

Mády Katalin – Mihajlik Péter:  
*Beszédatbázisok és leiratozási kutatások az NYTK–BME kooperációban*

Az elmúlt években egyre fontosabbá vált a spontán kommunikáció és a természetes beszélővariabilitás vizsgálata. Ezzel egyidőben a beszédtechnológiai alkalmazások is látványos fejlődésen mentek keresztül a mély neurális hálóknak köszönhetően, ami jelentősen megnövelte az adatmennyiség iránti igényt. Az előadásban szót ejtünk néhány nyelvi jelenségről, amik eltérő típusú korpuszok létrehozását teszik szükségessé. Ezután a legnagyobb magyar beszélt nyelvi korpusz, a Beszélt nyelvi Adatbázis (BEA) példáján bemutatjuk, milyen nehézségekkel találkoznak fonetikusok és beszédtechnológusok a hanganyag felvételétől a leírt változat kialakításáig. Ismertetjük a BEA-Base adatbázist, melyet kifejezetten gépi beszédfelismerő rendszerek tanítására és kiértékelésére alakítottunk ki. Megmutatjuk, hogy ezen adatbázis alapján, a legkorszerűbb, mesterséges mély neuronháló segítségével hatékony gépi beszédleiratozás valósítható meg a magyar nyelvű spontán beszédre. A legjobb eredményeket olyan, ún. felügyelet nélkül előtanított beszédakusztikai modellek alkalmazásával értük el, melyek előzetes betanítása csupán hanganyagot igényelt, leiratot nem. A legpontosabb, magyar, finn és észt európai parlamenti felszólalások hanganyagán előtanított, majd a BEA-Base tanítóhalmazán finomhangolt modellt beszélt nyelvi korpuszsal tanított neurális nyelvmodellel egészítettük ki, és kutatók számára ingyenesen elérhetővé tettük.

A kutatás az NKFIH K135038-as projekt keretében történik.