

## Záróbeszámoló az OTKA 108762 sz. pályázatról

A beszédatadabázis felvételeinek befejező annotálási munkálatait elvégeztük, megtörtént az átiratok többszörös ellenőrzése, szükség esetén a korrekciójuk a Praat programban. A teljes, rögzített beszédanyag már az elmúlt évben (a negyedik év lezárultával) hozzáférhetővé vált a kutatók számára; a hozzáférés folyamata tapasztalataink szerint jól működik.

A pályázati tervek teljesítése megtörtént, ezeket az alábbi pontokban összegezzük.

1. Adatbázis-fejlesztés és a beszédanyagok annotálása. 120 beszélő (84 nő, 36 férfi) felvételét rögzítettük a projekt időtartama alatt. A beszélők életkora 18 és 78 év között változott. A felvételek teljes időtartama 87,56 óra; az interjúkészítő az esetek többségében ugyanaz a személy volt, a felvételek átlagidőtartama mintegy 45 perc. Kialakítottuk a felvételekhez tartozó adattáblázatot (excelben), amely tartalmazza a beszélők anonimitást biztosító számát, életkorát, iskolai végzettségét, természetét, súlyát, dohányzási szokásait, esetleges beszédhibáját, valamint a spontán beszédfelvételek témaköreit. Az általunk kialakított kritériumok mentén létrehoztuk 120 beszédfelvétel annotált szöveganyagát a Praat programban.

2. Open access – ingyenes kutatói hozzáférés a létrehozott anyaghoz. A pályázat egyik fő célját – a tervek megfelelően – a múlt év végén teljesítettük azzal, hogy a BEA adatbázis részeként létrejött új részadatbázis lekérdezhető és ingyenesen hozzáférhető a kutatók számára a megfelelő feltételek teljesítésével. (Az open access működtetéséhez a tárhelyet az MTA Nyelvtudományi Intézet szervere biztosítja.)

3. Tervezett nemzetközi workshop. A 2017-es évre terveztük egy nemzetközi workshop megszervezését; ez a *CAPSS (Challenges in Analysis and Processing of Spontaneous Speech)* elnevezésű szakmai esemény 2017. május közepén zajlott le.

4. Kutatások. A pályázat időtartama alatt számos, különféle kutatás folyt az adatbázison a zöngképzéstől, a tempóvizsgálatokon, avagy a frázisvégi nyúlás jelenségének kutatásán át a narratívák sajátosságainak elemzéséig. A még le nem zárult, folyamatban lévő kutatások egy része befejeződött az ötödik évben, más részük befejezés előtt áll. A kutatásaink eredményességét az elfogadott előadások és a megjelent tanulmányok és könyvek mutatják.

5. Előadások és publikációk. Az utolsó évben három hazai és három nemzetközi konferencián tartottunk előadást, utóbbiak egyike plenáris előadás volt. A tárgyévben további két hazai és két külföldi konferenciára fogadtak el előadásunkat. Ugyancsak az ötödik évben fogadták el a javasolt előadássorozatunkat a Magyar Tudomány Ünnepeinek rendezvényeként az Akadémia székházában. Ezek az előadások az adatbázis anyagán készült kutatások eredményeit mutatják be.

Terveztük a nemzetközi workshop előadásaiból írt, lektorált tanulmányok publikálását egy önálló, angol nyelvű kötetben és a *The Phonetician* c. nemzetközi folyóiratban. A kéziratok feldolgozását, bírálatási folyamatát (peer-review process) elvégeztük. A tervezett angol nyelvű kötet megjelent. Egyúttal elkészült a *The Phonetician* 115. (megjelenik 2018-ban) száma, amely több tanulmányt közöl a workshopon elhangzott előadások anyagából.

Az ötödik évben további 8 tanulmány, közülük 6 angol nyelvű és 2 magyar nyelvű könyv látott napvilágot a projekt keretében. Mindkét kötet monográfia, az egyik a *Morfémák időzítési mintázatai a beszédben*, a másik a *Megakadásjelenségek a magyar spontán beszédben* címet viseli. Elemzéseik a pályázat keretében fejlesztett beszédanyagokon folytak. A harmadik megjelent könyv a (már említett) nemzetközi workshop előadásaiból válogatott, külföldi szakemberek bírálta, angol nyelvű szerkesztett kötet, a címe: *Challenges in analysis and processing of spontaneous speech*. A publikált kötetek a honlapunkon is elérhetők és letölthetők, így az érdeklődők számára ingyenesen hozzáférhetők.

Az OTKA-pályázatunk időtartama alatt összesen 51 tanulmány és 4 könyv látott napvilágot. A tanulmányok közül 23 angol nyelven jelent meg. Folyamatban van továbbá öt tanulmány megjelenése (ezek tartalmazni fogják a pályázat számát). Hasonlóképpen a megjelentetett kötetekhez, a tanulmányaink jelentős része is hozzáférhető a kutatók számára (hacsak azt a kiadó nem tiltja). Mindösszesen 35 magyar és 33 külföldi szakmai eseményen tartottunk előadást az elmúlt évek alatt.

Az a tény, hogy ez az új részadatbázis létrejött, és ingyenesen hozzáférhető a kutatóknak, rengeteg kutatási energia megtakarítását jelenti. A relatíve nagy beszédanyag, a különböző beszédműfajok, az adatközlők széles életkori skálája nagyon sokféle kutatás elvégzését teszik lehetővé. Tekintettel arra, hogy egyetlen percnyi beszédanyag annotálása mintegy két órányi időt vesz igénybe, a létrejött annotált beszédanyagok jelentős időmegtakarítást biztosítanak a célzott vizsgálatokhoz a kutatóknak.

A pályázat kitűzött célját elértük, a vállalt feladatokat maradéktalanul teljesítettük. Fontosnak tartjuk, hogy két, most elnyert NKFI-pályázat közül az egyik teljes egészében, a másik részben a jelenleg befejeződött projekt beszéd- és szöveganyagait fogja felhasználni kutatásaiban.