

# Graphical query interface for Hungarian Read Speech Precisely Labelled Parallel Speech Corpus Collection

## F-view-hu

This tool supports the visual study and correction of the formant values derived from the ParallelSpeech-hu database. It gives the picture of the spectrum of the given sentence, the sound boundaries and sound symbols, the place of measured formant values in 5 measuring points inside the vowel. The tool makes possible of searching according to sound environment, filtering the database items (eg. male /female) and interactive correction of the numerical formant values (in case of wrong formant measurement). The tool gives only visual information, sound not.

### Development

The construction of the tool started in 2010 at the Budapest University of Technology and Economics, Hungary, Department of Telecommunications and Media Informatics. The tool requires the R statistical program.

### The content of the database

The visual tool works from premade static picture files and from the txt file where the formant data are stored. On the visualised picture the user sees the spectrum of the sentence, the sound symbols and sound boundaries and the F1, F2, F3 formant values (in a dot format) on the 5 measuring point inside the vowel. The formant dots can be moved by hand, so a wrong placed format can be modified to the right place. This corrected formant value will be stored in the formant (txt) file for request.

The number of visualised formant values: 4 992 735

### Directory and file structure

The files needed to the tool are stored in two directories (spectrograms and the textgrids). The numeric data for the formant values is stored in a txt file.

### File types

Three file types are used

1. Spectrogram picture file: xx.png (a spektrogram picture of the sentence),
2. The text file of sound symbols and sound boundaries: xx.TextGrid
3. A text file of numerical formant data: xx.txt

### Sound symbols inTextGrid files

Special sound symbols are used in the phonemic description.

Hungarian vowels (letter= sound symbol): a=a, á=a1, e=e, é=e1, i=i, í=i1, o=o, ó=o1, ö=o2, ő=o3, u=u, ú=u1, ü=u2, ű=u3

Hungarian consonants (letter= sound symbol): b=b, d=d, g=g, gy=gy, p=p, t=t, k=k, ty=ty, m=m, n=n,

ny=ny, j=j, h=h, v=v, f=f, sz=sz, z=z, zs=zs, s=s, dz=dz, dzs=dzs, c=c, cs=cs, l=l, r=r. The phonological short/long contrast by the consonants is represented by the „:” character for long consonants (letter= sound symbol; for example: ssz=sz:, rr= r:)