

SFSWin을 이용한 음성분석 방법

동의대학교 영어영문학과

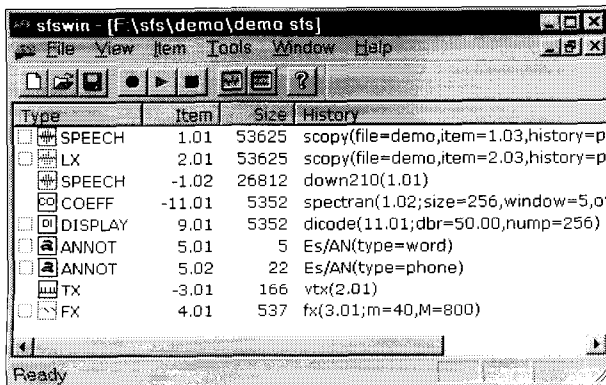
양 병 곤

머릿말

일상생활 중에 많은 사람들이 과도한 음성을 사용하게 되어 자신도 모르는 사이에 발생기관의 이상이 생기는 경우를 본다. 이러한 병적인 음성을 분석하기 위해서는 정상적인 음성의 특징을 먼저 객관적으로 분석하고 이들의 기준치를 잡은 다음 병적인 음성의 특징이 그 기준에서 얼마나 벗어났는지를 밝히는 과정이 필요하다. 사람의 음성을 지문과 같이 나타내는 음성분석 소프트웨어가 다양하게 개발되어왔다. 시중에는 너무 간단하여 사용하기는 편하나 학술논문을 작성할 때 사용할 만큼 정확한 데이터를 제공하지 못하는 프로그램이 있고, 반면에 매우 복잡한 기능들이 들어가 있어서 단순한 분석을 위해 그 모든 기능들을 익혀야 하는 불편한 프로그램도 나와 있다. 특히, 비싼 하드웨어를 동반한 소프트웨어들은 가격에 비해 사용자의 편의를 고려하지 않는 프로그램들도 있다. 한편, 인터넷에는 음성분석도구를 개발하여 이를 누구나 무료로 다운로드하여 소프트웨어만으로 모든 분석을 하고 학술논문에 이용할 만큼 좋은 도구들이 있다. 그 중에서 아주 정교하고 다양한 도구들이 포함된 패키지로는 암스테르담 대학의 Paul Boersma와 David Weenink가 음성분석 및 변형을 할 수 있도록 만든 Praat와 영국런던대학의 Mark Huckvale이 개발한 SFSWin (Speech Filing System for Windows)이 있다. 이 두가지 프로그램에는 각기 다른 메뉴와 연산방식으로 음성파일을 처리하지만 대체적인 음성분석 메뉴는 성대의 특성인 피치값과 성도의 특성인 포먼트 측정기능 등이 포함되어 되어 있다. 이미 프라트라는 음성분석 도구에 대한 설명은 필자의 홈페이지를 통해 소개해 두었기 때문에, 이번에는 SFS를 사용하는 방법에 대해 살펴보기로 한다.

SFSWin의 설치와 메뉴개관

SFSWin은 음성분석소프트웨어와 서브루틴 라이브러리, 독자적으로 작동하는 그래픽, 프로그램과 음성분석을 자동화하기 위한 스크립트 방식의 SML(Special Programming Language)로 구성되어 있다. 주로 음성입력, 재생, 발음표기, 스펙트로그램, 피치, 포먼트 분석 등을 할 수 있다. 실행파일은 인터넷을 통해 다운로드해야 하는데 홈페이지는 <http://www.phon.ucl.ac.uk/resource/sfs/index.html>로 들어가서 좌측열의 Download를 클릭한 뒤 최신 윈도우즈 버전(6.5Mb)을 받으면 된다. 설치는 sfs-440a.exe를 클릭하면 자동으로 파일이 생성되며 다음 그림은 실행한 뒤 demo.sfs라는 파일을 분석한 모양을 나타내고 있다. 사용법 설명이 인터넷을 통해 제공되며 동시에 실행프로그램의 Help부분에 자동으로 링크되어 있는 사용자 매뉴얼이 포함되어 있기 때문에 쉽게 필요한 정보를 찾아볼 수 있다. 이



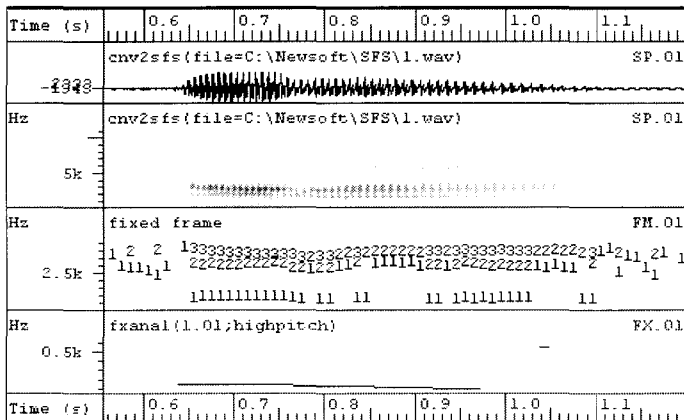
글에서는 간략하게 각 메뉴의 구성과 기능에 대해 살펴보고 기본적인 음성녹음과 재생, 분석 등에 대해 살펴보기로 한다.

먼저 File기본메뉴에 대해 살펴보면, 새로운 분석을 할 때는 New를 클릭하면 되고 이미 녹음하여 저장해둔 파일을 열려면 Open을, 파일을 닫을 때는 Close를 사용한다. 파일의 이름을 다른 이름으로 저장할 때는 Save As를 사용한다. Properties는 음성파일에 대한 화자이름, 녹음 날짜나 데이터베이스 이름, 단어 이름 등 필요한 정보를 입력할 수 있다. 인쇄는 Print기능을 이용한다. 네번째 그림

에서는 이미 열어본 최근의 파일목록이 보관되어 있어 빨리 열어볼 수 있다. Control키와 메뉴의 머리글자를 이용하면 키보드만으로 빠르게 실행할 수 있다. 마지막 Exit은 작업을 마칠 때 사용한다. 두번째의 View기본 메뉴아래에는 Toolbar와 Statusbar가 있는데 그림의 기본메뉴아래의 아이콘으로 나타난 도구들(Tool)을 나타내거나 가릴 때 사용한다. Statusbar를 체크해두면 그림의 맨 밑에 위치한 막대에 각각의 도구나 메뉴를 선택할 때마다 세부적인 기능 설명이 영어로 간략히 표시된다. 세번째의 Item기본메뉴아래에는 성문파형인 라링고그래프의 신호나 피치정보 파일 등을 가져오기 위한 Import와 분석결과를 삭제하는 Delete, 일부분을 복사하는 Copy, 파일끼리 서로 연결시키거나 복사하여 붙이기 등의 기능이 있고, 녹음기능인 Record, 재생할 때는 Replay, 재생을 일시에 중지하는 것은 Stop Replay, 분석결과를 모두 보이는 Display All, 또는 선택한 항목만 보이는 Display가 있다. 네번째의 Tools 기본메뉴 아래에는 Generate 하위메뉴에 정현파 등 시험용 음성파일을 생성하는 Test signals, 두개의 신호를 순서대로 합치는 Concatenate signals, 배음합성, MBROLA합성, 포먼트 합성 등의 메뉴가 있다. 이어서 첫번째 Speech 항목에는 편집하기 위한 Edit메뉴와 기본적인 스펙트로그램과 파형을 나타내주는 게시(Display)기능이 있고 세번째의 Process에는 저주파 통과, 고주파 통과 여과기를 사용하거나 음성의 표본속도를 변형하거나, 소음을 첨가하고, 피치를 변화시키는 기능들이 포함되어 있다. 네번째의 Analysis에는 음성의 강도의 포락선을 나타내는 Energy envelope, 기본주파수, 포먼트, LPC, 스펙트럼 분석과 8, 19, 26채널의 청각 척도로 된 여러가지 필터뱅크가 마련되어 있고, MFCC 분석, 유성성분석, 소음분석, 소리의 크기를 나타내는 에너지계적 등이 있다. Annotate는 음성파일에 필요한 발음기호나 단어를 입력할 수 있다. 영어는 물론이고 한글도 입력이 가능하다. 마지막에 있는 Export 기능을 사용하면 sfs파일을 wav형태로 저장할 수 있다. 긴 음성파일을 단어별로 표기한 다음 각각의 sfs 파일로 저장하거나 AIFF, AU 등의 다른 파일포맷으로 저장할 수도 있다. 피치분석을 한 뒤에 4번째의 Fx항을 이용하면 피치계적을 평균과 표준편차값을 이용해서 부드럽게 바꾸거나 MOMEL이란 모델로 피치값을 정확히 구하거나 단조로운 음성(buzz)의 변화로 바꿀 수 있다. 덧붙여, 측정된 여러 개의 피치값을 텍스트로 따로 저장하여 엑셀 등에서 불러와 겹쳐 그릴 수 있다. 11번 Coefficients메뉴에서는 각 계수값을 그레이스케일로 그림으로 나타내거나 필터뱅크나 19채널합성기로 변환하거나 텍스트파일로 저장할 수 있다. 19채널합성기로 음성을 재합성하려면, 먼저 음성을 선택하고 Tools의 Speech→Analysis→Fundamental frequency를 실행하여 피치값(Fx)을 구하고 이어서 Speech→Analysis→Filterbank→19-channel auditory filterbank를 선택하여 계수를 구한 뒤, COEFF와 FX의 앞에 있는 박스에 체크를 하고 Tools→11.Coefficients→19-channel synthesizer을 클릭하면 합성음(SPEECH vocsyn)이 생성된다. 이것을 선택한 뒤 재생단추를 누르면 완전하지는 않지만, 그런대로 좋은 합성음을 들어볼 수 있다. 찌그러진 듯한 합성음의 원인은 주로 잘못 측정된 포먼트값을 사용했기 때문이다. 12번 Formant메뉴에서는 포먼트 합성기에 입력할 자료로 변환되며 이것을 Display하여 실제 포먼트값으로 합성한 음성을 들어볼 수 있다. 그 아래에 있는 메뉴 가운데 Run SML Script는 분석과정을 스크립트로 작성하여 자동으로 분석시킬 때 사용한다. Display Tree는 지금까지 분석한 자료들의 관계를 쉽게 알 수 있도록 나타내어준다. List Items as Text는 분석된 파일들을 헤더정보와 함께 텍스트파일로 저장해준다. 다섯번째 기본메뉴인 Window에는 새로운 창을 열거나 이미 열려진 창을 차례로 가지런히 정돈하여 볼 수 있는 기능이 있으며 열려있는 창의 목록에서 선택할 수 있게 해 준다. 마지막으로 Help에

는 사용자메뉴얼이 항목별로 분류되어 있으며 클릭하면 자동으로 내용을 보여 주기 때문에 매우 편리하다.

다음 그림은 SFSWin을 이용하여 분석한 필자의 “일”라는 발음의 음성파형, 넓은대역스펙트로그램, 포먼트 추정값, 피치값을 나타내고 있다. 이 그림에서도 알 수 있듯이 피치값이 끝부분에서 갑자기 올라가 있는 것은 측정상 에러로 여겨지며, 제 1 포먼트 추정값도 중간부분에서 에러가 있는 것을 알 수 있다. 이러한 에러는 분석시 파라미터 설정을 조절하여 피할 수 있지만 비전문



가로서 측정값을 통계적인 비교에 사용할 때는 매우 주의하여야한다.

맺 음 말

지금까지 음성분석도구인 SFSWin의 각 메뉴와 사용법에 대해 설명했다. 지면상 기본적인 설명만 하였지만 프로그램이 매우 직관적으로 되어있으므로 사용하는데 큰 어려움은 없을 것으로 여겨진다. 이 프로그램에서도 파라미터 설정에 따라 측정상의 에러를 보이고 있기는 하지만 채널합성기능 등과 같은 다른 프로그램에서 볼 수 없는 기능이 있으므로 한번 사용해 보기를 권장한다. 이 프로그램을 이용하면 음성언어학 분야에서는 분석과 합성기능으로 연구에 활용할 수 있을 것이며, 특히, 음성공학 분야에서는 시험용 파일을 만들어 자신이 개발한 음성분석 프로그램이 믿을만한 분석결과를 보이는지 점검하거나, 음성분석 도구를 만들 때 사용자의 편의를 마련하기 위한 아이디어를 얻을 수 있을 것이다. 보다 더 전문적인 내용은 프로그램 저자가 실행 파일에 첨부해둔 사용자 메뉴얼을 이용하기 바란다.

본 연구는 보건복지부 보건의료기술진흥사업의 지원에 의하여 이루어진 것임(02-PJ1-PG10-31401-0005).

REFERENCES

- 양병곤. 음성학 실험의 실제. *The Dongeui International Journal* 1996; (1):111-21 (<http://www.deu.ac.kr/~bgyang/96prac.htm>).
- 양병곤. *Signalize*를 이용한 음성 분석법. 부산: 진영문화사;1998 (<http://www.deu.ac.kr/~bgyang/sigmanual.PDF>).
- 양병곤. 프리트를 이용한 음성 분석의 이론과 실제. 부산: 진영문화사;2003 (<http://www.deu.ac.kr/~bgyang/praat/praatdown.htm>).
- 양병곤. 청각 실험 도구 *Alvin* 사용법. 한국언어치료학회 제12회 학술발표회. 경주코오롱호텔;2003 (<http://www.deu.ac.kr/~bgyang/alvin.PDF>).