

# 기동무기체계에서의 통신을 위한 음성신호 포착 연구

A Study of Voice signal Capture for communication in the AFV

김 석 봉\*  
Kim Suk-Bong

이 성 태\*\*  
Lee Sung-Tae

## ABSTRACT

In the military communication environment, it is very difficult to obtain clear voice signal due to the high level noise. The purpose of this study is to find out the best body spot to get the vocal chords signal by measuring the skin or the bone conducting vibrations of different body positions within the noise environment. Based on the experimental study, it was found out that the measurement of sound signal within the ear is the best way to get the voice which comes from the vocal chords and this method can prevent the interruption of noise. This study will give the effective voice communication method in the high noise environment and be applicable to military purpose.

## 1. 서 론

현대전에서 전장환경이 전쟁의 승패에 미치는 영향은 지대하며, 그 중에서도 음성에 의한 의사소통은 전장환경을 결정하는 중요한 요소 중 하나이다<sup>[1]~[3]</sup>. 음성을 통한 의사소통은 지휘관의 의도대로 전쟁을 수행한다는 측면과 전장에 있는 전투원들이 상부에 현장감 있는 정보를 보고하는 측면 등 여러 측면에서 보장되어야 한다. 이러한 음성에 의한 의사소통은 여러 종류의 소음에 의하여 방해받게 되는데, 특히 전차 승무원의 사격통제, 헬기 조종사의 통신, 보병 지휘라인 상의 통신체계 등은 기동장비의 자체 소음

에 의하여 의사소통에 심한 방해를 받으며, 그 피해의 심각성이 매우 크다.

높은 레벨의 소음이 상존하는 군의 기동장비 내에서의 음성통신의 확보는 주변 소음의 영향을 제거하여야만 가능하며, 이러한 전장환경이 확보되어야만 음성을 통한 지휘 및 통제를 보장받을 수 있다.

본 논문은 이러한 높은 레벨의 소음이 주는 영향으로 인한 의사소통의 제한 요소를 해결하기 위하여 성대에서 발생시키는 음성신호를 신체의 여러 부위에서 포착하여 신호를 획득하고자 하였다. 만일 소음에 의한 영향이 있기 전에 성대의 신호를 신체의 어떤 부위에서 획득이 가능하다면, 어떠한 형태의 소음환경에서도 깨끗한 음질로 상호 의사소통이 가능하여 지휘의 제한요소가 해소될 것이다.

성대의 진동에 의한 음성 신호를 가장 효과 적으로

\* 육군사관학교 토목과

\*\* 육군사관학교 물리과

포착하기 위해서는 인체의 진동 및 음성신호를 여러 지점에서 포착 비교하여 입 앞에서 포착한 음성신호를 기준으로 비교 분석하여 가장 비슷한 음성신호를 획득 가능한 부위 및 방법을 선정하여 활용하는 것이 필요하다.

실험의 목적을 위하여 인체의 여러 부위에서 성대의 신호를 포착하였는데, 측정부위는 그동안의 자료조사에 의하여 성대의 진동인 음성 신호를 잘 포착하는 곳으로 알려진 위치들로 선택하였다. 또한 선택된 부위에서 사용하는 센서도 그 부위에서의 신호 형태에 적합하게 반응하여, 음성 신호를 최대한 효과적으로 포착하는 방법을 선택하였으며, 또한 소음환경을 고려할 때 노출되는 소음에는 민감하지 않도록 하여 소음 환경에서도 활용이 가능한 형태로 선택하였다.

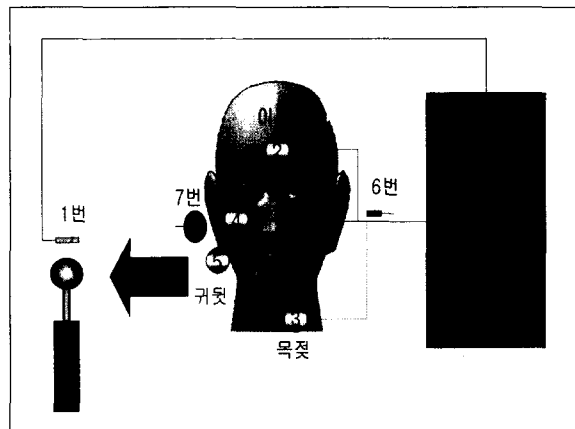
실험의 신뢰도를 높이기 위하여 실험자를 3명 선택하였으며 그 실험자들의 입 앞에서 포착되는 신호를 기준으로 하여 각 부위별로 포착된 음성신호를 비교 분석하였다.

## 2. 실험

먼저 성대의 진동에 의하여 발생하는 음성 신호를 가장 효과적으로 포착하기 위해서는 인체의 여러 위치에서 음성신호를 포착하고 비교하여 입 앞에서 포착한 음성신호와 비교 분석하여야 한다. 이러한 실험을 수행하기 위하여 6곳의 위치가 선정되었으며, 성대 신호의 기준이 되는 입 앞에서의 신호를 포착하는 것을 포함하여 총 7곳에서 음성신호를 동시에 포착한다. 입 앞에서 측정된 신호와 나머지 위치에서의 측정된 신호를 비교하여 어떠한 부분에서 포착한 신호가 감도가 좋고, 신호의 왜곡이 적고 가능한 넓은 주파수

대역에서 정확하게 음성을 재생할 수 있는가를 비교 분석하였다. 물론 본 논문에서 실시한 위치가 아닌 다른 위치에서 다른 종류의 센서를 사용하여 음성신호를 포착하는 것이 가능할 수도 있다. 그러나 그동안의 자료조사<sup>[4]~[7]</sup>와 연구 경험을 바탕으로 판단한 바에 의하면 그러한 가능성은 낮다고 판단하여 7곳의 위치에서 센서를 설치하여 음성신호를 포착하였다. 센서의 선정에 있어서도 종류가 다르고 측정 방법도 다르기 때문에 신호의 직접적인 비교는 어려움이 있으나 신체의 선정된 부위에서는 가장 잘 음성신호를 포착하는 신호포착 방법을 사용하였기 때문에 그 결과만을 상호 비교하여 가장 성능이 좋은 것을 선택하여 활용하고자 하였다.

실험을 실시한 상태를 그림으로 표현하면 아래와 같으며, 이들의 위치와 사용된 센서들을 그림 아래에 다시 설명하였다. 사람이 음성신호를 발생시킬 때 동일한 음성신호를 포착한 결과를 비교 분석할 수 있도록 하였으며 소음계를 사용하여 어느 정도 크기의 음성신호가 발생하였는가를 확인하도록 소음계를 동시에 사용하였다.



[그림 1] 성대의 음성신호를 포착하기 위해 센서를 설치한 위치들

1번 채널은 입 앞에서 컨덴서 마이크로 음성신호를 포착한다. 입에서 방출되는 공기의 유동에 의한 신호를 피하기 위하여 입과 마이크 사이에 충분한 간격을 주었다.

2번 채널은 이마 중앙에 PZT(Piezo-electric) 압전소자를 설치하여 골도 진동을 통한 이마의 진동을 음성신호로 포착할 수 있도록 하였다.

3번 채널은 목의 성대가 위치한 외부 피부에 PZT 압전소자를 설치하여 성대의 기계적 진동을 음성신호로 포착할 수 있도록 하였다.

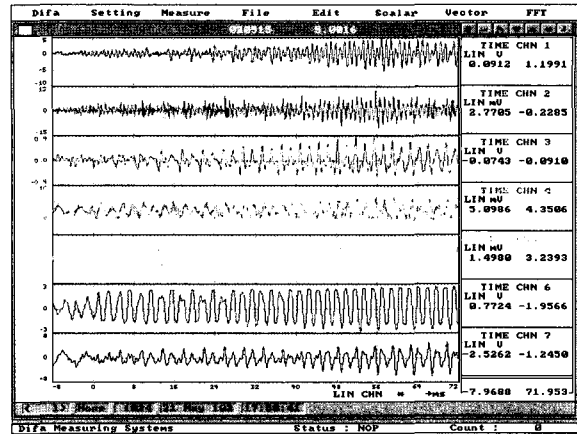
4번 채널은 귀바퀴 바로 앞 뼈가 있는 위치에 PZT 압전소자를 설치하여 골도를 통한 진동을 음성신호로 포착하도록 하였다.

5번 채널은 귀바퀴 바로 뒤 뼈가 있는 위치에 PZT 압전소자를 설치하여 골도를 통한 진동을 음성신호로 포착하도록 하였다.

6번 채널은 귀속 외이도에서 캔드레버 방식의 진동 센서를 이용하여 귀속의 진동 또는 음성신호를 포착하도록 하였다.

7번 채널은 귀속 외이도에서 고막의 진동을 컨덴서 마이크로 음성 신호를 포착하도록 하였다.

이와 같이 위치와 센서를 선정한 이유는, 귀속에서 고막의 진동을 측정하는 것은 공기 진동의 신호를 가장 정확하고 정밀하게 측정하는 컨덴서 마이크 형태가 좋다고 판단하였기 때문이며, 귀속에서 음성과 귀속 뼈의 진동을 동시에 포착하기 위해서는 고감도의 캔트레바 방식이 좋다고 판단하였고, 성대 밖의 피부 진동이나 이마, 귀 주변의 뼈의 진동을 측정하는데는 PZT 압전소자에 의한 진동의 포착이 정확하고 동시에 높은 소음에는 민감하게 반응하지 않는다고 판단되어<sup>[8],[9]</sup> 위에서 언급된 방식들로 음성신호를 동시에 포착한 후 비교 평가하였다.

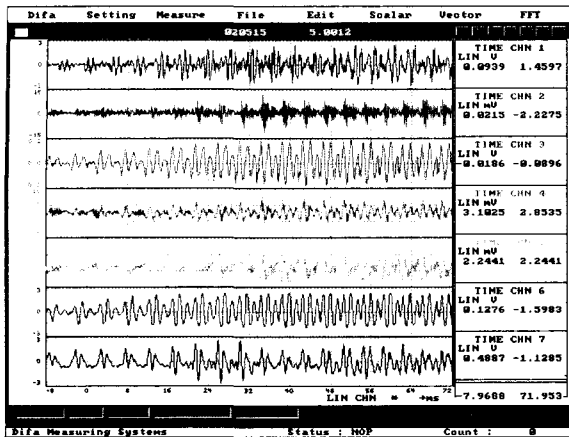


[그림 2] 인체의 각 부위별 신호의 포착 그래프

그림 2는 이러한 실험 결과를 보여주는 대표적인 그래프로 각 채널은 그림에서 표시된 1번에서 7번까지의 위치에서 포착한 신호를 순서대로 보여주고 있다. 그림에서 x 축은 시간을 나타내는 것으로 -8ms에서 72ms 까지 총 80ms 동안의 신호의 크기변화를 보여주는 것으로 y 축은 mV 또는 V를 나타내고 있으며 x축과 y 축은 서로 선형으로 표시한 것이다. 여러 번의 실험을 실시하면서 모두 위에서 결정한 것과 같이 동일한 위치에서 신호를 포착하였으며, 모든 채널은 주파수 분석을 통하여 주파수별 신호 왜곡의 정도와 포착하는 신호의 크기를 비교하여 보았다. 실험 결과는 실험자 3명을 선정하여 각 실험자가 “가, 토, 피, 러”를 발음하게 하여 각 센서에서 포착한 신호를 분석하였다.

### 3. 결과 및 분석

그림 1과 같은 상태에서 포착한 성대 진동에 의한 시간대별 신호를 측정하였으며 주파수 분석을 한 결과를 그림으로 나타내었다. 실험은 3명의 실험자에 대하여 동일하게 실시되었다.



[그림 3] 실험자 1의 '가' 발음의 시간대별 음성신호

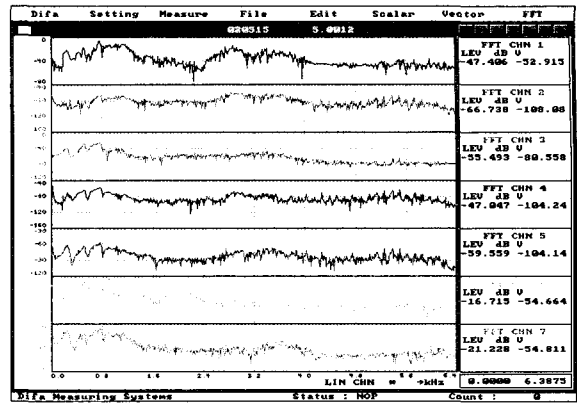
가. 실험자 1의 시간대 신호(1)

시간대 신호의 그림에 대한 설명은 위에서 언급하였으 며, 그림 3에서 볼 수 있는 것과 같이 맨 위에 나타나있는 1번 채널의 음성신호는 입 앞에서 포착된 것으로 음성포착 신호의 기준이 되는 것이며, 그 밑으 로 각각 다른 위치에서 음성 신호를 포착한 것으로 신호의 크기가 다르고, 신호의 형태도 약간씩 다른 것 을 확인할 수가 있다. 만일 모든 위치에서 포착한 신 호가 동일하다면 모든 채널에서 포착한 신호가 정확 히 일치할 것이다. 신호의 일치 정도를 주파수 별로 정확하게 알기 위해서는 주파수 분석을 하여야 하며 그 결과는 아래와 같다.

나. 실험자 1의 '가' 발음에 대한 주파수 분석 결과(1)

그림 4와 같은 주파수 분석 결과는 x 축이 주파수 를 나타내는 것으로 0에서 6.4kHz 까지를 나타내고 y 축은 V를 dB 값으로 나타낸 것으로 1 V를 기준값으 로 나타낸 것이다. 이와 같은 것은 모든 주파수 분석 결과를 나타내는 그림에서 동일하다.

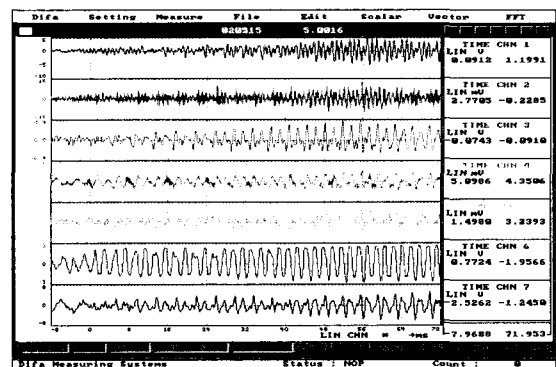
맨 위의 채널을 기준으로 볼 때, 2번째의 2번 채널



[그림 4] 실험자 1의 '가' 발음 음성신호에 대한 주파수 분석

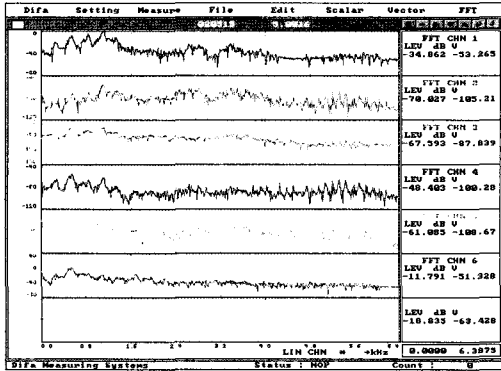
은 0.8 kHz 대역이 1번보다 신호가 작게 나타나 있 다. 이것은 이마에서 포착한 방식이 저주파 대역에서 음성을 제대로 포착하지 못하고 저주파 음성왜곡이 발생한 것을 보여주고 있으며, 3번 채널의 경우는 0.8 kHz 부근까지는 1번 채널과 유사한 신호 형태를 보 여주다가 그 이상에서는 신호를 아주 작게 포착하고 있음을 보여주고 있다. 특히 1번 채널과 비교할 때 2 ~3 kHz 부근의 신호를 거의 포착하고 있지 못하는 것을 볼 수 있으며 고주파 대역에서 음성왜곡이 발생 하고 있음을 알 수 있다.

다. 실험자 2의 시간대 신호(1)



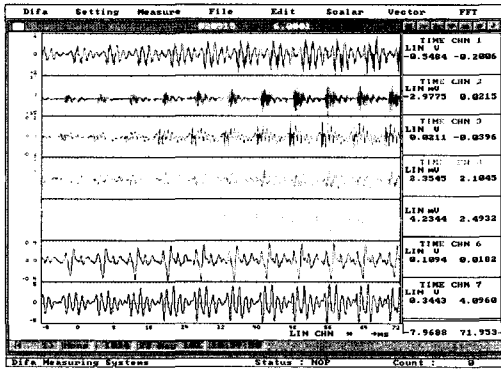
[그림 5] 실험자 2의 '가' 발음의 시간대별 음성신호

라. 실험자 2의 주파수 분석 결과(1)

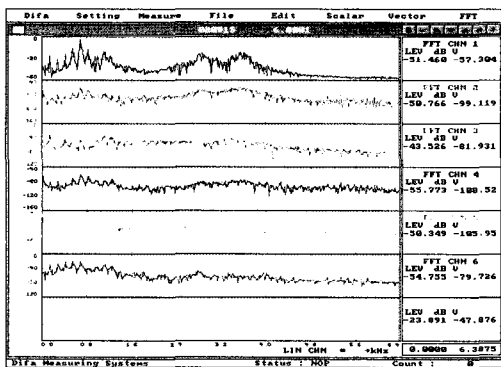


[그림 6] 실험자 2의 '가' 발음 음성신호에 대한 주파수 분석

마. 실험자 3의 시간대 신호(1)



[그림 7] 실험자 3의 '가' 발음의 시간대별 음성신호



[그림 8] 실험자 3의 '가' 발음 음성신호에 대한 주파수 분석

바. 실험자 3의 주파수 분석 결과(1)

그림 3, 5, 7은 실험자들이 '가' 발음을 실시할 때 그 음성신호를 시간대별로 포착한 음성신호이다. 이 음성신호를 가지고 주파수 분석한 결과를 그림 4, 6, 8에 각각 나타내었다. 그림을 바탕으로 실험 결과를 분석하면 아래와 같다.

그림 3에서 8가지에서 채널 1은 입 앞에서의 음성 신호와 주파수 분석을 한 것으로 각 채널 분석에 대한 기준을 제시한다.

2번 채널은 1번 채널의 신호를 기준으로 볼 때 이마의 진동은 입 앞에서 포착한 음성신호와 유사한 것을 확인 할 수 있으며, 특히 음성대역 중 고주파 성분인 1~3KHz의 재생성이 높은 것으로 측정되었다. 이러한 것은 이마 부위의 골도 진동을 포착하는 것이 음성신호를 포착하기에 적합한 위치라는 것을 말하고 있으며, 이것은 진동자를 이용한 이마의 골도 진동을 포착하였기 때문에 외부 또는 주변의 Noise에 대한 감도는 낮아서 소음에 쉽게 노출되지 않는다는 장점이 있다. 그러나 단점으로는 저주파 신호의 포착은 감도가 많이 떨어지는 문제점을 갖고 있으며 전체적으로 포착되는 신호가 작아서 mV 단위로 포착되는 것을 볼 수 있다.

3번 채널에 포착된 음성신호는 성대 위의 목 부분에서 음성을 포착한 것으로 그래프에서 보는 바와 같이 저주파 성분의 포착이 우수하나 고주파 성분의 음성신호 포착이 제대로 되지 않고 있는 것을 알 수 있는데 음성의 왜곡이 심하여 대화를 쉽게 할 수 없으며 대화하는 상대를 음색에 의하여 판단하기가 어려운 문제가 있다. 그러나 포착되는 신호의 크기는 가장 커서 V 단위로 표시되는 것을 볼 수 있어서 소음이 극단적으로 심한 곳에서 활용이 가능한 것으로 판단되나 전기적으로 증폭과 보상이 갈수록 쉬워지는 현

실에서는 음성신호를 정확하게 측정하는 이마 부분의 측정 신호보다는 신호의 효용성이 떨어진다.

4번 채널 및 5번 채널에 포착된 음성신호는 귀 앞 중앙의 뼈가 많은 부위와 귀 뒤 중앙의 뼈가 많은 부위에 진동센서를 사용하여 측정한 신호이다. 신호의 전체적인 크기가 mV 단위로 포착되는 것으로 보아 신호의 전달정도가 약하며, 신호의 왜곡도 심하여 사용하기가 어려운 부분이라 할 수 있다.

6번 채널의 경우는 귀 속 외이도 내에 캔트레버 형태의 센서를 삽입하여 귀 구멍 내의 골도 및 피부 진동과 고막에 의한 공기 진동에 의한 음성신호를 캔트레버의 진동형태로 포착한 그림을 보여주고 있으며, 고주파 영역의 대부분 신호를 손실하고 있으며 저주파 영역에서는 높은 감도로 감지하여 V 단위의 크기를 보여주고 있다.

7번 채널의 신호는 귀 속 외이도 내에 컨덴서 형태의 마이크를 삽입하여 고막의 진동을 공기의 진동으로만 측정한 것으로 감도도 높으며 동시에 모든 주파수 대역에서 음성신호의 재생성이 높은 것으로 나타난다. 물론 입 앞에서 측정한 신호와는 동일하지 않으며 고주파 영역에서 약간의 감소가 발생하나 1번과 가장 유사한 음성신호의 포착을 보여주고 있다. 또한 귀속의 외이도는 헤드셋으로 감싸기가 쉽기 때문에 감도가 높은 컨덴서 형태의 마이크를 사용하여도 주변의 Noise 와 음성신호를 분리해 낼 수 있어 효율성이 높은 감지 방법으로 판단되었다.

위에서 보여주는 그래프는 대표적인 것만을 선별한 것으로 3명의 실험자와 각 실험자들이 각기 다른 음을 방출하며 신호를 포착한 결과는 부록에 첨부하였다.

참조 : 부록(각 인체 부위별 성대의 진동을 측정한 신호의 비교 data)

가. 실험자 1의 시간대 신호(2~4)

나. 실험자 1의 주파수 분석 결과(2~4)

다. 실험자 2의 시간대 신호(2~4)

라. 실험자 2의 주파수 분석 결과(2~4)

마. 실험자 3의 시간대 신호(2~4)

바. 실험자 3의 주파수 분석 결과(2~4)

부록을 포함하는 모든 측정 및 분석 결과를 바탕으로 종합하여 판단할 때 다음과 같은 결과를 획득하였다. 입 앞에서 측정되는 음성신호를 기준으로 판단할 때, 이마에서 측정하는 진동신호는 고주파 영역의 음성을 왜곡이 적게 포착하는 것으로 측정되었으며 저주파 영역도 그리 나쁘지는 않았다. 그러나 전체적으로 감도의 저하가 심한편이어서 전 주파수 대역에서의 높은 증폭도를 갖는 음향장치가 필요하게 된다.

성대위에서 측정하는 방법과 외이도에 캔트레바 타입의 센서를 사용하는 것은 저주파 영역에서는 상당히 높은 감도로 신호를 포착하는 것이 가능하나 고주파 영역에 있어서는 신호를 대부분 잃어버려 음성의 왜곡이 심한 것으로 나타났다. 또한 귀바퀴 앞과 뒤는 전 주파수 대역에서 감도가 떨어지므로 사용이 어려운 것으로 나타났다.

귀 속 외이도에서 컨덴서 마이크를 이용한 방식이 다른 신체부위에서 포착하는 방식에 비하여 음성신호의 포착이 우수한 것으로 나타났으며, 고주파(2KHz~3KHz)와 저주파(100Hz)대역에서의 약간의 음색 왜곡을 제외하고는 입 앞에서의 음성신호와 동일한 음성신호를 포착하는 것이 가능한 것을 확인하였다.

#### 4. 결 론

성대에 의한 음성 신호의 방출을 신체 7곳의 위치

에서 음향 및 진동을 동시에 측정하여 음성신호를 포착 및 분석한 결과 각 위치별 그리고 그 위치에서 사용이 적합하다고 판단되는 감지방식의 장단점을 비교 분석 할 수 있었다.

귀 속 외이도에서의 음성신호의 포착은 성대에 의한 음성신호의 포착이 가장 효과적인 뿐 아니라, 귀에 귀덮개를 사용하여 소음이 귀 속으로 입력되는 것을 차단하면 음성이 소음에 의한 영향을 받지 않고 음성 신호만을 포착하여 재생할 수 있는 가장 적합한 것으로 판단하였다.

이러한 음성신호의 포착 방식 및 그 위치는 열악한 소음 상황에서 음성 통신을 가능하게 하여 여러 상황의 전장환경에서 소음에 의한 지휘통제상의 문제점을 해소할 수 있는 효율적인 방법이라 사료된다. 특히 탱크, 전함 그리고 전투기 등과 같이 음성신호에 비하여 높은 수준의 소음을 방출하는 기동장비에 있어서는 이러한 방법의 음성통신 시스템이 전투력 증강과 승무원의 안락함을 동시에 제공할 수 있는 방법이라 확신한다.

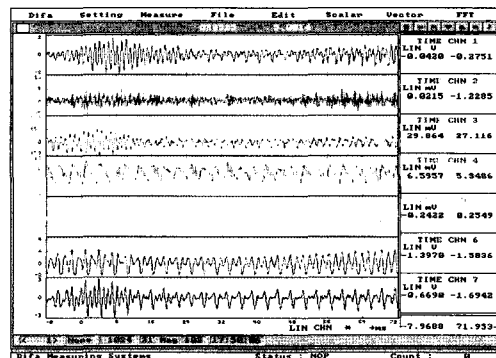
참 고 문 헌

[1] 김강녕, “미래전에 대비한 군사혁신”, 군사연구 117집, pp.195~226, 2001.  
 [2] "C4I systems-servant, not master", Asian defence journal, pp.32~37, 2002.  
 [3] 통신운용 야전교범 5-1 ; 1998.  
 [4] Hansen, M. O. and Stinson, M. R., "Air conducted and body conducted sound produced by own voice", Canadian Acoustical, v.26 no.2, pp.11~24, 1998.

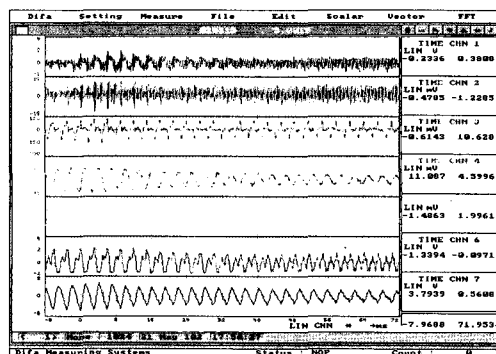
[5] 강한구, 이근수, 남창희, “군용 항공기 소음대책 연구”, 2000.  
 [6] 오병무, “골도전도용 헤드셋 BH-20M”, 평화, pp 86~91, 2002년 12월.  
 [7] Thales acoustics sales literature ; "sound sensor, throat microphone". 2002.  
 [8] 김동하 “신편 센서공학”, 2001, pp.41~50, 태훈출판사.  
 [9] 노병천 역 “센서공학”, 1987, pp.91~98, 동일출판사.

부 록

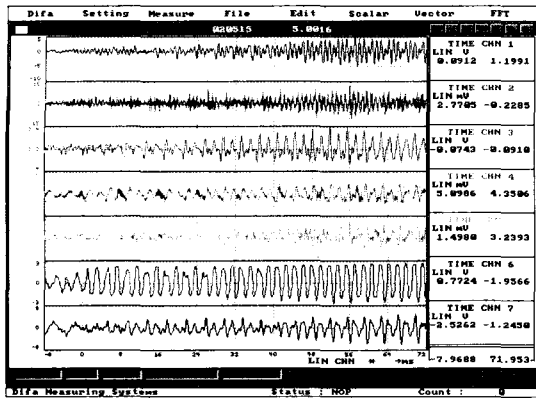
가. 실험자 1의 시간대 신호(2~3)



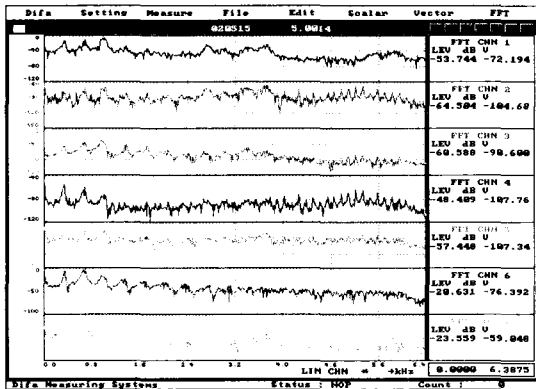
[그림 9] 실험자 1의 '토' 발음의 시간대별 음성신호



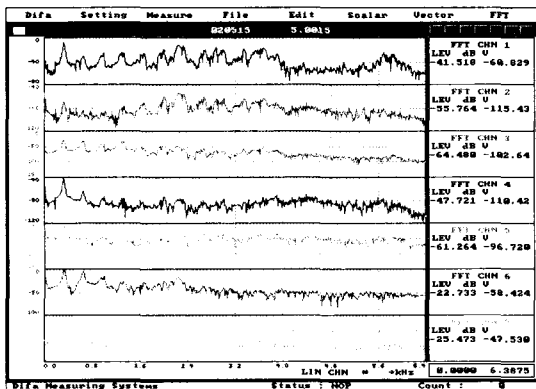
[그림 10] 실험자 1의 '피' 발음의 시간대별 음성신호



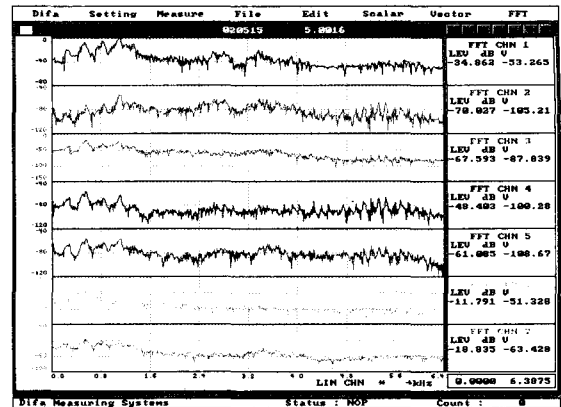
[그림 11] 실험자 1의 '러' 발음의 시간대별 음성신호  
나. 실험자 1의 주파수 분석 결과(2~3)



[그림 12] 실험자 1의 '토' 발음 음성신호에 대한 주파수 분석

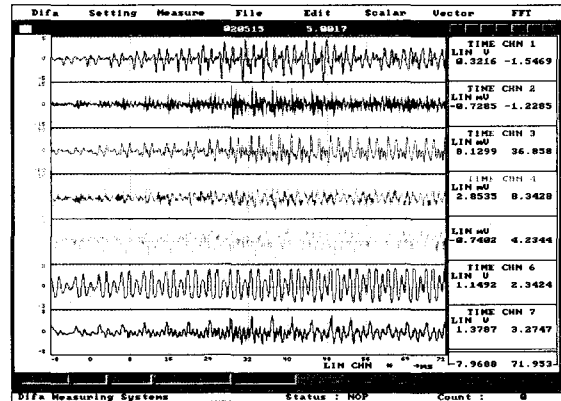


[그림 13] 실험자 1의 '피' 발음 음성신호에 대한 주파수 분석

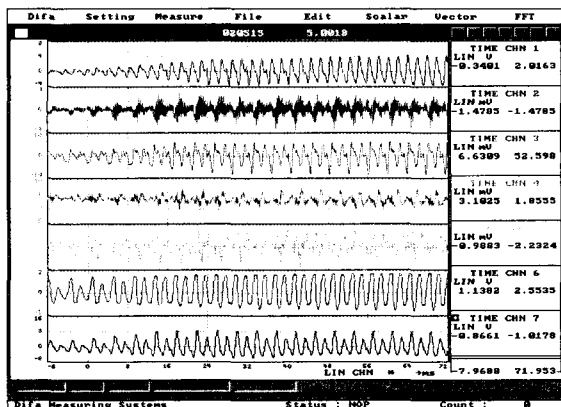


[그림 14] 실험자 1의 '러' 발음 음성신호에 대한 주파수 분석

다. 실험자 2의 시간대 신호(2~3)

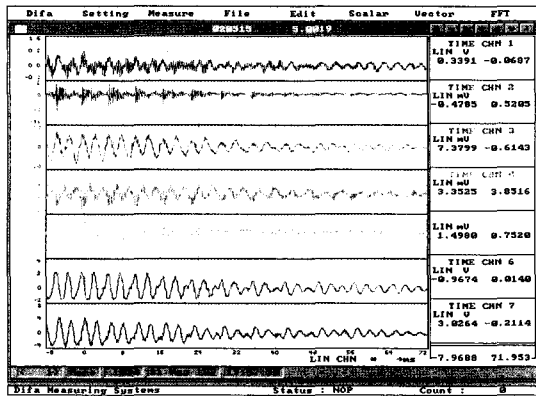


[그림 15] 실험자 2의 '토' 발음의 시간대별 음성신호

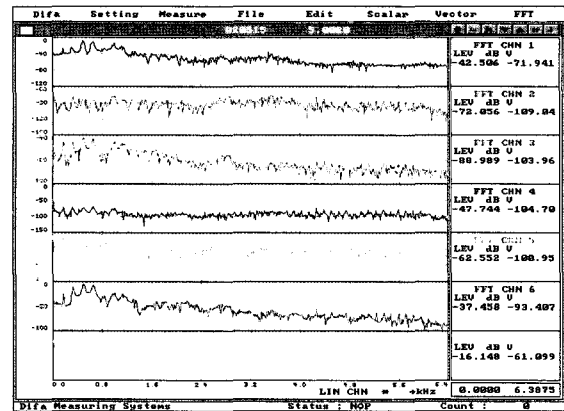


[그림 16] 실험자 2의 '피' 발음의 시간대별 음성신호

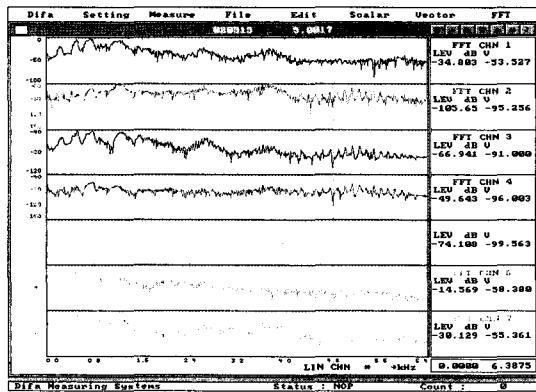




[그림 17] 실험자 2의 'ㄹ어' 발음의 시간대별 음성신호  
라. 실험자 2의 주파수 분석 결과(2~4)

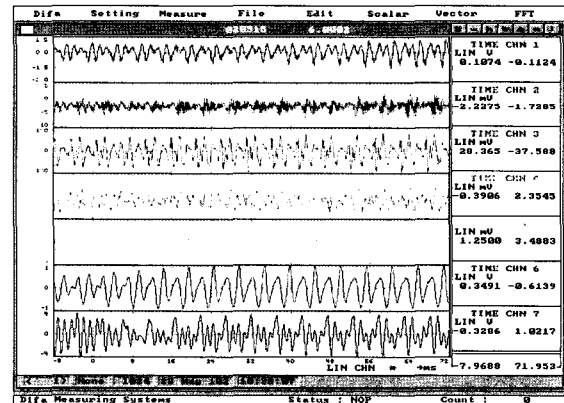


[그림 20] 실험자 2의 'ㄹ어' 발음 음성신호에 대한 주파수 분석

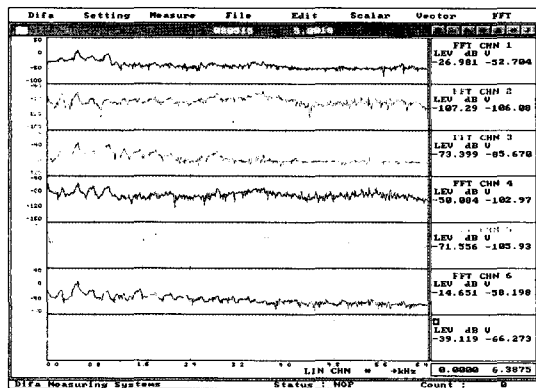


[그림 18] 실험자 2의 '토' 발음 음성신호에 대한 주파수 분석

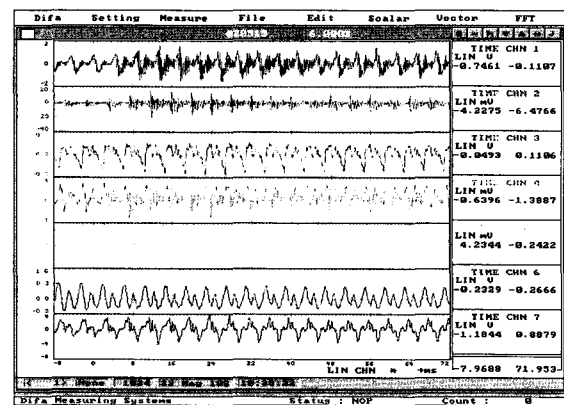
마. 실험자 3의 시간대 신호(2~4)



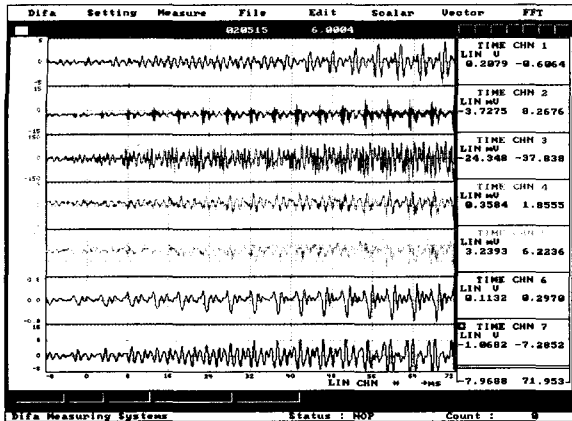
[그림 21] 실험자 3의 '토' 발음의 시간대별 음성신호



[그림 19] 실험자 2의 '피' 발음 음성신호에 대한 주파수 분석

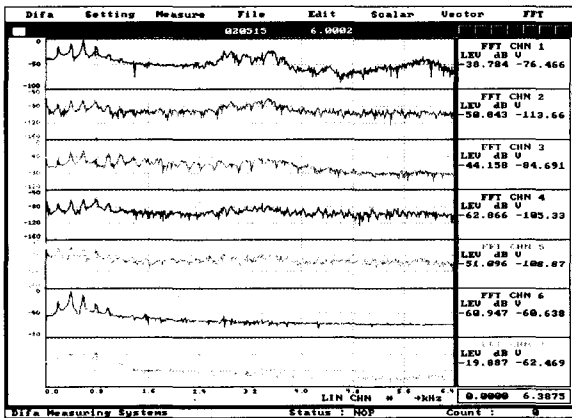


[그림 22] 실험자 3의 '피' 발음의 시간대별 음성신호

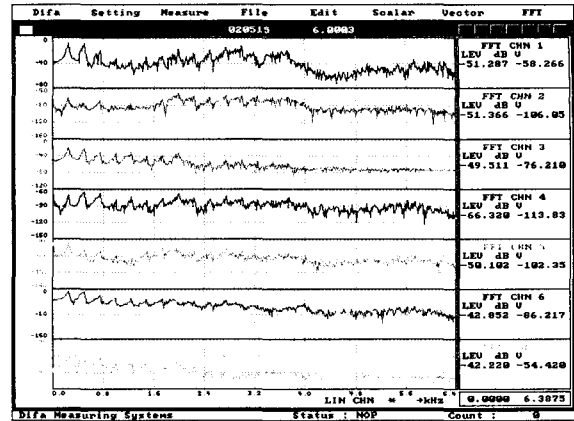


[그림 23] 실험자 1의 '러' 발음의 시간대별 음성신호

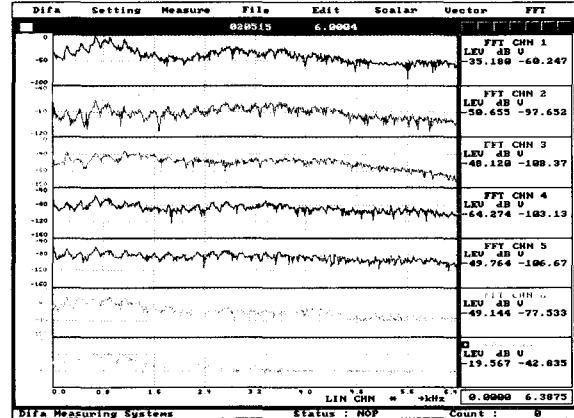
바. 실험자 3의 주파수 분석 결과(2~4)



[그림 24] 실험자 3의 '토' 발음 음성신호에 대한 주파수 분석



[그림 25] 실험자 3의 '피' 발음 음성신호에 대한 주파수 분석



[그림 26] 실험자 3의 '러' 발음 음성신호에 대한 주파수 분석