

## 디지털 보청기 사용자를 위한 압신 알고리즘의 성능 연구

황운수<sup>1</sup> · 한종희<sup>1</sup> · 지윤상<sup>1</sup> · 홍성화<sup>2</sup> · 이상민<sup>3</sup> · 김동욱<sup>4</sup> · 김인영<sup>1</sup> · 김선일<sup>1</sup>

<sup>1</sup>한양대학교 의용생체공학과, <sup>2</sup>성균관대학교 의과대학 이비인후과학교실,  
<sup>3</sup>인하대학교 전자공학과, <sup>4</sup>삼성종합기술원 바이오헬스연구실

### A Study on the Performance of Companding Algorithms for Digital Hearing Aid Users

Y.S. Hwang<sup>1</sup>, J.H. Han<sup>1</sup>, Y.S. Ji<sup>1</sup>, S.H. Hong<sup>2</sup>, S.M. Lee<sup>3</sup>,  
D.W. Kim<sup>4</sup>, In Young Kim<sup>1</sup> and Sun I. Kim<sup>1</sup>

<sup>1</sup>Department of Biomedical Engineering, Hanyang University

<sup>2</sup>Department of ORL-HNS, Sungkyunkwan University

<sup>3</sup>Department of Electronic Engineering, Inha University

<sup>4</sup>Bio & Health Lab, Samsung Advance Institute of Technology

(Received February 16, 2011. Accepted June 22, 2011)

218

**Abstract:** Companding algorithms have been used to enhance speech recognition in noise for cochlea implant users. The efficiency of using companding for digital hearing aid users is not yet validated. The purpose of this study is to evaluate the performance of the companding for digital hearing aid users in the various hearing loss cases. Using HeLPS, a hearing loss simulator, two different sensorineural hearing loss conditions were simulated; mild gently sloping hearing loss(HL1) and moderate to steeply sloping hearing loss(HL2). In addition, a non-linear compression was simulated to compensate for hearing loss using national acoustic laboratories-non-linear version 1(NAL-NL1) in HeLPS. In companding, the following four different companding strategies were used changing Q values(q1, q2) of pre-filter(F filter) and post filter(G filter). Firstly, five IEEE sentences which were presented with speech-shaped noise at different SNRs(0, 5, 10, 15 dB) were processed by the companding. Secondly, the processed signals were applied to HeLPS. For comparison, signals which were not processed by companding were also applied to HeLPS. For the processed signals, log-likelihood ratio(LLR) and cepstral distance(CEP) were measured for evaluation of speech quality. Also, fourteen normal hearing listeners performed speech reception threshold(SRT) test for evaluation of speech intelligibility. As a result of this study, the processed signals with the companding and NAL-NL1 have performed better than that with only NAL-NL1 in the sensorineural hearing loss conditions. Moreover, the higher ratio of Q values showed better scores in LLR and CEP. In the SRT test, the processed signals with companding(SRT = -13.33 dB SPL) showed significantly better speech perception in noise than those processed using only NAL-NL1(SRT = -11.56 dB SPL).

**Key words:** Speech recognition, Companding, NAL-NL1, Hearing loss simulator

### 1. 서 론

감각신경성난청(sensorineural hearing loss)은 와우

(cochlea)의 유모세포(hair cell) 손상으로 역동범위(dynamic range)가 감소하여 정상 청력에 비하여 잡음이 있는 상황에서 어음인지력이 더욱 떨어진다[1]. Plomp의 연구에 의하면 감각신경성난청을 가진 사람들은 언어형태의 배경잡음(speech-shaped noise)이 있는 경우 시간 분해능과 주파수 분해능의 저하로 언어인지능력이 저하되고 문장청취 역치가 난청정도에 따라 2.5~7 dB 정도 상승한다고 하였다[2]. 유모세포가 대부분 손상된 고도난청의 경우에는 보청기의 착

Corresponding Author : 김인영

서울시 성동구 행당동 17번지 한양대학교 제 2의학관 의공학교실

TEL: +82-2-2291-1713 / FAX: +82-2-2220-4949

E-mail : iykim@hanyang.ac.kr

본 연구는 지식경제부 바이오 의료기기 전략기술개발사업의 지원으로 이루어졌음(과제번호 : 10031741).

용으로 청각기능의 보정이 불가능하기 때문에 청각신경에 전기적 자극을 주어 중추신경이 소리를 인지할 수 있도록 인공와우장치(cochlear implant)를 삽입하여 청력을 복원할 수 있다. 특히 인공와우장치를 사용하는 난청자들은 주파수 분해능의 저하로 인해 어음이해도가 크게 떨어지기 때문에 이를 해결하기 위한 다양한 스펙트럼 콘트라스트 향상 알고리즘(spectral contrast enhancement algorithm)들이 연구 개발되어 왔다[3-12]. 그 중 Turicchia 와 Sarpeshkar 에 의해 개발된 압신 알고리즘(companding algorithm)은 다른 알고리즘처럼 고속푸리에변환(fast fourier transform, FFT)을 거치지 않고 시간영역에서 처리되며 저전력의 아날로그 초집적회로(very large-scale integration, VLSI)를 사용하였기 때문에 인공와우장치에 효과적이라고 알려져 있다[11]. 실제 인공와우환자를 대상으로 압신 알고리즘의 성능을 실험한 결과 압신 알고리즘의 사용시 0~5 dB SNR 상황에서 12.1~21.3% 향상된 언어인지를 보였다[12]. 이와 같은 결과는 음성신호에 압축기법(compression)과 신장기법(expansion)이 차례로 적용되어 two-tone suppression 효과[13]를 만들어냄으로써 특정 주파수영역의 스펙트럼 피크를 강화시키고 주변의 불필요한 스펙트럼 에너지를 약화시키기 때문이다. 즉, 음성의 스펙트럼 콘트라스트를 향상시키고 배경잡음을 제거하기 때문에 결과적으로 음성을 강화시켜준다. Baer 등은 스펙트럼 콘트라스트 강화는 인공와우환자 뿐만 아니라 보청기 사용자에게도 잡음상황에서의 언어인지를 높여준다고 보고하였다[7]. 보청기 사용자의 경우, 다채널 압축(multi-channel compression)을 하여 좁아진 역동범위(dynamic range) 안에 소리를 압축하여 넣음으로써 난청보상을 어느 정도 해결할 수 있지만 이로 인한 스펙트럼 상의 뭉개짐(spectral smearing)은 음성의 왜곡을 발생시키기 때문에 스펙트럼 강화기술이 필요하다는 연구가 소개되었다[3]. 이러한 연구결과들을 배경으로 스펙트럼 콘트라스트 강화기술은 보청기 사용자들의 어음인지력 향상에 기여할 수 있다고 가정하고, 본 연구에서는 스펙트럼 콘트라스트 강화기술의 대표적인 기술인 압신 알고리즘에 대한 검증 수행하였다.

지금까지 보청기 알고리즘에 대한 평가를 하기 위해서는 필수적으로 난청인들을 대상으로 임상실험을 해야 했다. 하지만 이러한 임상연구는 실제 난청인들을 대상으로 해야 하기 때문에 고령자가 많아 원활한 실험 진행에 어려움이 있고 대상자 모집을 하는데 많은 시간과 노력이 필요하다. 본 논문에서는 다양한 난청유형에 대한 임상결과를 간접적인 방식으로 예측하고 평가하기 위해 난청시뮬레이터(HeLPS, SENSIMETRICS, USA)를 사용하였다[14]. 난청시뮬레이터를 이용하여 경사형 경도난청(mild gently sloping hearing loss)과 경사형 중·고도난청(moderate to steeply sloping

hearing loss)의 두 가지의 난청상황을 만들어 유모세포가 손상된 와우를 시뮬레이션 하였고, 난청에 대해 각각 NAL-NL1으로 비선형적 이득을 주어 보상하여 기본적인 보청기의 기능을 시뮬레이션 하였다[14-15].

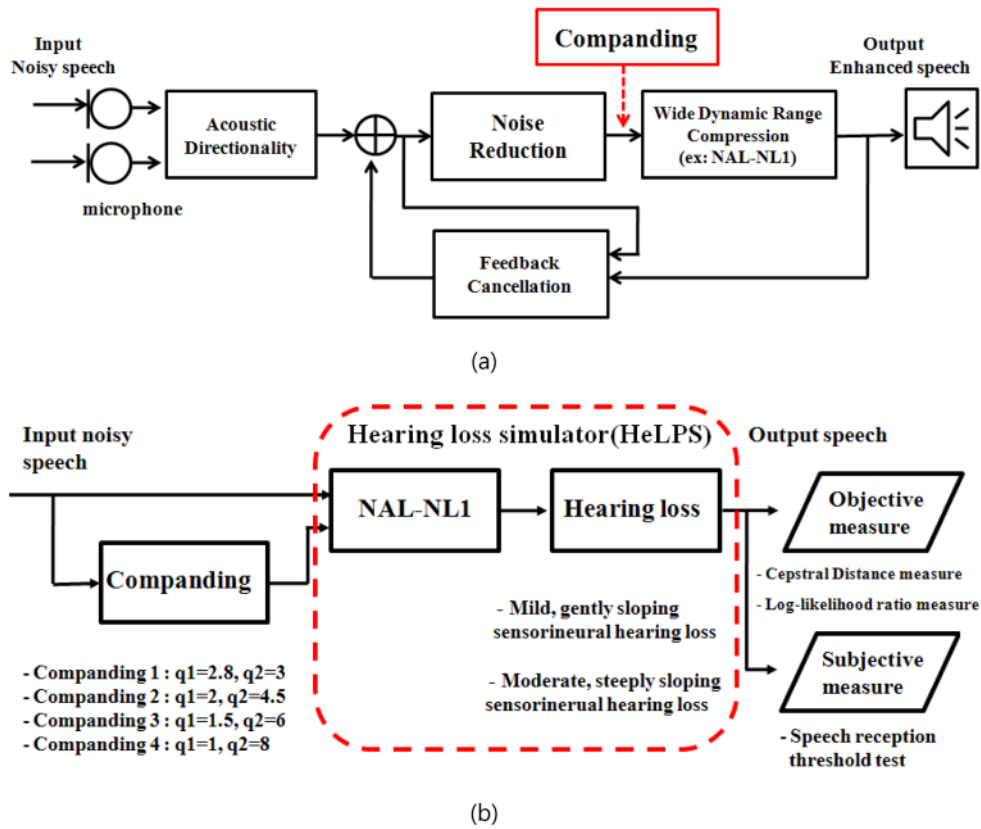
또한 보청기에 적합한 최적의 압신 알고리즘을 찾기 위해 밴드폭이 서로 다른 4가지의 필터뱅크를 이용하여 다양한 압신 알고리즘을 구현하여 이를 난청시뮬레이터에 각각 적용하였을 때와 적용하지 않았을 때에 대해서 객관적 평가(objective measure)를 하였다. 또한 청각이 정상인 14명의 남성들을 대상으로 speech reception threshold(SRT) 테스트를 실시하여 주관적 평가(subjective measure)를 수행하였다[16]. 객관적 평가와 주관적 평가를 통하여 압신 알고리즘이 인공와우장치 뿐만 아니라 보청기에서도 사용될 수 있으며, 어떠한 난청상황일 때 잡음상황에서 언어인지를 향상시킬 수 있는지에 대해서도 확인하였다.

## II. 본 론

### 1. 시스템 구성

보청기 전체시스템의 일반적인 알고리즘 구성은 그림 1(a)와 같다. 먼저 잡음이 전·후방 송화기(microphone)로 들어오면 지향성 알고리즘(acoustic directionality) 블록에서 신호의 방향을 감지하여 잡음이 없는 음성부분의 지향성(directionality)을 높여준다. 그 뒤 잡음을 추정하여 제거(noise reduction)하고 음향의 누출과 송화기의 재증폭된 소리로 인한 음향되울림을 제거하는 단계(feedback cancellation)가 있다. 그리고 난청인의 역동범위(dynamic range)를 높여주고 난청을 보정(NAL-NL1)하기 위한 광역동범위압축(wide dynamic range compression)을 거쳐 강화된 음성을 듣게 된다.

본 연구의 검증 시스템은 크게 그림 1(b)와 같이 잡음입력 신호(input noisy speech)가 압신 알고리즘을 거치지 않고 난청시뮬레이터를 통과하는 경우와 압신 알고리즘을 거쳐 난청시뮬레이터를 통과한 경우의 출력신호들을 각각 객관적 평가와 주관적 평가를 하는 단계로 구성된다. 여기서 NAL-NL1 블록에서는 보청기의 기본적인 난청보상 알고리즘을 시뮬레이션 하였고, hearing loss 블록에서는 유모세포가 손상된 와우에 의한 난청을 시뮬레이션 하였다. hearing loss는 두 가지 난청 상황(경사형 경도난청, 경사형 중·고도난청)이 시뮬레이션 되었고, 입력 신호가 보청기를 통해(NAL-NL1 블록) 손상된 와우(hearing loss 블록)를 통과해서 들리게 함으로써 실제 보청기를 착용한 난청환자의 상태를 모사하였다. 본 시스템의 구성에서는 압신 알고리즘만의 성능 비교를 위하여 일반적으로 보청기에 탑재되는 지향성 알고리즘(acoustic directionality)과 잡음 제거 알고리즘(noise reduction)은 포함하지 않았다.



220 그림 1. 시스템 구성 (a) 일반적인 디지털 보청기 시스템, (b) 압신성능 검증시스템

Fig. 1. System configuration (a) Block diagram of the common digital hearing aids, (b) Block diagram of the companding performance evaluation method

2. 압신 알고리즘(Companding Algorithm)

압신 알고리즘은 인간의 청각특성 중 하나인 two-tone suppression현상의 특성을 모델링하여 인공와우환자들의 언어인지향상을 위해 개발된 알고리즘이다[11]. 압신 알고리즘

은 그림 2(a)와 같이 고속푸리에변환(fast fourier transform, FFT)을 거치지 않고 시간도메인 상에서 N개의 채널로 나누어 처리되며, 광대역 필터(F filter), 압축블록(compression block), 협대역 필터(G filter), 팽창블록(expansion

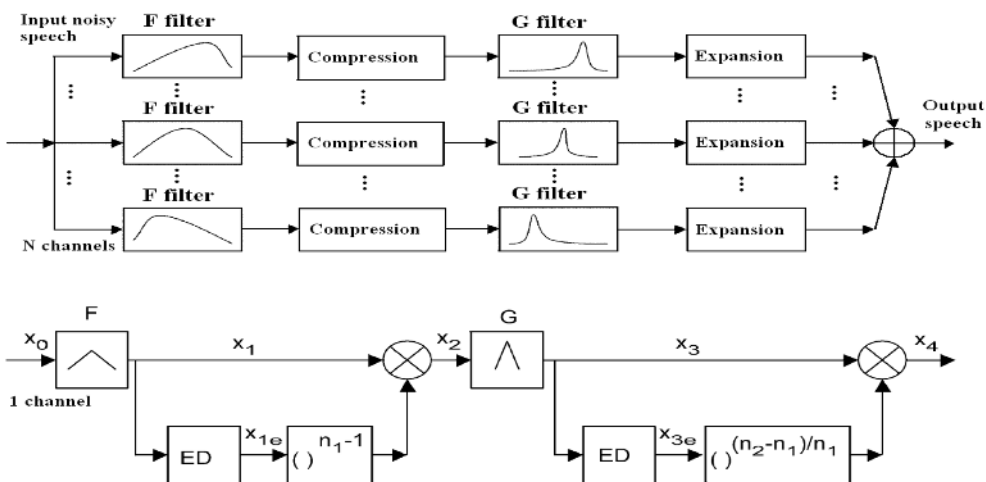


그림 2. 압신 알고리즘의 구성 (a) 다채널 압신 알고리즘, (b) 단일채널 압신 알고리즘

Fig. 2. Block diagram of the companding algorithm (a) A multichannel of the companding algorithm, (b) A single channel of the companding algorithm(F: F filter, G: G filter, ED: envelope detector,  $n_1$ : compression index,  $n_2$ : expansion index)

block)으로 구성된다[11-12]. 각 채널의 전처리 필터를 거친 신호는 전파 정류기(full wave rectification)와 1차 저대역통과필터(low-pass filtering)를 통과하고, 이 신호에 힐버트 변환(hilbert transform)을 적용하여 포락선(envelope)이 추출된다[17-18]. 그림 2(b)는 단일채널의 압신 처리과정을 나타낸다. 압축(compression)과 팽창(expansion)은 포락선 검출기(ED), 비선형 블록, 곱셈기로 수행된다. 포락선의 시상수는 각 채널의 중심 주파수에 의해 조절되며, 입력신호( $X_0$ )는 채널의 광대역 필터(F filter)를 통과한 후 포락선 검출기(ED)에서 강한 신호성분에 기반하여 결정된 압축비( $n_1$ )에 의해 압축된다. 그리고 압축된 신호( $X_2$ )는 채널의 협대역 필터(G filter)를 통과하게 되고 협대역 내의 신호 중 포락선 검출기(ED)의 강한 신호성분에 기반하여 결정된 팽창비( $n_2/n_1$ )에 의해 팽창되어 출력( $X_4$ )된다.

i번째 채널의 전처리 필터와 후처리 필터의 전달함수  $F'_i(s)$ 와  $G'_i(s)$ 는 Eq. [1], Eq. [2]와 같이 나타낼 수 있고, 필터의 밴드폭은 전달함수(transfer function)의 필터계수  $q_1, q_2$ 에 의해서 결정된다.

$$F'_i(s) = \left( \frac{2\left(\frac{\tau_i}{q_1}\right)s}{\tau_i^2 s^2 + 2\left(\frac{\tau_i}{q_1}\right)s + 1} \right)^2 \quad (1)$$

$$G'_i(s) = \left( \frac{2\left(\frac{\tau_i}{q_2}\right)s}{\tau_i^2 s^2 + 2\left(\frac{\tau_i}{q_2}\right)s + 1} \right)^2 \quad (2)$$

3. 필터뱅크 설계

본 연구에서는 표 1과 같이 압신 파라미터 중 필터의 밴드폭을 결정하는 광·협대역 필터계수( $q_1, q_2$ )를 다양하게 설정하여 4가지의 필터뱅크를 구현하였다. 그리고 필터의 밴드폭 변화에 따른 결과를 보기 위하여 다른 압신 파라미터 ( $n_1 = 0.3, n_2 = 1$ )값들은 모두 고정하였다. 압신 알고리즘은  $q_1, q_2$ 의 비율에 따라 two-tone suppression이 되는 정

표 1. 4가지의 광·협대역 필터뱅크 설계

Table 1. Four types of the F, G filterbank

Types of the F, G filterbank	F, G filterbank( $q_1 : q_2$ )	
	$q_1$	$q_2$
Companding 1	2.8	3
Companding 2	2	4.5
Companding 3	1.5	6
Companding 4	1	8

도가 다르게 나타난다. 즉, 그림 3과 같이  $q_1$ 은 작을수록  $q_2$ 는 클수록 광대역 필터폭은 넓어지고 협대역 필터폭은 더욱 좁아지게 되어 압신 효과가 커진다. Bhattacharya 와 Zeng 은 압신 알고리즘의 필터의 채널 수를 변경하여 압신이 적용된 모음(vowel)의 스펙트럼을 비교해 본 결과 필터 채널 수가 40보다 적어지면 비교적 제 2포만트와 제 3포만트에서 원하지 않은 감쇄가 일어나기 때문에 50채널이 가장 적당하다고 하였다[12]. 따라서 본 연구에서도 그림 3과 같이 50채널의 광·협대역 필터뱅크를 로그척도(100~8000 Hz)로 설계하였다.

III. 실험 및 평가

1. 실험자료

음질의 객관적 평가를 위해 난청시물레이션에 사용된 음원은 25 kHz의 샘플링 주파수를 갖는 IEEE sentence(5개의 음성데이터)이며, 이를 16 kHz로 다운샘플링 하여 사용하였다[19]. 잡음데이터는 언어형태의 배경잡음(speech-shaped noise,  $F_s = 16 \text{ kHz}$ )을 사용하여 입력신호대 잡음비(input SNRs = 0, 5, 10, 15 dB)로 섞은 음성잡음(noisy speech)을 만들었다[20]. 주관적 평가에 사용된 데이터는 객관적 평가에 사용된 동일한 잡음데이터이며, 어음청각검사시 사용되는 한국표준 일반용 이음절표(KS-BWL-A3)에서 제공된 12개의 이음절 단어를 사용하였다[21]. 그림 4는 잡음이 섞이지 않은 모음 /아/의 스펙트럼(가는 실선)과 모음 /아/를 0 dB SNR

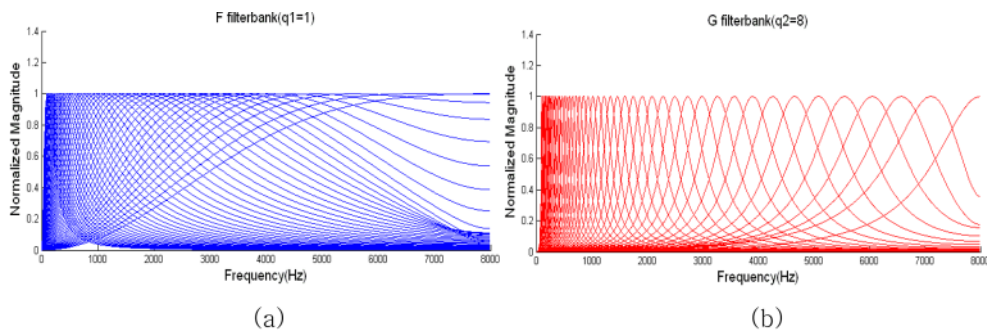


그림 3. 광·협대역 필터뱅크의 설계( $q_1 = 1, q_2 = 8$ ) (a) 광대역 필터뱅크, (b) 협대역 필터뱅크  
Fig. 3. Implementation of the F, G filterbank( $q_1 = 1, q_2 = 8$ ) (a) F filterbank, (b) G filterbank

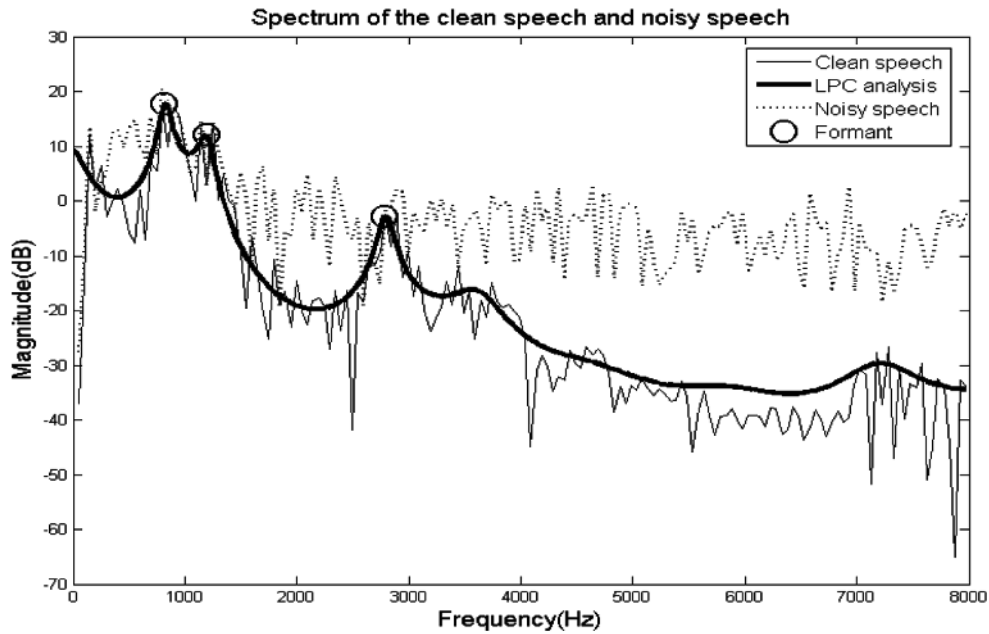


그림 4. 모음/아/와 언어형태의 배경잡음(speech-shaped noise)이 섞인 음성잡음(0 dB SNR)의 스펙트럼  
 Fig. 4. The spectrum of the vowel /a/ and noisy speech with speech-shaped noise at input SNR(0 dB SNR)

의 언어형태의 배경잡음(speech-shaped noise)으로 섞은 음성잡음(noisy speech)의 스펙트럼(점선)이다. 굵은 실선은 원음의 포만트(formant) 성분을 찾기 위해 18차 선형 예측 부호화(linear predictive coding, LPC)를 이용하여 포락선(envelope)을 추출한 것이다[22]. 음성잡음(noisy speech)의 스펙트럼을 보면 언어형태의 배경잡음(speech-shaped noise)으로 인하여 "o" 표시된 포만트 성분(F = 820 Hz, F2 = 1200 Hz, F3 = 2800 Hz)들의 사이가 뭉개진 것을 확인할 수 있다.

2. 난청 시뮬레이션

난청시뮬레이션은 표 2와 같이 6가지 상황을 고려하였다. 난청 정도 및 난청 유형은 여러 가지가 있는데 본 연구에서는 표 3과 같이 경사형 경도난청(HL1), 경사형 중 · 고도난청(HL2)을 선택하였다. 그 이유는 보청기를 사용하는 대부분의 난청인들이 2~4 kHz의 주파수 범위에서 청력이 급격히 나빠지는 경사형 난청유형이기 때문에 이를 선택하였다 [1]. 표 3은 두 가지 난청(HL1, HL2)에 대한 각 주파수 별 청력손실 정도를 나타낸 것이다.

손실된 청력을 보상하기 위한 보청기 처방에는 Berger처

표 2. 6가지 난청시뮬레이션 상황

Table 2. Six conditions of hearing loss simulation

No.	Method
1	mild gently sloping hearing loss(HL1)
2	moderate to steeply sloping hearing loss(HL2)
3	HL1+NAL-NL1(NHL1)
4	HL2+NAL-NL1(NHL2)
5	HL1+NAL-NL1+Companding 1~4(CNHL1)
6	HL2+NAL-NL1+Companding 1~4(CNHL2)

방법[23], POGO처방법[24], Libby처방법[25], NAL처방법[26], DSL처방법[27-28] 등 다양한데 본 연구에서는 NAL-NL1(national acoustic laboratories-non-linear version 1)처방법을 사용하였다[15]. NAL-NL1처방법은 NAL처방법의 비선형증폭방식 버전으로 음성명료도지수(speech intelligibility index, SII)[29]를 인용하여 정상청력이 인지하는 만큼 말소리의 명료도를 높임으로써 상향차폐(upward spread of masking)나 하향차폐(downward spread of masking)를 피하고 정상적인 음량을 회복하도록 이득을 산출한 것이

표 3. 경사형 경도난청(HL1)과 경사형 중 · 고도난청(HL2)의 청력 손실 정도

Table 3. The hearing threshold of the mild gently sloping hearing loss(HL1) and the moderate to steeply sloping hearing loss(HL2)

주파수(kHz)	0.25	0.5	1.0	2.0	3.0	4.0	6.0	8.0
HL1 threshold(dB HL)	20	20	30	40	45	50	50	50
HL2 threshold(dB HL)	25	30	55	65	80	85	90	90

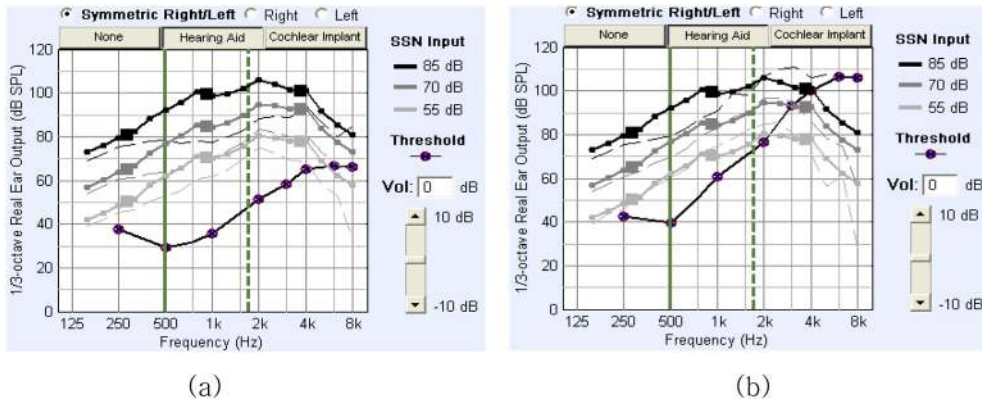


그림 5. NAL-NL1의 적용 시 출력레벨 (a) 경사형 경도난청, (b) 경사형 중·고도난청  
 Fig. 5. The output level of the hearing aid applying NAL-NL1 insertion gain (a) mild gently sloping hearing loss, (b) moderate to steeply sloping hearing loss

다. 그림 5(a), 그림 5(b)는 HL1과 HL2에 NAL-NL1처방법을 각각 적용하였을 때의 출력레벨을 나타낸 것이다. 난청정도에 따라 각 주파수의 청력역치(hearing threshold)가 다르게 표시("O")되며, NAL-NL1계산법에 의해 각 주파수의 출력레벨이 결정된다. NAL-NL1의 삽입이득(insertion gain)은 125~8000 Hz까지 각 표준 1/3옥타브에 해당하는 주파수에 의해 계산된다. 즉, 각 주파수의 이득은 그 주파수의 역치, 세 주파수의 평균역치, 500~2000 Hz까지 청력도의 기울기 등에 의해 결정된다. 또한 입력 음압이 작은 소리에서 큰 소리로 갑자기 변하거나 이와 반대현상이 일어날 때 출력 음압이 안정상태가 될 때까지 걸리는 시간(압축시간, 해제시간)을 각각 설정해야 하는데 시간이 너무 짧으면 사람이 말하고 있는 동안에 이득이 자주 바뀌게 되어 신호의 파형이 왜곡되고, 신호의 요철현상(pumping)이 발생한다. 따라서 이를 방지하기 위해 보청기의 압축시간을 10 msec, 해제시간을 200 msec를 통상적으로 사용하고 있으며, 본 연구에서도 압축시간 및 해제시간을 이와 동일하게 설정하였다[30].

3. 객관적 평가

음성의 객관적 평가방법 중 log-likelihood ratio(LLR), itakkura-satio distance(IS), cepstral distance(CEP) 등은 all-pole 모델을 기반으로 입력음성과 출력음성의 스펙트럼 상의 레벨차이를 보아 출력음성이 원음성과 비교하여 얼마나 일치하는가를 통해 음질을 평가한다[31]. 본 연구에서는 언어형태의 배경잡음(speech-shaped noise)을 입력신호 대 잡음비(input SNRs = 0, 5, 10, 15 dB)인 경우의 음성신호를 HL1과 HL2인 경우에 대해 각각 NAL-NL1만 적용하였을 때(NHL1, NHL2)와 NAL-NL1과 companding 1~4(4가지 필터뱅크)를 함께 적용하였을 때(CNHL1, CNHL2)에 대해 LLR, CEP를 각각 측정하였다. LLR은 입력신호와 출력신호의 LPC 계수 스펙트럼 차이로 표시되며, Eq. [3]과 같다.

$$d_{LLR}(\vec{a}_p, \vec{a}_c) = \log \left( \frac{\vec{a}_p^T R_c \vec{a}_p}{\vec{a}_c^T R_c \vec{a}_c} \right) \tag{3}$$

$\vec{a}_c, \vec{a}_p$ 는 음성입력신호와 출력음성신호의 LPC 계수 벡터 값이고  $R_c$ 는 입력신호의 자기상관함수(autocorrelation matrix)를 의미한다.

CEP는 LPC 기반의 여러 가지 거리측정방법 중의 하나로서 Eq. [4]와 같이 왜곡되지 않은 입력음성신호(clean signal)의 스펙트럼과 출력음성신호(processed signal)의 스펙트럼 상의 로그스펙트럼 차이를 나타내며, Eq. [5]의 캡스트럼 상관계수  $c(m)$ 을 갖는다.

$$d_{CEP}(\vec{c}_c, \vec{c}_p) = \frac{10}{\log_e 10} \sqrt{2 \sum_{k=1}^p [c_c(k) - c_p(k)]^2} \tag{4}$$

$$c(m) = a_m + \sum_{k=1}^{m-1} \frac{k}{m} c(k) a_{m-k} \quad 1 \leq m \leq p \tag{5}$$

$P$ 는 LPC 분석차수,  $\vec{c}_c$ 와  $\vec{c}_p$ 는 입력음성신호와 출력음성신호의 캡스트럼 상관벡터 값을 각각 의미하며,  $a_m$ 은 LPC 상관계수이다. LLR과 CEP는 각각 0~2, 0~10의 범위를 가지며 두 방법 모두 값이 적을수록 원음과 상관관계가 높아 음질이 향상되었다는 것을 의미한다.

4. 주관적 평가

객관적 평가결과를 검증하기 위해 주관적 평가를 실시하였다. 실험대상은 표 4와 같이 청력이 정상인 22~31세의 성인남성 14명을 대상으로 어음취척역치검사(speech reception threshold, SRT)를 실시하였다[16]. SRT 테스트는 피험자의 쾌적수준(most comfortable level, MCL)에서 2음절의 단어를 들려주고 피험자가 정답을 맞출 경우 2 dB SPL를 낮추

표 4. 정상청력의 피험자 정보

Table 4. Information of the normal hearing listeners

Subject	Age	Sex	Hearing threshold
NH(14 명)	22~34 세	Male	Normal

\*NH : normal hearing, \*Mean age  $\pm$  SD :  $27 \pm 2.54$

표 5. SRT 테스트에 사용된 4가지 방법

Table 5. Four methods of the SRT test

No.	Method
1	normal hearing(NH)
2	moderate to steeply sloping hearing loss(HL2)
3	HL2+NAL-NL1(NHL2)
4	HL2+NAL-NL1+Companding 3 (CNHL2, $q_1=1.5, q_2=6$ )

고, 오답일 경우 2 dB SPL를 높여주어 피험자가 50%를 정확하게 응답할 수 있는 최소 어음강도를 찾는 청각검사이다. 실험에 사용된 음성데이터는 표 6과 같이 한국어음성각감사에서 한국표준 일반용 이음절표(KS-BWL-A3)에 수록된 12개의 단어를 사용하였다[21]. 실험에 사용된 잡음데이터는 언어형태의 배경잡음(speech-shaped noise)이며 input SNRs  $\pm 2$  dB SPL간격의  $-20 \sim 20$  dB SPL 음압 세기로 원음과 섞어주었다[20]. 표 5는 SRT 테스트에 사용된 4가지의 시뮬레이션 방법이며 번호순서대로 실험을 진행하였으며, 표 2의 6가지 난청시뮬레이션 상황 중 HL2, HL2+NAL-NL1,



그림 6. SRT 테스트 평가용 Matlab GUI

Fig. 6. The Matlab GUI for SRT test

HL2+NAL-NL1+companding 3( $q_1 = 1.5, q_2 = 6$ )의 경우에 대해서만 평가하였다. 그 이유는 객관적 평가 시 HL1인 경우보다는 HL2일 때 압신 알고리즘의 성능이 더 좋게 나왔으며, 압신 알고리즘의 특성상 F, G 필터계수의 차이가 너무 작으면 음성정보에 대한 손실은 적으나 주변잡음으로 인해 원하는 음성을 구분하기 어렵고 반대로 차이가 너무 커지면 주변잡음은 어느 정도 제거되지만 음성정보의 손실은 커져 결국 음질이 떨어지게 되는 trade-off 관계가 성립되기 때문이다. SRT 테스트는 그림 6과 같이 실험평가를 위해 Mathworks사의 Matlab GUI를 이용하였으며 측정헤드폰은 insert 이어폰(Compumedics 10 ohm)을 사용하여 모든 음압을 대화음 크기(70 dB SPL)로 맞추어 주었다. 또한 모

224

표 6. SRT에 사용된 12개의 한국표준 이음절 단어(KS-BWL-A3)

Table 6. Two syllables of 12 Korean for SRT test

No.	1	2	3	4	5	6	7	8	9	10	11	12
단어	그림	아들	팔죽	동생	목표	냄새	바다	자연	접시	권투	방식	느낌

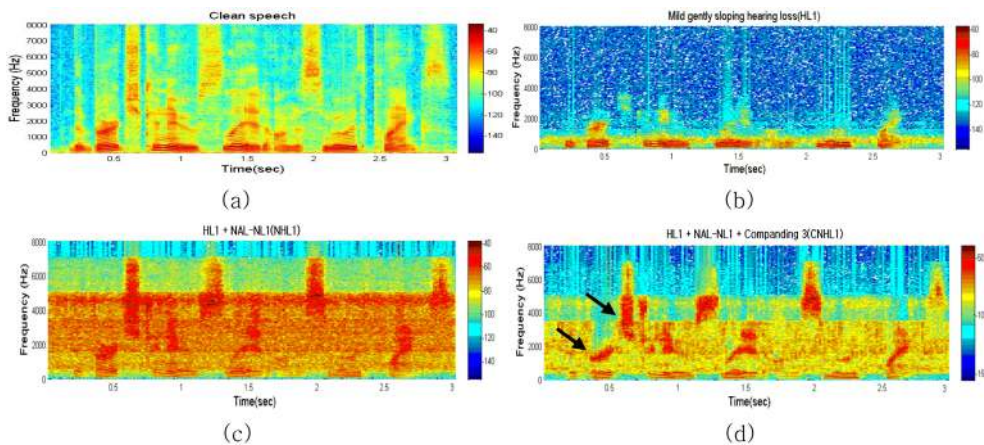


그림 7. 10 dB SNR에서 상황별 스펙트로그램 비교(HL1인 경우)

Fig. 7. The comparison of the spectrogram for each condition at 10 dB SNR(HL1) (a) clean speech, (b) mild gently sloping hearing loss(HL1), (c) HL1+NAL-NL1(NHL1), (d) HL1+NAL-NL1+Companding 3(CNHL1,  $q_1 = 1.5, q_2 = 6$ )

든 실험은 SONTEK사의 무향실(anechoic room)에서 측정하였다.

### IV. 결 과

#### 1. 스펙트로그램 분석

시간변화에 따른 주파수의 변화를 보기 위해 스펙트로그램(spectrogram)분석을 하였다. 그림 7(a), 8(a)는 원음의 스펙트로그램을 나타내며, 그림 7(b), 8(b)는 10 dB SNR 상황에서 경사형 경도난청, 경사형 중·고도난청이 있을 때 HeLPS에서 출력된 음성의 스펙트로그램이다. 원음의 스펙트로그램과 비교해 볼 때 그림 7(b), 8(b)는 모두 청력손실로 인하여 제 1포먼트 주파수 성분( $400 \pm 50$  Hz)만 제외하고 2 kHz이상의 고주파대역 음성에너지는 모두 감쇄되어 들을 수 없음을 알 수 있다. 즉, 이러한 경우에 출력된 음성웨이브 파일의 소리를 정상인이 듣게 되면 언어를 정확하게 인지하기 힘들며, 실제 난청인들도 이와 유사하게 청취할 것으로 예측할 수 있다. 또한 그림 7(c), 8(c)는 난청을 보정하기 위해 NAL-NL1으로 청력이 손실된 고주파 대역에 비선형

적 이득을 주었다. 하지만 잡음성분은 제거되지 않고 함께 증폭되는 결과를 보였다. 그림 7(d), 8(d)에서는 NAL-NL1의 비선형 증폭을 통해 고 대역의 주파수 성분을 들을 수 있을 뿐만 아니라 압신의 two-tone suppression 성능으로 강한 에너지를 갖는 포먼트 주파수(화살표 표시)가 강조되고 그 주변의 잡음은 감쇄됨을 확인할 수 있다.

#### 2. 객관적 평가결과

평가결과 그림 9(a), 그림 10(a)의 HL1인 경우 NAL-NL1만 적용하였을 때보다 압신을 함께 적용하였을 때의 LLR, CEP가 0~15 dB SNR에서 각각 0.001~0.14, 0.002~0.32 정도 차이의 값을 보였다. 또한 HL2의 경우도 같은 상황에서 LLR, CEP가 각각 0.002~0.40, 0.08~1.34로 HL2가 HL1일 때보다 스펙트럼 콘트라스트가 강화되어 원음과 더욱 일치함을 알 수 있었다(그림 9(b), 그림 10(b)). 그림 11은 이에 따른 통계적 분석 결과이다. 통계적인 분석은 student's t-test를 이용하여 두 집단 간의 평균간의 차의 통계적 유의성을 비교하였다(\*:p < 0.05, \*\*:p < 0.01)[32]. 또한 압신의 유무에 의한 유의성을 보기 위해 다음의 4가지 경우(I~IV)에

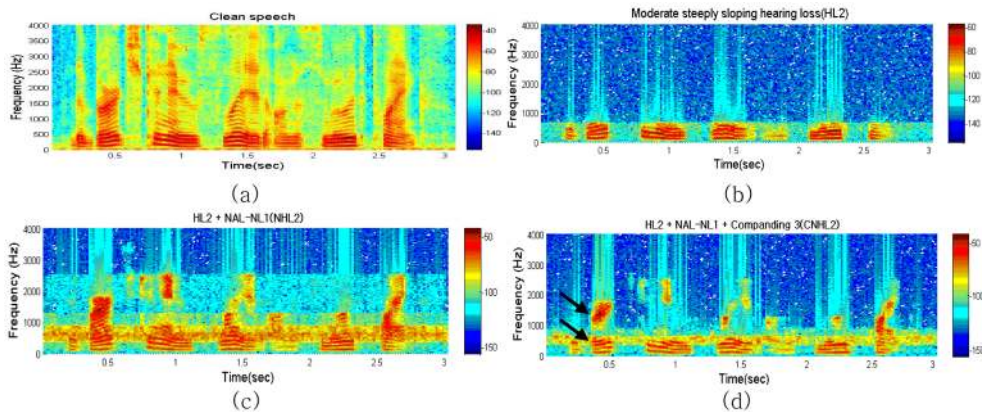


그림 8. 10 dB SNR에서 상황별 스펙트로그램 비교(HL2 인 경우)

Fig. 8. The comparison of the spectrogram for each condition at 10 dB SNR(HL2) (a) clean speech, (b) moderate to steeply sloping hearing loss(HL2), (c) HL2+NAL-NL1(NHL2), (d) HL2+NAL-NL1+Companding 3(CNHL2,  $q_1 = 1.5, q_2 = 6$ )

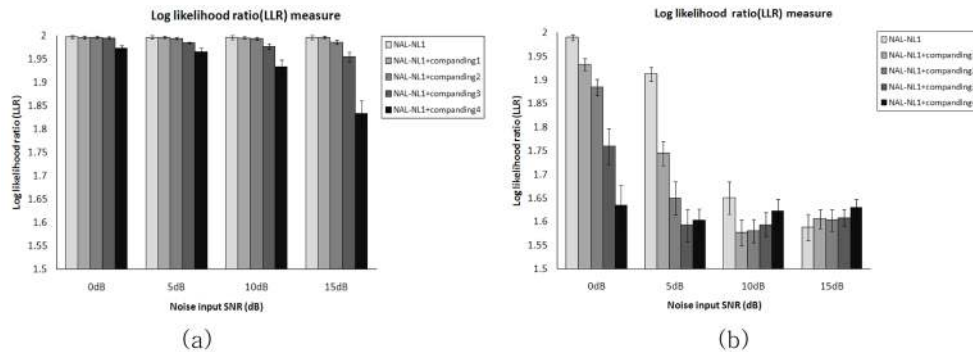


그림 9. 여러 SNR에 따른 LLR 측정결과 (a) HL1인 경우 LLR 측정결과, (b) HL2인 경우 LLR 측정결과  
Fig. 9. The result of the LLR measure at different SNRs(0, 5, 10, 15 dB) (a) HL1, (b) HL2



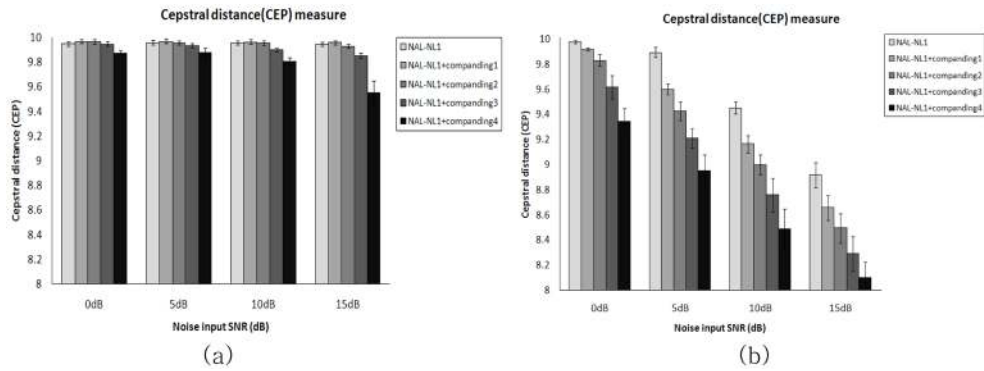


그림 10. 여러 SNR에 따른 CEP 측정결과 (a) HL1인 경우 CEP 측정결과, (b) HL2인 경우 CEP 측정결과  
 Fig. 10. The result of the CEP measure at different SNRs(0, 5, 10, 15 dB) (a) HL1, (b) HL2

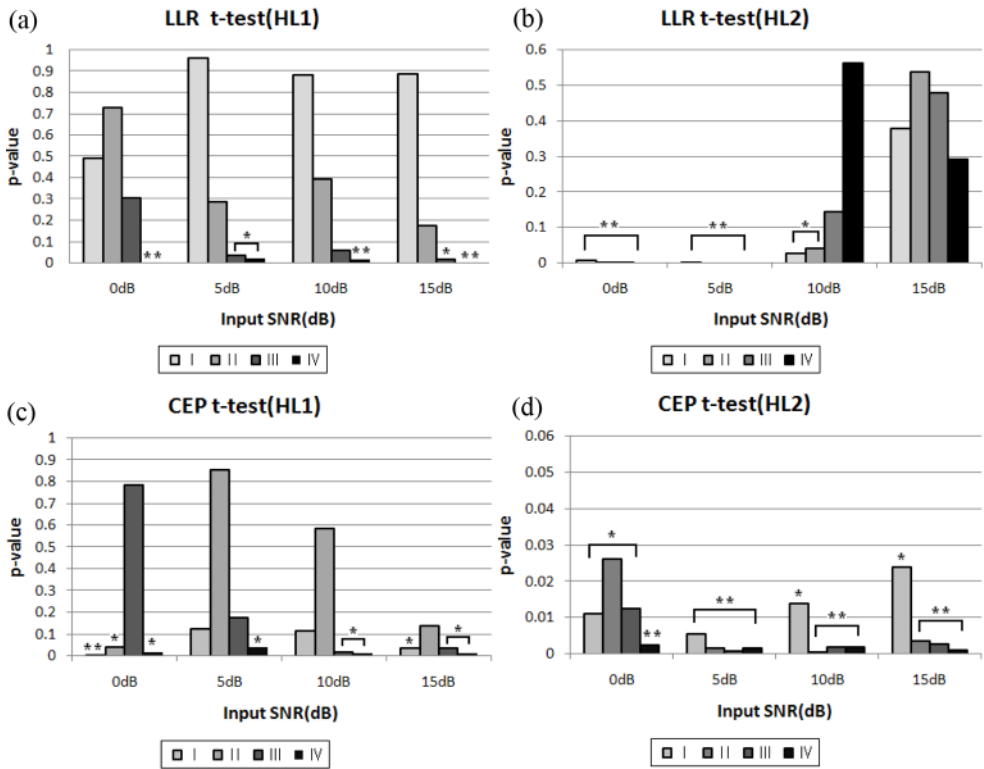


그림 11. 여러 SNR에 따른 난청유형별(HL1, HL2) LLR, CEP의 student's t-test 결과(\*:  $p < 0.05$ , \*\*:  $p < 0.01$ , I: NAL-NL1과 NAL-NL1+Companding 1 비교, II: NAL-NL1과 NAL-NL1+Companding 2 비교, III: NAL-NL1과 NAL-NL1+Companding 3 비교, IV: NAL-NL1과 NAL-NL1+Companding 4 비교) (a) HL1인 경우 LLR의 t-test 결과, (b) HL2인 경우 LLR의 t-test 결과, (c) HL1인 경우 CEP t-test 결과, (d) HL2인 경우 CEP의 t-test 결과  
 Fig. 11. The result of LLR, CEP student's t-test at different SNRs(0, 5, 10, 15 dB) (a) LLR(HL1), (b) LLR(HL2), (c) CEP(HL1), (d) CEP(HL2)

대하여 비교하였다. 즉, (1) NAL-NL1과 NAL-NL1+Companding 1의 비교(I), (2) NAL-NL1과 NAL-NL1+Companding 2의 비교(II), (3) NAL-NL1과 NAL-NL1+Companding 3의 비교(III), (4) NAL-NL1과 NAL-NL1+Companding 4의 비교(IV)를 하였다. HL1인 경우 LLR 결과 0 dB SNR(IV), 5 dB SNR(III, IV), 10 dB SNR(IV), 15 dB SNR(III, IV)일 때 유의미한 결과( $p < 0.05$ )를 얻었다. CEP는 0 dB

SNR(I, II, IV), 5 dB SNR(IV), 10 dB SNR(III, IV), 15 dB SNR(I, III, IV)일 때 유의미한 값을 보였다. 또한 HL2인 경우 LLR은 0 dB SNR(I~IV), 5 dB SNR(I~IV), 10 dB SNR(I, II)에서 통계적 유의성이 있는 반면 15 dB SNR에서는 모두 통계적으로 유의하지 않은 결과를 보였다. 그리고 CEP는 0~15 dB SNR에서 I~IV 방법 모두 유의미한 결과를 보였다. 이를 통해 HL1일 경우 압신계수비율( $q1 : q2$ )이

가장 높은 IV일 때만 모든 SNR에 대해 통계적으로 유의미한 결과를 보이고 나머지의 경우에는 유의미하지 않은 결과를 보여 압신의 성능이 좋다고 볼 수 없었다. HL2인 경우는 압신은 0~5 dB의 낮은 SNR에서 좋은 성능( $p < 0.05$ )을 보였으나 그 이상의 SNR에서는 성능이 좋지 않았다.

### 3. 주관적 평가결과

주관적 평가는 그림 12와 같이 정상인 14명을 대상으로 NH(normal hearing), HL2(moderate to steeply sloping hearing loss), NHL2(HL2+NAL-NL1), CNHL2(HL2+NAL-NL1+Companding 3)에 대해서 SRT 테스트를 각각 측정된 결과이다. 테스트 결과 NH(SRT = -13.73 dB), HL2(SRT = -5.7 dB), NHL2(SRT = -11.56 dB), CNHL2(SRT = -13.33 dB)으로 NH가 잡음상황에서 어음인지력이 가장 높고 HL2가 가장 낮게 측정되었으며, CNHL2는 NHL2보다 1.77 dB 낮은 SNR 상황에서 언어인지를 할 수 있었다. 압신의 적용여부(NHL2, CNHL2)에 따른 통계적 유의성을 확인한 결과( $p = 0.0328$ ) 서로 유의미한 값을 확인하였다(그

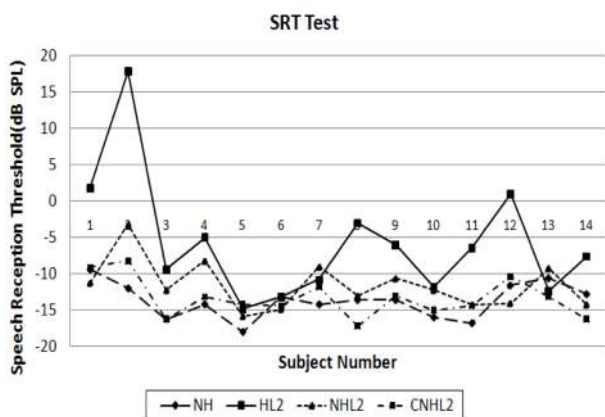


그림 12. 개인별 SRT 테스트 결과  
Fig. 12. The result of the SRT test for each subject

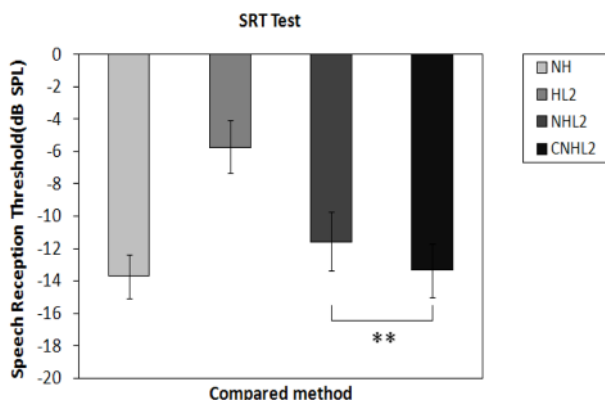


그림 13. SRT 테스트 결과(\*\*:  $r < 0.05$ )  
Fig. 13. The result of the SRT test

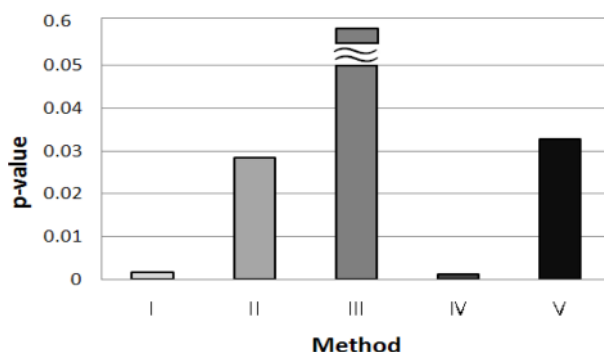


그림 14. SRT 테스트 결과 student's t-test 비교(I: NH와 HL2 비교, II: NH와 NHL2 비교, III: NH와 CNHL2 비교, IV: HL2와 CNHL2 비교, V: NHL2와 CNHL2 비교)  
Fig. 14. The comparison of the SRT test(student's t-test)

림 13, 그림 14). 그림 14는 SRT 테스트 결과(NH, HL2, NHL2, CNHL2)에 대한 통계적 분석결과이다(student's t-test)[32]. 통계적 분석을 위한 비교대상은 다음과 같다. (1) NH와 HL2 비교(I), (2) NH와 NHL2 비교(II), (3) NH와 CNHL2 비교(III), (4) HL2와 CNHL2 비교(IV), (5) NHL2와 CNHL2 비교(V)이며, III 방법( $p = 0.549$ )을 제외한 나머지 방법 들 모두 통계적으로 유의하였다( $p < 0.05$ ).

## V. 결론 및 토의

압신 알고리즘은 음성의 스펙트럼 콘트라스트를 향상시킴으로써 잡음환경에서 인공와우환자의 언어인지향상에 도움이 되어왔다[11-12]. 이러한 스펙트럼 콘트라스트 향상기술은 보청기를 착용하는 난청인들에게도 잡음상황에서 언어인지를 향상시키기 위해 필요한 기술이기 때문에 압신이 보청기를 사용하는 난청인들에게도 효과가 있는지에 대해서 연구해야 할 필요가 있다. 본 연구를 통하여 압신이 인공와우환자 뿐만 아니라 보청기를 사용하는 난청인들에게도 잡음환경에서 언어인지를 하는데 상당한 효과가 있음을 객관적(LLR, CEP), 주관적 평가결과(SRT 테스트)를 통하여 확인할 수 있었다. 즉, 실제 보청기를 착용한 난청환자의 상태를 모사하기 위해 난청시뮬레이터(HeLPS)를 사용하여 경도난청(HL1), 경사형 중·고도난청 상황(HL2)을 만들어주었으며, 보청기에 적합한 최적의 압신 알고리즘을 찾기 위해 압신의 필터계수비를 또한 변화시켜 시뮬레이션 하였다. 난청보상은 NAL-NL1을 이용하였으며, 단순히 NAL-NL1만 적용하였을 때와 압신을 함께 적용하였을 때의 경우를 서로 비교해 보았다.

난청시뮬레이션 결과 HL1인 경우보다 HL2의 경우에 잡음상황에 대한 객관적 평가지표가 더 높게 나타나 압신 알고리즘이 HL1보다 HL2의 경우에 더욱 성능이 뛰어난 것을 예측할 수 있었다. 또한 필터의 계수비율을 높일수록 스펙트럼

콘트라스트가 향상되고 주변 잡음이 더욱 제거되어 낮은 SNR에서 객관적 평가지표가 높게 측정되었다. 하지만 너무 강한 two-tone suppression에 의한 스펙트럼 콘트라스트는 SNR이 높은 상황에서는 결과적으로 음질의 왜곡(speech distortion)이 발생하여 평가지표가 낮게 측정되었다. 임상평가(SRT 테스트)결과에서는 압신 없이 NAL-NL1만 적용하였을 때보다 NAL-NL1과 압신을 함께 적용했을 때 잡음상황에서 1.77 dB의 언어인지향상을 보였다. 그리고 압신을 적용할 경우 HL2인 경우보다 7.63 dB의 언어인지향상을 보였다. Loizou와 Poroy의 연구결과에 의하면 고·심도난청인 경우 대화음 영역에서 4~6 dB 이득상승효과만 있어도 어음을 청취하고 대화를 판별하는데 큰 효과가 있다고 한다[33]. 따라서 본 연구의 결과는 압신이 보청기 사용자의 언어인지향상에 상당한 도움을 줄 수 있다고 볼 수 있다.

지금까지 난청 시뮬레이션 및 임상실험을 통해 압신이 인공와우환자 뿐만 아니라 보청기 사용자에게도 잡음상황에서의 언어인지력에 상당한 효과가 있다는 것을 알 수 있었다. 하지만 향후, 실제 보청기를 사용하는 난청인에게 적용을 하기 위해서는 실시간 환경(real time)에서 압신 알고리즘을 보청기에 탑재하였을 때 어느 정도 효과를 주는 알고리즘인지에 대한 검증과정이 추가적으로 반드시 필요하다. 또한 잡음 상황에서 난청인들의 음성명료도(speech intelligibility)를 향상시키기 위해 스펙트럼 콘트라스트 강화에 따라 발생하는 음질의 저하문제 등도 함께 고려해봐야 할 것으로 사료된다.

### 참고문헌

[1] H. Dillon, *Hearing aids*: Thieme Medical Pub, 2001. pp. 2-6, 244-245.

[2] R. Plomp, "Noise, amplification, and compression: Considerations of three main issues in hearing aid design," *Ear and Hearing*, vol. 15, pp. 2, 1994.

[3] R. Plomp, "The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation transfer function," *The Journal of the Acoustical Society of America*, vol. 83, pp. 2322, 1988.

[4] H.T. Bunnell, "On enhancement of spectral contrast in speech for hearing impaired listeners," *The Journal of the Acoustical Society of America*, vol. 88, pp. 2546, 1990.

[5] A. Simpson, *et al.*, "Spectral enhancement to improve the intelligibility of speech in noise for hearing-impaired listeners," *Acta oto-laryngologica. Supplementum*, vol. 469, pp. 101, 1990.

[6] M. Stone and B. Moore, "Spectral feature enhancement for people with sensorineural hearing impairment: effects on speech intelligibility and quality," *Journal of rehabilitation research and development*, vol. 29, pp. 39, 1992.

[7] T. Baer, *et al.*, "Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: effects on intelligibility, quality, and response times," *Journal of rehabilitation research and development*, vol. 30, pp.

49, 1993.

[8] B.A.M. Franck, *et al.*, "Evaluation of spectral enhancement in hearing aids, combined with phonemic compression," *The Journal of the Acoustical Society of America*, vol. 106, pp. 1452, 1999.

[9] J. Lyzenga, *et al.*, "A speech enhancement scheme incorporating spectral expansion evaluated with simulated loss of frequency selectivity," *The Journal of the Acoustical Society of America*, vol. 112, pp. 1145, 2002.

[10] J. Yang, *et al.*, "Spectral contrast enhancement: Algorithms and comparisons\* 1," *Speech Communication*, vol. 39, pp. 33-46, 2003.

[11] L. Turicchia and R. Sarpeshkar, "A bio-inspired companding strategy for spectral enhancement," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, pp. 243-253, 2005.

[12] A. Bhattacharya and F.G. Zeng, "Companding to improve cochlear-implant speech recognition in speech-shaped noise," *The Journal of the Acoustical Society of America*, vol. 122, pp. 1079, 2007.

[13] J.M. Kates, "Two-tone suppression in a cochlear model," *Speech and Audio Processing, IEEE Transactions on*, vol. 3, pp. 396-406, 1995.

[14] P.M. Zurek and J.G. Desloge, "Hearing loss and prosthesis simulation in audiology," *The Hearing Journal*, vol. 60, pp. 32, 2007.

[15] H. Dillon, "NAL-NL1: A new prescriptive fitting procedure for non-linear hearing aids," *Hearing Journal*, vol. 52, pp. 10-16, 1999.

[16] R. Carhart and J.F. Jerger, "Preferred method for clinical determination of pure-tone thresholds," *Journal of Speech and Hearing Disorders*, vol. 24, pp. 330, 1959.

[17] J. Claerbout, "Fundamentals of geophysical data processing," *Geophysical Journal of the Royal Astronomical Society*, vol. 86, pp. 217-219, 1986.

[18] L. Marple Jr, "Computing the discrete-time analytic signal via FFT," *Signal Processing, IEEE Transactions on*, vol. 47, pp. 2600-2603, 2002.

[19] E. Rothauser, *et al.*, "IEEE recommended practice for speech quality measurements," *IEEE Transactions on Audio Electroacoustics*, vol. 17, pp. 227-246, 1969.

[20] D. Byrne, *et al.*, "An international comparison of long-term average speech spectra," *Journal of the Acoustical Society of America*, vol. 96, pp. 2108-2120, 1994.

[21] J.H. Lee, *et al.*, *Korean speech audiometry*, 1st ed. Seoul: Hakjisa Publisher, 2010.

[22] A.M. Kondoz, *Digital speech: coding for low bit rate communication systems*: John Wiley & Sons Inc, 2004, pp. 65-67.

[23] K.W. BERGER, "Prescription of hearing aids: A rationale," *Ear and hearing*, vol. 2, pp. 71, 1976.

[24] G. McCandless and P. Lyregaard, "Prescription of gain/output (POGO) for hearing aids," *Hearing Instruments*, vol. 34, pp. 16-21, 1983.

[25] E. Libby, "The 1/3-2/3 insertion gain hearing aid selection guide," *Hearing Instruments*, vol. 37, pp. 27-28, 1986.

[26] D. Byrne and H. Dillon, "The National Acoustic Laboratories(NAL) new procedure for selecting the gain and frequency response of a hearing aid," *Ear and hearing*, vol. 7, pp. 257, 1986.

[27] R. Seewald and M. Ross, "Amplification for young hearing-impaired children," *Amplification for the hearing-impaired*,

- pp. 213-271, 1988.
- [28] R. Seewald, *et al.*, "Computer-assisted implementation of the desired sensation level method for electroacoustic selection and fitting in children: Version 3.1, user's manual," *Hearing Health Care Research Unit Technical Report*, vol. 2, 1993.
- [29] ANSI, "ANSI S3. 5-1997, Methods for the calculation of the speech intelligibility index," ed: American National Standards Institute New York, 1997.
- [30] D. Preves, "Output limiting and speech enhancement," *The Vanderbilt hearing-aid report II*, pp. 35-51, 1991.
- [31] Y. Hu and P.C. Loizou, "Evaluation of objective quality measures for speech enhancement," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 16, pp. 229-238, 2008.
- [32] T.N.K. Raju, "William Sealy Gosset and William A. Silverman: Two "Students" of Science," *Pediatrics*, vol. 116, pp. 732, 2005.
- [33] P.C. Loizou and O. Poroy, "Minimum spectral contrast needed for vowel identification by normal hearing and cochlear implant listeners," *The Journal of the Acoustical Society of America*, vol. 110, pp. 1619, 2001.