Evaluation of adaptive dynamic range optimization in adverse listening conditions for cochlear implants

Hussnain Ali and Oldooz Hazrati[a)]
Department of Electrical Engineering, The University of Texas at Dallas, Richardson, Texas 75080
hussnain.ali@utdallas.edu, hazrati@utdallas.edu

Emily A. Tobey
School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, Texas 75080
etobey@utdallas.edu

John H. L Hansen[b)]
Department of Electrical Engineering, The University of Texas at Dallas, Richardson, Texas 75080
john.hansen@utdallas.edu

Abstract: The aim of this study is to investigate the effect of Adaptive Dynamic Range Optimization (ADRO) on speech identification for cochlear implant (CI) users in adverse listening conditions. In this study, anechoic quiet, noisy, reverberant, noisy reverberant, and reverberant noisy conditions are evaluated. Two scenarios are considered when modeling the combined effects of reverberation and noise: (a) noise is added to the reverberant speech, and (b) noisy speech is reverberated. CI users were tested in different listening environments using IEEE sentences presented at 65 dB sound pressure level. No significant effect of ADRO processing on speech intelligibility was observed.

PACS numbers: 43.64.Me, 43.71.Ky [DO]
Date Received: June 4, 2014 **Date Accepted:** August 5, 2014

1. Introduction

Cochlear implants (CI) have enabled sound perception and speech identification in a vast majority of individuals with severe-to-profound hearing loss. However, electric hearing possess challenges in terms of mapping the input dynamic range of the acoustic signal (~90 dB) to the limited output electric dynamic range (the range between threshold levels and the maximum comfort levels which could be as low as 5 dB). This emphasizes the need to perform intelligent compression to optimally place the characteristic features of speech in the available limited output range for better intelligibility and quality of coded sounds. Commonly used CI coding strategies such as Continuous Interleaved Sampling (CIS) (Wilson *et al.*, 1991) and Advanced Combination Encoder (ACE) (Vandali *et al.*, 2000) use a global compression scheme (e.g., logarithmic compression) at the output level to compensate for the loudness growth. Adaptive Dynamic Range Optimization (ADRO), on the other hand, adaptively adjusts gains in each frequency band prior to the global compression to optimally utilize the limited output range based on the signal statistics.

ADRO is a multichannel signal equalization strategy to improve the audibility, comfort, and intelligibility of sounds for individuals who use CIs and/or hearing

---

[a)] Author to whom correspondence should be addressed.
[b)] Also at School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, Texas 75080.

aids (HA) (Blamey *et al.*, 1999; Blamey, 2005). The strategy uses statistical analysis to select the most information-rich section of the input dynamic range in multiple frequency channels, and adaptively adjusts the channel gains based on a set of fuzzy logic rules to optimally place the signal in the users' available hearing range. Thus, ADRO aims to make soft sounds more audible and loud sounds more comfortable, and is used in conjunction with sound processing in clinical HA and CI processors as a preprocessing strategy.[1] Clinical studies indicate preference for ADRO over alternative amplification strategies in quiet and various noisy conditions with HA and CI subjects (Martin *et al.*, 2001; James *et al.*, 2002; Dawson *et al.*, 2004; Iwaki *et al.*, 2008).

James *et al.* (2002) tested 9 adult cochlear implantees using ACE/SPEAK speech processing strategies (with and without ADRO) in quiet and in noise (multi-talker babble, SNR = 10 and 15 dB). Although significant speech perception improvement (16%) was observed using ADRO for low input level [50 dB sound pressure level (SPL)] in quiet, no significant improvement was seen in noise. Moreover, the environmental sound loudness tests indicated a 59% quality preference for the ADRO program in a majority of the conditions, where only 10% of the time the program without ADRO was preferred (31% of the time, sounds with and without ADRO programs were judged to have the same loudness level).

In a later study by Dawson *et al.* (2004), children with CIs (mean age: 10.6 yr) were tested with and without ADRO to establish if young implantees benefit from ADRO preprocessing in the same way as adults. A smaller mean group improvement was observed when testing children with CIs in quiet (50 dB SPL) compared to adults studied by James *et al.* (2002). They concluded that differences in microphone sensitivities for the two groups could be a contributing factor for this observed difference. Although speech perception scores for sentences in noise were not significantly different with and without ADRO for adults (James *et al.*, 2002), the speech perception scores of children improved significantly (single-digit percentage improvement) when using ADRO in noise (Dawson *et al.*, 2004). This may be due to the wider dynamic ranges and consequently steeper mapping functions seen in children as compared to adults (Hughes *et al.*, 2000). Children preferred sound coding with ADRO in 46% of the conditions, which is relatively smaller than the preference by adults (59% of conditions).

All studies conducted so far have evaluated ADRO in quiet and/or noisy (multi-talker babble) conditions (e.g., James *et al.*, 2002; Dawson *et al.*, 2004). However, these two conditions do not represent the naturalistic everyday situations where CI users are challenged to understand speech in the presence of reverberation and noise, individually and in combination. Speech perception scores of CI users drop substantially in reverberant environments when early and late reflections of the direct sound are added to speech, thereby blurring both temporal and spectral characteristics of speech (Hazrati, 2012). Unlike reverberation, noise is additive and affects speech in a different and complimentary fashion. Noise masks weak consonants to a greater degree than higher intensity vowels, but unlike reverberation this masking does not depend on the energy of the preceding segments (Nabelek *et al.*, 1989). Therefore, the combined effects of reverberation and noise affect speech intelligibility to a greater degree than either reverberation or noise alone (Hazrati and Loizou, 2012).

In the present study, we compare speech intelligibility scores obtained from ten adult CI users in quiet, noisy, reverberant, and noisy + reverberant (where noise and reverberation are simultaneously present) conditions. The main goal of the present study was to evaluate the effect of ADRO pre-processing on speech perception of CI users in adverse listening conditions in terms of intelligibility.

2. Methods

A. Subjects and material

Ten adult post-lingually deafened CI recipients participated in this study. All participants were native speakers of American English who received no benefit from hearing aids pre-operatively. All subjects were paid for their participation. CI users were fitted with the Nucleus 24 multichannel implant devices manufactured by Cochlear Corporation. All listeners used their devices routinely and had a minimum of 3 years experience with their devices. All participants were experienced users of ADRO as it was locked into their everyday MAPs. The detailed biographical data of the CI participants is presented in Table 1. All subjects had at least 20 active electrodes and a stimulation rate of 900 Hz per channel (except S5 and S6 with 1200 and 500 Hz stimulation rates, respectively).

IEEE sentences (IEEE, 1969) were used as the speech stimuli for testing. The root-mean-square (RMS) level of all sentences was equalized and presented at 65 dB SPL. The reverberant stimuli were generated by convolving the clean signals with measured room impulse responses (RIR) recorded in a 227.46 m^3 room (Hazrati and Loizou, 2012) with a reverberation time equal to 0.6 s, which is allowable in U.S. classrooms according to ANSI S12.60 (2002) standard. The direct-to-reverberant ratio (DRR) of the RIR was -1.8 dB. The distance between the single-source signal and the microphone was 5.5 m, which is beyond the critical distance ($\simeq 1$ m).

Speech-shaped noise (SSN) with the same long-term spectrum as the test sentences from the IEEE corpus was used as a continuous (steady-state) masker to generate the noisy signals at a 10 dB SNR level.

The noisy reverberant stimuli were generated using the following model [the masker was added to the reverberant stimuli at a 10 dB reverberant speech signal to noise ratio (RSNR)[2]]:

$$y(n) = \{x(n) * h(n)\} + m(n), \tag{1}$$

where $y(n)$, $x(n)$, $h(n)$, and $m(n)$ denote corrupted signal (by noise and reverberation), anechoic clean signal, RIR, and additive noise, respectively.

The reverberant noisy speech stimuli were generated using the following model (the noise-masked speech at 10 dB SNR was reverberated):

$$y(n) = \{x(n) + m(n)\} * h(n). \tag{2}$$

B. Signal processing

All CI participants used ACE speech coding strategy in their clinical processors (clinical processors were programmed with the users' clinical MAP and configured with and without ADRO for each listening condition). In the ACE coding strategy, the acoustic signal is split into 22 frequency bands by a combination of coefficients produced from an FFT analysis. ADRO dynamically applies channel gains to the output of the frequency bands every 2 ms. Next, "*n* maxima" (bands with highest energy, e.g., eight

Table 1. Demographic data of the CI participants.

Subjects	Gender	Age	Years implanted	CI processor	Etiology of hearing loss	Sensitivity level	Average dynamic range
S1	M	60	3	N5	Noise	9	38
S2	F	62	7	N5	Unknown	12	21
S3	F	54	4	N5	Unknown	12	48
S4	F	56	3	N5	Hereditary	12	39
S5	M	80	8	N5	Hereditary	12	30
S6	F	60	3	N5	Hereditary	10	10
S7	F	65	4	Freedom	Antibiotics	12	51
S8	M	61	3	N5	Meniere's Disease	12	45
S9	M	65	3	N5	Hereditary	12	52
S10	M	70	8	N5	Unknown	12	5

bands) are selected and compressed through a compression scheme (typically logarithmic compression) to generate current levels in the output dynamic range of the selected (active) electrodes.

ADRO uses four rules to continuously vary the input signal gain in each frequency band. The channel gain adjustments are conducted based on comfort, background noise, audibility, and maximum gain rules. The rules are applied based on the long-term calculated output levels (every 2 ms) using a percentile level estimator with a time constant of 20 dB/s. Three target levels (comfort, background, and audibility) define the dynamic range at each frequency band. The comfort rule reduces the gain if the 98th percentile of the long-term output level is greater than the target comfort level. The background noise rule decreases the gain if the 40th percentile of the long-term output level is greater than the background target level. If the 70th percentile of the long-term output is below the audibility target level, then the audibility rule increases the gain. Finally, the maximum gain rule limits the gain in order not to exceed a pre-determined maximum value [for more details on ADRO algorithm see James *et al.* (2002)].

C. Procedure

Subjects were tested using a clinical CI processor in a double-wall sound-proof booth (Acoustic Systems, Inc.). Recorded sentences were presented in free field at 65 dB SPL. CI listeners were tested unilaterally using the ear with the best performance. Bilateral/bimodal listeners were asked to remove the CI/hearing aid of the contralateral ear during test. The clinical processor was programmed with each individual subject's everyday clinical MAP (e.g., stimulation rate, microphone sensitivity, comfort, and threshold levels) using Custom Sound software developed by Cochlear Limited, and was configured with and without ADRO. All CI listeners used similar compression function with a base level of 4 and Q value of 20. Participants selected their sensitivity settings based on experience with their processors. Institutional review board (IRB) approval and informed consent were obtained from all participants prior to testing.

Subjects participated in a total of ten listening conditions: (1) Anechoic quiet ($T_{60} \approx 0.0$ s), (2) reverberant ($T_{60} = 0.6$ s), (3) noisy (SNR = 10 dB SSN), (4) noisy reverberant ($T_{60} = 0.6$ s, RSNR = 10 dB), and (5) reverberant noisy (SNR = 10 dB, $T_{60} = 0.6$ s) conditions (each with and without ADRO). Twenty IEEE sentences (two lists) were used per condition.[3] None of the lists used was repeated across conditions. The sequence of test conditions was randomized across subjects to minimize any order effects. To achieve a balance test order, half the CI users were tested with ADRO (ACE + ADRO) first, and the other half without ADRO (standard ACE). Evaluations were blind so subjects were not aware which was the ADRO condition. For each testing condition, 20 training sentences (not used in the test sessions) were played to the listener in order to familiarize them with the new condition. Participants were instructed to repeat as many words as they could identify. The responses of each individual were collected and scored off-line based on the number of words correctly identified. All words were scored. The percent correct scores for each condition were calculated by dividing the number of words correctly identified by the total number of words. To avoid listener fatigue, participants were given a 15 min break every 60 min during the test session. The entire test duration for each subject was approximately 4 h.

3. Results

Intelligibility listening tests were conducted in five different environments with and without ADRO. The individual as well as mean speech intelligibility scores for all five conditions are presented in Fig. 1. The intelligibility scores progressively declined with the level of difficulty in test condition, ranging from 96% for clean to 23% in reverberant noisy environment. The mean speech intelligibility difference between ADRO and non-ADRO test conditions varied from a minimum absolute value of 0.44% in clean (ADRO > non-ADRO) to a maximum absolute value of 4.76% in reverberant noisy condition (non-ADRO > ADRO). However, individual variations between ADRO and

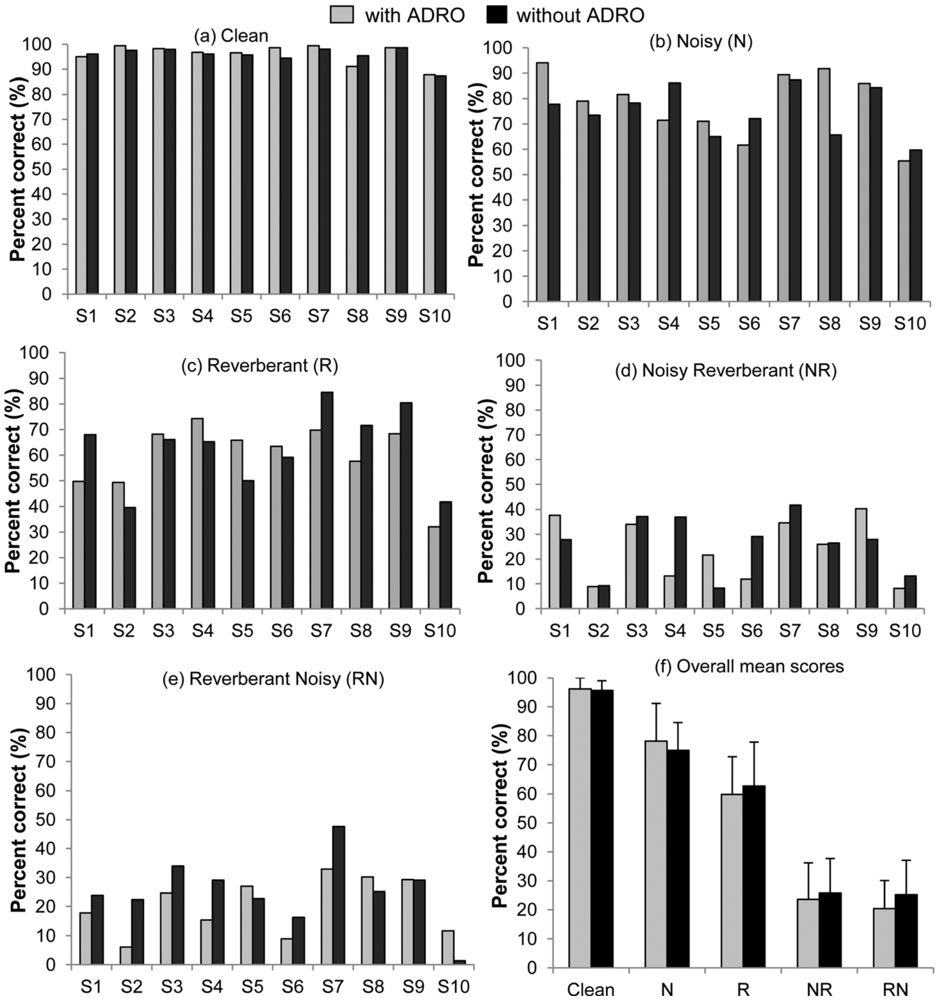

FIG. 1. Individual speech intelligibility scores of ten CI users in (a) anechoic quiet (clean), (b) noisy (N, SNR = 10 dB), (c) reverberant (R, $T_{60} = 0.6$ s), (d) noisy reverberant (NR, $T_{60} = 0.6$ s, RSNR = 10 dB), and (e) reverberant noisy (RN, SNR = 10 dB, $T_{60} = 0.6$ s) conditions. Panel (f) demonstrates average scores in all conditions. The error bars in panel (f) indicate standard deviations.

non-ADRO conditions ranged from −26% to +24%. On average, non-ADRO program performed slightly better (3.23%) than the ADRO program in the most challenging listening conditions (R, NR, and RN from Fig. 1).

Repeated-measures analysis of variance (ANOVA) was performed to assess the effect of environment type and program (ADRO/non-ADRO) on the intelligibility scores with an α factor set to 0.05. Subjects were considered a random (blocked) factor, while environment type and ADRO/non-ADRO conditions were used as the main analysis factors. No statistically significant difference in speech intelligibility was found between ADRO/non-ADRO conditions (F[1,9] = 0.656, $p = 0.439$). The interaction between the environment type and ADRO/non-ADRO conditions was not significant (F[4,36] = 0.900, $p = 0.474$). However, a significant main effect of environment type on speech intelligibility was observed (F[4,36] = 333.937, $p < 0.001$). The *post hoc* Bonferroni test for pairwise comparisons between the five environment types indicated significant differences between all, with the exception of reverberant-noisy and noisy-reverberant environments ($p = 1.000$).

In order to assess the effect of CI users' MAP parameters on speech intelligibility, correlations between the subjects' average electric dynamic range and speech intelligibility scores were computed for the five environment types for ADRO/non-ADRO programs. The results are presented in Table 2. Speech intelligibility was positively correlated with average electric dynamic range in all five conditions.

4. Summary and discussion

The main goal of this study was to assess the effect of ADRO pre-processing on speech intelligibility for CI users in various listening environments (anechoic quiet, noisy, reverberant, noisy reverberant, and reverberant noisy).

The ADRO strategy was initially developed for bimodal listening and has been previously validated for hearing aids and cochlear implants (Blamey, 2005). Studies by James *et al.* (2002) and Dawson *et al.* (2004) concluded that sound quality and speech perception performance were improved using ADRO as compared to fixed channel gains in both adults and children. The later study with children suggested ADRO to be locked into the processor for young children whose MAPs have been stabilized and may be left as an option for the older ones. In line with previous studies, Iwaki *et al.* (2008) reported significantly improved speech intelligibility with ADRO for six adult CI users in clean and noisy conditions using Japanese hearing in noise test (JHINT). However, all studies assessing the effect of ADRO pre-processing on speech intelligibility of CI users only considered anechoic quiet and noisy environments. The current study aimed to assess the potential ADRO benefit in everyday realistic environments where reverberation and/or noise can exist individually or in combination.

For all five environment types, our results indicate non-significant speech intelligibility benefit of ADRO over standard ACE program when speech material at 65 dB SPL were presented to CI users. Due to the subjective variability in scores, no clear trend in the pattern of results for either condition/program was found. On average, intelligibility scores for standard ACE program (non-ADRO) were only 1.23% higher than the ACE + ADRO program.

On average, the standard ACE program performed better than the ACE + ADRO program in R, NR, and RN conditions by 3.23%. Seven out of ten subjects had equal or better scores for the non-ADRO program in NR and RN conditions. One of the potential causes which could be attributed to this is that low energy late reflections of the reverberant sound may become amplified by the ADRO strategy as it tends to amplify low-intensity sounds. In such a scenario, ADRO programming may not be beneficial in reverberant environments. Further investigation into how late reflections of the sound are processed in ADRO is required to establish the exact explanation.

Eight out of ten subjects had similar sensitivity settings (12) in their processors. Because of the limited dataset, no relationship between subjects' intelligibility scores and sensitivity settings could be determined in the current study. Positive correlation between the subjects' electric dynamic range and intelligibility scores was observed in all tested conditions, indicating that subjects with a wider dynamic range could be expected to perform better in various listening conditions. This is in line with studies conducted by Loizou *et al.* (2000) and Fu and Shannon (2000).

The present study could not establish any significant benefit with ADRO pre-processing on speech intelligibility in the specific tested conditions. Due to the limited

Table 2. Correlation coefficients between speech intelligibility and electric dynamic range of CI users in different listening conditions. "ACE" and "ACE + ADRO" stand for standard ACE strategy without and with ADRO program, respectively. Significant correlation values are marked with "*."

Condition	Clean	N	R	NR	RN	Mean
ACE	0.76*	0.66*	0.83*	0.62*	0.85*	**0.74***
ACE + ADRO	0.36	0.82*	0.65*	0.83*	0.81*	**0.69***

number of participants, their highly variable performance, and similarity in their MAP parameters, no clear trend between the intelligibility scores and their processing parameters (such as stimulation rate and sensitivity level) could be determined. Given that a CI user may or may not benefit with ADRO in different listening environments, ADRO may be left as an optional setting which could be turned on or off according to personal preference of the implant user. Further research is warranted to investigate long-term benefits of ADRO in practical listening environments (reverberation + noise) as well as the effect of ADRO strategy on intelligibility of reverberant speech at both soft and loud presentation levels.

Acknowledgments

This work was supported by Grant R01 DC010494 awarded from NIH (PI: E.A.T.). The authors would like to thank the CI users for their time.

[1] The ADRO pre-processing strategy is available in CI devices manufactured by Cochlear Limited and many digital HAs (e.g., HAs manufactured by Interton, Siemens).

[2] For generating noisy reverberant stimuli, the reverberant signal served as the target signal in the SNR computation. Hence, we refer to the SNR values in this condition as reverberant signal to noise ratios (RSNR).

[3] The following website includes sample audio files for different conditions: www.utdallas.edu/hussnain.ali/AudioFiles/.

References and links

American National Standards Institute. (**2002**). ANSI S12.60. 2002. *Acoustical Performance Criteria*, Design Requirements and Guidelines for Schools (ANSI, New York).

Blamey, P. J. (**2005**). "Adaptive Dynamic Rango Optimization (ADRO): A digital amplification strategy for hearing aids and cochlear implants," Trends Amp. **9**(2), 77–98.

Blamey, P. J., James, C. J., Martin, L., McDermott, H. J., and Wildi, K. (**1999**). "Adaptive dynamic range optimization sound processor," International Patent Application PCT/ AU99/00076, U.S. Patent Application 09/478,022.

Dawson, P. W., Decker, J. A., and Psarros, C. E. (**2004**). "Optimizing dynamic range in children using the Nucleus cochlear implants," Ear Hear. **25**(3), 230–241.

Fu, Q. J., and Shannon, R. V. (**2000**). "Effects of dynamic range and amplitude mapping on phoneme recognition in nucleus-22 cochlear implant users," Ear Hear. **21**(3), 227–235.

Hazrati, O. (**2012**). "Development of dereverberation algorithms for improved speech intelligibility by cochlear implant users," Ph.D. dissertation, University of Texas, Dallas, TX.

Hazrati, O., and Loizou, P. C. (**2012**). "The combined effect of reverberation and noise on speech intelligibility by cochlear implant listeners," Int. J. Audiol. **51**(6), 437–443.

Hughes, M. L., Brown, C. J., Abbas, P. J., Wolaver, A. A., and Gervais, J. P. (**2000**). "Comparison of EAP thresholds with MAP levels in the Nucleus 24 cochlear implant: Data from children," Ear Hear. **21**(2), 164–174.

IEEE (**1969**). "IEEE recommended practice for speech quality measurements," IEEE Trans. Audio Electroacoust. **AU-17**, 225–246.

Iwaki, T., Blamey, P., and Kubo, T. (**2008**). "Bimodal studies using adaptive dynamic range optimization (ADRO) technology," Int. J. Audiol. **47**(6), 311–318.

James, C. J., Blamey, P. J., Martin, L., Swanson, B., Just, Y., and Macfarlane, D. (**2002**). "Adaptive Dynamic Range Optimization for cochlear implants: A preliminary study," Ear Hear. **23**(1S), 49S–58S.

Loizou, P. C., Dorman, M., and Fitzke, J. (**2000**). "The effect of reduced dynamic range on speech understanding: Implications for patients with cochlear implants," Ear Hear. **21**(1), 25–31.

Martin, L. F. A., Blamey, P. J., James, C. J., Galvin, K. L., and Macfarlane, D. (**2001**). "Adaptive dynamic range optimization for hearing aids," Acoust. Australia **29**, 21–24.

Nabelek, A. K., Letowski, T. R., and Tucker, F. M. (**1989**). "Reverberant overlap- and self-masking in consonant identification," J. Acoust. Soc. Am. **86**(4), 1259–1265.

Vandali, A. E., Whitford, L. A., Plant, K. L., and Clark, G. M. (**2000**). "Speech perception as a function of electrical stimulation rate using the Nucleus 24 cochlear implant system," Ear Hear. **21**(6), 608–624.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (**1991**). "Better speech recognition with cochlear implants," Nature **352**, 236–238.