

Háromfős társalgások annotálása a BEA-adatbázisban: elvek és kihívások

Horváth Viktória¹, Krepsz Valéria¹, Gyarmathy Dorottya¹,
Hámori Ágnes¹, Bóna Judit², Dér Csilla Ilona³, Weidl Zsófia²

¹Nyelvtudományi Intézet

²Eötvös Loránd Tudományegyetem BTK

³Károli Gáspár Református Egyetem BTK

The aim of this paper is to demonstrate the annotating method of the triadic conversations in the Hungarian Spontaneous Speech Database. The protocol of the annotation takes into consideration the phonetic and the pragmatic aspects of the analysis of the conversations supported by the project NKFI K-128810. The aim of the project is to analyse speech units, such as turns, turn-takings or overlapping speech, from multiple aspects. The large data sample allows describing synchronous changes such as the formation of discourse marker combinations. Furthermore, the aim of the project is a description of speech accommodation. The research has further scientific benefits as a result of the cooperation of several linguistic disciplines, such as a novel description of the structure of triadic conversations. Theoretically, if discourses are performed according to a particular system, they can be modeled technologically. So, in addition to the theoretically findings, the results may be applied in speech technology.

Keywords: conversations, annotation, Hungarian Spontaneous Speech Database

Kulcsszavak: társalgások, annotáció, BEA Spontán Beszéd Adatbázis

Bevezetés

Jelen tanulmány középpontjában a BEA-adatbázis (Gósy et al. 2012) társalgási moduljának mint a magyar nyelv jelentős méretű beszélt nyelvi adatbázisának bemutatása és annotálási elveinek ismertetése áll, különös tekintettel ezeknek a fonetikai és pragmatikai szempontokat ötvöző, társalgási jelenségekre is kiterjesztett annotálására, amely az NKFI K-128810 projekt keretében történik. A tanulmány célja, hogy a BEA társalgási adatbázis ismertetése mellett betekintést nyújtson az annotációs rendszerébe, szintjeibe és speciális szempontjaiba; egyben ráirányítsa a figyelmet a magyar nyelvű társalgások korpuszalapú vizsgálatának – ezen belül a fonetikai és társalgáselemzési, illetve pragmatikai szempontok összekapcsolásának, valamint a formális és funkcionális felfogások összehangolására törekvő megközelítés – fontosságára és módszertani lehetőségeire.

Az emberi nyelvhasználat egyik alapvető formáját a két vagy több személy között folyó kötetlen beszélgetések, társalgások jelentik, melynek során a beszéd-

partnerek jellemzően közvetlen interakciós helyzetben vannak, és a résztvevői szerepek (megnyilatkozó és befogadó) folyamatosan cserélődnek (a társalgások jellemzésében itt elsősorban a konverzációelemzés fogalmaira támaszkodunk, l. Schegloff 1968, Sacks–Schegloff–Jefferson 1974, Boronkai 2009, Eshghi–Healey 2009). A beszélők váltakozása során jönnek létre a társalgások elemi egységének tekinthető beszédfordulók („turns”, Sacks–Schegloff–Jefferson 1974: 696), amelyekhez a beszédjoggal bíró beszélő fogalmát is kapcsolhatjuk; ezek mellett a beszédjoggal nem bíró résztvevők részéről is megjelenhetnek közlések mint háttéracsatorna-jelzések („backchannels”) (Yngve 1970, magyarra Bata 2009, Dér 2012, Hámori 2014, Gyarmathy et al. megj. e.). A társalgások jellegzetességei között említhető a beszélőváltási rendszer gördülékenysége, a fordulóváltás gyorsasága és többnyire rövid szünetek vagy rövid egyszerre beszélések megjelenése (“overlaps”, magyarra Bata 2009, Boronkai 2009, Hámori–Horváth 2019). A társalgások szerkezeti jellemzői kapcsán említhető a beszélgetés lokális jellegű, szekvenciális szerveződése, vagyis egyes fordulónak minimálpárokká és/vagy hosszabb szekvenciákká való összekapcsolódása (pl. kérdés–válasz, köszönés–visszaköszönés-párok, vö. Schegloff 1968), valamint egy átfogó vagy makroszintű, nagyobb funkcionális egységek alapján szerveződő (többféleképpen is jellemezhető) globális szerkezet (pl. bevezetés/belebonyolódás, középrész, lezárás vö. Schegloff 1968, Clarke–Argyle 1997). Ezek mellett számos más társalgási jelenség is leírható (pl. javítások, lehetséges beszélőváltási pontok, beszédaktusok használata stb.), melyekkel az elmúlt évtizedekben nagyszámú konverzációelemző, interakcionális szociolingvisztikai, beszélésnéprajzi és pragmatikai kutatás foglalkozott. Ezekben meghatározó volt a társalgásnak a társas interakciók felől történő vizsgálata és a felismerés, hogy a társalgás ugyanúgy szabályok mentén szerveződik, ahogy az interakció más – verbális vagy nem verbális – formái is; szerveződése és jelenségei elválaszthatatlanok a társas cselekvések és jelentések síkjától, valamint az ezek mögött álló szociokulturális normáktól, illetve pszichológiai tényezőktől. Ezek a különböző szempontú kutatások mára jelentős részben beépültek a társalgáselemzés alapjaiba: kutatások ezekre vonatkozóan számos – részben eltérő fókuszú, de egymással összekapcsolódó – jellegzetességet tártak fel (pl. univerzális vonások és kulturális különbségek a beszélőváltásban; a társalgásokon belüli különböző beszédműfajok pl. történetmesélés, vicc, pletyka; résztvevői szerepek és társas jelentések stb.).

A társalgás általános jellemzése kapcsán végül lényeges kitérni arra is, hogy a prototipikus társalgásokban – ilyenek például a családi vagy baráti beszélgetések – meghatározó vonás a spontaneitás, vagyis az előre el nem tervezettség vagy a tervezettség alacsony foka, valamint az informális. Léteznek kevésbé spontán és formálisabb, illetve intézményes társalgások is (pl. orvos-beteg beszélgetések, bírósági tárgyalások): a két végpont (teljesen spontán – nem spontán, informális – formális) között számos átmenet létezik. A társalgás fogalmát

itt kiterjesztett értelemben használjuk (Schegloff nyomán, l. Schegloff 1968, Hámori 2006), amely a fenti átmeneteket is magában foglalja; az itt tárgyalt BEA-társalgások ezen belül egy szűkebb típust képviselnek.

Nemzetközi társalgási beszédkorpuszok és -adatbázisok¹

A társalgások empirikus vizsgálatához fontos segítséget adnak a különféle beszéd- és társalgáskorpuszok, illetve beszédatbázisok. Ezek többféle kutatási keretben és céllal (pl. szociolingvisztikai, fonetikai, korpusznyelvészeti vagy pszichológiai kiindulással), különböző módszertani elvek mentén jöhetnek létre, ami kihat tágabb felhasználási lehetőségeikre. Több beszédatbázist eleve olyan célra hoztak létre, illetve protokolljukat úgy alakították, hogy például a narratívák és felolvasások mellett valódi társalgások is legyenek bennük, így fonetikai, szintaktikai és társalgáselemző kutatásokra egyaránt alkalmasak. Az egyik legismertebb külföldi beszédatbázis, amely társalgásokat is tartalmaz, a London–Lund korpusz. Ez félmillió brit angol szóból áll, és többféleképpen – a beszédpartnerek ismertségét, a beszédhelyzetet és a csatornát variálva – rögzített társalgásokat tartalmaz (Svartvik ed. 1990). Az eredeti korpuszt 1959 és 1990 között rögzítették, a London–Lund korpusz 2 készítése folyamatban van. Beszédfelismerő rendszerek fejlesztéséhez készült az USA-ban a szintén angol anyanyelvűekkel rögzített CallHome (Canavan et al. 1997) beszédatbázis, amely 120 telefonbeszélgetést foglal magában. A Switchboard telefonos beszédkorpusz (Godfrey–Holliman 1993) 543 amerikai beszélő mintegy 260 órás dialógusát tartalmazza. Specifikus beszédhelyzetekben rögzített társalgási korpuszok is ismeretesek. Ilyen például a térképmódszerrel skót dialógusokat rögzítő HCRC Map Task Corpus, a repülőjegy-rendeléseket tartalmazó ATIS adatbázis (Hemphill et al. 1990), avagy a célzott társas feladatokat rögzítő Kiel Corpus of Spontaneous Speech (összefoglalásukra lásd Gósy et al. 2012). Spontán, több résztvevős társalgásokból álló korpuszokból, például konferenciamegbeszélések, illetve egyetemi órák és interaktív beszélgetések felvételeiből dolgoztak az AMI (Augmented Multiparty Interaction), a CHIL (Computers in the Human Interaction Loop) és a CALO (Cognitive Assistant that Learns and Organizes) projektek (összefoglalva lásd pl. Bechet 2012). Szintén felsőoktatási környezetben rögzített társalgásokat tartalmaz a MICASE (Michigan Corpus of Academic Spoken English, vö. Simpson et al. 2002). Többféle beszédhelyzetben (így társalgási helyzetben is), egynyelvű és kétnyelvű gyermekekkel készültek felvéte-

¹ A *korpusz* és az *adatbázis* fogalmak mind a hazai, mind a nemzeti szakirodalomban keveredve használatosak, a különböző tudományterületi háttértől, megközelítéstől, elméleti-módszertani keretektől függően (pl. Sass 2016). A jelen tanulmányban törekedtünk az eredeti megnevezésnek megfelelő fogalom/megnevezés használatára. A BEA hagyományosan *adatbázis* megnevezéssel szerepel, de a társalgási modul anyaga annotált korpusznak tekinthető.

lek a SALT adatbázisban (<http://www.saltsoftware.com/salt/databases/#>). Számos (jelenleg 26) nyelvből származó gyermeknyelvi diskurzusrészleteket tartalmaz az 1984-ben alapított CHILDES (Child Language Data Exchange System, <https://childes.talkbank.org/>). Célzott nyelvészeti elemzésekhez, például a megakadások vizsgálatához is rögzítettek nagy mennyiségű kétszereplős társalgást. Bortfeld et al. (2001) például az életkor, a nem, a beszélők közötti kapcsolat, illetve a beszédtema hatását vizsgálta a megakadások előfordulására 48 társalgás mintegy 192000 szavában. Gyakori az is, hogy rádió- vagy tévéműsorok beszélgetéseit használják elemzésre, ilyenkor nagyon változatos lehet a társalgásban résztvevők száma.

Az utóbbi időben a korpusznyelvészetben egyre inkább nő az érdeklődés a kisebb, speciális téma vagy műfaj köré szerveződő, kontextuálisan gazdagabb feldolgozást lehetővé tevő korpuszok iránt, amelyek eddig ilyen módszerekkel kevésbé kutatott területek és jelenségek vizsgálására alkalmasak; a társalgás jellegzetességei, egyes pragmatikai jelenségek vagy fonetika és pragmatika kapcsolata kifejezetten ilyennek számít (vö. Szirmai 2005, O’Keeffe et al. 2007, Vaughan–Clancy 2013, Walsh 2013).

Hazai társalgási beszédkorpuszok és -adatbázisok

A magyar nyelvű kötetlen beszélgetések korpuszalapú vizsgálata felé az első lépéseket az ELTE BTK Mai Magyar Nyelvi Tanszéke által az 1970–80-as évektől gyűjtött beszéltnyelvi felvételek jelentették (több beszédműfajt is tartalmazva, pl. előadások, baráti beszélgetések, bizottsági ülések, kollégiumi társalgások, vö. Keszler 1983), ezekből azonban csak később és kisebb részletek kapcsán születtek elemzések. Fontos állomást jelentett az 1980-as évek második felében a gazdagréti lakóközösség kábeltelevíziójának műsorában elhangzó többszereplős beszélgetések rögzítése és elemzése, amelyek átiratában az intonációs jelölések is központi szerepet kaptak (Varga 1987; Kontra 1988, 1990). 1987 és 1989 között jött létre az első magyar szisztematikusan létrehozott beszélt nyelvi korpusz a BUSZI keretében (Budapesti Szociolingvisztikai Interjú, I. Váradi 2003), a szociolingvisztikai-stilisztikai változatosság vizsgálata céljából. A korpuszt 50 szociolingvisztikai interjú alkotja, amelyek labovi mintára előzetesen megtervezett kérdésekkel irányított társalgásokat tartalmaznak, különböző szociolingvisztikai csoportba sorolható beszélők részvételével. A kérdések alkalmazásának fontos célja volt, hogy a beszélgetés során eltereljék az interjúalany/adatközlő figyelmét a saját beszédtevékenységéről, és előhívják a hétköznapi nyelvhasználatát, társalgási fordulatait. A korpusz annotálása elsősorban szociolingvisztikai jelenségekre vonatkozik, de korlátozott mértékben fonetikai, stilisztikai társalgási jelenségek vizsgálatára is alkalmas.

A magyar nyelvű társalgások fonetikai elemzésének egyik legfontosabb hazai mérföldköve Markó Alexandra 2005-ös doktori disszertációja volt, amelyhez

egy négyszereplős társalgást rögzített. A hanganyagon többféle vizsgálat is készült, főként temporális és intonációs elemzések (Markó 2005), valamint pragmatikai szempontúak is (Dér–Markó 2010). A nyelvészeti vizsgálatokhoz rögzített korpuszok közül megemlítendő még a Szögedi Szociolingvisztikai Interjú (SZÖSZI), amelynek irányított interjúin készültek pragmatikai elemzések is (Schirm 2018; 2019). Emellett specifikus helyzetekben rögzített társalgásokat tartalmaznak az egyetemi kollégiumokban készített, illetve osztálytermi beszédkorpuszok (Budapesti Egyetemi Kollégiumi Korpusz (BEKK, Bodó et al. 2017, Antalné 2006, Asztalos 2019; megj. e.), valamint egyes gyermekbeszédadatbázisok mint a MONYEK (Magyar Óvodai Nyelvi Korpusz, Mátyus–Orosz 2014) és a GABI (Gyermeknyelvi beszédAdatBázis és Információtár, Bóna et al. 2014). Ezek egy része dialógust, más részük többszereplős társalgást is tartalmaz. Asztalos (2019 megj. e.) multimodális korpuszának különlegessége, hogy nemcsak kamerákkal rögzítette a vizsgált tanórákat, hanem a tanulók közé elhelyezett diktafonok segítségével a tanulók közötti beszélgetésekről is készített felvételt. Emellett léteznek más, kisebb, speciális céllal gyűjtött társalgási korpuszok is (pl. orvos-beteg interakciók, politikusi viták anyagából), ezekre itt nincs mód kitérni.

A fenti korpuszok létrehozásában, valamint az ezeken végzett kutatásokban több nyelvészeti irányzat is megjelent; elsősorban a szociolingvisztika, társalgáselemzés és pragmatika szerepe emelhető ki, de születtek pszicholingvisztikai, stilisztikai és fonetikai elemzések is. Akusztikai-fonetikai vizsgálatra azonban a legtöbb említett korpusz nem, vagy csak korlátozottan használható.

Az első jelentős méretű, fonetikai mérésekre is alkalmas, sőt multimodálisan is vizsgálható magyar nyelvű párbeszédkorpusz a Debreceni Egyetemen a HuComTech projekt keretében jött létre (Hunyadi et al. 2012), kommunikációelméleti szakemberek, digitális képfeldolgozók és számítógépes nyelvészek közreműködésével. A projekt központjában az ember-gép kommunikáció fejlesztése és az ember-ember-kommunikáció ezt segítő vizsgálata áll, ennek érdekében egy olyan empirikus adatbázis létrehozása volt a cél, amely alkalmas alapanyagot ad a nyelvtechnológiai és robotikai kutatásokhoz és fejlesztésekhez, valamint más, pl. kommunikációelméleti, nyelvészeti, pragmatikai és viselkedéskutatási vizsgálatokhoz is.

A HuComTech korpusz összesen 50 órányi anyagot tartalmaz, 112 beszélő részvételével készült, részben formális, részben informális kontextusú dialógusok audio- és videofelvételei alapján. A korpusz annotálásának célja elsődlegesen a számítógépes feldolgozhatóság volt, ehhez többféle megközelítésben (vizuális jelek, nyelvi egységek és kommunikációs események megfigyelése) címkézték fel a vizsgált jelenségeket: készültek unimodális (videó és audio), valamint multimodális (videó + audio) annotációk, kézi és automatikus módszerekkel is. A modell elméleti hátterét a kommunikáció egy elméleti-technológiai

modellje adja, amely egy generatív grammatikai, moduláris nyelvszemléletre épül (Hunyadi et al. 2012). Az annotálás a vizuális jelzések (pl. tekintet, gesztusok, érzelmkifejezés), és multimodális jelenségcsoportok mellett tartalmazza az audiofelvételek prozódiai szintű ("IP-szint", intenzitás, hangsúlydetektálás), valamint szintaktikai szintű címkézését (pl. tagmondatok belső szerkezete, hierarchiája) (Hunyadi et al. 2012, Kiss 2011). Az annotáció kiterjed továbbá a – moduláris keretben értelmezett – pragmatikai szintre is, ennek kapcsán olyan jelenségek címkézésére, mint az egyetértés, a figyelem, az információstruktúra, a deixisek, ún. kommunikatív események, valamint a társalgási fordulószerveződés néhány alapvető jellemzője (pl. beszélés kezdete, vége, közbevágás, szóátvételi szándék). A különböző szintű és modalitású annotációk számos elemzést és fontos multimodális együttállások megállapítását tették lehetővé (Abuczki 2011, Németh T. (szerk) 2011, Pápay 2011), és sok új információt kínálnak a magyar társalgások leírásához (pl. a beszélőváltások vagy az információstruktúra és a nonverbális jelzések kapcsolata terén) és számítógépes modellezéséhez is. Bár az elméleti háttérmodell nyelvtechnológiai meghatározottsága, valamint generatív alapú, moduláris pragmatikafelfogása miatt a korpusz más, eltérő elméleti háttérű nyelvészeti vagy pragmatikai kutatásokra (pl. funkcionális pragmatikai, társas interakciós, társas kognitív elemzésekre) való alkalmazhatósága korlátozott lehet, a magyar nyelvű társalgások megismeréséhez egyértelműen jelentős segítséget kínál. A HuComTech-korpusz és a BEA társalgási korpusza több aspektusban hasonlóságokat mutat (l. alább), ami lehetővé teszi az eredmények összemérhetőségét és összekapcsolását, ugyanakkor lényeges különbségek is vannak köztük (elsősorban multimodalitás/unimodalitás, a társalgó felek száma, illetve az annotációk elméleti háttere, fókusza és részletei alapján), amelyek révén fontos kiegészítéseket kínálnak egymás eredményeihez.

A BEA Spontán Beszéd Adatbázis

Az MTA Nyelvtudományi Intézet Fonetikai Osztályán a modern fonetikai kutatási és beszédrögzítési követelmények kiszolgálásának céljával 2007-ben indult meg a BEA Spontán Beszéd Adatbázis fejlesztése (Gósy et al. 2012). Ennek elsődleges célja többféle típusú spontán beszéd rögzítése volt, de a fonetikai kutatási célok (összehasonlíthatóság) érdekében mondat- és szövegfelolvasásokat, illetve mondatismélteléseket is tartalmaz. Az adatbázis elsődlegesen nyelvészeti alap- és alkalmazott kutatások céljára készült: elsősorban olyan fonetikai vizsgálatok elvégzésére, amelyekhez megfelelő minőségben rögzített felvételek szükségesek, de az egyes részfeladatok során olyan multifunkcionális szövegek is létrejöttek, amelyek alkalmasak több különféle nyelvészeti terület (pl. pragmatika, szocio- és pszicholingvisztika) számos jelenségének vizsgálatához is. Az adatbázis fejlesztői kiemelten ügyeltek a tudományetikai és adatvédelmi elvek

betartására², ennek megfelelően a hanganyagok anonimizáltak, az adatközlők kódokkal szerepelnek az adatbázisban, személyük azonosíthatatlan. Minden felvétel esetén rögzítésre kerültek olyan metaadatok, kiegészítő információk, amelyek kutatásmódszertani szempontból elengedhetetlenek: ilyen az adatközlő neme, életkora, iskolai végzettsége, foglalkozása, súlya, magassága, dohányzási szokásai, a véleménykifejtés és a társalgás központi témája, a felvétel időtartama, illetve a felvételvezető és a beszélgetésben részt vevő harmadik partner személye. Ezek az adatok lehetőséget nyújtanak az adatbázis széleskörű kutathatóságára. Fontos sajátosság továbbá, hogy a felvételek mintegy 85%-ában ugyanaz a személy az interjúkészítő, ami kiváló lehetőséget nyújt további, speciális vizsgálatokra, például a beszédalkalmazkodás vagy a beszélői szerepek dinamikus alakulásának kutatására is.

Az adatbázis hangfelvételei egységes protokoll szerint készülnek, amely meghatározza a rögzítendő beszédanyagok tartalmi vonatkozásait, technikai jellemzőit, illetve a beszélők számát. Az adatbázis főként különféle típusú spontán beszédanyagokat tartalmaz (a spontán fogalmat fonetikai értelmezésben használva), de a szélesebb kutathatóság érdekében mondatisméltéseket és felolvasásokat is magában foglal. A tartalmi protokoll a következő hat részből áll: mondatisméltés, narratíva, véleménykifejtés, tartalomösszegzés, társalgás, felolvasás. Az egyes részek tartalma a következő:

1. A mondatisméltés 25 egyszerű és összetett mondatot tartalmaz, melyek összeállításánál a fejlesztők ügyeltek a grammatikai szerkezet, a szórend és a koartikulációs szabályok változatosságára, például: *Az ügyfeleknek kompromisszumot kellett kötniük.; Nem lehetett teljes bizonyossággal megítélni a vádlott elmeállapotát.* A mondatok átlagosan 8–12 szóból állnak, melyek átlagosan 3–4 szótagosak.
2. A narratíva az adatközlő életéről, családjáról, munkájáról, hobbijáról szóló, többé-kevésbé összefüggő monologikus szöveg.
3. A véleménykifejtés (jobbára szintén narratíva) az interjúkészítő által megadott téma véleményezése. A felvetett témák általánosak és hétköznapiak, aktuális társadalmi, közéleti kérdésekkel kapcsolatosak, a róluk történő véleménykifejtés sem konkrét ismereteket, sem háttértudást nem kíván. Példák: a házasság vs. együttélés; eutanázia; közlekedés a fővárosban; ittasan vezetés; a magántulajdon védelme; új adók, új szabályozások; az emberi szervek felajánlása.
4. A tartalomösszegzés irányított spontán beszéd. Az adatközlőnek egy-egy felvételről meghallgatott szöveget kell a saját szavaival elmondania. Az

² A Nyelvtudományi Intézetben érvényben lévő a *Humán vizsgálatokon alapuló nyelvészeti kutatások etikai szabályozása*, az „1995. évi CXIX. törvény a kutatás és a közvetlen üzletszerzés célját szolgáló név- és lakcímadatok kezeléséről”, valamint az „1992. évi LXII. törvény a személyes adatok védelméről és a közérdekű adatok nyilvánosságáról”.

- egyik egy rövid tudományszerűsítő szöveg (174 szavas; 1 perc 37 mp tartamú), a másik egy anekdotaszerű történet (270 szavas; 2 perc 5 mp tartamú). Ez a rész teljesen monologikus, a kísérletvezető nem szólal meg közben.
5. A társalgási modulban három személy részvételével tematikailag bizonyos mértékig irányított, ugyanakkor alapvetően informális és sok szempontból kötetlen beszélgetésre kerül sor, amelyben az adatközlőn és az interjúkészítőn kívül egy további személy (szintén a projekt munkatársa) is részt vesz. Az interjúvezető által felvetett téma változó, az élet mindennapjaihoz kapcsolódik; ugyanazon adatközlő esetében mindig különbözik a véleménykifejtés témájától. A társalgás témáiból: karácsony, húsvét ünnepelése; mobiltelefon kisgyermekeknek; új KRESZ és a biciklisek; halálbüntetés; dohányzás; éjszakai élet, szórakozási lehetőségek Budapesten; állattartási kultúra Magyarországon. A téma kiválasztása az adatközlő életkorának és érdeklődési körének figyelembevételével történik. Az e részben rögzített beszélgetések egyben egy önálló társalgási (al)korpuszt is alkotnak, ez áll "BEA társalgási korpusz" néven jelen tanulmány középpontjában is (l. alább).
 6. Az utolsó részben az adatközlő kétféle szöveget olvas fel; az egyik a huszonöt, korábban ismételt mondat, a másik egy 291 szóból álló tudományszerűsítő cikk felolvasása.

A beszédrögzítés technikai paramétereit tekintve a legfontosabb követelmény a BEA adatbázis esetében a megfelelő jel/zaj viszonyú, széles frekvenciatartományú, magas dinamikájú, torzításmentes felvételek elkészítésének biztosítása; amely akusztikai szempontból megköveteli a zajszigetelt és visszhangmentesített helyiséget, elektronikus tekintetben pedig a megfelelő hangfelvevő és hangrögzítő rendszert. Mindezek miatt a korpusz hangfelvételei az MTA Nyelvtudományi Intézet Fonetikai Osztályán található zajszigetelt szobában készülnek; a hangcsillapítás mértéke a külső környezethez képest 50 Hz-en 35 dB, 250 Hz fölött pedig ≥ 65 dB. A kritériumainknak megfelelő technikai háttérrel az Audio-technika AT4040 típusú kardioid kondenzátor mikrofonok, a Phonic AM642D típusú 10 csatornás phantomtápos kompakt keverőpult, továbbá a GoldWave szoftverrel, 44,1 kHz-es mintavételezéssel közvetlenül a számítógépre történő digitális, .wav formátumú rögzítés biztosítják.

A BEA adatközlői egynyelvű, köznyelvet beszélő budapestiek, jelenleg 461 beszélő, közülük 281 nő és 180 férfi, átlagéletkoruk 36,8 év (18 évestől 90 évesig). A beszélők iskolai végzettségét tekintve a felsőfokú végzettségűek vannak többségben (54,66%), illetve magas az érettségizettek aránya is (42,08%). Az adatközlők kis százaléka (3,25%) rendelkezik csak általános iskolai végzettséggel. Foglalkozásukat tekintve beszélőink között található általánosan ismert munkaköröket betöltők, például pedagógusok, gyógypedagógusok, orvosok, au-

tószerelők, informatikusok, adminisztrátorok, színészek; de akadnak közöttük viszonylag ritkább foglalkozásokat űzők, például: cserépkályha-készítő, jelmeztervező, katolikus pap, evangélikus lelkész, orgonaépítő, pókerjátékos, forgatókönyvíró, sajtótitkár és kaszkadőr.

A BEA-adatbázis jelenlegi 461 felvétele 367 óra, 28 percnyi hanganyagot jelent. A legrövidebb felvétel 19 perc és 17 másodperc hosszú, míg a leghosszabb ennek a hétszerese, 2 óra 24 perc 47 másodperc. Egy felvétel átlagosan 47 perc 50 másodperc időtartamú.

A jelen projektről

A BEA felvételeken belül speciális csoportot alkotnak a három személy részvételével folyó társalgások, amelyek együttesen egy önálló társalgási korpuszként is vizsgálhatók. Az NKFI K-128810 (A beszédegységek fonetikai jellemzőinek összefüggései) pályázat célja a társalgási modul alapján a fonetikai paraméterek vizsgálata és változásaik leírása valós kommunikációs helyzetben folyó beszélgetésekben, fonetikai és pragmatikai szempontokat ötvözve, valamint az ehhez szükséges főbb annotációk létrehozása. A projekt és az annotáció célja kettős: egyrészt a fonetikai kutatások számára kíván új szempontokat és újszerű megközelítésű mérési eredményeket adni, másrészt a társalgáselemzés (és/vagy pragmatikai kutatások) számára néhány, a területeken kiemelt jelenség vizsgálatát fonetikai adatokkal és elemzésekkel segíteni. A társalgási szituációban rögzített felvételek fonetikai jelentősége többek között, hogy jobban tükrözik a mindennapi nyelvhasználatot, mint a monologikus, más kontextusú felvételek. A beszélgetések többszintű, társalgáselemzési jelenségekre is kiterjedő annotálása és az ezeken alapuló célzott elemzések révén pedig alkalom nyílik olyan alapvető – fonetikai keretben magyar nyelven eddig keveset vizsgált – társalgási jelenségek árnyalt és több szempontú vizsgálatára, mint a beszédforduló, a beszélőváltás vagy az egyszerre beszélés fonetikai jellemzői és ezeknek különféle társas és kognitív tényezőkkel való összefüggései. A magyar nyelvre eddig sporadikus vizsgálatok születtek a beszélőváltásokkal kapcsolatban, a szupraszegmentális szerkezet és a szintaxis együttes működését vizsgálva, a projekt egyik fő célja ezek kapcsán kiterjedtebb elemzések lehetővé tétele illetve végzése. A verbális társalgásszervező elemek közül a diskurzusjelölők több szempontú elemzése szintén fontos kutatási cél; a korpuszon lehetőség nyílik olyan szinkron változás leírására, mint a diskurzusjelölő-társulások létrejötte, illetve a véleményjelölő diskurzusjelölők realizációja és ezek kapcsolata a szupraszegmentális jellemzőkkel. A projektben cél továbbá a beszédalkalmazkodás jelenségének szélesebb körű leírása, amely fonetikai, szociolingvisztikai és pragmatikai szempontból is fontos jelenség, ugyanakkor mindegyik keretben nagyon kevés hazai kutatási előzménye van: így beszédkutatási szempontból, valamint a különböző megközelítések összekapcsolása révén is nővumnak számít. Továbbá más társalgási

jelenségek is szerepet kapnak mind az annotációban, mind az ezekre épülő elemzésekben (pl. háttércsatorna-jelzések, nevetés, szünetek, együttbeszélések stb.).

A projekt a társalgási jelenségek annotálásában a fonetikai szempontok mellett elsősorban a nyelvészeti konverzációelemzés és az interakciós szociolingvisztikai társalgáselemzések elméleti és fogalmi háttérére támaszkodik, egy funkcionális pragmatikai szemlélet keretében. Ennek során a projekt fontos újdonsága – ugyanakkor jelentős elméleti és módszertani kihívás is – a fonetika hagyományosan formális, strukturalista szemléletének és a társalgáselemzés, illetve az újabb szociolingvisztikai és funkcionális pragmatikai elméletek gyökeresen funkcionális megközelítéseinek az összehangolása, illetve a fonetikában szokásos kvantitatív igényű, felülről lefelé, előzetesen meghatározott és objektív kategóriákon alapuló módszertan összekapcsolása a társalgáselemzés, funkcionális pragmatika és a kortárs szociolingvisztika funkcionális alapú, az interakcióból és közös jelentéskonstruálásból kiinduló szemléletével és nem eleve adott, hanem az interakcióban konstruálódó jelentésekre és kategóriákra épülő, alulról felfelé történő, kvalitatív módszertanával. Ennek kapcsán a módszertani megoldásokban is – a nyelvészet legújabb tendenciáival összhangban – a formális és funkcionális szempontok reflektált összehangolására és közös érvényesítésére törekedtünk, azzal a céllal, hogy az annotált adatok később mind formális, mind funkcionális háttérű további kutatásokhoz is használhatóak legyenek: ehhez az annotációs módszertani elvek elméleti alapjainak széleskörű újragondolására, valamint egyes annotációs eseteknek többszöri és több szempontot is bevonó értékelésére támaszkodtunk.

A projekt első évében a kutatások többek között a társalgások belső szerveződésére, a globális és lokális struktúrájára; a háttércsatorna-jelzésekre, a társalgások szünetezésére vonatkoztak. Fontos eredmények születtek azzal kapcsolatban, hogy a társalgások bizonyos paraméterei lokálisan konstruálódnak és dinamikusan változnak például a belső szerveződés, a beszélői szerepek függvényében (vö. pl. Hámori–Horváth 2019, Gyarmathy et al. megj. e., Horváth et al. 2019, Krepesz et al. 2019).

A BEA-társalgások meglévő annotációja

Az annotálás a nemzetközileg elfogadott Praat programban, .TextGrid formátumban történik (Boersma–Weenink 2018) a leiratok kettős ellenőrzésével történik. A program képernyőjén felül található a hullámforma (rezgéskép vagy oscillogram) és alatta a hangszínekép (spektrogram). Ez alatt találhatóak a címkesorok. Két szegmentumhatárral jelölt rész egy szegmentum, amely tartalmazza az adott egység tartalmát. A lejegyzők előre meghatározott nagyságú, azonos ablakmérettel dolgoznak. Ez azért szükséges, hogy a lejegyzők a rezgéskép és a hangszínekép közel azonos információi alapján döntsenek a szegmentumhatárokról.

meghatározásában. A beszédszakaszok és a szavak lejegyzése során javasolt a felbontást 5 másodperces ablakra beállítani.

A beszélők jelölése egységes. IK jelöli az interjút készítő személyt, B a kísérleti személyt (beszélő) és TP (társalgó partner) a harmadik személyt, aki IK kollégája.

A társalgások címkézése során a korábban elkészült BEA-lejegyzésből (Gyarmathy et al. 2014) megtartjuk a 3 beszélőtől a beszédszakasz és szószintű lejegyzést, illetve ezeket szükség esetén korrigáljuk is – a társalgási felvételek eredetileg sem tartalmazzak beszédhang szintű átiratot. A beszédszakasz szinten a lejegyzés helyesírásban (fonéma alapon), és központozás nélkül történik. Egy beszédszakasz a beszélő által tartott (néma vagy kitöltött) szünettől szünetig tart. A szószintű lejegyzés szintén helyesírásban (fonéma alapon) történik. A szóhatár is a helyesírási határok szerint van jelölve, tehát az összetett szavak esetében a szóösszetételt alkotó szavak egy szakaszban vannak címkézve. A szünetek időtartamának nincs alsó határa, minden hosszúságú szünetet jelölünk. Ezeket a jelenségeket az adott beszélő szakasz és szószintű címkesorában feltüntetjük (lásd 1. ábra). A néma szüneteken belül meg vannak különböztetve az egyes beszélők saját beszédében tartott jelkimaradások (jele: SIL) és a hallgatások (PAUSE): az a néma rész, amikor a résztvevő nem beszélői szerepben vesz részt, hanem a partner(ek)e)t hallgatja. A kitöltött szüneteket szintén külön címkében tüntetjük fel, nagybetűvel és azzal a hanggal/hangkapcsolattal, aminek hallja a lejegyző (Ö, ÖM, M). A szünet a szóban jelenség esetén a szünet külön szegmentumba kerül. A nem beszéd jellegű hangok külön szegmentumban szerepelnek (pl. sóhaj: SÓH, torokköszörülés: TKS, köhögés: KÖH), az adott beszélő mindkét címkesorában. Ha előttük és/vagy utánuk szünet szerepel, akkor attól külön szegmentáljuk a jelenséget. Amennyiben a beszélő nevetős beszédet produkál, a címkében a szöveg előtt NEV rövidítés szerepel (szakaszszinten az egész beszédszakaszra érvényes a NEV jelölés; szószinten csak arra a szóra/szavakra, ahol elhangzik).

Társalgási jelenségek címkézése

Az előzőekben ismertetett 6 címkesorból álló meglévő annotációt további 5 címkesorral egészítjük ki a társalgások annotálása során. Ezekben – elsősorban a nyelvészeti társalgáselemzés hagyományaira építve – a társalgás néhány alapvető jelenségét jelöljük (pl. háttérzsoltár-jelzések, egyszerre beszélés, l. alább). Ezen jelenségek annotálása nemcsak ezek vizsgálatát segíti, hanem lehetővé teszi számos további társalgáselemzési, szociolingvisztikai és pragmatikai jelenség kiterjesztett kutatását is.

A meglévő annotáció természetesen tovább bővíthető, a pályázat egyik fő témája többek között a véleményjelölő diskurzusjelölők elemzése, így azok is külön címkesorba kerülhetnek.

A hetedik címkesorba kerülnek a **háttércsatorna-jelzések**. Yngve (1970) definíciója alapján a háttércsatorna-jelzés megnevezést használjuk, és azokat a rövid megjegyzéseket értjük rajta „amelyeket egy beszélgetés során az egyik résztvevő a másik beszélése (fordulója) közben tesz” (Ward–Tsukahara 2000), más szavakkal, azokat a verbális vagy nonverbális jelzéseket, amelyet egy társalgásban az egyik fél (a hallgató) a beszélő beszéde közben ad, a figyelem vagy megértés jelzésre, vagy tovább-beszélést ösztönző céllal, pl. *üüm, aha, igen, az tök jó*. Fel van tüntetve a szegmentumban, hogy mi hangzik el: verbális elem (*üüm, ja, igen igen, aha*) vagy nonverbális vokális jelzés (nevetés: NEV, torokköszörülés: TKS, sóhaj: SÓH), esetleg ezek kombinációja (NEV+ *igen*); illetve azt is tartalmazza a címke, hogy melyik beszélőtől hangzott el a háttércsatorna-jelzés, például: NEV, IK vagy *aha*, TP.

A nyolcadik címkesor tartalmazza az **egyszerre beszéléseket**. (Korábban az egyszerre beszélések esetén minden érintett beszélő mindkét címkesorába EB jelölést tüntettek fel függetlenül attól, hogy érthető volt-e amit a beszélők mondtak. A mostani annotációban minden érthető beszédet annotálunk szakasz és szószinten is; illetve egy új címkében azt, hogy kik beszélnek egyszerre). Jelöljük a címke határait (az egyszerre beszélés teljes időtartamát), benne pedig azt, hogy melyik beszélők érintettek, vesszővel elválasztva. A beszélők sorrendjének is szerepe van: elsőként a fő beszélő szerepel, a vesszőt követi annak a beszélőnek a jele, aki másodikként szólalt meg (B, IK – a B beszélő beszélt, IK bekapcsolódott átfedő beszédet eredményezve). Ha teljesen egyszerre történt a megszólalás, akkor a meglévő annotáció eredeti sorrendjét követjük felülről lefelé (IK, TP – IK és TP egyszerre szólaltak meg, de az eredeti annotációban IK szerepel felül, TP legalul).

A kilencedik címkesorban jelöljük a **beszélőváltásokat**. A beszélőváltást megelőzheti szünet, vagy egyszerre beszélés. Az előbbi esetben a beszélőváltás időtartamának azt az időtartamot jelöljük, ami az egyik beszélő elhallgatása és a másik beszélő megszólalása között telik el; előfordul olyan eset is, amelyben a beszélőváltás azonnal, 0 ms szünettel valósul meg, ezt is jelöljük. Az egyszerre beszéléssel megvalósuló szóátvétel esetén az egyszerre beszélés teljes időtartamát jelöljük beszélőváltásnak ebben a címkesorban. A szegmentum tartalmazza, hogy melyik beszélőtől ki vette át a szót, alsó vonással elválasztva a két jelölést, pl.: B_TP (az adatközlő beszélőtől a társalgó partner vette át a szót).

A tizedik címkesor tartalmazza az egyes beszélőkhöz tartozó **fordulók** jelölését. Fordulónak itt a konverzációelemzés “turn” fogalmát és legelterjedtebb értelmezését követve egy beszélőnek a szóátvételét követően elmondott szövegét tekintjük a következő fordulóváltásig, más szóval a beszélők cseréjéig (Sacks–Schlegloff–Jefferson 1974). A fordulók hosszának pontos meghatározását a gyakorlatban néha megnehezíti, hogy egyes esetekben (pl. rövid egyszerre beszélés, vagy szünetben elhangzó háttércsatorna-jelzés) nem dönthető el egyértelműen, történt-e beszélőváltás; ezekben az esetekben több (tartal-

mi/szemantikai, szintaktikai, fonetikai és funkcionális) szempont együttes mérlegelésével döntöttünk (pl. szintaktikai lezártság megléte, megnyilatkozás funkciója).

A tizenegyedik címkesor a **globális szerkezethez** kapcsolódik: a társalgás szakaszait aszerint osztályozzuk, hogy döntően monologikus vagy dialogikus-e az adott rész. Bár a globális szerkezet nagyobb funkcionális egységek megvalósulásával jön létre, módszertani okokból a jelen annotáció nem az egyes részletek funkcionális jellegét, hanem a szövegszerveződés módját jelöljük, ami a jellegzetes szövegfajta (narratíva vagy dialógus) megjelenésével a különböző funkcionális egységeket is jelzi.

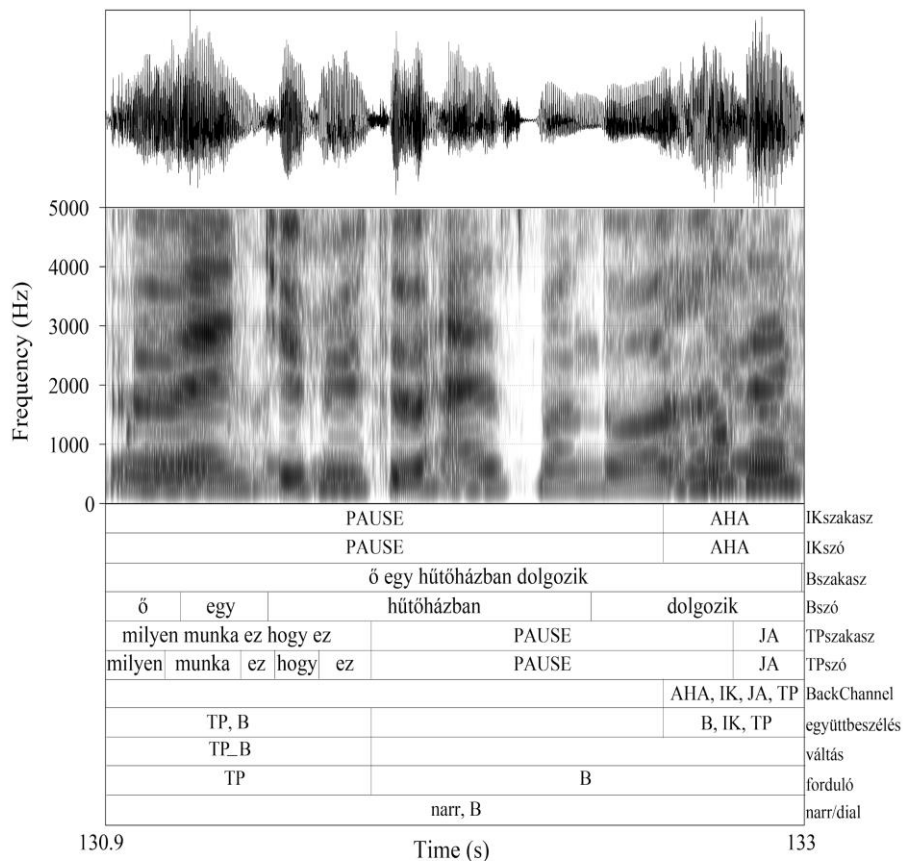
A tizenegy címkesor elnevezése a következő (1. ábra):

1. interjúkészítő beszédszakasz szintű lejegyzése (IKszakasz)
2. interjúkészítő szószintű lejegyzése (IKszó)
3. adatközlő beszédszakasz szintű lejegyzése (Bszakasz)
4. adatközlő szószintű lejegyzése (Bszó)
5. harmadik beszélő beszédszakasz szintű lejegyzése (TPszakasz)
6. harmadik beszélő szószintű lejegyzése (TPszó)
7. háttércsatorna-jelzések (BackChannel)
8. egyszerre beszélések (együttbeszélés)
9. beszélőváltások (váltás)
10. fordulók
11. globális szerkezet (narratív/dialogikus szakaszok; narr/dial)

A „társalgási korpusz” jelentősége és felhasználhatósága

A BEA-társalgások nagy tudományos jelentőséggel bírnak a konverzációelemzés és a beszédkutatás számára: nemcsak azért, mert új és újszerű lehetőséget adnak a magyar társalgások korpuszalapú vizsgálatára, hanem azért is, mert nemzetközi szinten is igen kevés olyan korpusz létezik, amely 3 résztvevő spon-tán társalgásait rögzíti egyidejűleg.

A korpusz jelentőségét egyrészt a kisebb és kontextualizáltabb, társalgási és pragmatikai jelenségek vizsgálatát segítő korpuszok iránt egyre növekvő, már említett érdeklődés adja. Ugyanilyen fontos az annotálás során érvényesülő komplex megközelítés, amely a fonetikai kiindulást pragmatikai és társalgás-elemzési fogalmakkal és szemlélettel ötvözi.



1. ábra: A tizenegy címkesor a Praatban

Ezek a nagy mennyiségű, szisztematikus és empirikus gyűjtésen alapuló elemzések elengedhetetlenek a természetes emberi kommunikációhoz a lehetőségek szerint leginkább hasonló ember-gép interakciók fejlesztéséhez, ami a nemzetközi kutatásokban egyre nagyobb figyelmet kap. Az ilyen jellegű kutatások elsődleges vizsgálati anyaga a valós multi-party típusú interakciók, leggyakrabban üzleti vagy baráti beszélgetések elemzése. Ilyen korpusz például az ICSI (Janin et al. 2003), amely 3–10 (átlagosan 6) résztvevő üzleti megbeszélésén elhangzott megnyilatkozásait tartalmazza, összesen körülbelül 72 órás audiofelvételt. Ehhez hasonló az AMI korpusz (Carletta et al. 2005), ami összesen 100 órányi anyagot, valamint a hanganyagok átíratát és számos különböző jelenség annotációját, például a társalgási jelenségek (dialog acts), és a fejmozgások leiratozását is tartalmazza.

A BEA társalgási korpuszában összesen 100 darab ilyen felvétel áll a kutatók rendelkezésére, ezek mintegy harmadában ugyanazok a személyek szerepelnek IK és TP szerepben, ami nagy mennyiségű anyagon több speciális társalgási jelenség (háttércsatorna-jelzések, egyszerre beszélések, beszédfordulók, nevetések stb.) összehasonlító vizsgálatát is lehetővé teszi, így például a beszédpartner személyétől, témától vagy érzelmek megjelenésétől függő változásoknak vagy partnerek egymáshoz való alkalmazkodásának elemzését a beszéd különböző paramétereiben.

A korpusz tudományos értékét erősíti a manuális annotáció: minden meglévő BEA-lejegyzést egy egyetemi hallgató ellenőriz, szükség esetén javít és kiegészíti a társalgási elemeket tartalmazó 5 címkesorral. Minden elkészült .TextGrid fájlt egy további, senior kutató manuálisan ellenőriz; egyes nehezen besorolható társalgási jelenségek annotálását további személyek ellenőrzésével és kvalitatív elemzéssel is kiegészítettük. A hanganyagot tartalmazó .wav fájlokat és a hozzájuk tartozó annotációt tartalmazó .TextGrid fájlokat egy, a kutatócsoport tagjai számára hozzáférhető felhőben tároljuk.

A társalgási modul tartalmazó anyag adatbázisként értelmezhető, mert nagy mennyiségű, összesen 100 felvételt (nagyjából 25 órányi beszéd) tartalmaz, amelyek a felvétel adatai (a beszélő életkora, neme, foglalkozása, végzettsége stb.); a felvétel időtartama; a beszélgetés témája alapján lekérdezhetők, a címkéből szöveges átiratok nyerhetők.

A korpusznak további tudományos jelentőséget ad, hogy az adatközlők részt vesznek a BEA-adatbázis többi részében is (érvelés, felolvasás, mondatvisszamondás stb.), így a beszédjellemzőik összevethetők a más beszéd típusokban rögzített adatokkal – erre nincs lehetőség a csak társalgásokat tartalmazó korpuszok esetében.

E korpusz segítségével, több nyelvészeti tudományág együttműködésének eredményeként, a nemzetközi trendeknek megfelelően a társalgások rendszere és a magyar társalgások jellegzetességei pontosabban és komplexen leírhatóvá válnak. Ez egyrészt lényeges mind a különféle kiterjedt beszélt nyelvi korpuszok automatikus feldolgozása, mind az ezekben végezhető adatbányászat szempontjából (vö. Bechet et al. 2012). Például Jurafsky és munkatársai (1997) a Switchboard-adatbázis (Godfrey–Holliman 1993) 1155 darab társalgásában, 42 különböző társalgási esemény (kérdés, válasz, háttércsatorna-jelzés, egyetértés stb.) vizsgálata alapján építettek modellt, amely az akusztikai információ alapján 65%-os eredményességgel detektál automatikusan társalgási eseményeket, szószintű annotáció alapján 70% feletti eredményességgel. Hasonlóképpen, ha a diskurzusok rendszere, szerveződési jellegzetességei leírhatóvá válnak (amiben grammatikai, szemantikai és pragmatikai-társalgási tényezők mellett bizonyos fonetikai mintázatoknak is lényege szerepe van), az megteremti az alapot a gépi modellezéshez is; ezek az eredmények tudományos hasznosságuk mellett be-

szédtechnológiai alkalmazásokban, robotikai és mesterségesintelligencia-fejlesztésekben is felhasználhatóak lesznek. Nemzetközi kutatásokban kiemelt figyelmet kap és már jelentős eredményeket mutat az ember–gép kommunikáció, különösen a társas-beszélgető robotok fejlesztése, sőt már ezek speciális alkalmazásai is (például idősokkal vagy autizmussal élő személyek esetében, Shimaya és tsai. 2019), bizonyos szempontok kiemelt figyelembevételével, mint pl. idősokkal való “társalgás” esetén az aktív figyelem biztosítása a robot részéről a beszélő felé (Bevacqua et al. 2012, DeVault et al. 2014, Johansson et al. 2016, Lala et al. 2017). Az Erica nevű robot fejlesztése során a fejlesztők hosszú távú célja az, hogy a robot képes legyen emberhez hasonló módon beszélni és gesztusokat tenni a kommunikáció során. A fejlesztéshez a háttéracsatorna-jelzések kerültek fókuszba: a korábbi kutatásokban a gép csak az embertől származó beszédforduló végén produkált háttéracsatorna-jelzéseket, Lala és munkatársai (2017) rendszerében a robot folyamatosan ad jelzéseket a beszélő számára arról, hogy figyelmesen hallgatja.

A tudomány szempontjából a jelen kutatássorozat célkitűzése azért nagyon fontos, mert magyar nyelven egyedülálló módon ötvözi a fonetikai és pragmatikai, a kvalitatív és kvantitatív szemléletű vizsgálatokat a társalgás, mint az egyik legtermészetesebb kommunikációs forma elemzése során. A társalgási korpuszoknak kiemelkedő szerepük van továbbá a jövőbeni elméleti és a közvetlenül felhasználható, alkalmazott kutatások szempontjából is, különösen a robotika és a gépi beszédfeldolgozás terén.

Irodalom

- Abuczki Ágnes (2011), Társalgási fordulók multimodális vizsgálata a HuComTech adatbázis informális dialógusaiban. In: Boda, I. Károly – Mónos K. (szerk.), MANYE XX. Az alkalmazott nyelvészet ma: innováció, technológia, tradíció (XX. Magyar Alkalmazott Nyelvészeti Kongresszus kötete). Debreceni Egyetemi Kiadó, Debrecen. 282–288.
- Antalné Szabó Ágnes (2006), A tanári beszéd empirikus kutatások tükrében. Magyar Nyelvtudományi Társaság, Budapest.
- Asztalos Anikó (2019), A csoportkommunikációs stratégia vizsgálata a tanulói beszédben magyar nyelvű tanórák alapján. In: Karlovitz, János Tibor – Torgyik, Judit (szerk.), Szakmódszertani és más emberközpontú tanulmányok. International Research Institute, Komárno, Szlovákia. 71–84.
- Asztalos Anikó (2019, megjelenőben), A tanulói szóátadások vizsgálata. In: Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Konferencia. MTA Nyelvtudományi Intézet, Budapest.
- Bata Sarolta (2009), Beszélőváltások a beszédpartnerek személyes kapcsolatának függvényében. *Beszédkutatás* 17: 107–120.

- Bechet, F. – Maza, B. – Bigouroux, N. – Bazillon, T. – El-Bèze, M. – De Mori, R. – Arbillot, E. (2012), DECODA: a call-center human-human spoken conversation corpus. In: International Conference on Language Resources and Evaluation (LREC) 1343–1347.
<https://pdfs.semanticscholar.org/ede5/d2d2557188c153272152c9c8117ce8b67898.pdf>
- Bevacqua, Elisabetta – Cowie, Roddy – Eyben, Florian – Gunes, Hatice – Heylen, Dirk – Maat, Mark – Mckeown, Gary – Pammi, Sathish – Pantic, Maja – Pelachaud, Catherine – De Sevin, Etienne – Valstar, Michel – Wollmer, Martin – Shroder, Marc – Schuller, Bjorn (2012), Building Autonomous Sensitive Artificial Listeners. *IEEE Transactions on Affective Computing* 3(2): 165–183.
- Boersma, Paul – Weenink, David (2018), Praat: doing phonetics by computer. http://www.fon.hum.uva.nl/praat/download_win.html.
- Bodó Csanád – Kocsis Zsuzsanna – Vargha Fruzsina Sára (2017), A Budapesti Egyetemi Kollégiumi Korpusz. Elméleti és módszertani kérdések. In: Benő Attila – Fazakas Noémi (szerk), *Élőnyelvi kutatások és a dialektológia*. Erdélyi Múzeum-Egyesület, Kolozsvár. 169–178.
- Bóna Judit – Imre Angéla – Markó Alexandra – Váradi Viola – Gósy Mária (2014), GABI – Gyermeknyelvi beszédAdatBázis és Információtár. *Beszédkutató* 2014. 246–251.
- Boronkai Dóra (2009), *Bevezetés a társalgáselemzésbe*. Ad Librum Kft., Budapest.
- Bortfeld, Heather – Leon, Silvia D. – Bloom, Jonathan E. – Schober, Michael F. – Brennan, Susan E. (2001), Disfluency rates in conversation: effects of age, relationship, topic, role, and gender. *Language and Speech* 44 (2): 123–147.
- Canavan, Alexandra – Graff, David – Zipperlen, George (1997), CALLHOME American English Speech LDC97S42. Web Download. Philadelphia: Linguistic Data Consortium.
- Carletta, Jean – Ashby, Simone – Bourban, Sebastien – Flynn, Mike et al. (2005), The AMI Meetings Corpus: a pre-announcement. In: Renals, Steve – Bengio, Samy (eds.), *Machine Learning for Multimodal Interaction*. Second International Workshop, MLMI 2005, Edinburgh, UK, July 11-13, Revised Selected Papers, Springer. 28–39.
- DeVault, David – Artstein, Ron – Benn, Grace – Dey, Teresa – Fast, Ed – Gainer, Alesia – Georgila, Kallirroi – Gratch, Jon – Hartholt, Arno – Lhomme, Margaux – Lucas, Gale – Marsella, Stacy – Morbini, Fabrizio – Nazarian, Angela – Scherer, Stefan – Stratou, Giota – Suri, Apar – Traum, David – Wood, Rachel – Xu, Yuyu – Rizzo, Albert – Morency, Louis-philippe (2014) SimSensei Kiosk: A Virtual Human Interviewer for Healthcare Decision Support. In *International Conference on Autonomous Agents and Multi-Agent Systems* 1: 1061–1068.
- Dér Csilla Ilona (2012), Beszélőváltások során használt diskurzusjelölők a magyar spontán beszédben. *Beszédkutató* 20: 130–141.
- Dér, Csilla Ilona – Markó, Alexandra (2010). A pilot study of Hungarian discourse markers. *Language and Speech* 53 (2): 135–180.
- Eshghi, Arash – Healey, Patrick (2009), What is conversation? Distinguishing dialogue contexts. In: Taatgen, N. A. – van Rijn, H. (eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society 2009*. 1240–1245.

- Godfrey, John – Holliman, Edward (1993), Switchboard-1 Release 2 LDC97S62. Web Download. Philadelphia: Linguistic Data Consortium. <https://catalog.ldc.upenn.edu/LDC97S62>
- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Gráczki Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter (2012), BEA: Beszélt nyelvi adatbázis. In: Gósy Mária (szerk.), *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 9–24.
- Gyarmathy Dorottya – Neuberger Tilda – Gráczki Tekla Etelka (2014), Lejegyzési útmutató a BEA Spontánbeszéd-Adatbázis háromszintű annotálásához. *Alkalmazott Nyelvtudomány* 14 (1-2): 35–44.
- Gyarmathy Dorottya – Krepsz Valéria – Dér Csilla Ilona – Hámori Ágnes – Horváth Viktória (megj. e.), Háttercsatorna-jelzések: határterület és új szempontok a diskurzusok elemzésében. In: Fóris Ágota – Bölskei Andrea (főszerk.), *Nyelv, kultúra, identitás. Alkalmazott nyelvészeti kutatások az információs térben*. MANYE – OFFI – KRE BTK Magyar Nyelvtudományi Tanszék, Budapest.
- Hámori Ágnes (2006), A társalgási műfajokról. In: Tolcsvai Nagy Gábor (szerk.), *Szöveg és típus. Szövegtipológiai tanulmányok*. Tinta Könyvkiadó, Budapest. 157–181.
- Hámori Ágnes (2014), Nevetés a társalgásban. In: Laczkó Krisztina – Tátrai Szilárd (szerk.), *Elmélet és módszer. Nyelvészeti tanulmányok*. Eötvös József Collegium, Budapest. 105–130.
- Hámori Ágnes – Horváth Viktória (2019), Társalgás, beszélőváltás és diskurzusszerveződés új megközelítésben – fonetikai jellemzők és pragmatikai tényezők összefüggései magyar társalgásokban (pilot study). *Beszédkutatás* 27: 134–153.
- Horváth Viktória – Gyarmathy Dorottya – Krepsz Valéria – Gráczki Tekla Etelka – Hámori Ágnes (2019), A társalgások szünetezése – különböző funkciók és realizációk (konferenciaelőadás). *A nyelv közösségi perspektívája*. Nagyvárad. 2019 június 28–29.
- Hemphill, Charles T. – Godfrey, John J. – Doddington, George R. (1990), The ATIS spoken language systems pilot corpus. *HLT '90 Proceedings of the workshop on Speech and Natural Language*. 96–101.
- Hunyadi László – Földesi András – Szekrényes István – Staudt Alexandra – Kiss Hermina – Abuczki Ágnes – Bódog Alexa (2012), Az ember–gép kommunikáció elméleti–technológiai modellje és nyelvtechnológiai vonatkozásai. *ANyT XXIV*.
- Janin, Adam – Baron, Don – Edwards, Jane – Ellis, Dan et al. (2003), The ICSI meeting corpus. *Proceedings. (ICASSP '03)*. 2003 IEEE International Conference. Vol. 1. <ftp://ftp.icsi.berkeley.edu/pub/speech/papers/icassp03-janin.pdf>
- Johansson, Martin – Hori, Tatsuro – Skantze, Gabriel – H'othker, Anja – Gustafson, Joakim (2016), Making turn-taking decisions for an active listening robot for memory training. In: Agah, Arvin – Cabibihan, John-John – Howard, Ayanna M. – Salichs, Miguel A. – Hongsheng He (eds.), *Social Robotics: 8th International Conference, ICSR 2016*. Springer International Publishing, Cham. 940–949.
- Jurafsky, Daniel – Bates, Rebecca – Coccaro, Noah – Martin, Rachel – Mateer, Marie – Ries, Klaus – Shriberg, Elizabeth – Stockle, Andreas – Taylor, Paul – Van Ess-Dykema, Carol (1997), Automatic detection of discourse structure for speech recognition and understanding. *Proceedings of the 1997 IEEE Workshop on Speech Recognition and Understanding*. Santa Barbara: IEEE. 88–95.

- Keszler Borbála (1983), Kötetlen beszélgetések mondat- és szövegtani vizsgálata. In: Rácz Endre – Szathmári István (szerk.), *Tanulmányok a mai magyar nyelv szövegtana köréből*. Akadémiai Kiadó, Budapest. 164–202.
- Kiss Hermina (2011), A HuComTech audio adatbázis szintaktikai szintjének elvei és szabályrendszerének újdonságai. *Magyar Számítógépes Nyelvészeti Konferencia* (6). 199–208. http://acta.bibl.u-szeged.hu/58802/1/msznykonf_008_199-208.pdf
- Kontra Miklós (szerk.) (1988), *Beszélt nyelvi tanulmányok*. MTA Nyelvtudományi Intézet, Budapest.
- Kontra Miklós (1990), Budapesti élőnyelvi kutatások. *Magyar Tudomány* 1990/5. 512–520.
- Krepsz Valéria – Hámori Ágnes – Gyarmathy Dorottya – Horváth Viktória (2019), A nevetés főbb változatai spontán társalgásokban (konferenciaelőadás). VI. Interdiszciplináris Humorkonferencia. Eszterházy Károly Egyetem, Eger. 2019. október 3–4.
- Lala, Divesh – Milhorat, Pierrick – Inoue, Koji – Ishida, Masanari – Takanasi, Katsuya – Kawahara, Tatsuya (2017), Attentive listening system with backchanneling, response generation and flexible turn-taking. In: *Proceedings of the SIGDIAL 2017 Conference*, Saarbrücken, Germany, 2017. 127–136. <https://www.aclweb.org/anthology/W17-5516>. [Az utolsó letöltés: 2019. június 6.]
- Markó Alexandra (2005), A spontán beszéd néhány szupraszegmentális jellegzetessége. PhD-értekezés. ELTE BTK, Budapest.
- Mátyus Kinga – Orosz György (2014), MONYEK – Morfológiailag egyértelműsített óvodai nyelvi korpusz. *Beszédkutatás* 2014: 237–245.
- Németh T. Enikő (szerk.) (2011), *Ember-gép kapcsolat. A multimodális ember-gép kommunikáció modellezésének alapjai*. Tinta Könyvkiadó, Budapest.
- O’Keeffe, Anne – McCarthy, Michael – Carter, Ronald (2007), *From corpus to classroom*. Cambridge University Press, Cambridge.
- Pápay Kinga (2011), A beszélő személy akusztikus és vizuális gesztuskészlet-használatának vizsgálata multimodális korpusz alapján – beágyazások, beékelések és adaptáció. In: Hunyadi L. (szerk.), *Rekurzió a nyelvben I. Prozódiai megközelítés*. Tinta Könyvkiadó. Budapest. 31–95.
- Sacks, Harvey – Schegloff, Emanuel A. – Jefferson, Gail (1974), A simplest systematic for the organization of turn-taking for conversation. *Language* 50: 696–735.
- Sass Bálint (2016), *Nyelvészeti szövegkeresők*, Nemzeti Korpuszportál. *Magyar Tudomány*. 2016 (7): 798–808.
- Schegloff, Emanuel A. (1968), Sequencing in conversational openings. *American Anthropologist* 70: 1075–1095.
- Schirm Anita (2018), Diskurzusjelölő-társulások a Szögedi Szociolingvisztikai Interjúban. *Alkalmazott Nyelvtudomány* 18/1: 1–16.
- Schirm Anita (2019), A diskurzusjelölő-használat életkori sajátosságai a nyelvi interjú szövegtípusában. *Beszédkutatás* 27: 187–205.
- Shimaya, Jiro – Yoshikawa, Yuichiro – Kumazaki, Hirokazu – Matsumoto, Yoshio – Miyao, Masumoto – Ishiguro, Hiroshi (2019), Communication support via a tele-operated robot for easier talking: case/laboratory study of individuals with/without autism spectrum disorder. *International Journal of Social Robotics* 11 (1): 171–184.

- Simpson, Rita C. – Briggs, Sarah L. – Ovens, Janine – Swales, John M. (2002). The Michigan Corpus of Academic Spoken English. Ann Arbor: The Regents of the University of Michigan. <http://www.helsinki.fi/varieng/CoRD/corpora/MICASE/>
- Svartvik, Jan ed. (1990), The London Corpus of Spoken English: Description and Research. Lund Studies in English 82. Lund University Press, Lund.
- Szirmai Monika (2005), Bevezetés a korpusnyelvészetbe. A korpusnyelvészet alkalmazása az anyanyelv és az idegen nyelv tanulásában és tanításában. Tinta Könyvkiadó, Budapest.
- Váradai Tamás (2003). A Budapesti Szociolingvisztikai Interjú. In: Kiefer Ferenc – Siptár Péter (szerk.), A magyar nyelv kézikönyve. Akadémiai Kiadó, Budapest. 339–59.
- Varga László (1987), Prozodémák a magyar beszédben és jelölésük az intonációs átíratban. Műhelymunkák a nyelvészet és társtudományai köréből III. Magyar Tudományos Akadémia, Budapest.
- Vaughan, Elaine – Clancy, Brian (2013). Small Corpora and Pragmatics In: J. Romero-Trillo (ed.). New Domains and Methodologies, Yearbook of Corpus Linguistics and Pragmatics 1. Springer. Dordrecht 53–77.
- Walsh, Steve (2013), Corpus linguistics and conversation analysis at the interface: theoretical perspectives, practical outcomes. In: Romero-Trillo, Jesús (ed.), New Domains and Methodologies, Yearbook of Corpus Linguistics and Pragmatics 1., Dordrecht. 37–53.
- Ward, Nigel – Tsukahara, Wataru (2000), Prosodic features which cue back-channel responses in English and Japanese. *Journal of Pragmatics* 32: 1177–1207.
- Yngve, Victor H. (1970), On getting a word in edgewise. *Papers from the Sixth Regional Meeting Chicago Linguistics Society*. 567–578.

Köszönjük az annotálók, Kondor Kitti és Száraz Bettina kitartó és magas színvonalú munkáját.

A projektet a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal K-128810 számú pályázata támogatja.