

## 청각 실험 도구 Alvin 사용법

동의대학교 영어영문학과 양병곤

### I. 서론

사람의 음성은 의사소통을 위해 사용된다. 의사소통 과정에서는 항상 화자의 발화를 청자가 듣고 어떤 의미인지 처리함으로써 음성의 기능을 다하게 된다. 이 때 발성된 목소리가 제대로 전달이 되지 않는다면 의사소통이 제대로 이뤄지지 않을 뿐만 아니라 음성으로서의 가치도 잃게 되므로 청자의 처리과정이 무엇보다도 중요한 부분이 된다. 이런 의미에서 청자 쪽의 반응을 살펴보는 청각실험은 음성에 대한 특징을 찾아보거나 파라미터의 중요성을 측정하는데 매우 중요하다.

일반적으로 청각실험을 하는 방법에는 다음과 같은 몇 가지를 들 수 있다. 먼저, 어떤 음성의 쌍을 들려주고 이들이 같은지 다른지 구별하는 구별실험(discrimination test)이 있고, 두 개의 서로 다른 음성을 연이어 들려준 뒤 어느 발음인지 확인하게 하는 확인실험(identification test)이 있다.

구별실험은 주로 자음이나 모음 쌍을 들려주어 차이를 지각할 수 있는지를 밝히는 것으로 국어와 영어의 모음을 들려주어 차이가 있는지 여부를 표시하게 하는 것이다. 보통 구별실험에서는 두 개의 연속된 쌍의 음성을 들려주어 같은지 다른지를 표시하는 방식(AX: A에 따라오는 음성 X가 A와 같으면 '같다'에 동그라미를 치고 다르면 '다르다'에 표시함)이다. 실제, 실험해 본 결과 청자 자신이 자극을 음성학적으로 판단할 수도 있고 순수하게 청각적인 차이로 판단할 수도 있기 때문에, 일반적으로는 두개의 쌍에 덧붙여 한 개의 음성을 더 들려주고 이를 첫 번째와 두 번째 가운데 어느 것과 유사한지 표시하게 하는 방식을 사용하기도 한다 (ABX: A와 B에 따라오는 음성 X가 A와 같으면 'A'에 동그라미를 치고 B와 같으면 'B'에 표시함). 이렇게 하면 순전히 소리의 차이만 구별하는 것이 아니라 A, B 두개의 음성에 대한 음성학적인 특징을 파악하고 기억하고 있다가 다음에 따라오는 소리의 특징과 어느 것이 일치하는지 판단하게 된다고 할 수 있다. 이 때 실험 총수에서 맞게 판단한 개수가 차지하는 비율을 구하여 50%보다 높은 값을 가지면 청자는 이를 실제 구별할 수 있다고 판단할 수 있다. 얼마의 값을 유의미한 것으로 받아들여야 할지는 주관적이라 할 수 있으나 연구의 목적에 따라 80%이상 등으로 정해볼 수 있다. 일반적으로 사람은 자극 쌍의 경계선에서 매우 잘 구별한다고 한다. 또 다른 방법은 4IAX로서 4개의 음성이 두개씩 이어져 제시되는데 두 쌍 중 어느 것이 같은 음성인지를 확인하는 것이다. AA-AB의 순서나 AB-AA로 제시하여 두 쌍 중 반드시 하나는 다르다고 설명하고 찾아내게 하는 것이다.

확인실험에서는 어떤 음성을 들려주고 몇 개 가운데 어느 것인지 직접 받아쓰게 하거나 해당 음성에 동그라미를 그리도록 한다. 이때 해당 음성 가운데 반드시 하

나를 택하도록 강제선택(forced choice)하거나 그 외의 음성으로 듣는다면 자유선택 (free choice)을 할 수 있는 방법을 취할 수 있다. 강제선택을 할 경우의 장점은 통계처리 등에서 화자의 반응을 제한된 수의 선택결과를 가지고 처리할 수 있는 반면 자유선택은 너무 광범위한 응답이 나와 일관성 있는 결론을 내리기 어려운 경우가 발생하기도 한다.

그 외에도 청각적인 판단 결과의 정확도나 음질의 수준 등을 임의의 잣대로 표시하는 방식이 있다. 예를 들어, 음질을 판단하기 위한 GRBAS 척도가 있는데, 모음의 특성에서 grave(둔감한 정도), roughness (거침 정도), breathiness(성대가 완전히 붙으면서 진동하지 않아서 숨이 새는 듯한 기식음 정도), asthenic(무기력하고 약한 정도), strained(긴장되고 억지 쓰는 듯한 목소리의 정도) 등을 들어서 7점 척도 또는 약간(mild), 보통(moderate), 심함(severe)의 어느 부분에 해당하는지 판단하는 것이다.

양이 자극에 대한 실험은 좌뇌와 우뇌의 언어처리에 대한 비교실험으로 사용되는데, 먼저 좌뇌는 우측귀의 정보를 처리하므로 양쪽 귀에 두개의 다른 음성을 들려주면 오른쪽 귀로 들어온 음성을 더 잘 듣는다고 한다. 이에 반해, 이중지각(duplex perception)에서는 과열음+모음을 합성할 때 한쪽 귀에는 모음의 제 2포먼트 전이부분만 들려주고 다른 쪽 귀에는 나머지 부분을 들려준다. 이렇게 되면 양쪽 귀에 들어온 정보가 합해져서 과열음의 전이와 일치할 만큼 충분한 정보이면 그 자음으로 듣고 그렇지 않으면 잡음으로 듣는 결과를 보인다고 한다.

이러한 다양한 청각실험을 할 수 있는 소프트웨어가 Alvin이다.

## II. Alvin 설치 및 청각실험

Alvin은 인터넷 웹사이트인 <http://homepages.wmich.edu/~hillenbr/>에 들어가 Alvin experiment-control software (~56 Mb)를 마우스 왼쪽으로 클릭하면 저장할 장소를 묻는 대화상자가 나오고 자신의 하드디스크에 저장하면 된다. 설치하는 Alvinsetup.exe를 클릭하면 자동으로 풀려서 하드디스크에 실행파일과 관련파일들이 함께 생성된다. 실행을 하려면 Alvin.exe를 누르면 그림 1과 같은 실험환경 조절 박스가 나타난다.

그림 1의 실험환경을 시작하려면 먼저 여러 가지 실험방식 가운데 하나를 마우스로 클릭하여 택하고, Initials(머리글자)에 실험대상자의 이름 첫 한 두 글자를 입력한 뒤 Begin이란 단추를 누르면 된다. 실제 예를 들어, w-r 확인을 시험하는 첫번째 목록을 클릭하면 그림 2와 같은 실험환경이 나타난다. 맨 가운데 위치한 Instructions 단추를 누르면 실험방법에 대한 자세한 설명이 영어로 나타난다.

그림 1. Alvin의 다양한 청각실험 목록

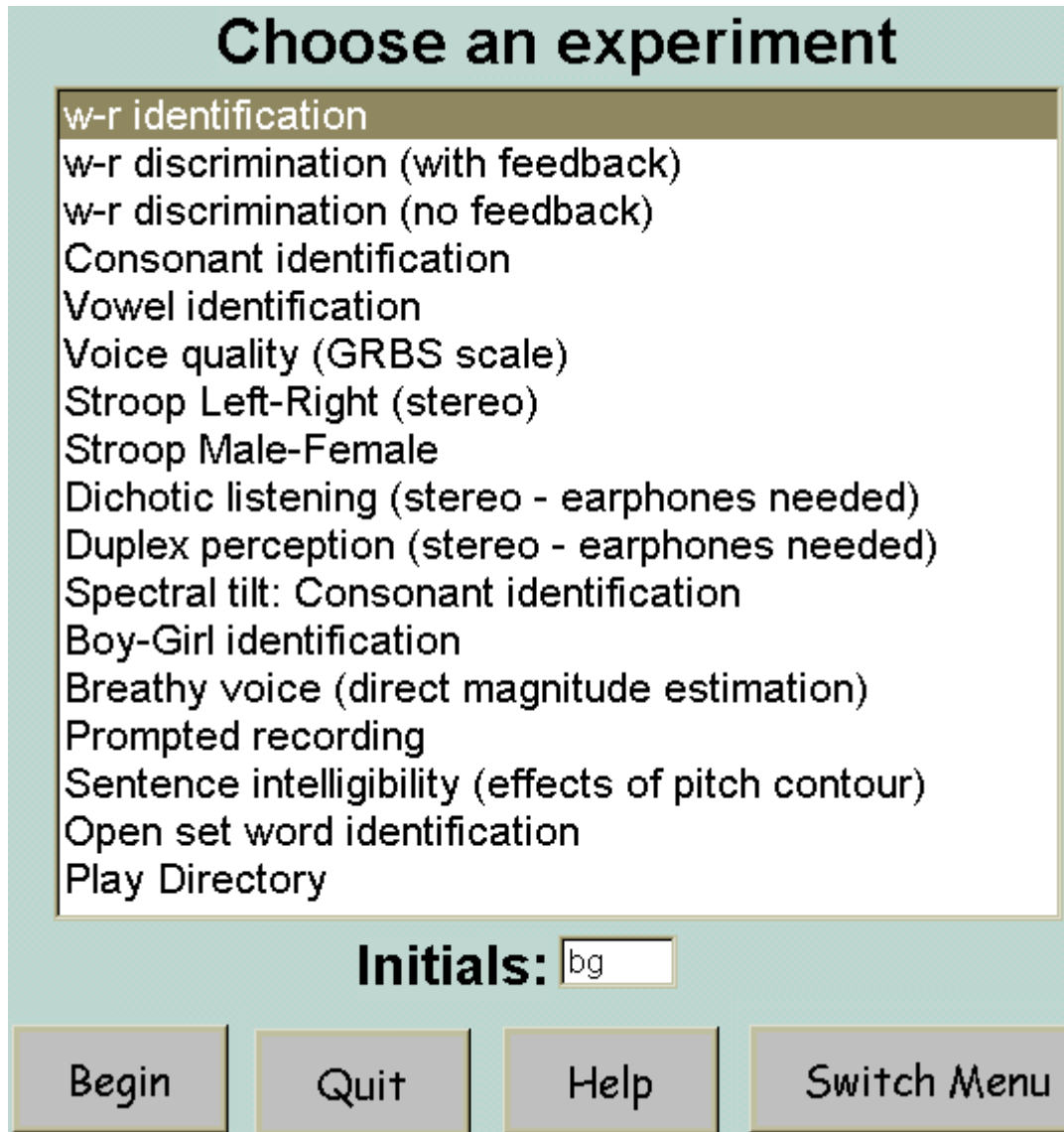


그림 2. w-r 구별실험

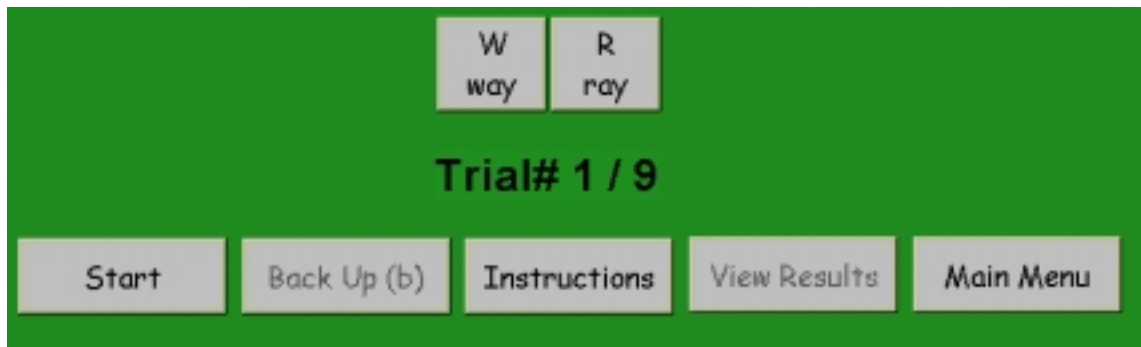


그림 2에서 Trial#1/9는 9개의 듣기 실험 가운데 첫번째이고 Start단추를 누르면 way나 ray 중 한 단어가 들리며, 하나를 클릭하면 다음 발음으로 이동된다. 되돌아가려면 Back Up(b) 단추를 누르거나 키보드로 b를 누르면 된다. 9번의 실험을 마치고 결과를 살펴보려면 View Results 단추를 클릭하면 된다. 원래의 첫 메뉴로 돌아가려면 Main Menu 단추를 클릭한다.

그림 1에 표시된 Alvin에 탑재된 각각의 실험방식을 간단히 설명하면 w-r identification은 두 발음을 듣고 어느 것인지 확인하는 실험이다. w-r discrimination은 ABX 구별실험으로 세 개의 발음이 연이어 나왔을 때 세번째 나온 발음이 첫번째 발음과 같으면 (1)을 선택하고 두번째 발음과 같으면 (2)를 선택해야 한다. (with feedback)은 응답을 한 직후에 바로 정답을 표시해주며 (no feedback)은 정답표시가 나타나지 않는다. 다음에 있는 Consonant identification은 b, d, g, z... 와 같은 여러 가지 영어 자음을 하나씩 듣고 각각에 해당하는 자음을 클릭하는 자음 확인실험이다. 다시 한번 들으려면 Replay단추를 누르면 된다. Vowel identification은 모음 확인 실험으로 재생된 모음을 듣고 해당 모음을 클릭한다. Voice quality(GRBS scale)는 ‘아’하고 길게 발음한 모음을 듣고서, 거친 정도(Roughness), 숨이 새는 듯한 기식성 정도(Breathiness), 긴장성 정도(Strain), 피치 비정상정도(Pitch abnormality)등을 약간(Mild), 중간정도(Moderate), 심함(Severe)의 세 가지 척도위에 마우스로 눈금자를 옮겨서 표시한다. Stroop Left-Right(stereo)는 <http://faculty.washington.edu/chudler/words.html>에 잘 소개되어 있는데, 성인과 같이 단어를 익힌 상태에서 그 단어가 지시하는 것과 반대인 것을 말하거나 반응하는데 처리되는 시간이 많아진다는 것이다. 어린아이와 같이 단어를 모르는 상태에서 쉽게 응답할 수 있다고 한다. 이 패키지에서는 Left와 Right라는 단어가 화면에 제시되어있고, 스피커의 왼쪽에서 나온 음성이면 Left를 클릭하고 오른쪽에서 나온 음성이면 Right를 클릭하면 된다. Left란 음성이 왼쪽으로 나타날 때는 바로 반응하는데 오른쪽으로 나올 때는 망서리게 되어 반응시간이 길어지게 된다. Stroop Male-Female도 두 명의 남녀화자가 발음하는 단어를 듣고 남자의 목소리면 Male을, 여자의 목소리면 Female을 클릭해야 한다. 발음 가운데는 남자가 ‘female’이라고 하거나 여자가 ‘male’이라고 발음하기 때문에 판단이 늦어지거나 잘못되기도 한다. Dichotic listening은 양이효과를 실험하는 것으로 ba와 da라는 발음을 왼쪽귀와 오

른쪽 귀에 동시에 들려주면 어떤 발음을 들었는지 표시한다. 이 때 오른쪽귀로 들 어간 발음을 선택하는 경향이 높는데 이는 오른쪽 귀와 연결된 왼쪽 뇌가 언어처리 과정에서 우뇌보다는 특성화되어있다는 기존의 연구를 뒷받침한다. 왼쪽귀와 오른 쪽 귀에 동시에 같은 소리를 들려주지는 않는다. 다음으로는 Duplex perception으로 서 11살 된 남녀아동의 발음을 듣고 남자인지 여자인지 판단하고 이를 1-5점 척도 에서 확실하면 5를 불확실하면 1을 체크하여 응답하는 실험이다.

Spectral tilt: Consonant identification은 자음의 음향적 특성 가운데 스펙트럼의 기울기를 다양하게 변형시켜 들려준 뒤 자음 가운데 어떤 발음이었는지 선택하게 한다. 그 결과를 보면 어떤 자음을 제대로 지각하지 못하는지 %로 나타난다. Boy-Girl identification은 발음을 듣고 남자인지 여자인지를 구별하고 얼마의 확신 을 가지고 판단했는지 1-5점 척도에서 선택한다. Breathly voice(direct magnitude estimation)은 앞서의 Voice quality와 같이 모음 발음을 하는데 이 발음의 성대가 제대로 닫히지 않아 발생하는 기식성이 얼마나 큰지 구체적인 숫자(10 또는55)로 나타내게 한다. 음성을 합성하거나 소음과 합친 뒤 어떤 요소가 소음성이 많다고 판단하게 만드는지 확인할 수 있을 것이다. 이 실험도구를 이용하여 영어나 한국어 발음의 유창도를 체크하게 하는 실험도 만들 수 있을 것이다. Prompted recording 은 미리 지정한 단어를 발음하고 컴퓨터에 저장하는 기능을 한다. 이것을 실행하면, Say the word 'rye'로 나타나고 실제 녹음된 영어발음이 들리면 Record라는 단추 를 누르고 녹음을 시작하고 끝나면 Stop 단추를 누르면 된다. 미리 지정된 단어발 음을 듣고 따라하게 한 뒤 이를 저장하여 음성분석을 하는데 매우 편리하다. 물론 다음 섹션에 설명하는 편집과정을 거치면 우리말 발음을 녹음시켜 재생하게 하고 피험자가 이를 따라 녹음하게 만들 수도 있다. Alvin 폴더의 녹음관련 내용을 살펴 보면, 그림 3과 같이 나타난다.

그림 3. 녹음하기위한 폴더의 파일 구성

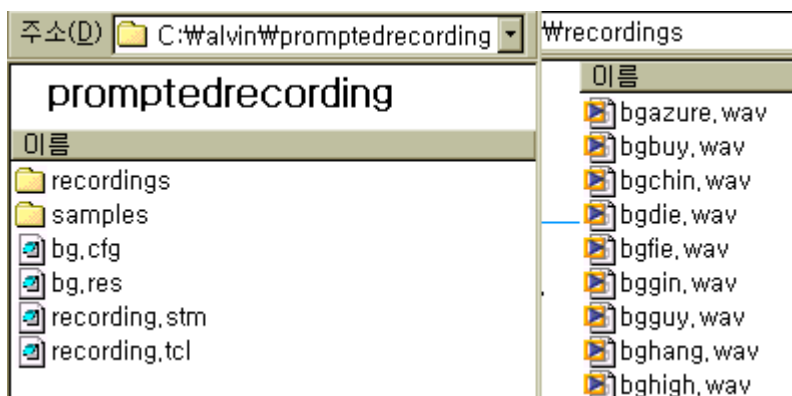


그림 3과 같이 Alvin폴더의 prompted recording아래 samples폴더는 미리 녹음할 음성을 들려주고자 할 때 녹음한 파일들을 보관해둔다. 녹음된 음성은 recordings에

피험자의 이니셜(여기서는 bg)과 (발음한 음성의 단어이름.wav)로 저장되기 때문에 여러 사람이 실험해도 폴더에서 각각의 음성파일을 구분하기 쉽다. bg.cfg는 피험자(bg)가 녹음할 때 어떤 순서로 samples 폴더에 녹음된 단어와 지시문이 재생될지를 나타내고, bg.res는 처리된 결과를 나타내는데 저장된 것은 1로 안된 것은 0으로 표기하여 동일한 피험자가 나중에 다시 녹음할 때, 지난번에 녹음한 단어는 건너뛰고 남은 부분부터 녹음을 재개하거나, 앞서 녹음한 것을 무시하고 새로 녹음할 수도 있다. recording.stm은 자극음을 제시할 때 화면에 표시될 문장과 발음 목록을 기록하는 것이다. 이 자극 목록을 뒤섞어서 bg.cfg에 저장하고 실제 실험에서는 bg.cfg의 제시된 자극 순서에 따라 녹음된다. 녹음된 파일은 16000 Hz의 표본속도로 윈도우즈의 웨이브파일로 저장되므로 어떤 음성분석프로그램에서나 쉽게 불러올 수 있다. 표본속도를 바꾸려면 recording.tcl 파일을 열어서 맨 마지막 줄에 있는 startrecording 16000 .meter의 16000을 원하는 숫자로 바꾸면 된다. 피험자가 녹음하기 전에 미리 발음해 주는 부분을 없애려면 recording.tcl파일의 소스 본문의 해당 부분에 #표를 붙이면 된다.

Sentence intelligibility(effect of pitch contour)라는 실험은 영어문장의 전체 피치값을 여러 가지 형태로 변환하여 들려준 뒤 피험자는 그 문장을 따라 발음하게 하여 컴퓨터에 녹음한다. Open set word identification은 정해지지 않는 임의의 단어를 듣고 키보드로 빈칸에 타자하는 실험이다. 실험에 걸린 총 지속시간을 인쇄해준다. 마지막으로 Play all은 음성파일을 한꺼번에 보여주고 따로따로 또는 한꺼번에 차례로 재생하는 기능을 그림 4와 같이 보여준다.

그림 4. 자음 확인 폴더에 들어있는 음성파일 재생



Path:란에 직접 파일을 찾을 경로를 입력하거나 그 곁에 있는 Browse 단추를 눌러 해당 파일이 있는 폴더를 선택하여 웨이브파일을 하나 클릭한 뒤, 오른쪽에 있는 Display를 눌러 모두 열어도 된다. Play All은 열려진 모든 파일을 재생해주고, 중간에 그만두려면 Stop 단추를 누르면 된다. Spec:은 파일 가운데 모든 웨이브파일은 \*.w\*로 표시하고 이때 \*은 개수에 관계없이 아무 문자가 와도 된다는 뜻이다. 만약 프라트 파일로 저장 되어있는 경우는 이곳에 \*.aifc로 입력하면 된다.

### III. 지시문과 메뉴 한글화

Alvin은 미국에서 개발되었기 때문에 영어로 모든 실험안내문, 화면지시문, 메뉴, 자극발음세트가 영어로 작성되어 있다. 이들을 모두 우리말로 변환시키려면 이 프로그램을 모두 파악하여 원래의 소스 가운데 화면에 뿌려주는 인용부호로 나타난 부분을 한글로 바꿔야한다. 필자와 같이 전문 프로그래머가 아닌 분들이 한국인 피험자를 대상으로 실험할 때 간단히 한글화할 수 있는 부분을 소개해 본다.

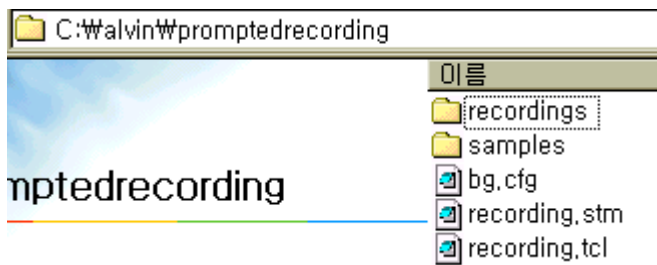
안내문을 바꾸려면 각 실험폴더의 instructions.txt를 찾아 한글로 내용을 편집하면 된다. 예를 들어, dichotic listening 실험의 안내문은 그림 5와 같이 alvin폴더 아래의 instructions.txt를 클릭하여 워드패드와 같은 텍스트편집프로그램으로 연 뒤 영어로 된 문장을 한글로 바꾸기만 하면 된다.

그림 5. 실험안내문 편집



메뉴부분은 Tcl/ Tk의 원래의 스크립트 파일의 메뉴부분을 한글로 표시하면 된다. 한 가지 예를 들어 보면, Prompted recording의 메뉴와 화면지시문을 한글로 바꾸려면 그림 6과 같이 Alvin폴더의 promptedrecording을 찾아서 recording.tcl을 연다.

그림 6. 녹음 화면 지시문 실행스크립트 파일 위치



아래와 같이 열려진 파일의 본문에서 button.start의 “Start” 대신에 한글로 “시작”이라고 입력하고 영문을 지운다. 나머지 재생부터 처음메뉴”는 이렇게 바꾼

모양을 보여준다. 그러나 녹음을 재시작하거나 중단시키는 메뉴는 원 소스파일에서 수정이 되어야하기 때문에 여전히 영어로 나타나게 된다.

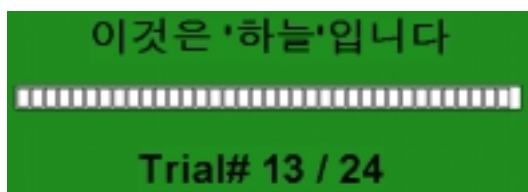
```
button .start -x 3 -label "Start시작" -command startClicked -key n
button .replay -x 18 -label "재생" -command replayClicked
button .record -x 33 -label "녹음(r)" -command startRecordingClicked -key r
```

덧붙여 녹음지시문을 한글로 바꾸려고 할 때는 바로 아래에 있는 recording.stm을 열면 다음과 같은 파일 본문이 나타난다.

```
"Say the word 'buy'" buy.wav
"Say the word 'rye'" rye.wav
==>"이것은 '하늘'입니다" sky.wav
```

여기서 앞의 인용부호 안의 부분은 녹음지시문으로 나타나므로 위와 같이 우리말로 바꾼 뒤 저장하고 녹음실험을 실시하면 섞여진 파일 제시 순서가운데 중간쯤에서 그림 7과 같이 한글로 나타난다.

그림 7. 한글 녹음화면 지시문의 예



이 때 sky.wav를 영어 발음 대신에 ‘하늘’이라는 발음으로 재생하여 피험자에게 들려주려면, promptedrecording을 이용하여 필요한 단어들을 녹음하고 이것을 해당 실험폴더의 samples폴더에 넣고 동시에 새로 녹음한 파일 목록을 recording.stm에 넣어주면 된다. 녹음 파일을 만들려면, 윈도우즈의 자체 녹음프로그램인 녹음기[시작-->보조프로그램-->엔터테인먼트-->녹음기]를 열어서 녹음하거나, 참고문헌에 있는 프라트 프로그램을 이용하여 16000 Hz의 표본속도로 녹음하고 윈도우즈 웨이브 파일 형태로 저장하여 samples폴더에 넣어도 된다.

이상에서 살펴본 Alvin의 기능은 매우 다양한 청각실험을 할 수 있도록 잘 만들어진 프로그램으로 평가할 수 있다. 누구나 개인용 컴퓨터만 있으면, 프라트의 음성 조작(Manipulation)기능을 활용하여 피치변환이나 재생속도 변환, 음성여과기능으로 다양하게 변형한 음성을 Alvin에 가져와서 청각실험을 할 수 있는 아주 완벽한 연구 환경을 구축할 수 있을 것으로 확신한다. 앞으로 많은 분들이 활용하여 다소 미개척 분야인 음성지각 부분에 대한 많은 발표가 있기를 기대한다.



## 참고문헌

- 고도홍, 구희산, 김기호, 양병곤 공역. 음성과학. 한국문화사. 2000.
- 양병곤. "합성한 한국어 단모음의 지각실험연구." 언어 20권 제 3호.  
pp. 127-146. 1995.
- 양병곤. "인간의 청각 척도에 관한 고찰." 음성과학 제 2권, pp. 125-134. 1997
- 양병곤. "성도 변형에 따른 모음 포먼트의 변화 고찰" 음성과학 제 3권,  
pp. 83-92. 1998.
- 양병곤. 프라트를 이용한 음성 분석의 이론과 실제. 만수출판사.  
2003. (<http://www.donggeui.ac.kr/~bgyang/praat/praatdown.htm>)
- Byunggon Yang. "Perceptual experiment on number production for  
speaker identification." Speech Sciences Vol. 8. No. 1. pp. 7-19.2001.

※본 연구는 한국과학재단 목적기초연구(RO1-1999-000-00229-0)지원으로 수행되었음.

홈페이지: <Http://www.donggeui.ac.kr/~bgyang>  
이메일: [bgyang@donggeui.ac.kr](mailto:bgyang@donggeui.ac.kr)  
전화: 051-890-1227