Riquimaroux H. Relative contributions of spectral and temporal resolutions to the perception of syllables, words, and sentences in noise-vocoded speech. *Acoust Sci Technol*, 2013; 34(4): 263–70.
11. Li X, Ning Z, Brashears R, Rife K. Relative contributions of spectral and temporal cues for speech recognition in patients with sensorineural hearing loss. *J Otol*, 2008; 3(2): 84–91.
12. Fortune TW, Woodruff BD, Preves DA. A new technique for quantifying temporal envelope contrasts. *Ear Hear*, 1994; 15: 93–9.
13. Jenstad LM, Souza PE. Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. *J Speech Lang Hear Res*, 2005; 48: 651–67.
14. Jenstad LM, Souza PE. Temporal envelope changes of compression and speech rate: combined effects on recognition for older adults. *J Speech Lang Hear Res*, 2007; 50: 1123–38.
15. Ricketts TA, Hornsby BWY. Sound quality measures for speech in noise through a commercial hearing aid implementing “Digital Noise Reduction”. *J Am Acad Audiol*, 2005; 16: 270–7.
16. Chung K, Zeng FG, Waltzman S. Utilizing hearing aid directional microphones and noise reduction algorithms to improve speech understanding and listening preferences of cochlear implant users. *Int Congress Series*, 2004; 1273: 89–92.
17. Walden BE, Surr RK, Cord MT, Edwards B, Olson L. Comparison of benefits provided by different hearing aid technologies. *J Am Acad Audiol*, 2000; 11(10): 540–60.
18. Nordrum S, Erler S, Garstecki D, Dhar S. Comparison of performance on the hearing in noise test using directional microphones and digital noise reduction algorithms. *Am J Audiol*, 2006; 15: 81–91.
19. Geetha C, Manjula P. Effect of compression, digital noise reduction and directionality on envelope difference index, log-likelihood ratio and perceived quality. *Audiol Res*, 2014; 4(1): 46–51.
20. Vinodhini P. Relationship Between Envelope Difference Index (EDI) and Sentence Recognition and Speech Quality in Individuals with Hearing Impairment. Unpublished Masters dissertation, University of Mysore, Mysore, 2015.
21. Walaszek J. Effect of compression in hearing aids on the envelope of the speech signal: signal based measures of the side-effects of the compression and their relation to speech intelligibility. Unpublished Masters thesis, Technical University of Denmark, Lyngby, 2008.
22. Kennedy E, Levitt H, Neuman AC, Weiss M. Consonant–vowel intensity ratios for maximizing consonant recognition by hearing-impaired listeners. *J Acoust Soc Am*, 1998; 103: 1098–114.
23. Geetha C, Kumar KSS, Manjula P, Pavan M. Development and standardization of sentence identification test in Kannada language. *J Hear Sci*, 2014; 4(1): 18–26.
24. Ramakrishna BS, Nair KK, Chiplunkar VN, Atal BS, Ramachandran V, Subramanian R. Some aspects of the relative efficiencies of Indian languages. Ranchi, India: Catholic Press, 1962.
25. Boothroyd A, Medwetsky L. Spectral distribution of /s/ and the frequency response of hearing aids. *Ear Hear*, 1992; 13: 150–7.
26. Winholtz WS, Titze IR. Conversion of a head-mounted microphone signal into calibrated SPL units. *J Voice*, 1997; 11: 417–21.
27. Korhonen P, Kuk F, Slugocki C. A method to evaluate the effect of signal processing on the temporal envelope of speech. *Hear Review*, 2019; 26(6): 10–18.
28. Ricketts TA. Impact of noise source configuration on directional hearing aid benefit and performance. *Ear Hear*, 2000; 21: 194–205.
29. Souza PE. Effect of compression on speech acoustics, intelligibility, and sound quality. *Trends Amplif*, 2002; 6: 131–65.
30. Aswathi S, Geetha C. Combined effect of compression and digital noise reduction algorithms on speech perception and speech quality. Articles based on dissertations done at AIISH, 2013; 10: 1–7.
31. Hoover EC, Souza PE, Gallun F. The consonant-weighted Envelope Difference Index (cEDI): a proposed technique for quantifying envelope distortion. *J Speech Lang Hear Res*, 2012; 55(6): 1802–6.

Acknowledgments

The authors extend their thanks to the Director, All India Institute of Speech and Hearing, Mysore, and all the participants. This data is part of a project funded by the AIISH Research Fund. We also thank Dr Arivudai Nambi, Associate Professor, KMC, Mangalore, for his help on the Matlab code for the calculation of EDI. This research was presented at the 9th KSB ISHACON 2016 at Palakkad, Kerala, on 23 October 2016.

Conflict of interest: None

32. Alexander JM, Masterson K. Effects of WDRC release time and number of channels on output SNR and speech recognition. *Ear Hear*, 2015; 36: e35–e49.
33. Gatehouse S, Naylor G, Elberling C. Linear and nonlinear hearing aid fittings. 2. Patterns of candidature. *Int J Audiol*, 2006; 45: 153–71.
34. Balakrishnan U, Freyman RL, Chiang YC, Nerbonne GP, Shea KJ. Consonant recognition for spectrally degraded speech as a function of consonant–vowel intensity ratio. *J Acoust Soc Am*, 1996; 99: 3758–69.
35. Ohde RN, Stevens KN. Effect of burst amplitude on the perception of stop consonant place of articulation. *J Acoust Soc Am*, 1984; 74: 706–14.
36. Hedrick MS, Rice T. Effect of a single-channel wide dynamic range compression circuit on perception of stop consonant place of articulation. *J Speech Lang Hear Res*, 2000; 43: 1174–84.
37. Hedrick MS, Younger MS. Perceptual weighting of relative amplitude and formant transition cues in aided CV syllables. *J Speech Lang Hear Res*, 2001; 44: 964–74.
38. Hedrick MS, Younger MS. Labeling of /s/ and /ʃ/ by listeners with normal and impaired hearing, revisited. *J Speech Lang Hear Res*, 2003; 46: 636–48.
39. Souza PE, Turner CW. Effect of single channel compression on temporal speech information. *J Speech Hear Res*, 1996; 39: 901–11.
40. Bentler R, Chiou L. Digital noise reduction: an overview. *Trends Amplif*, 2006; 10: 67–82.