

Di-phone database for text-to-speech conversion in Hungarian

Di-phone-hu

The Di-phone-hu contains sound sequences from which the diad elements can be extracted for a Hungarian TTS system.

Development

Collection of words and preparation for demonstration: Mátyás Bartalis, Tamás Gábor Csapó, Géza Németh, Gábor Olaszy, Bálint Pál Tóth, Csaba Zainkó at the Budapest University of Technology and Economics, Hungary, Department of Telecommunications and Media Informatics.

The content of the database

The database contains short sound sequences read by a female speaker. The goal is to cover all Hungarian sound combinations. Each sound combination is represented by one read sound sequence (like *aboka=bo* sound combination). The diphone waveform can be extracted from the sound sequence. The CV, VC, VV, CC diphone set serves as the sound element database for the TTS system. Connecting the appropriate diphones the synthetic baseline speech waveform can be concatenated.

Sound sequence types

Nonsense sound sequences incorporate the CV, VC, VV diphones (like *aboka=bo* sound combination). For CC combinations short words incorporate the given CC combinationa (like: *labda(ball)=bd* consonant cluster).

Voice: female

Number of sound sequences: for 37 speech sound and the pause, altogether 1455 sound sequences

Annotation: the annotation and labelling was carried out by hand.

Directory and file structure

Main directory's name: `di_phone_hu` (contains 8 subdirectories).

Each subdirectory contains the database units in different file formats described below. The filenames are always the same for the given unit. (eg. `diphone_female_01_000.wav`, `diphone_female_01_000.TextGrid`, `diphone_female_01_000.txt`).

File types

the wave form (`xx.wav`),

the orthographic text of the word with Hungarian letters ISO-8859-2 (`xx.txt`)

the Praat textgrid file containing the sound boundaries and the sound symbols (`xx.TextGrid`)

Sound symbols

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, 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:)

Other symbols:

start_sil= the end of start_sil is the position of first speech sound of the sound sequence,

end_sil= the beginning of the end_sil is the end of the last speech sound of the sound sequence