HTK (Hidden Markov Model Toolkit)

HTK:

 - nyílt forráskódú beszédfelismerő rendszer

 - htk.eng.cam.ac.uk

 - alapvetően C-ben írva

 - eredetileg linux alá írták, de némi macerával Windows alatt is megy

 - sok kis futtatható parancsból áll, ezek scriptekből hívhatók

 - dokumentáció: https://www.inf.u-szeged.hu/~tothl/speech/htkbook.pdf

 - elsőként egy egyszerű számfelismerőt fogunk megcsinálni:
 https://www.inf.u-szeged.hu/~tothl/speech/Numbers-Demo.zip

A HTK fő moduljai:
 - előfeldolgozás: HCopy

 - HMM akusztikus modellek tanítása: HInit, HRest, HeRest

 - nyelvi modell készítése: HParse, HBuild

 - kiértékelés: HResults

Egy beszédfelismerő alapvetően az alábbi komponensekből fog állni:

 - akusztikus modell (a tanító adatból tanuljuk)

 - nyelvi modell (megadható nyelvtanként vagy tanítható szöveges adaton)

 - kiejtési szótár (elkészíthető kézileg)

A HTK által használt fájlformátumok:

 - HTK szereti a szöveges fájlokat, ezekbe könnyen bele tudunk nézni

 - input jellemzők: .mfc format (ez bináris)

 - címkefájlok: szöveges formátum

 - kiejtési szótár: szöveges formátum

 - akusztikus modell: szöveges vagy bináris

 - nyelvi modell: szöveges vagy bináris

Előfeldolgozás:

 - HCopy.exe -C config input.wav output.mfc

 - Sok fájl feldolgozásához fájllista is átadható a -S filelist kapcsolóval

 - példa config file: preprocess.config --> MFCC features

Akusztikus modell tanítása:

 - ha van fonetikai szegmentálásunk: --> HInit.exe, HRest.exe

 - lásd pl. runhinit.bat és runhrest.bat

 - ha csak szószintű átirat van: HeRest.exe

Felismerés:

 - HVite.exe vagy HDecode.exe

 - lásd pl. runhvite.bat és runhvitep.bat

Kiértékelés:

 - HResults.exe

 - lásd pl. runhresults.bat és runhresultsp.bat