

## A magyar nyelv ejtésvariáció vizsgálata gépi beszédfelismerés segítésére

Szaszák György<sup>1</sup> – Vicsi Klára<sup>1</sup>

<sup>1</sup> Budapesti Műszaki és Gazdaságtudományi Egyetem, Távközlési és Médiainformatikai Tanszék, Beszédakusztikai Kutatólaboratórium. 1111 Budapest, Sztoczek u. 2.  
szaszak@tmit.bme.hu , vicsi@tmit.bme.hu  
<http://alpha.ttt.bme.hu/speech/>

**Kulcsszavak.** Beszédfelismerés, ejtésvariáció-modellezés, nyelvi modell.

Előadásunkban a számítógépes beszédfelismerés fonológiai szintjén használható ejtésvariáció modellezésről szeretnénk áttekintő képet nyújtani. Az emberi beszédet – különösen annak spontán, társalgásbeli formáját – a kiejtés szempontjából nagyfokú variáltság jellemzi, azaz ugyanazon szöveg(rész) többféle, de egyaránt helyes kiejtéssel is előfordulhat. A beszédfelismerők tévesztéseiben ennek a jelenségnek is nagy szerepe van, az ejtésvariáció modellezés célja tehát a felismerés hatékonyságának növelése.

Magyar nyelvre ejtésvariáció vizsgálatokat a Magyar Telefonbeszéd Adatbázis anyagán végeztünk, adatalapú, azaz statisztikai megközelítést használva, de célunk ennek alapján kiejtési szabálybázis megkonstruálása is. A vizsgálati metodika röviden a következő: elegendően nagyszámú beszélő által bemondott és speciálisan megszerkesztett szöveg hanganyaga van rögzítve, amely reprezentálja a magyar köznyelv bemondási variációit. A vizsgálat során a bemondott szöveg kanonikus ejtésének megfelelő fonetikus átírást összehasonlítjuk a szegmentálás során szakértők által elkészített audio-vizuális fonetikai átírással. Az eredményeket mátrixos formában kiértékelve kaphatjuk meg. Kanonikus kiejtés alatt egy referenciaként elfogadott (célszerűen a leggyakoribb) kiejtés értendő. A vizsgálatot beszédhangokra, majd hangkapcsolatokra végeztük el.

Az ejtésvariáció-vizsgálatok során nyert eredmények felhasználási lehetőségeiről az alábbiakat mondhatjuk el. Amellett, hogy az ejtés variáltságának mint jelenségnek beszédképzésbeli, illetve nyelvi, sőt információelméleti összefüggéseit is megismerjük, az eredmények közvetlenül két fő vonalon használhatóak fel. Az első lehetőség a felvételek szegmentálása és címkézése során használt fonetikai átírási szabályrendszer, illetve szimbólumkészlet kontrollálása, a második lehetőség a szerzett ismeretek beszédfelismerés során történő hasznosítása, a felismerés hatékonyságának növelése az ejtésvariációk modellezése révén. Mindennek jelentősége különösen a nyelvi modellezés során van, hiszen ehhez pontosan ismernünk kell az egyes kiejtésalakok gyakoriságát vagy előfordulásuk szabályszerűségeit.