A Study of Adaptive WDRC in Hearing Aids under Noisy Conditions

Ying-Hui Lai[1], Yu Tsao[1,*] and Fei Chen[2]

[1]*Research Center for Information Technology Innovation, Academia Sinica, Nankang, Taipei 11529, Taiwan*

[2]*Division of Speech and Hearing Sciences, The University of Hong Kong, Prince Philip Dental Hospital, 34 Hospital Road, Hong Kong*

Abstract: Background noise poses a great challenge to the speech recognition capability of hearing-impaired patients fitted with hearing aid (HA) devices. In an HA system, a speech enhancement unit is generally employed to enhance the signal-to-noise ratio (SNR) of noisy speech in order to yield better speech understanding for HA users in noisy conditions. However, previous studies reported that a subsequent static amplification scheme, such as wide-dynamic-range compression (WDRC), may deteriorate the enhanced speech and thus decrease the speech recognition capability. This work examines the performance of a recently proposed adaptive WDRC (AWDRC) amplification scheme when used in conjunction with a speech enhancement method in HA signal processing. Experimental results demonstrate that when integrated with the same speech enhancement method, AWDRC outperforms WDRC, in terms of long-term SNRs, at several typical hearing loss conditions. The results suggest that AWDRC can be a better choice than WDRC when combining with speech enhancement to improve speech recognition capabilities for HA users in noisy conditions.

Keywords: Adaptive WDRC, GMAPA, speech enhancement, hearing aids, NAL-NL1.

1. INTRODUCTION

Sensorineural hearing loss, which is the most common type of hearing impairment, introduces several deficits that need to be overcome in sound perception, such as decreased audibility, narrowed dynamic range, and degraded frequency and temporal resolutions [1]. Hearing aid (HA) devices are currently the most popular methods for improving the communication ability of hearing-impaired (HI) people. Since the last two decades, HAs have undergone a technological transformation in the sense of no longer processing sound in an analog way but digitally with an electronic processor. Digital signal processing has made possible many technological innovations for HAs, such as speech enhancement (or noise reduction) and amplification scheme, to enhance HA users' abilities to listen, communicate, and participate more fully in daily activities. These two algorithms were designed to reach different goals: speech enhancement is to improve output SNR in noisy conditions while amplification scheme is to keep the output sound in the audibility and comfort range. In digital HA systems, speech enhancement and amplification scheme are both integrated and used jointly.

Speech enhancement has been long utilized in the pre-processing stage in HAs to provide an improved

signal-to-noise ratio (SNR) [2] through HA signal processing. This is achieved by identifying segments where noise is particularly intense relative to speech, and applying less amplification to these segments than to other segments where the SNR level is large [3, 4]. On the other hand, amplification scheme is also an important signal processing unit in HAs, as it amplifies or limits the input sounds between the audible and comfortable listening levels (or hearing dynamic range) of HI patients. So far, the wide-dynamic-range compression (WDRC) scheme is the most common method used to amplify the sound in HAs [5]. The design of WDRC is similar to human outer hair cells in cochlea in that its effect is to non-linearly amplify quiet sounds more than those with large levels [6, 7]; hence, WDRC provides more amplification for the low-level sounds and less amplification for the high-level sounds. In addition, WDRC also provides different amounts of amplification for sounds in different frequency regions. While WDRC simultaneously fulfills the audibility and comfort requirements, it often reduces the performance of speech recognition in noise [8] and may worsen the output SNRs of the processed speech in HAs [9-12]. Chung showed that speech enhancement algorithms greatly enhanced the modulation depth of the noise-suppressed signals when an HA was set to linear amplification (i.e. the speech enhancement algorithm reduced the noise level and enhanced the modulation depth of the speech envelope) [4]. However, when the HA was set with a larger compression ratio (CR) (e.g. 3:1), the modulation depth of the speech envelope was greatly reduced compared to that processed by the

*Address correspondence to this author at the Research Center for Information Technology Innovation, Academia Sinica, No 128, Sec. 2, Academia Road, Nankang District, Taipei city 11529, Taiwan (R.O.C.); Tel: 886-2-2787-2390; Fax: 886-2-2787-2315; E-mail: yu.tsao@citi.sinica.edu.tw

linear setting. This indicated that the noise level was increased and the speech level was decreased by WDRC, and the benefit of speech enhancement could be affected by the compression inherent in WDRC [13].

Many approaches have been developed in order to improve the output long-term SNRs of the processed speech in HAs. Lai *et al.* recently proposed a novel adaptive amplification scheme (i.e. adaptive WDRC or AWDRC) by using the input short-term dynamic range to adjust the CR of WDRC to maximize the audibility and comfort of the processed speech, and to decrease the negative effect of large CR on the output long-term SNRs [12]. Figure **1** shows the block diagram of the AWDRC-based signal processing in one channel. The signal processing steps in AWDRC are similar to those used in WDRC, but with an extra feedback pathway to modify the WDRC parameters of low-level gain (LLG) and high-level gain (HLG) in order to decrease the compression as much as possible (i.e. close to linear compression or CR=1) in each channel. In the feedback pathway of AWDRC in Figure **1**, the first task is to perform a boundary-check calculation, whereby the output level of WDRC is estimated on a frame basis and stored in a first-in-first-out queue. When a new estimate is obtained, the maximum and minimum estimates in the queue are used to drive the operations according to the AWDRC rules by Δg to change the CR value of WDRC. Based on the results of boundary-check calculation, three AWDRC processing rules are applied independently in each frequency channel to control the CR value of WDRC to keep the output speech in a satisfactorily comfortable and audible level. The three rules are: (1) the decrease-CR rule, which keeps the WDRC processing as close to linear processing as possible; (2) the comfort rule, which ensures the output sound not to be amplified to an uncomfortable level; (3) the audibility rule, which ensures the output sound audible to HI individuals. More information on these three rules can be found in [12]. As a result, AWDRC does not use a static CR value as that used in WDRC for sound perception in complex listening environments; instead, its operation is based on the characteristics of the input sound signals to adaptively adjust the CR value by following the above-mentioned three rules. Previous studies indicated that AWDRC can provide better output long-term SNRs than the static WDRC method in listening conditions simulating those experienced by HI individuals fitted with HAs [12].

The main purpose of this study is to further examine the effects of AWDRC when it is used in conjunction with speech enhancement processing in HAs. More specifically, we intend to investigate whether the advantage of the speech enhancement processing could be preserved when integrated with a subsequent AWDRC, and how this advantage would be influenced by the factors regarding speech recognition (e.g. hearing loss audiogram, input level, and input SNR level) for HI patients fitted with HAs under noisy conditions.

2. METHODS

This section presents the speech enhancement and amplification scheme that are used in this study. In addition, the methods of measurement setup and procedures, and SNR evaluation are described.

2.1. Speech Enhancement

Generally speaking, speech enhancement algorithms can be divided into three categories, as: (1) filtering method, which designs a filter to attenuate
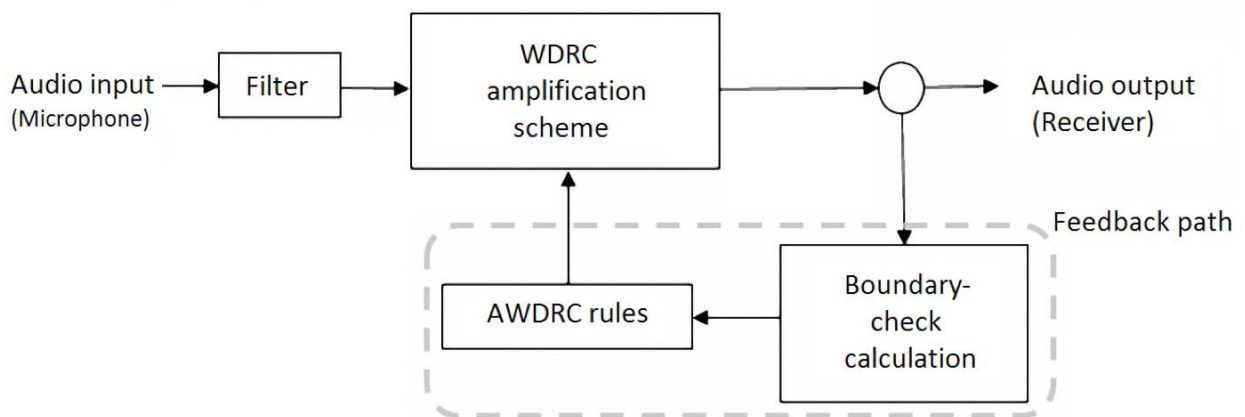
Figure 1: Block diagram of a hearing aid signal processing with adaptive dynamic compression (AWDRC) amplification in one channel.

noise components and generate a clean speech; (2) spectral restoration, which estimates a gain function to perform noise reduction in the frequency domain to obtain clean speech spectrum from the noisy speech spectrum; (3) speech model method, which combines human speech production models and speech reduction functions to perform speech enhancement [14, 15]. This study uses a recently developed generalized maximum *a posteriori* spectral amplitude (i.e. GMAPA) algorithm for speech enhancement [16]. The GMAPA algorithm dynamically adjusts the scale of *apriori* information to calculate the gain function for spectral restoration. More specifically, at conditions with high SNR levels, GMAPA adapts a small scale to prevent over-compensation that may result in speech distortions and decrease sound quality. On the other hand, at conditions with low SNR levels, GMAPA uses a large scale to enable the gain function to more effectively remove noise components from the noisy speech. The gain function (G_{GMAPA}) in the GMAPA algorithm is derived as:

$$G_{GMAPA} = \frac{\xi + \sqrt{\xi^2 + (2\alpha - 1)(\alpha + \xi)\xi / \gamma}}{2(\alpha + \xi)}, \qquad (1)$$

where ξ denotes *a priori* SNR, γ denotes *a posteriori* SNR, and coefficient α is dependent to the SNR level of testing conditions. The training data is used to find a mapping function that determines the correlation between the optimal α and SNR of the testing signals. When performing GMAPA, a sigmoid function is used to optimally determine the scale factor α for the gain function G_{GMAPA} in Eq. (1) for each utterance, as:

$$\alpha = \frac{\alpha_{max}}{1 + \exp[-b(\bar{\gamma} - c)]}, \qquad (2)$$

where α_{max} is the maximum value (upper bound) for α; b and c are coefficients of the sigmoid function;and $\bar{\gamma}$ is the mean of *a posteriori* SNR for a given utterance. Figure **2** shows the overall speech enhancement system, which can be decomposed into FFT, noise

tracking, gain estimation, and IFFT stages. FFT stands for fast Fourier transform, which converts the noisy speech signal, *y*, from time-domain to frequency-domain, *Y*. Next, the noise-tracking stage computes the noise power spectral density from noisy speech, *Y*, to obtain *a priori* SNR and *a posteriori* SNR statistics. Then gain estimation involves calculating a gain function, G_{GMAPA}, based on the computed *a priori* SNR and *a posteriori* SNR statistics, to obtain the enhanced speech components, $\hat{S}$, by filtering *Y* through G_{GMAPA}:

$$\hat{S} = G_{GMAPA} \cdot Y. \qquad (3)$$

Finally, inverse FFT (IFFT) is applied to convert $\hat{S}$ to obtain the time-domain signal, $\hat{s}$.

In this study, the parameters of GMAPA are set to $\alpha_{max} = 2$, $b=-1$, and $c=11$. More detailed descriptions on the GMAPA algorithm can be found in [16, 17]. Lai *et al.* recently used this GMAPA algorithm in conjunction with WDRC (denoted as G+WDRC hereafter) to evaluate the output long-term SNRs of the processed speech. They found that G+WDRC could provide better output long-term SNR performance than WDRC alone (i.e. without GMAPA pre-processing) for speech in noise [17].

2.2. Five Typical Audiograms of Sensorineural Hearing Loss

Figure **3** shows the five audiograms to be used in the present study to compare the efficacies of G+WDRC and GMAPA with AWDRC (denoted as G+AWDRC hereafter). These five audiograms were taken from a series of studies conducted by National Acoustic Laboratories (NAL) to compare the gain/frequency response of National Acoustic Laboratories – Nonlinear Fitting, version 1 (NAL-NL1) and other fitting strategies [18, 19]. As shown in Figure **3**, audiograms 1 to 5 represent a flat loss, a reverse sloping loss, a moderately sloping high-frequency loss, a steeply sloping high-frequency loss with a normal low frequency threshold, and a steeply sloping high-
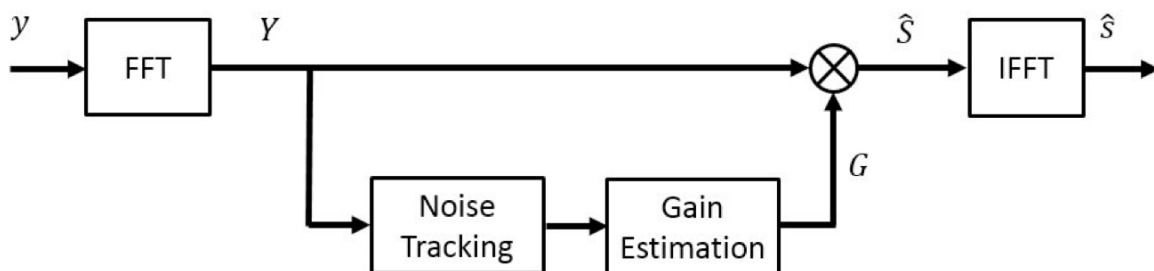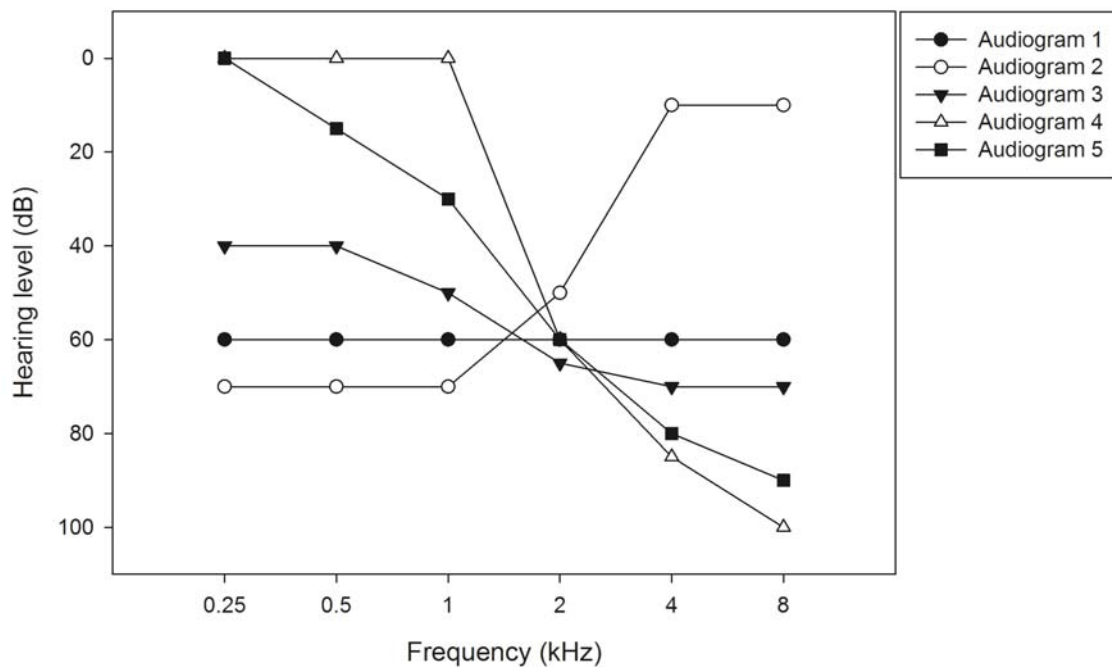

Figure 2: Block diagram of a speech enhancement system.

Figure 3: Five typical audiograms of sensorineural hearing loss used for comparing the output long-term SNRs between the G+WDRC and G+AWDRC processed speech. Audiograms 1 to 5 are a flat loss, a reverse sloping loss, a moderately sloping high-frequency loss, a steeply sloping high-frequency loss with a normal low frequency threshold, and a steeply sloping high-frequency loss with a mild low-frequency hearing loss, respectively.

frequency loss with a mild low-frequency hearing loss, respectively.

2.3. Measurement Setup

A platform was developed using LabVIEW and Matlab to simulate the effects of speech enhancement algorithm (i.e., GMAPA in this study) and amplification schemes (i.e. WDRC and AWDRC) to the output long-term SNRs of the processed speech. When implementing the amplification scheme, the crossover frequencies between the four channels were 0.56, 1.43, and 3.56 kHz. Based on individual HI audiogram (see Figure **3**), the fitting strategy of NAL-NL1 [18, 20] was used to calculate the parameters of WDRC and AWDRC. The attack and release time of WDRC were set as 5 and 26 msec, respectively, which were similar to those used by Naylor and Johannesson [10]. The boundary of discomfort level (DCL) in the AWDRC scheme was set according to the prediction from the fitting software, and Δg was set to 1dB SPL.

2.4. Methods of SNR Evaluation

The separation technique of the long-term SNR developed by Hagerman and Olofsson [21] was used to extract the speech and noise components from the WDRC- and AWDRC-amplified speech. Three types of speech-plus-noise files were created, as: (1) the

original speech (S) and the original noise (N); (2) the inverted original speech (–S) and the original noise (N); (3) the original speech (S) and the inverted original noise (–N). In each case the speech and noise were combined to produce the desired input SNR level by adjusting the sound level in each channel of a two-channel waveform file, with speech and noise in each channel, respectively. The speech-plus-noise files were processed by the WDRC and AWDRC amplifiers, after which the speech or noise files were ready to be extracted. Two files were used to obtain the waveform of the speech after processing: signal plus noise "S+N", and signal plus inverted noise "S–N". The noise was canceled by adding these files, leaving only the speech. This extracted speech-alone file included speech at double intensity, and hence the speech level was reduced by 6 dB SPL to obtain accurate speech levels. The waveform of the noise after processing was obtained similarly by using signal plus noise "S+N", and inverted signal plus noise "–S+N", with the level of the extracted noise also being reduced by 6 dB SPL to obtain accurate noise levels. The long-term SNR is defined by the difference between the levels of speech and noise in decibels.

2.5. Measurement Procedures

Figure **4** shows the block diagram of the experiment framework. A 15-sec Mandarin sentence spoken by a
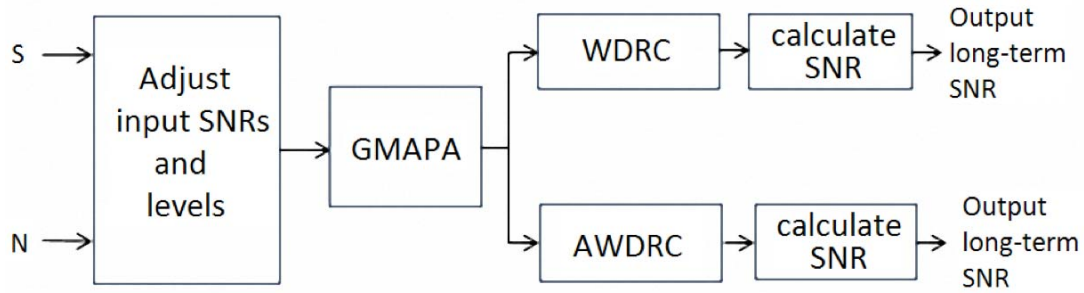
Figure 4: Block diagram of the overall processing to obtain the output long-term SNRs, where S and N denote the speech and noise signals, respectively.

female newscaster was used as "Signal (S)", and a 15-sec pink noise was used as "Noise (N)". The S and N were combined at four input SNR levels (i.e. –2, +2, +6, +10 dB), that were typical in everyday situations and produced a range of intelligibility scores with ceiling or floor effects [9]. Three input sound levels were prepared, i.e. loud (75 dB SPL), moderate (65 dB SPL) and soft (50 dB SPL). The speech and noise signals were adjusted simultaneously by an identical absolute amount to produce different input SNRs. For instance, when the SNR value was set to 6 dB SPL, the level of speech signal was increased by 3 dB SPL, and the level of noise was decreased by 3 dB SPL. Each input noisy signal was processed by speech enhancement (i.e. the GMAPA algorithm in this study), and followed by the amplification scheme of WDRC or AWDRC, as shown in Figure **4** (denoted as G+WDRC and G+AWDRC, respectively). Finally, the separation technique of the long-term SNR developed by Hagerman and Olofsson [21] was used to extract the speech and noise components from the output of
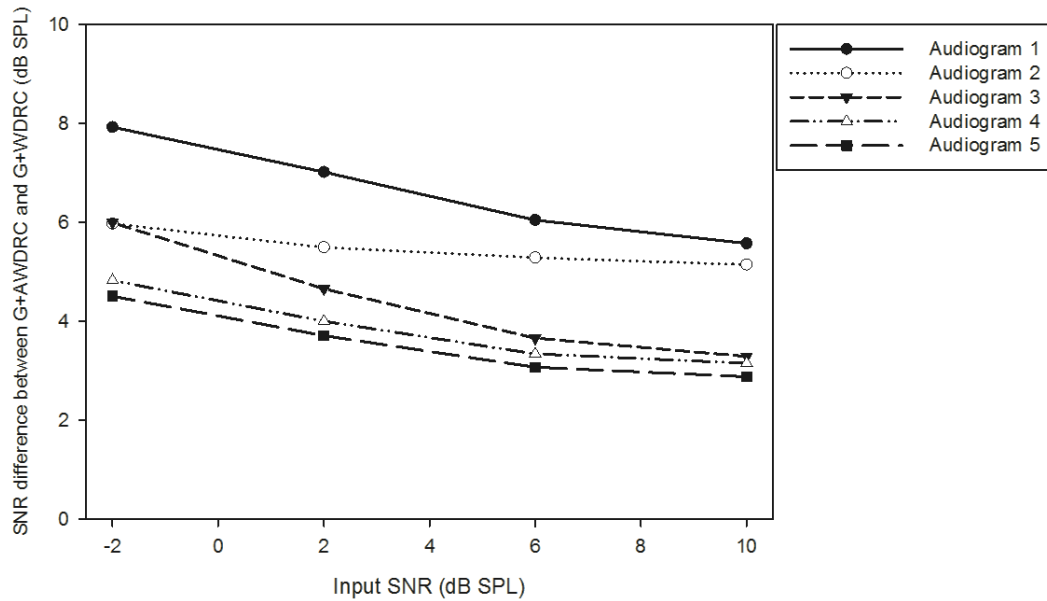
G+WDRC processed and G+AWDRC processed signals. In total, twelve (=4 input SNR levels × 3 input sound levels) types of test signals were used to compare the output long-term SNR performance of the G+WDRC and G+AWDRC processed speech.

3. RESULTS

Figures **5**, **6** and **7** show the output long-term SNR differences between the G+WDRC and G+AWDRC processed speech for the three input sound levels, i.e. loud, moderate and soft, respectively. In these figures, a positive difference means that the G+AWDRC method yields a better output long-term SNR performance than the G+WDRC method. For audiograms 1 and 2, the respective output long-term SNRs in the loud, moderate and soft input levels are 9.15 to 6.49 dB, 7.93 to 5.15 dB, and 2.87 to 1.22dB; meanwhile, the corresponding differences for audiograms 3, 4 and 5 are 6.29 to 0.72 dB, 6.00 to 2.88 dB, and 1.55 to 0.77 dB, both showing that the long-term SNRs of the G+AWDRC method are higher than
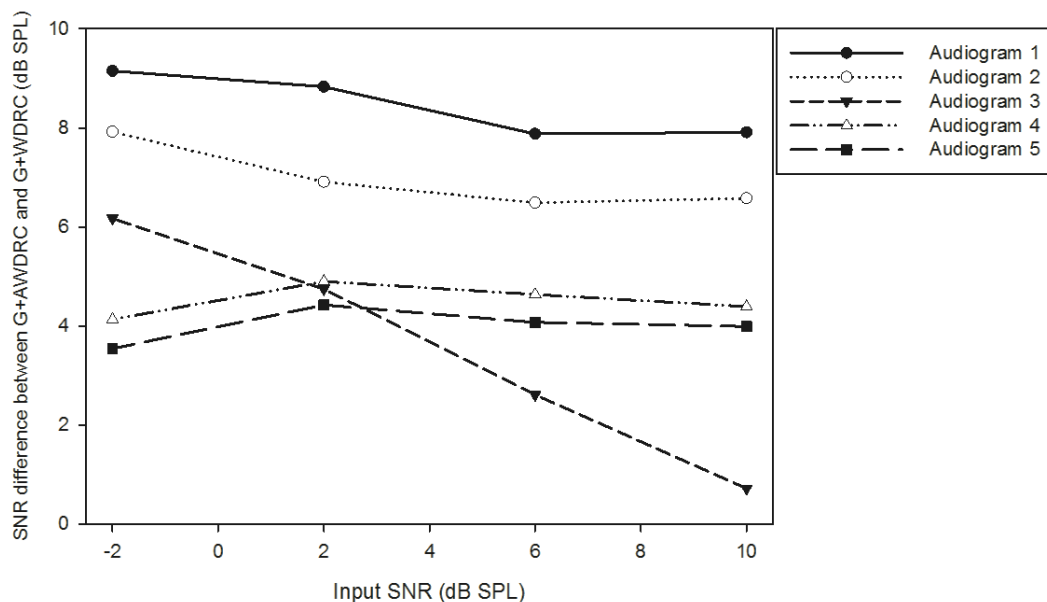
Figure 5: Long-term SNR difference between the G+WDRC and G+AWDRC processed speech under loud input sound level.

Figure 6: Long-term SNR difference between the G+WDRC and G+AWDRC processed speech under moderate input sound level.
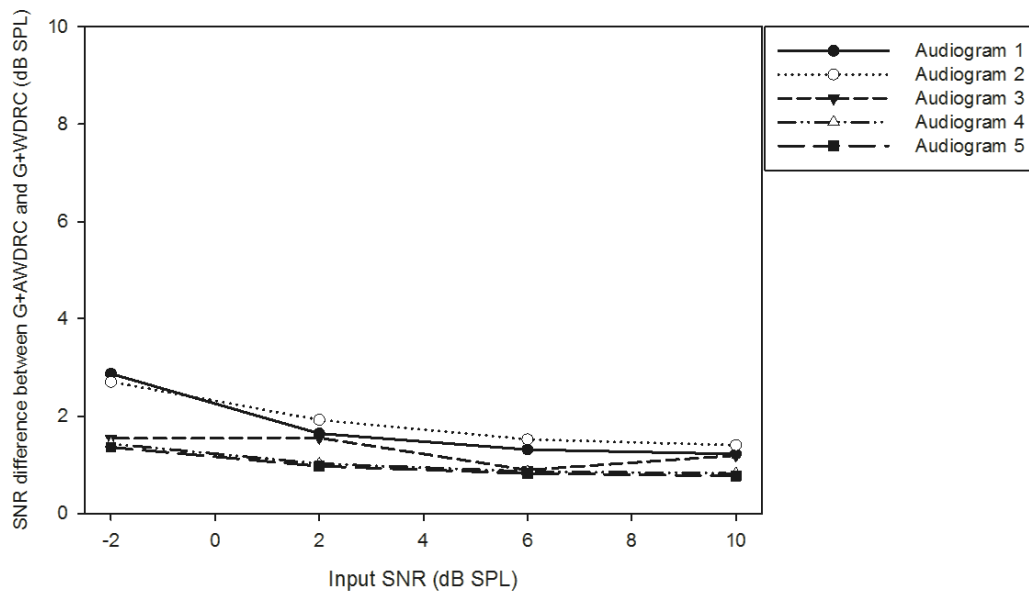
Figure 7: Long-term SNR difference between the G+WDRC and G+AWDRC processed speech under soft input sound level.

those of the G+WDRC method for the input SNR levels ranging from −2 to 10 dB. The overall mean ± standard deviation values are 10.36 ± 5.27 dB and 14.25 ± 4.65 dB for the G+WDRC and G+AWDRC methods, respectively. To confirm that the improvement of G+AWDRC over G+WDRC is significant, a paired-sample t-test was conducted. The t-test result indicated that the output long-term SNR difference between the G+WDRC and G+AWDRC methods was significant (t=12.8, $p<0.001$).

Table **1** lists the results of one-way analysis of variance (ANOVA) and Tukey post hoc comparison for the effects of three factors, i.e. hearing-loss audiogram, input sound level, and input SNR level. The results of ANOVA testing reveal that the preference differs significantly (F=3.946, p=0.007) across the five groups of hearing-loss audiogram, and the Tukey post hoc comparison of the five audiogram groups indicates significant difference for group pairs of (type 1 audiogram, type 3 audiogram), (1,4), (1,5), and (2,5). The preference also differs significantly (F=40.522, $p<0.001$) across the three groups of input sound level, and the Tukey post hoc comparison of the three groups indicates significant difference for group pairs of (loud,

Table 1: **Mean Differences of the Long-Term SNRs between the G+AWDRC and G+WDRC Processed Speech, whereas the Effects of Three Factors (i.e., Hearing-Loss Audiogram, Input Sound Level, and Input SNR Level) are Analyzed by One-Way ANOVA and Post-Hoc Tukey Test. Asterisk Indicates Statistically Significant (*p*<0.05) Difference between the Selected Pair of Groups**

Group notation	Variables	*n*	Mean difference[1]	SD	*F*	*p*	Post-hoc comparison* (group$_i$, group$_j$)
	Audiogram				3.946	0.007	
1	1	12	5.62	3.04			(1,3)
2	2	12	4.78	2.29			(1,4)
3	3	12	3.09	1.97			(1,5)
4	4	12	3.13	1.64			(2,5)
5	5	12	2.84	1.46			
	Input sound level				40.522	<0.001	
L	Loud	20	5.50	2.20			(L,S)
M	Moderate	20	4.78	1.41			(M,S)
S	Soft	20	1.39	0.57			
	Input SNR level				1.083	0.364	
10	10 dB	15	3.27	2.30			
6	6 dB	15	3.50	2.24			
2	2 dB	15	4.12	2.38			
−2	−2 dB	15	4.67	2.49			

[1]Mean difference = the mean output long-term SNR of the G+AWDRC processed speech minus the mean output long-term SNR of the G+WDRC processed speech (in dB).

soft) and (moderate, soft). Finally, the preference does not show significant difference across the four groups of input SNR level (*F*=1.083, *p*=0.364), but the mean differences among the four groups are at least 1.56 dB.

4. DISCUSSION AND CONCLUSION

This study compared the output long-term SNRs of the G+WDRC and G+AWDRC processed speech. For all of the testing conditions, G+AWDRC achieved notably higher output long-term SNRs than G+WDRC consistently under various input sound levels, input SNR levels, and five typical types of hearing-loss audiograms. The results also showed that the type of audiogram and input sound level significantly affected the output long-term SNR difference between G+AWDRC and G+WDRC; on the other hand, the input SNR level did not cause clear output long-term SNR difference. When the same GMAPA speech enhancement algorithm was used, AWDRC provided higher long-term SNR than WDRC with 6.49, 5.15 and 1.22 dB for type1 and 2 audiograms, 3.54, 2.88 and 0.77 dB for type 4 and 5 audiograms, and 0.72, 3.29 and 0.89 dB for type 3 audiogram, in loud, moderate and soft input levels, respectively.

Figure **8** shows the example waveforms of the original noisy speech (at −2 dB input SNR level and loud input level), GMAPA processed, G+WDRC processed and G+AWDRC processed speech. It can be seen from Figure **8b** that GMAPA effectively reduces the background noise and enhances speech segments, leading to a noise-suppressed speech with higher SNR level comparing to the original noisy speech at −2 dB SNR in Figure **8a**. However, when the noise-suppressed speech is further processed by WDRC, its output long-term SNR is considerably decreased (Figure **8c**). The decrease may be owing to a larger gain for the low-intensity sounds than for the high-intensity sounds, and the noise level is generally lower than the speech level. Hence, the noise components are increased more by WDRC (which uses a static CR value), yielding a decreased output SNR level. In contrast, the rules of AWDRC always try to decrease the CR value of WDRC in each channel, so that the low-level noise during the pause segments are only amplified with a small gain. This thereby improves the output long-term SNRs (Figure **8d**) relative to those from WDRC (Figure **8c**).

Many studies have suggested that the modulation depth of the speech envelope is important for speech
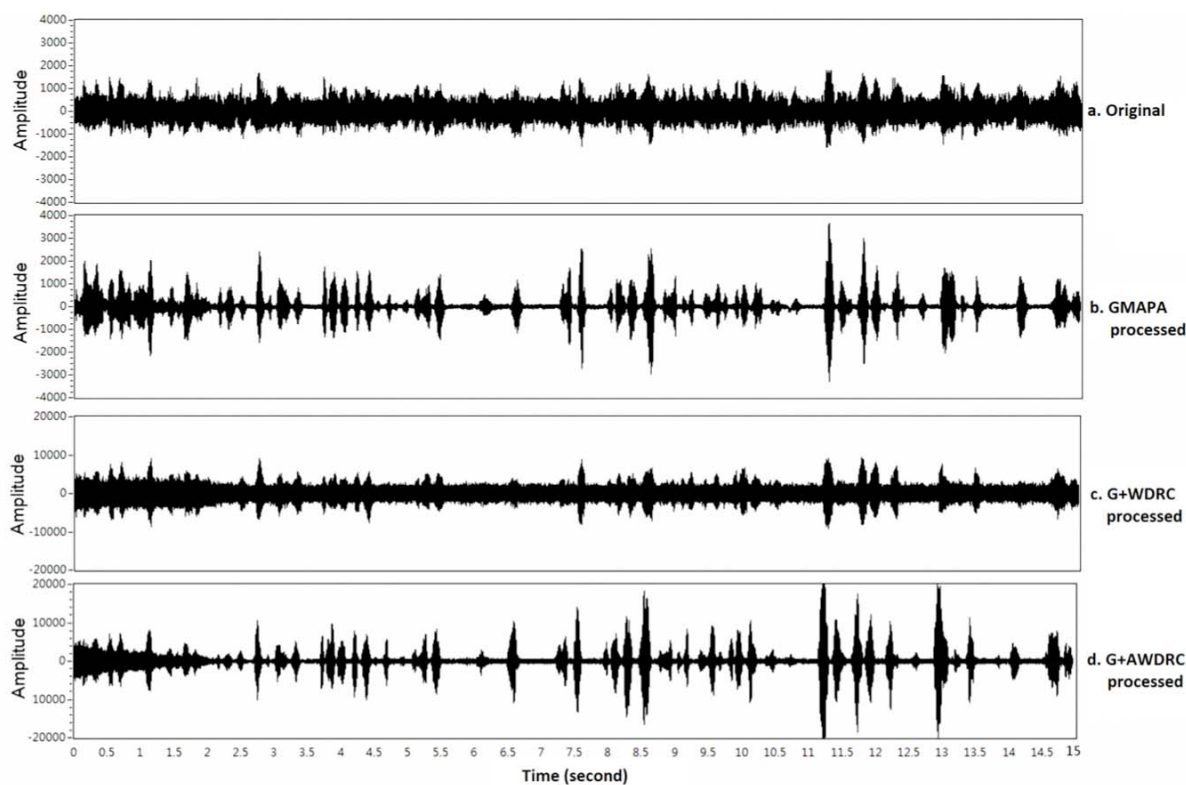
Figure 8: Example waveforms of (**a**) unprocessed noisy sentence, at −2 dB input SNR level and loud input level, (**b**) GMAPA processed sentence, (**c**) G+WDRC processed sentence, and (**d**) G+AWDRC processed sentence. In (**a**) to (**d**), a y-axis magnitude of 3270 corresponds to 94 dB SPL.

perception, especially at low SNR levels [22-27]. Chung *et al*. found that, in general, the higher the modulation depth in transmitted or processed signal, the higher the speech transmission index or the predicted speech intelligibility [13]. They also showed that speech enhancement algorithms greatly increased the speech intelligibility index, which implied that the modulation depth of the temporal envelope was greatly enhanced [13]. In addition, the study of Kates *et al*. indicates that when the processed envelope was more similar to clean speech, the coherence speech intelligibility index score will be higher, suggesting better speech intelligibility for hearing loss individuals [28]. From the example of Figure **8c** and **d**, we can see that G+AWDRC provides better modulation depth than G+WDRC, and the envelope of G+AWDRC processed was more similar to clean speech than that processed by G+WDRC. This implies that G+AWDRC could potentially provide better speech intelligibility than G+WDRC for HI individuals under noisy conditions.

The results in this study indicate that AWDRC can preserve more benefits of improved output long-term SNRs of the enhanced speech from the previous speech enhancement stage than WDRC does. This implies that AWDRC could potentially provide more

intelligibility benefits of speech recognition than WDRC when combining with speech enhancement algorithm in noise for HI individuals fitted with HAs. However, a limitation of this study is that the experiments involved software simulations to demonstrate the possible benefits of long-term SNR and speech intelligibility in noise. The characteristics of microphone, receiver, and recording space may also affect the final performance. Therefore, the overall effectiveness of GMAPA in adjunction with AWDRC in real HA systems needs to be further investigated. In addition, the objective and subjective benefits of G+AWDRC, such as speech intelligibility and sound quality, should be evaluated in clinical trials. These two parts will be conducted in our future study.

In conclusion, consistent with previous findings regarding the advantage of AWDRC to provide better long-term SNRs against WDRC, the present study further suggests that AWDRC could maintain the long-term SNR advantage when it is used in conjunction with a speech enhancement process. This advantage is achieved under different types of audiogram, input SNR levels, and input sound levels. In addition, G+AWDRC may preserve better modulation depth than G+WDRC, which implies that AWDRC could potentially maintain more SNR benefits from the enhanced

speech obtained from the previous stage for HA users in noisy listening conditions.

ACKNOWLEDGEMENT

REFERENCES

[1]　Dillon H. Hearing aids (2nd), Chapter 1: introductory concepts, Thieme 2012; pp. 1-19.

[2]　Chen F, Loizou PC. Impact of SNR and gain-function over- and under-estimation on speech intelligibility. Speech Commun 2012; 54: 272-81.
http://dx.doi.org/10.1016/j.specom.2011.09.002

[3]　Levitt H. Noise reduction in hearing aids: an overview. J Rehabil Res Dev 2001; 38: 111-21.

[4]　Chung K. Challenges and recent developments in hearing aids part I. Speech understanding in noise, microphone technologies and noise reduction algorithms. Trends Amplif 2004; 8: 83-124.
http://dx.doi.org/10.1177/108471380400800302

[5]　Schaub A. Digital hearing aids, Chapter 4: basic amplification schemes, Thieme 2008; pp. 39-50.

[6]　Hudspeth A. Making an effort to listen: mechanical amplification in the ear. Neuron 2008; 59: 530-45.
http://dx.doi.org/10.1016/j.neuron.2008.07.012

[7]　Venema T. Compression for clinicians, Chapter 4. Compression and the DSL and NAL-NL1 fitting methods. Thomson Delmar Learning 2006; pp. 64-92.

[8]　Souza PE. Effects of compression on speech acoustics, intelligibility, and sound quality. Trends Amplif 2002; 6: 131-65.
http://dx.doi.org/10.1177/108471380200600402

[9]　Souza PE, Jenstad LM, Boike KT. Measuring the acoustic effects of compression amplification on speech in noise. J Acoust Soc Am 2006; 119: 41-4.
http://dx.doi.org/10.1121/1.2108861

[10]　Naylor G, Johannesson RB. Long-term signal-to-noise ratio at the input and output of amplitude-compression systems. J Am Adad Audiol 2009; 20: 161-71.
http://dx.doi.org/10.3766/jaaa.20.3.2

[11]　Goldstein JL. Hearing aids based on models of cochlear compression using adaptive compression thresholds, Google Patents 2005.

[12]　Lai YH, Li PC, Tsai KS, *et al.* Measuring the long-term SNRs of static and adaptive compression amplification techniques for speech in noise. J Am Acad Audiol 2013; 24: 671-83.
http://dx.doi.org/10.3766/jaaa.24.8.4

[13]　Chung K. Effective compression and noise reduction configurations for hearing protectors. J Acoust Soc Am 2007; 121: 1090-1101.
http://dx.doi.org/10.1121/1.2409859

[14]　Haykin SS. Adaptive Filter Theory, 4th ed., Chapter 2: Pearson Education India 2005.

[15]　Chen J. Fundamentals of Noise Reduction in Spring Handbook of Speech Processing. Springer 2008.

[16]　Su YC, Tsao Y, Wu JE, Jean FR. Speech enhancement using generalized maximum a posteriori spectral amplitude estimator. In ICASSP, Vancouver, BC, Canada 2013; pp. 7467-71.

[17]　Lai YH, Su YC, Tsao Y, *et al.* Evaluation of generalized maximum a posteriori spectral amplitude (GMAPA) speech enhancement algorithm in hearing aids. In ISCE, Hsinchu, Taiwan 2013; pp. 245-46.

[18]　Byrne D, Dillon H, Ching T, *et al.* NAL-NL1 procedure for fitting nonlinear hearing aids: Characteristics and comparisons with other procedures. J Am Acad Audiol 2001; 12: 37-51.

[19]　Keidser G, Brew C, Peck A. How proprietary fitting algorithms compare to each other and to some generic algorithms. Hearing J 2003; 56: 28-38.
http://dx.doi.org/10.1097/01.HJ.0000293014.56004.ee

[20]　Dillon H. NAL-NL1: A new procedure for fitting non-linear hearing aids. Hearing J 1999; 52: 10-2.
http://dx.doi.org/10.1097/00025572-199904000-00002

[21]　Hagerman B, Olofsson A. A method to measure the effect of noise reduction algorithms using simultaneous speech and noise. Acta Acust United Ac 2004; 90: 356-61.

[22]　Drullman R. Temporal envelope and fine structure cues for speech intelligibility. J Acoust Soc Am 1995; 97: 585-92.
http://dx.doi.org/10.1121/1.413112

[23]　Hohmann V, Kollmeier B. The effect of multichannel dynamic compression on speech intelligibility. J Acoust Soc Am 1995; 97: 1191-5.
http://dx.doi.org/10.1121/1.413092

[24]　Shannon RV, Zeng F-G, Kamath V, *et al.* Speech recognition with primarily temporal cues. Science 1995; 270: 303-4.
http://dx.doi.org/10.1126/science.270.5234.303

[25]　Smith ZM, Delgutte B, Oxenham AJ. Chimaeric sounds reveal dichotomies in auditory perception. Nature 2002; 416: 87-90.
http://dx.doi.org/10.1038/416087a

[26]　Healy EW, Warren RM. The role of contrasting temporal amplitude patterns in the perception of speech. J Acoust Soc Am 2003; 113: 1676-88.
http://dx.doi.org/10.1121/1.1553464

[27]　Souza PE, Boike KT. Combining temporal-envelope cues across channels: Effects of age and hearing loss. J Speech Lang Hear Res 2006; 49: 138-49.
http://dx.doi.org/10.1044/1092-4388(2006/011)

[28]　Kates JM, Arehart KH. Coherence and the speech intelligibility index. J Acoust Soc Am 2005; 117: 2224-37.
http://dx.doi.org/10.1121/1.1862575