GYARMATHY DOROTTYA – NEUBERGER TILDA – GRÁCZI TEKLA ETELKA

MTA Nyelvtudományi Intézet

gyarmathy.dorottya@nytud.mta.hu – neuberger.tilda@nytud.mta.hu –

graczi.tekla.etelka@nytud.mta.hu

Lejegyzési útmutató a BEA Spontánbeszéd-adatbázis háromszintű annotálásához

The development of the largest Hungarian spontaneous speech database, BEA (the abbreviation stands for the letters of the original name of the database: BEszélt nyelvi Adatbázis) started at the Phonetics Department of the Research Institute for Linguistics in 2007. This database involves a great number of speakers who speak relatively lengthily long, contains various styles of speech materials, and has various levels of transcription. This paper presents the criteria and rules of the third type of transcription which is being done at present using Praat software.

Due to its recording protocol and the transcription, this database provides a challenging material for various comparisons of segmental structures of speech also across languages.

1. Bevezetés

A beszédadatbázisok elemzése és felhasználása a nyelvészetben mintegy három évtizedes múltra tekint vissza. A beszédkorpuszoknak különösen azokon a területeken van nagy jelentősége, ahol az adott nyelvi, nyelvhasználati jelenség nem vizsgálható, avagy a kitűzött cél nem érhető el megfelelő mennyiségű, körültekintően gyűjtött adathalmaz nélkül. A fonetika, a beszédtechnológia és a pszicholingvisztika számos ilyen kutatási területtel rendelkezik. A beszéd artikulációs és akusztikai sajátosságainak vizsgálata adatbázis híján csaknem elképzelhetetlen. Ilyen célból jöttek létre a világ számos nyelvében korpuszok, illetve beszédadatbázisok, például az UPSID adatbázis (UCLA Phonological Segment Inventory Database; vö. Maddieson 1984). Ez utóbbi jelenleg 451 nyelv adatait tartalmazza.

A magyar beszédtudományos kutatások hosszú időn keresztül főleg felolvasott vagy előre betanult szövegek vizsgálatán alapultak. Különféle beszédadatbázisok keletkeztek a mesterséges beszéd előállításához, melyek mind az aktuális szintézis szükségleteit tükrözik. Léteznek ún. diádos adatbázisok, amelyben két hang kapcsolódása alkot egy-egy elemet (pl. Profivox szövegfelolvasó szoftver), kötött szótáras rendszerekhez készített elemtárak (pl. hangposta, pályaudvari utastájékoztató), továbbá kevert felépítésű beszédadatbázisok (Profivox legújabb változata) (Olaszy 1999). A magyar beszédtechnológiai kutatások és fejlesztések támogatására készült a vezetékes és mobiltelefonról rögzített, 500 adatközlő által felolvasott szövegeket tartalmazó MTBA magyar telefonbeszéd-adatbázis (Vicsi et al. 2002).

A Magyar Tudományos Akadémia Nyelvtudományi Intézetének Élőnyelvi Osztályán 1987–89 között elkészült Budapesti Szociolingvisztikai Interjú (BUSZI adatbázis) 250 adatközlője a budapesti lakosság szociológiailag reprezenta-

tív mintáját adja (Váradi 2003). A felvételek egyaránt tartalmaznak spontán és nem spontán beszédet. A BUSZI előmunkálataiként 1985-ben elkészült a gazdagréti televízió már sugárzott adásaiból válogatott felvételek több szempontú elemzése. A felvételek intonációs átiratát Varga László készítette, melynek felhasználásával a kutatók elemezték a beszéd logikai struktúráját, mondattani szerkezetét, a témaismétlő névmásokat, a spontán beszéd és az írott nyelv különbségét, továbbá a nonverbális kommunikáció, azaz a gesztusnyelv eszközeit (Kontra 1988). 1998-ban keletkezett az első nemzetközi szabvány alapján készült, felolvasásokat rögzítő adatbázis, a BABEL, melynek célja a magyar hivatalos köznyelv hanganyaggal való reprezentálása (Vicsi–Víg 1998).

A beszédkutatás új feladatai, illetőleg a spontán beszéd fonetikai elemzésének igénye szükségessé tette egy a modern korpuszépítés szabályainak megfelelő, a minőségi hangrögzítés minden kritériumát teljesítő, nagy mennyiségű spontán beszédet tartalmazó hangtár létrehozását, amely egyaránt megfelel mind a fonetikai, az alkalmazott fonetikai, illetve a pszicholingvisztikai kutatások kritériumrendszerének.

Az MTA Nyelvtudományi Intézetének Fonetikai Osztályán 2008-ban kezdődött meg a BEA spontánbeszéd-adatbázis (Gósy et al. 2012) feltöltése, ami jelenleg is tart. A korpusz elsődleges célja többféle típusú spontán beszéd rögzítése, de a fonetikai célok kielégítése (összehasonlíthatóság) érdekében mondat- és szövegfelolvasásokat, illetve mondatismétléseket is tartalmaz.

A BEA hanganyagainak lejegyzése kezdetben egy elsődleges írásos tükröztetés volt. Ezek az elsődleges átiratok a Microsoft Office Word programjában .doc formátumban, helyesírásban, központozás nélkül készültek, a későbbi feldolgozás szempontjából fontosnak ítélt adatok, mint például a megakadásjelenségek, illetve a fiziológiai hangadások jelölésével. A helyesíráson alapuló lejegyzés nem jelölte a kiejtés és a helyesírás eltéréseit, tehát nem érvényesítette a hasonulási, összeolvadási szabályokat (a *szabadság* szó nem *szabaccság*-ként lett lejegyezve). A megakadásjelenségek jelölése a vastagon szedett alak volt, és ha nem hangzott el javítás, a helyes szóalak []-ben lett megadva, például: *kell még tenyér meg kej [kenyér meg tej]*. A lejegyzésben jelölve voltak a néma és kitöltött szünetek, nyújtások, a spontán beszédben gyakran használt, nem szótári alakban előforduló szavak (pl. *asszem, nem tom),* az idegen szavak, rövidítések, betűszók, mozaikszók, illetőleg a lejegyző számára értelmezhetetlen szóalakok. Az elsődleges lejegyzések a kutatók munkáját megkönnyíteni hivatott durva átiratok voltak. A későbbiekben a kutatási igények, illetve az automatikus, gépi beszédfelismerésben való felhasználás szükségessé tették a lejegyzési elvek újragondolását, és a szoftveres háttér megváltoztatását. 2010 októberétől a Word dokumentumban történő átírást felváltotta a Transcriber szoftverrel történő lejegyzés, és ezzel egyidejűleg megkezdődött a már elkészült Wordben lejegyzett anyagok Transcriberbe való átkonvertálása. A Transcriber programban elkészült

lejegyzések nagy előnye, hogy az írott szöveg és a hanganyag egyszerre látható és hallható; a program lehetővé teszi a beszéd szegmentálását, címkézését és leírását (Barras et al. 1998). A BEA Transcriberes lejegyzéséről bővebben lásd: Gyarmathy–Neuberger 2011.

Noha az imént röviden bemutatott lejegyzési mód az adatbázis beszédfelismeréshez történő felhasználásához kiválóan megfelel, a kutatásban való felhasználhatósága korlátozott. Nem alkalmas ugyanis fonetikai mérésekre, elemzésekre. Ahhoz, hogy az adatbázis eredeti céljainak megfelelően minél szélesebb körben felhasználhatóvá váljék, a Fonetikai Osztály munkatársai kidolgoztak egy újabb, a Praat szoftverrel (Boersma–Weenink 2013) történő lejegyzési rendszert.

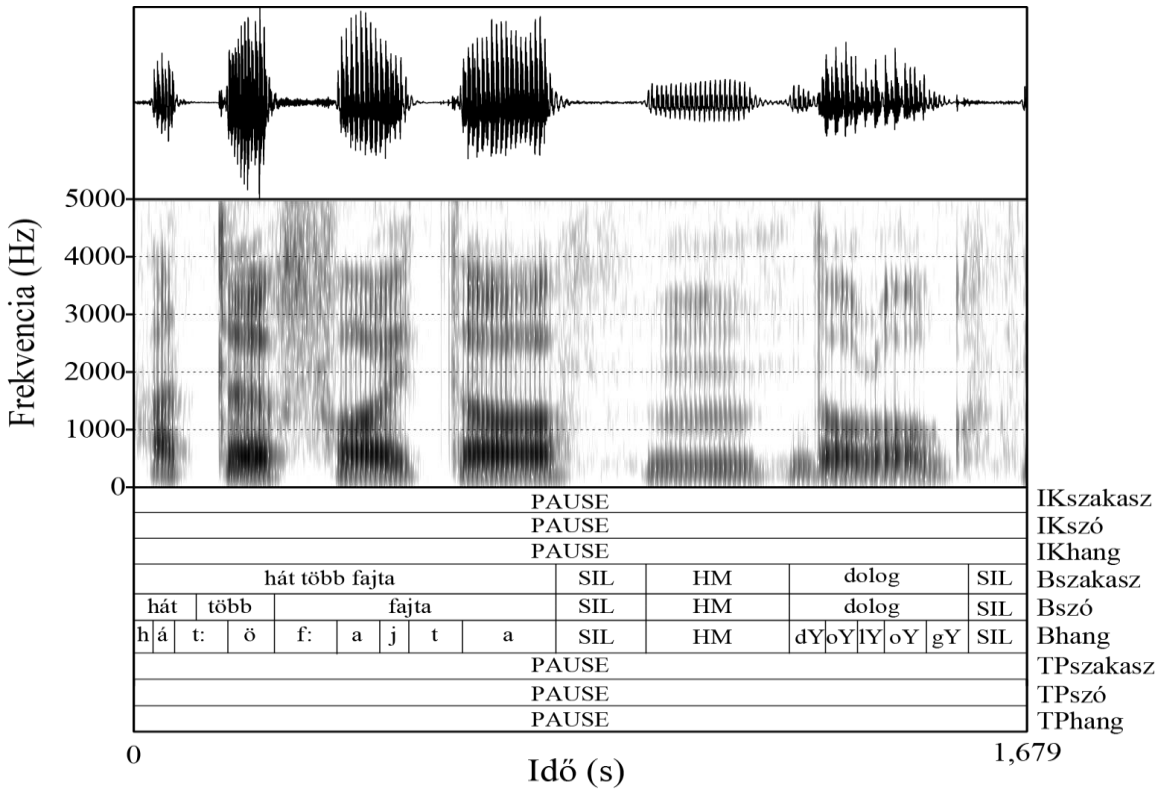
A jelen tanulmányban bemutatott lejegyzési útmutató a 108762. számú OTKA-pályázat keretében jött létre. A pályázat célja, hogy 20 és 90 év közötti magyar felnőtt beszélők beszédanyagának (különböző spontánbeszéd-típusokat, társalgást, felolvasást és mondatismétlést) rögzítése, valamint a beszédanyagok több szinten való szegmentálása és annotálása (beszédhangok, szavak, beszélőváltások).

2. A lejegyzés menete a Praat szoftver segítségével

A lejegyzés során a hangfájl mellé egy szöveges fájl készül, amely egyrészt az elhangzott szöveget tartalmazza, másrészt a hanganyagot kisebb szegmentumokra bontja időbélyegek alkalmazásával. Az annotálás a nemzetközileg elfogadott Praat programban (Boersma–Weenink 2013) a leiratok kettős ellenőrzésével, illetve statisztikai kontrolljával történik. A Praat komplex akusztikai jelfeldolgozó, amely lehetőséget nyújt az annotálásra is.

A BEA spontánbeszéd-adatbázis lejegyzése három szinten történik: beszédszakaszok, szavak és beszédhangok szintjén. A felvételben résztvevők száma határozza meg, hogy a Praatban az annotálás során hány címkesorral dolgozunk. A beszélők száma a felvételi protokoll egyes részeiben eltér: a mondatismétlés, a narratíva, a véleménykifejtés, az interpretált beszéd és a felolvasás részekben a felvételvezető és az adatközlő van jelen, a társalgási részben pedig három beszélő vesz részt. Ennek megfelelően a Praatban két beszélő esetén 6, három beszélő esetén 9 címkesort rendelünk a hanganyaghoz. A kilenc címkesor elnevezése a következő (1. ábra):

1. interjúkészítő beszédszakasz szintű lejegyzése (IKszakasz)
2. interjúkészítő szószintű lejegyzése (IKszó)
3. interjúkészítő hangszintű lejegyzése (IKhang)
4. adatközlő beszédszakasz szintű lejegyzése (Bszakasz)
5. adatközlő szószintű lejegyzése (Bszó)
6. adatközlő hangszintű lejegyzése (Bhang)
7. harmadik beszélő beszédszakasz szintű lejegyzése (csak társalgás)(TPszakasz)
8. harmadik beszélő szószintű lejegyzése (csak társalgás) (TPszó)
9. harmadik beszélő hangszintű lejegyzése (csak társalgás) (TPhang)



1. ábra: A kilenc címkesor a Praatban

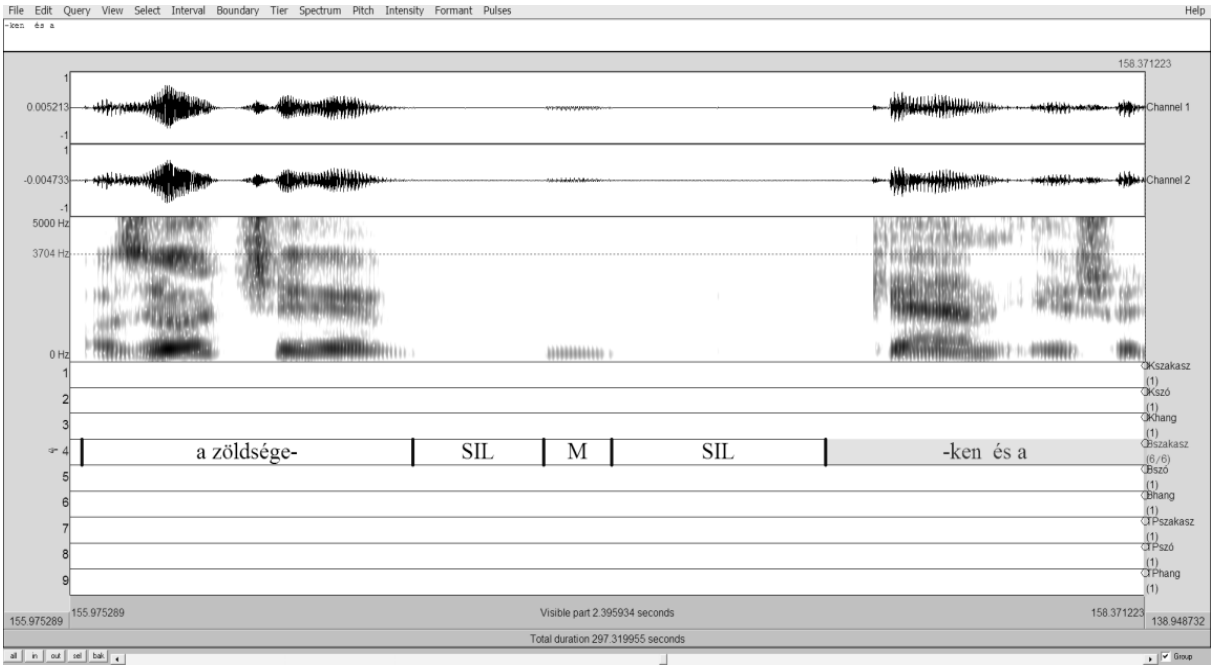
A program képernyőjén felül található a hullámforma (rezgéskép vagy oszcillogram) és alatta a hangszínkép (spektrogram). Ez alatt találhatóak a címkesorok, amelyekben a lejegyzők dolgoznak. Két szegmentumhatárral jelölt rész egy szegmentum, amely a háromszintű lejegyzésben beszédszakaszt, szót vagy beszédhangot (továbbá szünetet, zajt stb.) jelölhet (1. ábra). A szegmentumba kattintva írhatjuk be a szegmentum tartalmát. Minden lejegyző egy előre meghatározott nagyságú, azonos ablakmérettel dolgozik. Ez azért szükséges, hogy a lejegyzők a rezgéskép és a hangszínkép közel azonos információi alapján döntsenek a szegmentumhatárok meghatározásában. A beszédszakaszok és a szavak lejegyzése során javasolt a felbontást 5 másodperces ablakra beállítani. A hangszintű átirathoz pedig minden lejegyzőnek 500 ms-os (Visible part: 0.5 seconds) ablakméretet kell használnia a beszédhangok szegmentálásához.

A lejegyzőnek minden, a felvételen elhangzó dolgot (a nem beszéd jellegű hangokat is) jelölnie kell a lejegyzésében. A továbbiakban rövid áttekintést adunk a lejegyzési szintekről, majd részletezzük a lejegyzés során alkalmazott jelölésrendszert.

2.1 Beszédszakasz szint

A beszédszakasz szinten a lejegyzés helyesírásban (fonéma alapon), nagybetű használata (mivel a nagybetűk bizonyos hangokat jelölnek hangszinten) és központozás nélkül történik. Egy beszédszakasz a beszélő által tartott (néma vagy

kitöltött) szünettől szünetig tart. A szünet a szóban jelenség esetén a szünet külön szegmentumba kerül (2. ábra).



2. ábra: A szünet a szóban jelenség jelölése (- = szótöredék, SIL = néma szünet, M = hezitálás)

2.2 Szószint

A szószintű lejegyzés az előzőhöz hasonlóan helyesírásban (fonéma alapon) történik. A szóhatárt is a helyesírási határok szerint jelöljük, tehát az összetett szavak esetében a szóösszetételt alkotó szavak mind ugyanazon szakaszban vannak jelölve.

2.3 Beszédhangszint

A beszédhangszintű lejegyzés (az előző két szinttől eltérően) már nem fonéma, hanem beszédhang alapon történik. Ezen a szinten már az 1 beszédhang = 1 karakter elve érvényesül. Az ehhez használandó átírási rendszert a Fonetikai Osztály kutatói a Sampa lejegyzési elveit alapul véve külön a BEA adatbázis beszédhangszintű lejegyzésére dolgozták ki (1. táblázat). A nyelvileg hosszú mássalhangzókat kettősponttal jelöljük.

1. táblázat: A hangszintű lejegyzés karakterkészlete

MAGÁNHANGÓKAT JELÖLŐ BETŰK	ÁTÍRÁS
a	a
á	á
e	e
é	é
i	i
í	í
o	o
ó	ó
ö	ö
ő	ő
ü	ü
ű	ű
u	u
ú	ú

MÁSSAL-HANZÓKAT JELÖLŐ BETŰK	ÁTÍRÁS
p	p
b	b
t	t
d	d
k	k
g	g
c	c
dz	D
cs	C
dzs	J
ty	T
gy	G
f	f
v	v
sz	S
z	z
s	s
zs	Z
h	h
r	r
l	l
j	j
m	m
n	n
ny	N

A hangszintű annotálás során a vizuális és az auditív információ minden esetben együttesen veendő figyelembe. A beszédhangok ejtése során az alábbi speciális eseteket jelöljük:

- Abszolút szókezdő zöngétlen zárhangok és zár-réshangok; pl. *pár, teve, kút, cica, csiga)* esetén a zárszakaszt egységesen 30 ms-osra jelöljük be.
- Nazalizált magánhangzó: a nazalizált magánhangzót egységesen X-szel jelöljük az adott magánhangzó címkéjében, pl. iX.
- Nazálisos mássalhangzó-kapcsolat esetében előfordulhat koartikulációs néma fázis. Ezeket úgy jelöljük, hogy a nazális kap egy 3-as karaktert. Például: *orvos nem* [n]-je: n3.
- Irreguláris magánhangzó (érdes zönge, amikor a rezgés aperiodikussá vá-

lik): egységesen Y-nal jelöljük az adott magánhangzó címkéjében, pl. aY.
- Levegős magánhangzó (breathy voice): egységesen W-vel jelöljük az adott magánhangzó címkéjében, hogyha a magánhangzó látványosan levegős, pl. aW.
- Hogyha hehezetes az abszolút szóvégi hang: egységesen H-val jelöljük az adott hang címkéjében, pl. tH.
- Gégezárhang (egyetlen periódust érintő irregularitás): a szünet része.
- Az ij/íj és ji/jí kapcsolatokat nem választjuk el, hanem egy közös hangcímkét kapnak a hangkapcsolat jelével.
- A hiátustöltés [j] hangját is a *j* karakter jelöli.
- Magánhangzók alulartikulált ejtése esetén is az alaprealizációt jelöljük, de ha megakadás (pl. egyszerű nyelvbotlás), akkor az elhangzott hangot írjuk.
- Mássalhangzók esetében a fonológiai koartikulációt jelöljük, a fonetikaiakat (*hamvas, ing*) nem.
- [r] hang: amikor hallható, de csak a magánhangzók formánsmenete közelít az r konfigurációjára jellemző tipikus akusztikai szerkezethez, akkor a formánsátmenet(ek)ből kell kivágni ki az r hangszintű címkéjét.
- Amennyiben a koartikuláció miatt szóhatáron egy hosszú hangon „osztozik" két szó, akkor szó szinten felezzük a hangot és mindkét szóhoz írjuk, amiből ered; de hangszinten egy hosszú hangot jelölünk.
- Artikulációs vagy koartikulációs hatásra megjelenő svá (semleges magánhangzó): a beszédhang címkéjében jelöljük egy 2-essel. Ha van r, akkor ahhoz, ha nincs (pl. *gnú*), akkor a svát az első mássalhangzóhoz kapcsoljuk. A címkében tehát a mássalhangzó és a 2 jele jelenik meg, pl. *gnú:* g2.

3. Minden szintet érintő jelölésrendszer

3.1 Szünetek

A szünetek időtartamának nincs alsó határa, minden hosszúságú szünetet jelölünk. Ezeket a jelenségeket az adott beszélő mindhárom címkesorában feltüntetjük (lásd 1. ábra).

- néma szünet jele**:** SIL (a silence szóból adódóan)
- hezitálás jele: Ö, M, ÖM, nagybetűvel és azzal a hang(g/kapcsolatt)al jelöljük, aminek hallja a lejegyző.
- szünet a szóban: A szünet a szóban jelenség esetén a szünet külön szegmentumba kerül.

Beszélőváltáskor fellépő néma szünet:

- hallgatás jele: PAUSE - A hallgatás nem egyenlő a beszéd közben tartott szünettel. Az a néma rész, amikor a résztvevő nem beszélői szerepben vesz részt, hanem a másika(ka)t hallgatja, illetve amelyik beszédlépésváltáskor (szóátadás/átvevéskor) jelenik meg.

3.2 A lejegyző számára érthetetlen vagy egyéb zaj miatt használhatatlan beszédrészek

Ezek a részek a KUKA jelölést kapják a lejegyzésben. Az érthetetlen beszédrészeknél a KUKA jelölés az adott beszélő mindhárom címkesorának megfelelő részére beírandó.

3.3 Egyszerre beszélés/átfedő beszéd

Egyszerre beszélés vagy átfedő beszéd esetén minden érintett minden sorába beírjuk az EB jelölést. Az egyszerre beszélést akkor sem jegyezzük le szakasz, szó, hang szintjén, ha érthető. Lejegyezzük azonban azokat a részeket a szó és a beszédhang szintjén, ami még nem átfedő beszédben hangzott el.

3.4 Nem beszéd jellegű hangok

Ezeket a jelenségeket nagybetűvel, külön szegmentumban jelöljük, az adott beszélő mindhárom címkesorában. Mindezeket, ha előttük és/vagy utánuk szünet áll, akkor attól külön szegmentáljuk mindhárom szinten.

- lélegzetvétel: nem jelöljük külön, csak néma szünetként: SIL
- sóhaj: SÓH
- hümmögések: úgy jelöljük, ahogy halljuk pl. ÜHÜM, ÜHM
- nevetés: NEV
- köhögés: KÖH
- nyelvcsettintés: CSET
- nyelés: NYEL
- torokköszörülés: TKS
- zaj, székrecsegés, gyomorkorgás: minden esetben a KUKA jelölést kapják
- nevetős beszéd: címkében a szöveg elé NEV-et írunk (szakaszszinten ilyenkor az egész beszédszakaszra érvényes a NEV jelölés; szószinten csak arra a szóra/szavakra, ahol elhangzik; beszédhangszinten csak azok elé a hangok elé írjuk be, amely hangokat érinti a nevetés). Hogyha nevetés hangzik el a másik beszéde alatt, akkor a nevető címkéjébe: EBNEV jelölést írunk.

3.5 Megakadásjelenségek

- téves kezdés, újraindítás, szünet a szóban, töredék

  Amennyiben a beszélő nem javítja ezeket a jelenségeket, akkor a szándékoltatnem tüntetjük fel sehogy. A beszédszakasz-szinten és szószinten a fragmentum (töredék) jelölése kötőjellel történik, pl. *karácsonykor ka- kacsa volt a menü*

  Hangszinten az egyes hangokat lejegyezve: k,a,k,a,C,a.
- nyelvbotlás

  Beszédszakasz szintjén azt jegyezzük le, amit hallunk, szögletes zárójelben megadva a szótári alakot, akkor is, hogyha a beszélő javítja, pl.

szeklény [szekrény]. Szószinten a szótári alakot jegyezzük le helyesírásban: *szekrény*. Hangszinten az elhangzott beszédhangokat: S,e,k,l,é,N.

- nyújtás

  Mivel a lejegyzés az objektivitásra törekszik, a nyújtás jelenségének megítélése pedig rendkívül szubjektív, az átirat ezt nem jelöli.

3.6 Köznyelvben használatos, de nem szótári alakjukban előforduló szavak, kifejezések

A beszédszakasz szintjén az elhangzott alakban jegyezzük le, például *nemtom*, *asszem, szal*, majd – a megakadásokhoz hasonlóan – feloldjuk a szótári alakkal: *[nem tudom], [azt hiszem], [szóval]*. A szószinten csak az elhangzott alakot tüntetjük fel, a hangszinten pedig a megvalósult beszédhangokat (n,e,m,t,o,m).

3.7 Kötőjeles és összetett szavak

Beszédszakasz- és szószinten helyesírásnak ellentmondóan sehol nincs kötőjel, egybeírjuk a teljes szót, pl. *spontánbeszédadatbázis*. (A kötőjelet a töredékes szavak jelölésére használjuk.)

3.8 Számok

A számokat mindig betűvel kiírva, a helyesírásnak ellentmondóan kötőjel nélkül egybeírva jegyezzük le. (pl. *kétezerháromszáztízben*)

3.9 Betűszavak

A betűszavakat kisbetűvel jegyezzük le, pl. *eltére*, a *magyénak* a vezetője. Beszédszakasz szinten feloldjuk őket a szótári alakkal, a toldalékolásra nem használunk kötőjelet, mert az kizárólag a töredékes szavakat jelöli: [ELTEre], [MAGYEnak].

3.10 Idegen szavak

Az idegen szavakat úgy írjuk le, ahogy a felvételen elhangzanak, pl. ***kvescsön***, ***kálgöriben*** *jártam a nyáron,* ***mencseszter junájtid*** *meccs volt tegnap az ertéel klubbon*. Beszédszakasz szinten feloldjuk őket a szótári alakkal, a toldalékolásra sem használunk kötőjelet, mert az kizárólag a töredékes szavakat jelöli: [question], [calgary].

3.11 Kérdés

Mindenféle (eldöntendő, kiegészítendő stb.) kérdő megnyilatkozás esetében az összes érintett szakasz elejére egy Q-t tegyünk.

4. Összegzés

A tervezett adatbázis lehetőséget nyújt arra, hogy a hangfájlok és az annotálások egyidejű elemzésével akusztikai fonetikai szempontból vizsgálhassuk a felnőtt

beszélők olvasott és spontán beszédének különféle jelenségeit. Olyan kutatások végzésére is alkalmas, amelyekhez nagy mennyiségű adat szükséges, és azok automatikusan lekérdezhetők, akár magyarul nem, vagy csak kevéssé tudó külföldi kutatók számára is.

Nagyméretű, strukturált és lekérdezhető beszédadatbázis annotálása és fejlesztése alapvető fontosságú nyelvészeti, fonetikai alapkutatásokhoz, beszédtechnológiai alkalmazásokhoz. Az ilyen adatbázison végzett kutatások új ismereteket nyújtanak a spontán beszéd sajátosságairól, és alapot adnak további nyelvészeti vizsgálatokhoz. Az adatbázis használatával idő takarítható meg, amely mérésekre és elemzésre fordítható. A fejlesztendő adatbázis – protokolljának, felvételi körülményeinek, az annotálásnak és mennyiségi jellemzőinek következtében (sok beszélő, beszélőnként relatíve nagy mennyiségű beszédanyag, különböző beszédtípusok) – jól használható nyelvek összehasonlításában, a beszédszintézis kutatásában, az adatbányászatban, a bűnügyi beszélőazonosításban, illetve klinikai diagnosztikai eljárásokban. Az annotált adatbázis nagy jelentőségű a beszédszintézisben és az automatikus beszédfelismerő rendszerekben, amelyek hozzájárulnak a hallássérült, siket és vak emberek kommunikációjának megsegítéséhez. A fonetikai kutatások a valós nyelvhasználatot írják le, új megközelítések válnak lehetővé, és újabb kérdések fogalmazhatók meg a szélesebb nyelvészeti kutatásokban is (pl. a spontán beszéd grammatikája, nyelvtipológia, univerzálék).

Irodalom

Barras, C., Geoffrois, E., Wu, Z., & Liberman, M. (1998) Transcriber: a Free Tool for Segmenting, Labeling and Transcribing Speech. *First International Conference on Language Resources and Evaluation (LREC)*. pp. 1373–1376.

Boersma, P. & Weenink, D. (2013) Praat: Doing phonetics by computer [Computer program]. Version 5.3.56. Letöltés: 2013. szeptember 15. http://www.praat.org

Gósy M., Gyarmathy D., Horváth V., Gráczi T. E., Beke A., Neuberger T. és Nikléczy P. (2012) BEA: Beszélt nyelvi adatbázis. In: Gósy M. (szerk). *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 9–25.

Gyarmathy D. és Neuberger T. (2011) A BEA adatbázis alkalmazásfüggő lejegyzései. *Beszédkutatás 2011*. 109–121.

Kontra M. (1988) Bevezető. In: Kontra M. (szerk.): *Beszélt nyelvi tanulmányok*. MTA Nyelvtudományi Intézet. Budapest. 1–4.

Maddieson, I. (1984) *Patterns of sounds*. *Cambridge studies in speech science and communication*. Cambridge: Cambridge University Press.

Olaszy G. (1999) Beszédadatbázisok készítése gépi beszédelőállításhoz. *Beszédkutatás 1999*. 68–89.

Váradi T. (2003) A Budapesti Szociolingvisztikai Interjú. In: Kiefer F. és Siptár P. (szerk.): *A magyar nyelv kézikönyve.* Akadémiai Kiadó, Budapest, 339–359.

Vicsi K. és Víg A. (1998) Az első magyar nyelvű beszédadatbázis. *Beszédkutatás 1998*. 163–178.

Vicsi K., Tóth L., Kocsor A., Gordos G. és Csirik J. (2002) MTBA – magyar nyelvű telefonbeszédadatbázis. *Híradástechnika* 8. 35–39.