

논문 2008-45SC-5-3

# 디지털 보청기에서의 포먼트 강조에 의한 마스킹 효과 연구 ( A Study of Acoustic Masking Effect from Formant Enhancement in Digital Hearing Aid )

전 유 용\*, 길 세 기\*\*, 윤 광 섭\*\*\*, 이 상 민\*\*\*\*

( Yuyong Jeon, Se-kee Kil, Kwang-sub Yoon, and Sangmin Lee )

## 요 약

청력 손실을 보상하고 난청인이 다른 사람들과 대화할 수 있도록 디지털 보청기 알고리즘은 개발 되어 왔음에도 불구하고, 디지털 보청기 사용자들은 음성을 듣는데 어려움이 있다고 불만을 토로한다. 그 이유는 피드백이나 잔여 노이즈 등에 의해 디지털 보청기를 통한 음성의 질이 이해하기에 불충분하기 때문이다. 또 다른 이유로 포먼트들 사이에서 일어나는 마스킹 현상이 될 수 있다. 이 연구에서 정상 청각 피험자와 노인성 난청을 갖고 있는 난청인 피험자의 마스킹 특성을 측정하여 음성에서의 마스킹에 의한 음성 인지 저하를 확인하기 위한 실험을 하였다. 실험은 순음검사, 어음 청취 역치 검사, 낱말 분별력 검사, 순음 마스킹 검사, 어음 마스킹 검사의 5개 테스트로 이루어졌다. 어음 마스킹 검사에서 각각 어음 세트에 25개의 어음이 사용되었다. 각 어음의 왜곡을 객관적으로 평가하기 위해서 log likelihood ratio (LLR)를 도입하였다. 결과적으로 포먼트 향상의 양을 늘리면 늘릴수록 어음 인지는 낮아졌고, 각 어음 세트에서 각각의 향상된 어음은 통계적으로 비슷한 LLR을 갖지만 어음 인지는 그렇지 않게 나타났다. 이것은 왜곡이 아닌 음향 마스킹이 어음 인지에 영향을 준다는 것을 의미한다. 실제로 피험자들 대부분이 맞추지 못한 음성을 주파수 분석한 결과 첫 번째와 두 번째 포먼트 사이의 레벨 차이가 약 35dB이며 이는 순음 마스킹 실험 결과(정상 청각 피험자:36.36dB, 난청인 피험자:32.86dB)와 비슷한 양상을 보였다. 실험 결과에서 볼 수 있듯이 음향 마스킹의 특성은 정상 청각인과 난청인 사이에서 다르게 나타난다. 그렇기 때문에 보청기 착용 전 마스킹 특성을 검사하고, 피팅 시에 적용해야 한다.

## Abstract

Although digital hearing aid algorithms have been developed to compensate hearing loss and to help hearing impaired people to communicate with others, digital hearing aid user still complain about difficulty of hearing the speech. The reason could be the quality of speech through digital hearing aid is insufficient to understand the speech caused by feedback, residual noise and etc. And another thing is masking effect among formants that makes sound quality low. In this study, we measured the masking characteristics of normal listeners and hearing impaired listeners having presbycusis to confirm masking effect in speech itself. The experiment is composed of 5 tests; pure tone test, speech reception threshold (SRT) test, word recognition score (WRS) test, puretone masking test and speech masking test. In speech masking test, there are 25 speeches in each speech set. And log likelihood ratio (LLR) is introduced to evaluate the distortion of each speech objectively. As a result, the speech perception became lower by increasing the quantity of formant enhancement. And each enhanced speech in a speech set has statistically similar LLR, however speech perception is not. It means that acoustic masking effect rather than distortion influences speech perception. In actuality, according to the result of frequency analysis of the speech that people can not answer correctly, level difference between first formant and second formant is about 35dB, and it is similar to result of pure tone masking test(normal hearing subject:36.36dB, hearing impaired subject:32.86dB). Characteristics of masking effect is not similar between normal listeners and hearing impaired listeners. So it is required to check the characteristics of masking effect before wearing a hearing aid and to apply this characteristics to fitting.

**Keywords :** hearing aid, acoustic masking, formant enhancement

\* 학생회원, 인하대학교 전자공학과 (Electronic Engineering, Inha University)

\*\* 정회원, 인하대학교 전자공학과 정보전자공동연구소 (Institute for Information and Electronics Research)

\*\*\* 정회원, \*\*\*\* 정회원-교신저자, 인하대학교 전자공학부 정보전자공동연구소

(Department of Electrical and Computer Engineering, Inha University)

\* This work was supported by the Korea Science and Engineering Foundation(KOSEF) grant funded by the Korea Government(R01-2007-000-10801-0) and by the Korea Health 21 R&D Project, Ministry of Health & Welfare, Korea (02-PJ3-PG6-EV10-0001)

접수일자: 2007년11월20일, 수정완료일: 2008년9월8일

## I. 서론

청력 손실을 보상하기 위해서 많은 디지털 보청기 알고리즘이 개발되어 왔음에도 불구하고 디지털 보청기 사용자들은 여전히 말소리를 알아듣기 힘들다고 불평한다. 이러한 문제는 보청기가 소리 증폭을 통해 보청기 착용자에게 큰 소리를 제공하는 과정에서 왜곡이나 피드백 혹은 잡음 제거 미흡 등의 원인 때문에 소리의 질이 좋지 않게 될 수 있기 때문에 발생한다. 이러한 좋지 않은 소리의 질은 소리에 대한 인지를 안 좋게 하여 난청인들의 음성에 의한 의사소통을 오히려 방해한다.

왜곡이나 피드백 외에 소리 인지를 나쁘게 만드는 요인 중의 하나는 음성의 포먼트(formant) 사이에서 발생하는 음향 마스킹(acoustic masking)이 될 수 있다<sup>[1]</sup>. 보청기에서 대부분의 음성 강조(speech enhancement)는 잡음 제거 알고리즘과 포먼트의 대비를 높이는 것으로 이루어진다. 포먼트의 대비를 높이기 위해서 포먼트를 찾아서 포먼트 향상(formant enhancement)을 한다. 포먼트 향상 과정에서 각 음성의 포먼트에 비중을 더 많이 주게 되고, 향상된 포먼트에 의해서 다른 포먼트나 다른 요소들이 마스킹 될 수 있다. 이렇게 되면 음성 인지를 높이기 위한 포먼트 향상으로 인해서 음성의 인지는 더 나빠질 수 있다.

음향 마스킹이란 주변 잡음으로 인해서 소리의 역치(threshold)가 변하는 현상이다<sup>[2]</sup>. 즉 어떤 소리에 의해 다른 소리가 들리지 않는 현상을 말한다. 마스킹은 크게 두 가지로 나눌 수 있는데 하나는 비동시적 마스킹(non-simultaneous masking)이고 다른 하나는 동시적 마스킹(simultaneous masking)이다. 비동시적 마스킹은 시간적으로 듣고자 하는 소리 이전 혹은 이후에 또 다른 소리가 존재 할 때 듣고자 하는 소리가 들리지 않게 되는 현상을 말한다. 동시적 마스킹은 다른 소리의 존재로 그와 동시에 존재하는 듣고자 하는 소리가 들리지 않게 되는 현상을 말한다. 여기서 우리가 듣고자 하는 소리를 프로브(probe)라고 하고 듣고자 하는 소리를 들을 수 없도록 하는 소리를 마스커(masker)라고 한다. Egan 등은 간단한 청각 자극의 마스킹 패턴을 측정했다<sup>[3]</sup>. 그들은 중심 주파수가 고정된 협대역 잡음(narrowband noise)을 마스커(masker)로 사용하고, 순음(pure tone)을 프로브(probe)로 사용했다. 그들의 결과는 주파수 영역에서 상대적으로 저주파가 고주파를 마스킹하는 상향 마스킹(upward masking)이 고주파가 저주파를 마스킹 하는 하향 마스킹(downward

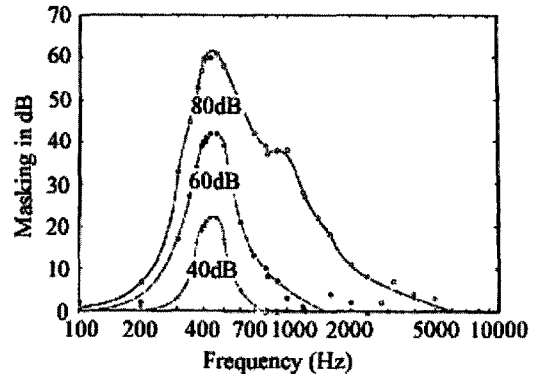


그림 1. 협대역 잡음 마스커(narrowband noise masker)와 순음 프로브(pure tone probe)를 이용한 마스킹 패턴

Fig. 1. Masking pattern using narrowband noise masker and pure tone probe.

masking) 보다 넓다는 것을 보였다. 그림 1은 그들의 실험 결과를 나타내는 그림으로 각 피험자 데이터를 평균한 값을 그래프로 그린 것이다. 그리고 Ehmer는 중심 주파수가 다른 순음을 마스커와 프로브로 사용해서 마스킹 패턴을 측정했다<sup>[4]</sup>. 각 주파수의 마스커에 의한 마스킹 패턴은 유사한 모양이 아니고 고주파 쪽으로 비대칭적으로 확장했다는 것을 확인했다.

Egan, Ehmer 외에도 마스킹에 대해서 많은 연구가 진행되면서, 여러 자극 및 여러 피험자를 대상으로 하는 마스킹 연구<sup>[5-8]</sup>, 마스킹에 영향을 줄 수 있는 요인에 대한 연구<sup>[9-10]</sup>, 마스킹을 이용하여 잡음을 제거 하거나 음성의 질을 높이는 연구<sup>[11-13]</sup> 등의 연구가 진행되어 왔지만 음성 강조, 특히 포먼트 강조에 의해 발생할 수 있는 마스킹에 대한 연구는 거의 찾아 볼 수 없다. 이 연구에서 2 음절 단어를 이용해서 음성에서 발생하는 마스킹을 측정했다. 또한 음성에서의 마스킹과 비교하여 마스킹을 정량화하고 파라미터화 시키기 위해서 순음과 협대역 잡음을 이용한 마스킹을 확인했다.

## II. 음성 포먼트 강조 방법

음성 강조를 위해 그림 2와 같은 알고리즘을 사용했다. 포먼트를 찾기 위해서 15차 선형 예측 부호화(linear predictive coding, LPC) 알고리즘을 사용했다. LPC로 음성의 포락선(envelope)를 찾으면 포먼트라 불리는 몇 개의 피크가 나타나고 이 포먼트에 가중치(Q)를 주기 위해서 변형된 해밍 윈도우(hamming window)를 사용하는 방법은 제안했다. 이는 포먼트의 주파수와 같은 중심주파수를 갖는 일정한 너비의 해밍 윈도우를 만들

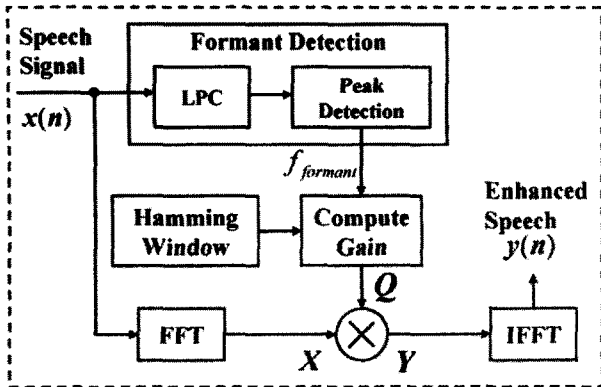


그림 2. 음성 강조(speech enhancement) 방법  
Fig. 2. Speech enhancement method.

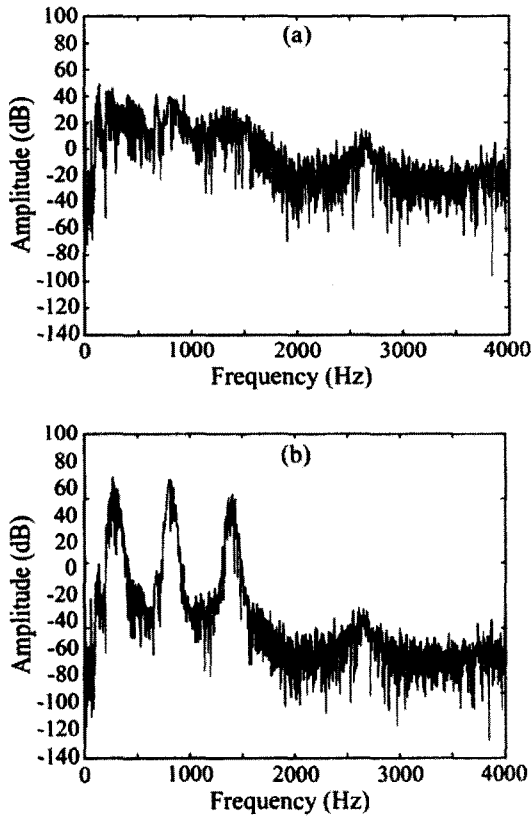


그림 3. 원본 음성의 주파수 분석 결과(a)와 포먼트 강조된 음성의 주파수 분석 결과(b)  
Fig. 3. FFT of original speech(a) and FFT of a formant enhanced speech (b).

고, 그 윈도우의 크기를 변화 시켜 포먼트에 주는 가중치를 조절하는 방법으로 포먼트를 강조시켰다. 식 1은 다음과 같은 과정을 식으로 나타낸 것이다.

$$Q(k) = G(h(N, FF) + 1)^w \quad (1)$$

G는 전체 이득(gain)을 조절하는 상수이고, h는 해밍 윈도우, N은 윈도우의 크기, FF는 포먼트의 주파수, k는 주파수의 색인(index), w는 가중 인자(weight

factor)이다. 포먼트라고 생각되는 부분에는 가중치를 주고 그렇지 않은 부분은 전체 이득을 조절해서 크기를 작게 만들었다. 주파수 영역에서 포먼트에 가중치를 주기 위해서 FFT(Fast Fourier Transform)를 사용했다. n이 이산신호 색인이고 x(n)을 음성이라고 하자. FFT는 다음과 같이 정의 된다.

$$X(k) = \sum_{n=0}^{l-1} x(n)e^{-j(2\pi/l)nk} \quad (2)$$

l은 음성의 길이를 나타낸다.

FFT 이후 포먼트 검색 알고리즘에 의해서 찾은 포먼트에 가중치를 주고 IFFT를 했다.

$$y(n) = IFFT[Y(k)] = IFFT[Q(k)X(k)] \quad (3)$$

그림 3은 원본 음성 “아들”의 주파수 분석 결과(a)와 그림 2에서 설명한 알고리즘에 의해서 포먼트를 향상시킨 음성의 주파수 분석 결과(b)를 나타낸다. 가로축은 주파수 영역을 나타내고, 세로축은 각 주파수에서의 크기를 데시벨(dB) 단위로 나타낸 것이다.

### III. 실험 조건

앞에서 설명한 포먼트 향상 알고리즘으로 처리한 2음절 음성이 실험에 사용되었다. 사용된 음성은 4 세트 로 구성되고, 각 세트에는 25개 음성으로 구성되었다. 각 세트의 음성에서 포먼트를 조절하기 위해서 각 세트에서 포먼트에 다른 가중치를 적용했다. 식 (1)에서 가중 인자(w)를 1번 세트부터 7, 6, 5, 4로 각각 주었고, 결국 100개의 음성을 사용했다. 음성은 병원에서 어음 청취 역치(speech reception threshold, SRT)에 쓰이는 강강격 단어(spondee word)를 사용하여 8kHz로 샘플링 했다.

표 1은 실험에 참가한 피험자에 대한 정보를 나타낸다. 실험에는 정상 청각 피험자 22명(여자 12명, 남자

표 1. 실험에 참가한 피험자 정보  
Table 1. Information about subjects.

	정상 청각 피험자		난청 피험자	
평균연령	23.2 세		72.6 세	
증상	정상 청각		노인성 난청	
숫자	남자	여자	남자	여자
	10 명	12 명	7 명	14 명

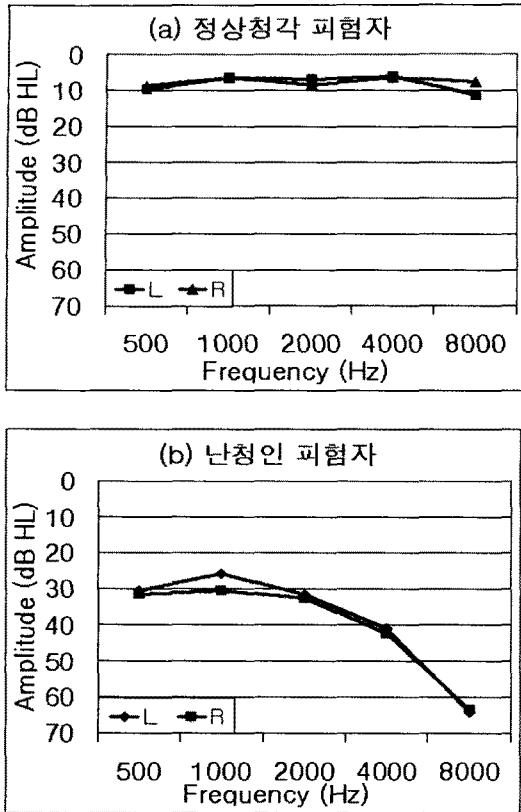


그림 4. 정상 청각 피험자와 난청 피험자의 주파수별 순음 검사 결과  
 Fig. 4. Result of pure tone test of normal listeners and hearing impaired listeners.

10명), 난청 피험자 21명(여자 14명, 남자 7명)이 참가하였고, 정상 피험자의 평균 연령은 약 23.2세였고 난청 피험자의 평균 연령은 약 72.6세였다. 그림 4는 순음 검사를 통해서 정상 청각 피험자에서와 노인성 난청을 갖고 있는 피험자에서의 각 주파수별 평균 청력을 나타낸 것이다. 이 그림을 보면 고주파에서 난청 피험자들의 청력이 급격히 감소하는 것을 볼 수 있는데 이로서 전형적인 노인성 난청이라는 것을 알 수 있다. 반면 정상 청각 피험자들은 청력 이상이 없었으며 피험자 스스로 질병이 없었다고 대답했다.

IV. 실험 과정 및 결과

1. 실험 과정

음성에서의 마스킹 현상을 확인하기 위해서, 그림 5와 같은 절차로 실험을 하였다. 먼저 각 피험자의 청각 특성을 확인하기 위해서 5dB 단위로 순음 검사를 실시하였다. 순음 검사 결과 정상 청각 피험자들은 모두 정상 청각을 갖고 있다는 것을 확인하였고, 난청 피험자

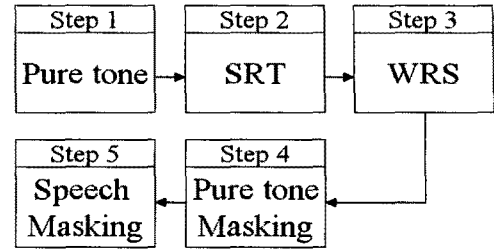


그림 5. 실험 절차  
 Fig. 5. Procedure of experiment.

는 고주파에서 특히 청각 특성이 좋지 않은 전형적인 노인성 난청을 갖고 있는 것이 확인 되었다. 순음 검사 이후 정답률 50%를 기준으로 어음 청취 역치(speech reception threshold, SRT) 검사를 실시하였고, 어음 청취 역치에서 30~40dB를 높인 소리 크기를 쾌적 역치(comfortable level)라고 임의로 정하고 피험자에게 소리 크기가 적당한지 물어보는 과정을 반복해서 쾌적 역치로 정했고, 쾌적 역치에서 1음절 단어를 이용해서 낱말 분별력(word recognition score, WRS) 검사를 실시했다. WRS 검사 시작 후 5개의 2음절 단어를 들려주고 소리의 크기가 듣기 편한지 다시 한 번 피험자에게 물어 쾌적 역치를 다시 한 번 정하였다. Cool Edit Pro 2.0으로 만든 협대역 잡음 마스커(narrowband noise masker)와 순음 프로브(puretone probe)를 사용해서 마스킹 현상이 얼마나 발생하는지 알아보는 순음 마스킹 검사(pure tone masking test)를 했다. 이 검사에서 두 음을 들려주는데, 한 음은 마스커만 출력한 소리이고 다른 한 음은 마스커와 프로브를 동시에 들려주는 소리이다. 두 소리를 한 세트로 해서 피험자에게 들려주고 두 소리가 같은지 다른지를 판단하도록 한다. 만약 피험자가 이 두 소리가 같다고 느낀다면 프로브가 들리지 않은 상태, 즉 마스커에 의해 프로브가 마스킹 된 상태라고 볼 수 있다. 순음 마스킹 검사에서 사용한 협대역 잡음은 중심 주파수가 250Hz, 770Hz이고, 사용한 순음 프로브는 770Hz, 2000Hz이다. 즉, 실험 세트는 250Hz 마스커와 770Hz, 2000Hz 프로브, 750Hz 마스커와 2000Hz 프로브로 구성되었다. 순음 마스킹에서 사용한 프로브와 마스커의 중심 주파수는 사용한 음성의 포먼트 주파수를 평균해서 결정하였다.

2. 기초 검사 결과

그림 4는 실험에 참여한 피험자 개개인의 순음 검사 결과를 정상 청각 피험자와 난청 피험자로 나누어 각각 평균한 결과이다. 그림 4에 따르면 일반 청각 피험자들

표 2. SRT 테스트, WRS 테스트의 결과  
Table 2. Result of SRT and WRS test.

(a) 정상 청각 피험자

SRT(dB HL)	CL(dB HL)	WRS(%)
6.625	45.125	97.1

(b) 난청 피험자

SRT(dB HL)	CL(dB HL)	WRS(%)
30.715	69.534	93.429

은 정상 청각을 유지하고 있으며 난청 피험자들은 전형적인 노인성 난청을 나타내고 있다.

표 2는 SRT와 WRS를 나타낸 것이다. 어음 청취 역치는 2음절 단어 4개 중에 반 이상 들을 수 있는 최소 소리 크기를 말한다. 낱말 분별력은 쾌적 역치에서 1음절 단어를 얼마나 알아듣는지를 나타내는 것이다. 어음 청취 역치와 낱말 쾌적 역치는 정상 청각 피험자와 난청 피험자 사이에 큰 차이를 보이지만 낱말 분별력은 큰 차이를 보이지 않았다. 이는 음성의 크기를 쾌적 역치에 맞춰 준다면 음성 인지(speech perception)에는 큰 차이가 없다는 것을 의미한다.

3. 객관적인 평가

음성의 질을 객관적으로 평가하기 위해서, log likelihood ratio(LLR)을 사용했다<sup>[14]</sup>. LLR은 원신호와 향상된 신호의 포먼트 사이의 불일치를 스펙트럼에서 측정한다. LLR 측정은 다음과 같이 계산된다.

$$LLR = \log_{10} \left( \frac{a_y R_x a_y^T}{a_x R_x a_x^T} \right) \quad (4)$$

$a_x$ 는 깨끗한 음성 프레임의 선형 예측 계수 벡터이

표 3. 각 세트 LLR의 최대값, 최소값, 평균값, 표준 편차

Table 3. Maximum, minimum, average and standard deviation of LLR in each set.

	세트 1	세트 2	세트 3	세트 4
평균	1.1205	0.9681	0.7873	0.5792
최대값	1.4138	1.1867	0.9532	0.6697
최소값	0.7626	0.7074	0.6137	0.4912
표준 편차	0.1634	0.1254	0.0863	0.0462

고,  $a_y$ 는 향상된 음성 프레임의 선형 예측 계수 벡터이고,  $R_x$ 는 깨끗한 음성 프레임의 자기상관(autocorrelation) 행렬이다. 낮은 LLR은 더 적은 왜곡을 의미하며, 더 나은 음성의 질을 의미한다. 표 3은 각 세트 LLR의 최대값, 최소값, 평균값, 분산을 나타낸다.

4. 주관적인 평가

표 4는 순음 마스크의 결과이다. 첫 번째 행은 마스크의 주파수를 나타내고 두 번째 행은 프로브의 주파수를, 세 번째 행은 마스크의 레벨과 프로브의 크기 차이를 나타낸다. 크기 차이가 많이 날수록 마스크가 적게 일어난다는 것을 의미한다. 표 4에서 보면 250Hz 마스크에 의한 770Hz 프로브의 마스크가 가장 잘 일어난다는 것을 확인할 수 있다. 즉 첫 번째 포먼트에 의한 두 번째 포먼트로의 마스크가 가장 쉽게 일어난다는 것이다. 또한 첫 번째 포먼트에서 세 번째 포먼트로의 마스크와 두 번째 포먼트에서 세 번째 포먼트로의 마스크는 덜 일어난다는 것을 확인할 수 있다.

순음 마스크 검사 이후, 위에서 설명한 음성 강조 방법을 통해 포먼트를 향상시킨 음성으로 마스크 실험을 했다. 이 실험에서는 피험자에게 음성을 듣고 따라하도록 하여 각 음성과 세트에 대한 전체 피험자의 정답률 평균을 구했다.

표 5는 일반 청각 피험자와 난청 피험자로부터 포먼트가 향상된 음성을 들려주고 따라하도록 해서 얻은 어음 마스크 테스트의 결과이다. 표 3에 따르면, 음성에 더 많은 가중치를 주면 정답률은 낮아진다. 정답률이 낮다는 것은 음성의 지각이 어렵다는 것을 뜻한다. 표

표 4. 피험자들의 순음 마스크 결과  
Table 4. Result of pure tone masking test.

(a) 정상 청각 피험자

Masker Frequency	250Hz		770Hz
Probe Frequency	770Hz	2000Hz	2000Hz
Level Difference(dB)	36.36 ± 5.73	53.86 ± 6.31	49.09 ± 4.20

(b) 난청 피험자

Masker Frequency	250Hz		770Hz
Probe Frequency	770Hz	2000Hz	2000Hz
Level Difference(dB)	32.86 ± 4.30	46.43 ± 6.13	36.19 ± 7.00

표 5. 어음 마스킹 실험 결과  
Table 5. Result of speech masking test.

(a) 정상 청각 피험자의 어음 마스킹 결과

가중 인자 (w)	Correct Response Percentage(%)		
	남자	여자	전체
4	86.80 ± 8.26	87.67 ± 8.68	87.27 ± 8.17
5	74.40 ± 14.22	77.00 ± 15.28	75.82 ± 14.58
6	68.40 ± 14.41	72.67 ± 13.90	70.73 ± 13.81
7	55.60 ± 17.20	57.33 ± 15.32	56.55 ± 15.35

(b) 난청 피험자의 어음 마스킹 결과

가중 인자 (w)	Correct Response Percentage(%)		
	남자	여자	전체
4	73.71 ± 13.31	75.71 ± 13.44	75.05 ± 12.69
5	60.00 ± 13.36	62.57 ± 13.77	61.71 ± 12.68
6	49.71 ± 15.58	46.86 ± 14.62	47.81 ± 14.17
7	29.71 ± 12.36	29.43 ± 15.02	29.52 ± 13.25

3과 표 5의 결과를 종합해 볼 때 정답률이 낮아진 것은 음성의 질이 안 좋기 때문이라고 생각 할 수 있다. 하지만 표 3의 값은 한 세트 음성의 평균 LLR을 나타낸 값이고, 각각 음성의 LLR을 보면 LLR 낮다고 해서 그 음성의 정답률이 높아지는 양상을 보이진 않는다.

그림 6은 네 번째 세트에서의 각 음성에 대한 피험자들의 정답률과 LLR 값을 나타낸 것이다. 만약 음성 인지에 영향을 미치는 요소가 오직 왜곡이라고 가정했을 경우, LLR이 커지면, 즉 왜곡이 커지면 정답률, 즉 음성

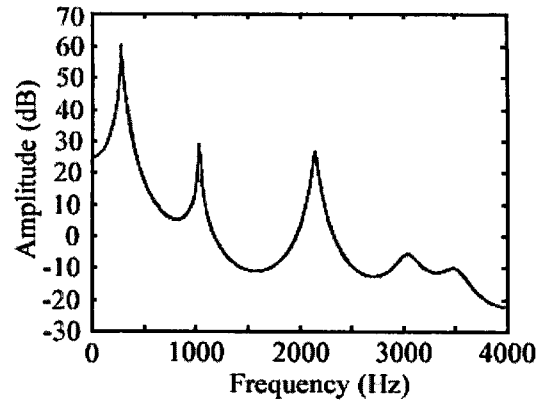


그림 7. 마스킹이 일어나는 단어의 예  
Fig. 7. Speech sample that masking is occurred.

인지율이 낮아져야 한다. 하지만 정상 청각 피험자나 난청인 피험자 모두 왜곡과는 상관관계를 보이지 않는 정답률을 보였다. 이러한 경향은 모든 음성 세트에서 볼 수 있으며 이로서 왜곡에 의한 영향은 적다고 말할 수 있다.

그림 7은 거의 대부분의 피험자들이 알아듣지 못한 단어(오리)의 주파수 분석 결과의 포락선을 나타낸 것이다. 실제 주파수 분석 결과를 보면 첫 번째 포먼트와 두 번째 포먼트 사이의 크기 차이가 약 35dB 정도라는 것을 볼 수 있다.

### V. 결 론

이 논문에서, 디지털 보청기에 의해서 강조된 음성에서 포먼트에 의해서 마스킹이 일어나고, 그로 인해서 음성 인지에 어려움이 생길 수 있다는 것을 말하기 위

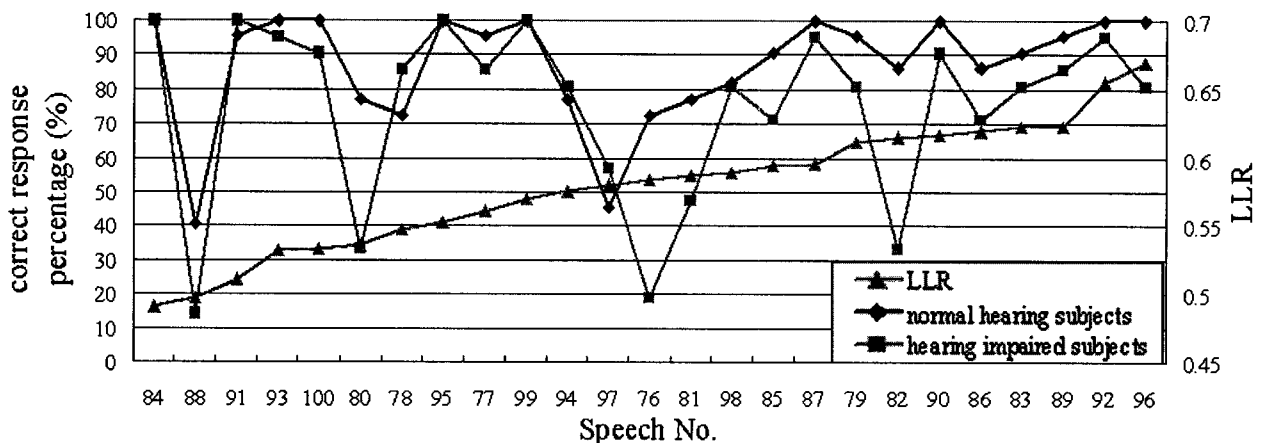


그림 6. 한 음성 세트에서 각 음성의 정답률과 LLR  
Fig. 6. Correct response percentage and LLR of speeches in one set.

해서 음성을 강조하여 인지율이 떨어지는 것을 보였다. 표 2에서 보면 난청인 피험자와 정상 청각 피험자 사이에서 낱말 분별력이 크게 다르지 않다는 것을 알 수 있다. 하지만 표 5에서 보면 난청인 피험자와 정상 청각 피험자 사이에서 많은 언어 인지율의 차이를 보인다. 또한 그림 6에 따르면 한 세트의 음성에서 왜곡이 작다고 해서 언어를 잘 인지하는 것이 아니라는 것을 알 수 있다.

순음 마스크링 검사와 어음 마스크링 검사로 디지털 보청기에서 포먼트에 의한 마스크링이 생길 수 있다는 것을 간접적으로 설명할 수 있다. 그림 7은 어음 마스크링 실험에서 피험자들의 대부분이 정답을 맞지 못하는 음성을 주파수 분석을 하고 선형 예측 부호화 알고리즘을 통한 결과이다. 첫 번째 포먼트(약 250Hz)와 두 번째 포먼트(약 770Hz)의 차이를 보면 약 35dB 정도의 차이가 난다는 것을 알 수 있고 이것은 순음 마스크링 검사의 결과와 거의 일치하는 값이라고 볼 수 있다. 이로서 순음 마스크링 검사 결과와 어음 마스크링 검사 결과가 연관성이 있다고 생각할 수 있다. 이로서 왜곡의 영향을 받지 않는 것은 아니지만 왜곡만으로 이런 결과가 만들어진 것이 아니라는 것을 알 수 있다. 즉 음성을 이해하는데 왜곡 외에 마스크링 현상의 영향을 받는다는 것을 의미한다.

이러한 마스크링 현상을 최소화 할 수 있는 첫 번째 방안으로 포먼트 향상 과정 중 각 포먼트에 일정한 가중치를 주기보다는 첫 번째 포먼트에 가중치를 덜 주는 방법이 있다. H. Timothy Bunnell은 마스크링이 일어나는 것을 방지하기 위해서 주파수 축에 해밍 윈도우로 저주파와 고주파에 덜 가중치를 주어 마스크링을 막고자 했다<sup>[15]</sup>. 두 번째 방안은 보청기 피팅 시에 주파수에 따른 보상 외에 마스크링을 유념한 피팅을 개별적으로 시행하는 것이다. 저주파에서 증폭도가 고주파에서의 증폭도보다 약간 크거나, 고주파에서 증폭도가 저주파에서 증폭도보다 과도하게 크게 되는 경우에 마스크링이 일어날 것으로 생각되기 때문에, 개인의 마스크링 특성을 이용하여 이득(gain)을 조절하는 것이 마스크링 억제에 도움이 될 것으로 생각된다.

## 참고 문헌

[1] Danaher, E. M., and Pickett, I. M., "Some masking effects produced by low-frequency vowel formants in persons with sensorineural

hearing loss," *J. Speech Hear. Res.*, 18, pp. 261-271, 1975.

- [2] Irwin Pollack, "Auditory informational masking", *The Journal of the Acoustical Society of America*, Volume 57, Issue 51, p.55, April 1975.
- [3] James P. Egan and Harold W. Hake, "On the Masking Pattern of a Simple Auditory Stimulus", *The Journal of the Acoustical Society of America*, Vol.22, no.5, pp.622-630, September 1950.
- [4] Richard H. Ehmer, "Masking Patterns of Tones", *The Journal of the Acoustical Society of America*, Vol.31, no.8, August 1959.
- [5] Hartley D. E. H., Wright B. A. and Moore D.R., "Age-Related Improvements in Auditory Backward and Simultaneous Masking in 6- to 10 -Year-Old Children," *Journal of Speech, Language, and Hearing Research*, Vol.43, pp.2-1415, December 2000.
- [6] Dubno, Judy R, Ahlstrom, Jayne B, "Psychophysical suppression effects for tonal and speech signals," *The Journal of the Acoustical Society of America*, Vol.110, Issue 4, pp.2108-2119, October 2001.
- [7] Hans A. Gustafsson and Stig D. Arlinger, "Masking of speech by amplitude-modulated noise," *The Journal of the Acoustical Society of America*, Vol.95, Issue 1, pp518-529, January 1994.
- [8] Van Summers Michelle R. Molis, "Speech Recognition in Fluctuating and Continuous Maskers," *Journal of Speech, Language, and Hearing Research*, Vol.47, pp245-256, April 2004.
- [9] Alcantara J.I., Moore B.C.J., "The relative role of beats and combination tones in determining the shapes of masking patterns," *Hearing Research*, Vol.165, Number 1, pp103-116, March 2002.
- [10] Ifat Yasin and Christopher J. Plack, "The Role of Suppression in the Upward Spread of Masking," *Journal of the Association for Research in Otolaryngology*, Vol. 6, Number 4, pp368-377, December 2005.
- [11] J. Thiemann, "Acoustic Noise Suppression for Speech Signals using Auditory Masking Effects," *M.S. thesis McGill University*, Montreal, Canada, May 2001.
- [12] Viraq, N., "Single Channel Speech Enhancement Based on Masking Properties of the Human Auditory System," *Speech and Audio Processing, IEEE Transactions on*, Vol. 7, Issue 2, pp126-137, Mar 1999.
- [13] Hesu Huang Kyriakakis, C., "Binaural Noise

Reduction Combining Binaural Analysis and Psychoacoustically Motivated Spectral Subtraction," *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on, IEEE, Vol. 2, pp2260-2264, Nov. 2004.*

[14] Quackenbush, S.R., Barnwell, T.P., Clements, M.A., "Objective Measures of Speech Quality", Prentice-Hall,NJ, pp.47-51, 1988.

[15] H. Timothy Bunnell, "On Enhancement of spectral contrast in speech for hearing impaired listeners", *The Journal of the Acoustical Society of America.* 88(6). December 1990.

저 자 소 개



**전 유 용(정회원)**  
 2006년 전북대학교 생체정보 공학부 학사 졸업.  
 2007년 3월~현재 인하대학교 전자공학과 석사 과정  
 <주관심분야 : 심리음향, 음성 신호처리, 보청기 시스템>



**윤 광 섭(정회원)**  
 1982년 2월 인하대학교 전자공학과 졸업  
 1983년 2월 Georgia Institute Inc, Technology 전자공학 석사 졸업  
 1989년 2월 Georgia Institute Inc, Technology 전자공학 박사 졸업

1984년 3월~1989년 2월 Georgia Institute of Technology Research Assistant  
 1989년 3월~1992년 2월 Silicon Systems Inc, Tustin Calif. U.S.A Senior Design Engineer

1992년 3월~현재 인하대학교 전자공학과 교수  
 <주관심분야 : 음성신호처리 집적회로 설계, 설계 자동화 및 소자/회로/시스템 모델링 등>



**길 세 기(정회원)**  
 1998년 인하대학교 전자공학과 학사 졸업  
 2000년 인하대학교 대학원 전자공학과 석사 졸업  
 2006년 인하대학교 대학원 전자공학과 박사 졸업

2006년 9월~2007년 9월 일본 와세다대학교 대학원 초청연구원

2007년 10월~현재 인하대학교 전자공학과 정보전자공동연구소 선임연구원

<주관심분야 : 생체신호처리, 음성신호처리, 유비쿼터스 헬스케어, 의용영상처리>



**이 상 민(정회원)**  
 1987년 인하대학교 전자공학과 학사 졸업.  
 1989년 인하대학교 전자공학과 석사 졸업  
 2000년 인하대학교 전자공학과 박사 졸업

1989년 1월~1994년 7월 LG이노텍 선임연구원

1995년 1월~2002년 3월 삼성종합기술원 책임 연구원

2002년 4월~2005년 2월 한양대학교 의공학교실 연구교수

2005년 3월~2006년 8월 전북대학교 생체정보공학부 조교수

2006년 9월~현재 인하대학교 전자전기공학부 조교수

<주관심분야 : Healthcare system design, Psycho-acoustic, Brain-machine interface>