Improvement of cochlear implant performance: changes in dynamic range

Ahmed Khater[a], Amira El Shennaway[b], Ahmed Anany[a]

[a]Audiology Unit, Departments of ENT, Faculty of Medicine, Zagazig University, Zagazig, [b]Cairo University, Cairo, Egypt

Correspondence to Ahmed Khater, MD, Department of ENT, Faculty of Medicine, Zagazig University, Zagazig, Egypt
Tel: 00966541498894; fax: 00966114803635; e-mail: khateraudio@gmail.com

Received 30 July 2014
Accepted 25 August 2014

The Egyptian Journal of Otolaryngology
2015, 31:36–41

Context
Theoretically, a wide input dynamic range (IDR) will capture more of the incoming acoustic signal than a narrow IDR, allowing the cochlear implant (CI) user to hear soft, medium, and loud sound. A narrow IDR may restrict the CI user's ability to hear soft speech and sound because less of the incoming acoustic signal is being mapped into the CI user's electrical dynamic range.

Aim
The overall goal of the study is to provide guidelines for audiologists to efficiently and effectively optimize performance of CI recipients for two difficult listening situations: understanding soft speech and speech in noise.

Settings and design
Two variables were studied; the independent variables were IDR and the electric dynamic range of the channels. The dependent variables were six Ling sounds, monosyllabic word test, and speech in noise test.

Materials and methods
Fourteen patients participated in the study. For each patient, seven programs were created. In each program, dependent variables were assessed in different independent ones.

Results
A restricted IDR resulted in poor speech recognition compared with the relatively wide IDR. Subjectively determined T level and most comfortable level (MCL) at the most, not the maximum, comfortable level appears to have a positive effect on both soft sound recognition and speech discrimination.

Conclusion
Dynamic range is an important factor -among others- to improve the ability of CI users to understand soft speech as well as speech in noise.

Keywords:
cochlear implant, input dynamic range, speech performance

Egypt J Otolaryngol 31:36–41

1012-5574

Introduction

Cochlear implant (CI) patients who perform well on word and sentence tests presented in quiet at a comfortable listening level often report considerable difficulty understanding in most noisy environments encountered in daily life [1]. Moreover, they report difficulty understanding soft speech spoken by children and individuals speaking from a distance. If optimizing patient performance in daily life is the goal, then it is essential that clinical fitting address the ability of CI users to understand soft speech as well as speech in noise [2].

The acoustic information carried by speech is quite complex and has many dynamic variations. Sounds are-by their nature-dynamic, changing over time in terms of level and spectral content [3]. It has been shown by formant analysis that the dynamic spectral variation in vowels provides reliable acoustic cues in fluent speech that contribute toward both consonant and vowel identification [4]. As the speech signal is highly variable in terms of its intensity, the relationship in consonant and vowel amplitude ratios play an important role in speech intelligibility.

A characteristic finding in individuals with sensorineural hearing loss, in addition to an increase in hearing threshold, is essentially a reduction in dynamic range. This reduction in dynamic range has its drawbacks in speech intelligibility. Dynamic range includes both input dynamic range (IDR) and electric dynamic range (C and T). The IDR is the range of the incoming acoustic signal that is mapped into the CI user's electrical dynamic range (range between minimum stimulation levels (T levels) and maximum stimulation levels (C levels) [5]. Theoretically, a wide IDR will capture more of the incoming acoustic signal than a narrow IDR, allowing the CI user to hear soft, medium, and loud sound. A narrow IDR may restrict the CI user's ability to hear soft speech and sound because less of the incoming acoustic signal is being mapped into the CI user's electrical dynamic range [6]. Jacquelyn *et al.* [7] reported that speech is not all the same loudness; consequently, lowering IDR too much

 DOI: 10.4103/1012-5574.152706

can decrease speech comprehension, even without any background noise.

The objectives of the present study were two-fold. The first was to compare two ways for the minimum threshold level (T level), a fixed value and a subjectively determined one, enabling audibility of soft speech cues. The second objective was to evaluate the effect of different IDRs on CI patient performance. The overall goal of the study is to provide guidelines for audiologists to efficiently and effectively optimize performance of CI recipients for two difficult listening situations: understanding soft speech and speech in noise.

Materials and methods

The study was approved by the ethical committee of Zagazig University, Faculty of Medicine, Zagazig, Egypt. Fourteen patients participated in the study. All patients were implanted with Advanced Bionics (AB) 90K CI devices in the ENT Medical Center, Kingdom of Saudi Arabia (Valencia, CA 91355, USA). Patients ranged in age from 14 to 27 years at the time of the study. The duration of hearing loss ranged from 2 to 9 years. Onset of hearing loss was perilingual in six patients and postlingual in eight patients. This category of patients was chosen to yield reliable results during the study. The length of implant use ranged from 2 to 4 years. Patients' data are presented in Table 1. Neural response image (NRI) confirmed well-functioning electrodes at all channels together with a good response for all patients in the study. Intraoperative plain radiography confirmed the correct position of the electrode array into the cochlea. The participants used HiRes 120 speech processing strategies and all had open-set speech recognition.

Table 1 Patients' data

Patient number	Age (years)	Duration of hearing loss (years)	Onset of hearing loss	Length of implant use (years)
P1	18	4	Postlingual	3
P2	14	8	Perilingual	3
P3	23	9	Postlingual	4
P4	19	2	Postlingual	2
P5	15	8	Perilingual	3
P6	26	5	Postlingual	2
P7	14	7	Perilingual	4
P8	16	9	Perilingual	4
P9	25	3	Postlingual	3
P10	27	2	Postlingual	2
P11	17	9	Perilingual	4
P12	21	5	Postlingual	3
P13	15	8	Perilingual	3
P14	22	4	Postlingual	2

Electric dynamic range (difference between T and C levels)

After a sufficient healing period, initial programming and activation of the sound processor was performed through Soundwave fitting software, the HiRes clinical fitting tool, version 2.2 for Advanced Bionics device (AB).

C level

C levels reflect the amount of electrical current needed for each electrode to elicit a comfortable loudness percept. The maximum (C) stimulation level for each electrode was programmed using ascending loudness judgments. The patient reports the loudness of the sound on a five-point scale (barely audible, soft, most comfortable, maximum comfortable, and uncomfortable). Each participant's preferred program with C level at the maximum comfortable level had been used for at least 1 month before being clinically evaluated for the current study. Another trial period was allowed with the C levels at all electrodes set at the most comfortable level. It was 10 current unit (CU) below the maximum comfortable one.

T level

T level represents the minimum amount of electrical current needed for each electrode to elicit a low-level or a soft percept for the recipient. When using the manufacturer's software (SoundWave) to set T levels for an Advanced Bionics speech processor, a 'default' setting can be selected. By doing so, T levels are calculated automatically as a level that represents 10% of the recipient's C levels. Alternately, T levels can be set manually on the basis of the patient's perception of minimally audible sounds. Two programs were created. In one program, T level, it was 10% of the C level. Another program involved behavioral assessment of the T level as the patient reports the loudness of the sound on a scale to be just below soft. It is more than what was calculated as 10% of the C level.

Input dynamic range

IDR is the range of the softest to loudest sounds that are detected by a sound processor. The wider the range, the more sounds the patient hears. The IDR, of a CI sound processor is the ratio between the loudest and the softest sounds that it will present at any given time.

Plan of the study

The independent variables tested in this study were input acoustic dynamic range (IDR) and the electric dynamic range of the channels (the range between T and C levels). Seven programs were created. In the first trial, two programs were created; by fixing C level to the maximum comfortable level, IDR at 60 dB (default), and sensitivity at zero, T level was tested with the six Ling sounds threshold comparing between two programs: program with T level at 10% of the C level and program with the T level just at barely audible sound to determine the program with the best detection threshold for the six Ling sounds. In the second trial, with the T level that yielded the best detection threshold for the six Ling sounds and all other parameters fixed, two more programs were created: one program with C level at the maximum comfortable level and another program with C level 10 cu below the maximum comfortable level (at the most comfortable level), and then testing the best C level by testing speech discrimination by a monosyllabic word list and establishing the program with the C level that yielded the best score. In the third trial, using the T level that yielded the best detection threshold for the six Ling sounds and using the C level that yielded the best score for monosyllabic words and all other parameters were fixed, three programs with three different IDRs (50, 60, and 80 dB) were created to compare the results of speech in noise test (SPIN) between the three IDRs.

The dependent variables tested in this experiment were six Ling sounds, monosyllabic word test, and SPIN. All speech tests were performed in a soundproof booth through a loudspeaker placed at the ear level at 0° azimuth and 1 m from the center of the participants' heads. The test materials were presented through an IBM compatible, Pentium II computer that controlled a mixing and attenuation network to present stimuli through a power amplifier and loudspeaker.

Speech tests

Six Ling sounds detection threshold levels

The Ling six sounds [8,9] represent different speech sounds from low to high pitch. It was developed as a quick and easy test that professionals can use to check the hearing of the patient. The test checks that the patient can hear (detection) and in time recognize each sound (identification) across the different speech frequencies. The Ling 6 sounds test uses isolated phonemes consisting of three vowels [(ah), (oo), (ee)] and three consonants [(m), (s), (sh)] that span the speech frequency range of 250–8000 Hz. They are uttered as follows: ah (as in father), oo (as in moon), ee (as in key), sh (as in shoe), s (as in sock), and m (as in mommy). These phonemes were recorded by a female speaker, were 800 ms in duration, and had a root mean square level within 1 dB of each other. Detection thresholds for recorded Ling sounds were obtained.

Ling sound frequency [10]

M	is a low-frequency sound
O	has low-frequency information
E	has some low-frequency and some high-frequency information
A	is at the center of the speech range
Sh	is in the moderately high-frequency speech range
S	is in the very high-frequency speech range

Speech discrimination

Speech discrimination was carried out using the Arabic monosyllabic word list according to Soliman [11]. It was presented at 65 dB sound pressure level (SPL). The patient's response was in the form of repetition of the word heard.

Speech in noise test

Arabic SPIN was used according to Tawfik *et al.* [12]. It is an open-set test that includes 25 items. The speech material was delivered to the patients through a front loudspeaker at zero azimuths while the background noise (multitalker babble) was delivered from a back speaker. The intensity of the signal was set at 65 dB SPL with 0 dB S/N ratios. The participant was instructed to ignore the noise and to repeat the speech signals.

Results

The results are shown in Tables 2–4 and reflecting that subjectively determined T values that were at levels slightly higher than the manufacturer's recommended setting of Ts (10% of Cs) resulted in statistically significant decrease in sound field threshold levels for the Ling six sounds. Monosyllabic word discrimination was significantly better with the most comfortable level

Table 2 Comparison of six Ling sounds detection threshold in all tested patients using two levels of T value: one program with T level 10% of C level and another program with a subjectively determined T level (with all other parameters fixed)

Items	Mean ± SD			Paired *t*-test	
	T 10%	T (subjective)	Difference	*t*	*P*-value
Ah	24.929 ± 2.814	14.714 ± 2.555	10.214 ± 3.262	11.715	0.000
M	29.286 ± 2.998	20.000 ± 3.138	9.286 ± 3.474	10.002	0.000
Oo	27.286 ± 2.555	19.857 ± 2.656	7.429 ± 2.533	10.972	0.000
Sh	28.286 ± 2.463	20.714 ± 2.431	7.571 ± 2.377	11.920	0.000
S	31.143 ± 2.179	23.857 ± 2.413	7.286 ± 2.164	12.599	0.000
Ee	32.286 ± 2.054	24.571 ± 2.138	7.714 ± 1.729	16.695	0.000

All comparisons showed highly statistically significant results.

Table 3 Comparison of monosyllabic word test scores in all tested patients using two levels of C value: one program with C level at the maximum comfortable level and another program with the most comfortable level (with all other parameters fixed)

Speech discrimination test	Mean ± SD			Paired *t*-test	
	C (maximum comfortable level)	C (most comfortable level)	Difference	*t*	*P*-value
Monosyllabic word test	69.714 ± 8.695	72.571 ± 6.630	−2.857 ± 3.302	−3.238	0.006

Comparison between C maximum and C at most comfortable levels was statistically significant.

Table 4 Comparison of speech in noise test scores in all tested patients using three different input dynamic ranges (50, 60, and 80) (with all other parameters fixed)

Input Dynamic range	SPIN test		ANOVA		Tukey's test	
	Range	Mean ± SD	*F*	*P*-value		*P*-value
IDR (50)	40.00–64.00	51.14 ± 7.22			50 and 60	<0.001*
IDR (60)	52.00–76.00	66.00 ± 7.32	16.54	<0.001*	50 and 80	<0.001*
IDR (80)	48.00–72.00	63.14 ± 7.22			60 and 80	0.56

Comparison between IDR 50 and 60 is highly statistically significant; IDR between 50 and 80 is highly statistically significant; comparison between 60 and 80 was nonsignificant; ANOVA, analysis of variance; IDR, input dynamic range; SPIN, speech in noise test.

than the maximum comfortable level. Comparison of SPIN test results between IDR at 50 and 60 dB was highly statistically significant, and at the same time SPIN with IDR between 50 and 80 dB was also highly statistically significant; however, comparison between SPIN at 60 and 80 dB was nonsignificant.

Discussion

The present study addresses the effects of IDR, T, and MCL levels on CI patient's performance. Within the AB SoundWave clinical software, a range of IDR settings from 20 to 80 dB is available with 60 dB as the default. The T level is determined as a fixed 10% of MCL according to the default value. A general trend noted in the present study was that a restricted IDR would produce poor speech recognition. Subjectively determined T level (slightly higher than the default fixed level, 10% of MCL) and MCL at the most, not the maximum, comfortable level appears to have a positive effect on both soft sound recognition and speech discrimination. The results of the present study are in agreement with that reported in the literature. Skinner *et al.* [13] found decreased sound-field thresholds and improved the perception of soft speech by increasing T levels so that low-level sounds were mapped to higher levels within Nucleus 22 CI users' electrical dynamic range. Zeng *et al.* [14] found that an IDR between 50 and 60 dB provided the best vowel and consonant recognition for 10 Clarion CI users. Spahr *et al.* [15] recommended an IDR of 60 dB for use with the AB CII BTE speech processor.

Maximum versus most comfortable level

The results of the present study indicated that the performance with the most, not maximum, comfortable level was statistically significant. Although higher stimulus levels exert positive effects that include better encoding of speech signals because of increased discharge rates and increased numbers of fibers carrying the signal [16], negative effects could include rate saturation and increased channel overlap. The loudness of a pulse train presented on a single electrode of a CI grows monotonically with the stimulus amplitude [17,18]. When multiple electrodes provide interleaved stimulation as is typical of most modern CI processing strategies, the loudness of the interleaved stimulus is greater than the loudness provided by the stimulation from any one of the individual electrodes presented in isolation. This phenomenon is known as loudness summation [19]. The mechanism by which electrical stimulation level affects perception is related to the spatial extent of neural excitation. With a higher level of stimulation, the degree of current spread is increased. Increasing the level of electrical stimulation causes larger activating-potential fields and thus leads to an increase in the number of the stimulated neural population [20]. With more neurons contributing toward the representation of temporal cues at higher electrical stimulus levels, increased psychophysical and speech perception performance are the outcome. However, with more increments of electrical stimulus level, degradation in the specificity of tonotopic stimulation is expected secondary to greater overlap of adjacent populations of stimulated neurons [16]. Site stimulation shifts as a function of stimulus level are assumed to be another explanation for the effects of higher stimulus level on speech performance. Such shifts in the site of stimulation can affect speech perception [21].

Input dynamic range

Good recognition of vowels and consonants — the constituent units of words — is dependent on the maintenance of spectral contrast. One set of the acoustic cues that specify manner and voicing are found in the gross shape of the amplitude envelope [14]. The present study showed that patients' performance with a small dynamic range was worse than performance with a large dynamic range on speech recognition evaluation.

Phoneme spectra are characterized by peaks and valleys with the vowel spectra are typically characterized by high-amplitude peaks and relatively low-amplitude valleys [22]. Although the frequencies of the spectral peaks are considered to be the primary cues to phoneme identity, the spectral contrast, that is, the difference between the spectral peak and the spectral valley, needs to be maintained to some extent for accurate phoneme identification as mentioned by Loizou and Poroy [22]. Normal-hearing listeners required a 1–2 dB peak-to-valley difference to identify four vowel-like harmonic complexes with relatively high 75% correct. Listeners with a flat, moderate hearing loss required a 6–7 dB peak-to-valley difference for vowel identification [14]. This was attributed to the lack of compression and the abnormally broad auditory filters associated with hearing loss. Spectral contrast is reduced when phonemes are processed through broad filters because of the shallow filter roll-off. As a result, the internal phoneme representation is 'blurred', leading to poorer identification [23].

Spectral contrast is reduced in CI listeners, not because of the abnormally broad auditory filters — which are bypassed with electrical stimulation — but primarily because of the reduced dynamic range and amplitude compression [16]. The large acoustic dynamic range is typically compressed in implant speech processors using a logarithmic function to a small electrical dynamic range, 5–15 dB [24]. Another factor that could potentially reduce spectral contrast is the steepness of the compression function used for mapping acoustic amplitudes to electric amplitudes [22]. A highly compressive mapping function would yield a small spectral contrast even if the dynamic range were large. A third factor of the effect of background noise could also reduce spectral contrast probably to a larger degree in CI listeners compared with normal-hearing listeners because of the limited electrical dynamic range [25].

Large IDR maintains a sufficient spectral contrast enough to make the peak in the channel amplitude spectrum more distinct and perceptually more salient, leading to a significant improvement in identification. A wider IDR may present a more complete picture of the sound environment. Narrower IDR may improve speech comprehension if there is trouble with background noise as it reduces unnecessary noise and limits the sound range to that of the normal variations in speech; however, this occurs at the expense of a feeling of isolation as patients are not hearing as many sounds around them.

Higher T level

As the consonant envelope distribution was about 20 dB lower than the vowel envelope distribution, the consonants are likely to be mapped into a less-than-optimal electric range [23]. First, some low envelope levels may be mapped into electric levels below threshold; second, some of the upper portion of the electric dynamic range may not be utilized because few amplitude envelope levels are present. Third, most envelope levels are likely mapped into the lower portion of the electric dynamic range, where both intensity discrimination and modulation detection are poor. Higher T level will raise previously inaudible low envelope levels above the threshold, reduce the unused portion of the electric dynamic range, and map more of the envelope into the upper electric dynamic range, where intensity discrimination and modulation are optimal [26]. One negative trade-off for the higher T level and the more compressive mapping is the possibility that low-level noise may become audible [27]. Another negative trade-off is the slightly distorted envelope level distribution; however, Kewley and Burkle [26] found in their study that this distortion should produce little, if any, decrease in consonant recognition.

Acknowledgements

Conflicts of interest

There are no conflicts of interest.

References

1 Donaldson GS, Chisolm TH, Blasco GP, Shinnick LJ, Ketter KJ, Krause JC. BKB-SIN and ANL predict perceived communication ability in cochlear implant users. Ear Hear 2009; 30:401–410.

2 Holden L, Reeder R, Firszt J, Finley Ch. Optimizing the perception of soft speech and speech in noise with the advanced bionics cochlear implant system. Int J Audiol 2011; 50:255–269.

3 Oxenham A, Bacon S. In: Bacon SP, Fay RR, Popper AN, editors. Psychophysical manifestations of compression: normal-hearing listeners. *Compression from cochlea to cochlear implants*. New York: Springer-Verlag; 2004. 62–106.

4 Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M. Perception of speech code. Psychol Rev 1967; 74:431–461.

5 Holden LK, Skinner MW, Fourakis MS, Holden TA. Effect of increased IIDR in the nucleus freedom cochlear implant system. J Am Acad Audiol 2007; 18:778–791.

6 Laura K, Holden A, Ruth M, Reeder A, Jill B, Firszt A, *et al.* Optimizing the perception of soft speech and speech in noise with the advanced bionics cochlear implant system. Int J Audiol 2011; 50:255–269.

7 Jacquelyn B, Jamie C, Jill B, Ruth M, Jerrica L. Optimization of programming parameters in children with the advanced bionics cochlear implant. J Am Acad Audiol 2012; 23:302–312.

8 Ling D. *Speech and the hearing-impaired child: theory and practice*. Washington, DC: Alexander Graham Bell Association for the Deaf; 1967.

9 Ling D. *Foundations of spoken language for the hearing-impaired child*. Washington, DC: Alexander Graham Bell Association for the Deaf; 1989.

10 Ling D. *Speech and the hearing impaired child*. 2nd ed. Washington, DC: Alexander Graham Bell Association for the Deaf and Hard of Hearing; 2002.

11 Soliman S. Speech discrimination audiometry using Arabic balanced words. Ain Shams Med J 1976; 27:27–30.

12 Tawfik S, Shehata W, Shalabi A. Development of Arabic speech intelligibility in noise (SPIN) test. Ain Shams Med J 1992; 3:677–682.

13 Skinner MW, Holden LK, Holden TA, Demorest ME. Comparison of two methods for selecting minimum stimulation levels used in programming the Nucleus 22 cochlear implant. J Speech Lang Hear Res 1999; 42:814–828.

14 Zeng F, Grant G, Niparko J, Galvin J, Shannon R, Jane Opie J, Phil Segel Ph. Speech dynamic range and its effect on cochlear implant performance. J Acoust Soc Am 2002; 111:377–386.

15 Spahr AJ, Dorman MF, Loiselle LH. Performance of patients using zdifferent cochlear implant systems: effects of input dynamic range. Ear Hear 2007; 28:260–275.

16 Franck K, Xu L, Pfingst B. Effects of stimulus level on speech perception with cochlear prostheses. J Assoc Res Otolaryngol 2002; 04:49–59.

17 Tong YC, Blamey PJ, Dowell RC, Clark GM. Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant. J Acoust Soc Am 1983; 74:73–80.

18 Shannon RV. Threshold and loudness functions for pulsatile stimulation of cochlear implants. Hear Res 1985; 18:135–143.

19 Padilla M, Landsberger D. Loudness summation using focused and unfocused electrical stimulation. J Acoust Soc Am 2014; 135:EL102.

20 Bierer JA, Middlebrooks JC. Auditory cortical images of cochlear implant stimuli: dependence on electrode configuration. J Neurophysiol 2002; 87:478–492.

21 Pfingst BE, Franck KH, Xu L, Bauer EM, Zwolan TA. Effects of electrode configuration and place of stimulation on speech perception with cochlear prostheses. J Assoc Res Otolaryngol 2001; 2:87–103.

22 Loizou Ph, Poroy O. Minimum spectral contrast needed for vowel identification by normal hearing and cochlear implant listeners. J Acoust Soc Am 2001; 110:1619–1627.

23 Fu Q, Shannon R. Effect of acoustic dynamic range on phoneme recognition in quiet and noise by cochlear implant users. J Acoust Soc Am 1999; 106:65–70.

24 Loizou P, Dorman M, Fitzke J. The effect of reduced dynamic range on speech understanding: implications for patients with cochlear implants. Ear Hear 2000; 21:25–31.

25 Kam AC, Yee IH, Cheng MM, Wong TK, Tong MC. Evaluation of the clear voice strategy in adults using HiResolution Fidelity 120 sound processing. Clin Exp Otorhinolaryngol 2012; 5:S89–S92.

26 Kewley-PD, Burkle TZ, Lee JH. Hyperlink "http://www.ncbi.nlm.nih.gov/pubmed/17902871" Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. J Acoust Soc Am 2007;122:2365-75.

27 Baudhuin J, Cadieux J, Firszt J, Reeder R, Maxson J. Optimization of programming parameters in children with the advanced bionics cochlear implant. J Am Acad Audiol 2012; 23:302–312.