

A PROZÓDIA SZEREPE A SPONTÁN BESZÉD TAGOLÁSÁBAN

Markó Alexandra

Bevezetés

A folyamatos beszéd feldolgozásának első lépéseként a percepció mechanizmus a folytonos akusztikus jelet egységekre bontja – ez elengedhetetlen feltétele az információ kinyerésének (a téma meghatározásának, a tartalmi elemzésnek). Míg az írott szövegek számos vizuális kulcsot szolgáltatnak a kisebb-nagyobb nyelvi egységekre történő tagolás számára (a szóközöktől a különféle központosási jelekig), az élőbeszédben jellemzően a prozódia segít(het)i ezt a folyamatot. Értelemszerűen a felolvasott szöveg esetében az írásképhez hű, általában nagyobb mértékben tervezett prozódia (Levelt 1989: 259) jobban segíti a percepció tagolást, mint a spontán beszédben, mivel ez utóbbiban a szupraszegmentális formák sokszor esetlegesen realizálódnak a tervezés és a kivitelezés kvázi egyidejűségéből adódóan. „A szupraszegmentumok létrehozása a [spontán] beszéd során bizonyos értelemben kevésbé tudatos, mint a beszédhangok sorozatáé [...] a szupraszegmentumok kontrollja sokkal kevésbé működik, a korrekciós folyamatok pedig szinte egyáltalán nem” (Gósy 2004: 182).

A szakirodalom szerint a tagolás és az értelmezés a percepció mechanizmusban egyszerre, egymásra utalva történik, egyazon folyamat két vetületeként (Chater–Christiansen 2001 – idézi Gósy–Kovács 2008). A tagolás az első lépése a különböző nyelvi egységek (szavak, mondategységek, mondatok) szintaktikai és szemantikai kapcsolata feltárásához vezető folyamatnak.

A tagolási folyamat működése és eredményessége azonban nemcsak a human percepció szempontjából jelentős kutatási kérdés. Az elméleti megközelítés számára is alapvető fontosságú annak tisztázása, hogy milyen egységekben írható le a (spontán) beszéd szupraszegmentális szerkezete. A szupraszegmentális szerkezet vizsgálatának elengedhetetlen feltétele annak meghatározása, hogy mit tekintünk a szupraszegmentális szint alapegységének. A szegmentális fonetikai megközelítésben vitán felül áll a beszédhang mint alapegység (annak ellenére, hogy a hangátmenet-hanghatár problematikája természeténél fogva megoldatlan marad), a szupraszegmentális megközelítésben azonban az alapegységet illetően sem elnevezésében, sem terjedelmében, tartalmában, sem definíciójában nincs konszenzus a kutatók között. A spontán beszéd vizsgálatában alapvető problémát jelent a hangfolyam tagolása, annak a közlésegyységnek a kijelölése, amely „mértékénél” és információtartalmánál

fogva alkalmas arra, hogy a beszéd szupraszegmentális alapegységének tekintsük. További kérdésként adódik, hogy ez az alapegység univerzálisnak tekinthető-e, vagy bizonyos nyelvi sajátosságok befolyásolják.

Elekfi 1962-es tanulmányában azt mondja, hogy „nagyobb értelmi és kritikai egységeket kell keresni, melyekből a mondat már viszonylag közvetlenül felépíthető. Alkalmas egységnek látszott erre a **beszédütem**” (25); ennek tekintette „minden mondatnak minden olyan szakaszát, mely szünetre következik, vagy nyomatékkal, illetőleg hangsúllyal kezdődik” (31), nagyjából tehát a klitikumos egységnek (egy főhangsúlyt viselő szócsoport) felel meg, azonban nem azonos a szólammal. Elekfi a *szólam* terminust többértelműsége miatt nem használja. Világosan látszik, hogy ez a megközelítés alapvetően a felolvasás hangzásának leírására alkalmas, hiszen a szünetek, hangsúlyok spontán beszédbeli esetlegességét nem kezeli.

Wacha (1988) a megnyilatkozást tekinti az élőbeszéd egységének, s a következőképpen definiálja: „**Megnyilatkozáson** [...] az élőszóbeli, a (spontán) beszélt nyelvi közlésegesznek (szövegnek, szövegegesznek, beszédműnek) azt a – pontosan többé-kevésbé – elkülöníthető kisebb részét/egységét értem, melyet az írott nyelvhasználatról szólva a mondat, szövegmondat terminussal szokás megnevezni. A megnyilatkozás a beszélt nyelvnek-nyelvhasználatnak olyan mondat értékű része tehát, melynek határait (kezdetét és végét) utólag – az elhangzó szöveg lejegyzésekor (átírásakor) – állapítottuk meg és jelöltük meg írásjelekkel, figyelembe véve az írásbeliség alapján kialakult mondatfelfogást (konvenciót) is” (152). A megnyilatkozások megnyilatkozásegységekből állnak (Deme 1971-ben definiált *mondategység* fogalmának analógiájára), így ezek tekintendők alapegységeknek, ugyanakkor Wacha használja a *fonemikus frázis* kifejezést is „hangképzési szünettől hangképzési szünetig tartó, csöndekkel határolt beszédszakasz” (109) értelemben.

Wacha problematikusnak tartja a megnyilatkozás határok megállapítását, végül az alábbi öt tényező közül valamely kettőnek az együttes jelenléte esetén tekint befejezettnek egy megnyilatkozást: „1. akusztikus zár, illetőleg ennek hiánya: a közlés dallama úgynevezett pont-hanglejtéssel zárul-e vagy sem, azaz a beszéd dallama mélyre szálló hanglejtéssel jelzi-e a megnyilatkozás befejezését, vagy esetleg nyitva tartott dallammal a folytatást ígéri (a beszédét vagy a megnyilatkozását); [Bekezdés.] 2. grammatikai zár: a megnyilatkozás grammatikailag befejezettnek, esetleg kereknek tekinthető-e vagy sem; [Bekezdés.] 3. követi-e az intonációs vagy grammatikai zárat új megnyilatkozást (új megnyilatkozás kezdetét) jelző intonációs indítás; [Bekezdés.] 4. követi-e vagy sem a megnyilatkozást valamiféle kiegészítő hozzáto-

1. Szende Tamásnál (1976) a *fonemikus frázis* – a fentiekől némiképp eltérően – szavak csoportját jelenti, amely hanglejtésformájával jellemezhető, és értelemmegkülönböztető szerepe van. „A magyarban fonemikus frázisok a beszéd folyamat grammatikai tagolódási egységeiben képződnek, illetőleg ezekben valósulnak meg” (161).

dás (pl. értelmező, valamilyen »mondatrész«); [Bekezdés.] 5. a megnyilatkozássorozatot megtöri, megszakítja-e szünet vagy sem” (107). Wach a megnyilatkozáshatár legbiztosabb jelzője az új megnyilatkozás kezdetét jelző intonáció megjelenése.

Bolla Kálmán a szupraszegmentális alapegység meghatározásakor (1992) ezt írja: „A szegmentális szerkezet struktúráképző alapeleme a beszédhang, míg a szupraszegmentális hangszövet legkisebb szerkezeti építőblokkját szupraszegmentális hangszerkezetnek nevezzük” (10), majd megjegyzi, hogy hasonló értelemben használják még az *intonációs szerkezet* kifejezést is. A bolliai *hangszerkezet* ez alapján tehát analóg a más szerzők által *intonációs frázis*-ként tárgyalt alapegységgel. Egy bekezdéssel lejjebb azonban ezt olvassuk: „a hangtest belső tagolódását, a hangfolyam szerkezeti felépítését állítom előtérbe [...] Ezáltal az alábbi egységekre tagolódik a hangsor: beszédhang, hangkapcsolat (szótag), fonetikai szó, beszédszakasz, megnyilatkozás, ... és a végén a beszédmű” (10); majd pedig: a „nagyobb nyelvi egységeket (szintagmákat, beszédszakaszokat) realizáló hangszerkezeti összetevőket, akusztikus formációt nevezünk **szupraszegmentális hangszerkezetnek**. A szupraszegmentális hangszerkezet a mondatfonetika alapegysége” (14). Alapegységnek tehát a szupraszegmentális **hangszerkezetet** (intonációs frázist) tekinti, de ugyanakkor bizonyos értelemben alapegységként kezeli az ezt „hordozó” beszédszakaszt is, amelynek mibenlétét, definícióját azonban nem adja meg. A beszédszakasz a fentiek alapján a szintagma és a megnyilatkozás közötti egységet jelentheti, tehát valószínűsíthetjük, hogy a wachai megnyilatkozásegységnek feleltethető meg. Bollánál azonban nem találunk olyan kritériumrendszert, amely alapján meghatározhatnánk az alapegység határait.

Németh T. Enikő (1996) a szóbeli diskurzusok **megnyilatkozáspéldányokra** tagolásakor arra a következtetésre jut, hogy figyelembe kell venni a dallam- és szünetprozodémák mellett a mondattal való kapcsolatba hozhatóságot, valamint pragmatikai szempontokat is. Ez utóbbiak között tekintetbe veszi az interperszonális funkciót, a beszédettsorozatok összetételét és a pragmatikai kötőszókat, valamint az attitudinális funkciót (és az azt mutató nyelvi eszközöket).

Varga László (1999–2001, valamint 2002) fonológiai szempontú megközelítése nem annyira a spontán beszéd szupraszegmentális tagolását célozza, mint inkább a fonológiai szabályrendszer működési hatóköréül szolgáló nyelvi egység mibenlétét határozza meg, de komplex megközelítése figyelembe vesz fonetikai szempontokat is. Szerinte a magyar **intonációs frázis** minimáldefiníciója: olyan szótagsorozat, amely vagy egy *a*) függelék-dallammal, vagy egy *b*) (előkészítő dallam +) karakterdallammal realizálódik (1999–2001: 7). Ugyanakkor az intonációs frázist olyan egységként is meghatározhatjuk, amely egynél több dallamhangsúlyt fog át, ezt nevezi Varga maximáldefiníciónak, amely szerint az intonációs frázis kétféle lehet: *a*) függelék típusú, amely csak egy függelék-dallamot tartalmaz és *b*) hangsúlyos, amely

legalább egy karakterdallamból áll (előtte állhat egy vagy több szünet nélküli féleső karakter, s ez(eke)t megelőzheti egy főhangsúlyt nem tartalmazó előkészítő dallam), és szünet követi (1999–2001: 8–9). Varga tehát egyszerre alkalmazza a minimális (Elekfihez hasonlóan) és a maximális (ahogy Wachá és Németh T. Varga [1981] ismeretében teszi) szekvenciákra tagolás elvét, s e kettő közül az utóbbit preferálja, az a szintaktikai szerkezetet is tükrözi. Ugyanakkor elismeri, hogy bizonyos spontán közlések esetében csak a minimaldefiníció alkalmazható.²

A percepció folyamatból indul ki Gósy Mária ún. virtuálismondat-kísérlete (2003). Következtetései szerint a spontán szöveg mondatokra tagolása nagymértékű (az adatok szerint mintegy 70%-os) egyöntetűséget mutat, és akusztikai fonetikai korrelátumai vannak. Így a szerző szerint joggal feltételezhető, hogy „a virtuális mondatok kognitív szinten egyértelműen léteznek – függetlenül [...] a grammatikai definíció nehézségeitől” (22). A kísérlet résztvevői a mondatvégek jelölésében elsődleges akusztikai kulcsként a szünet megjelenését és még inkább annak hosszát használták fel, másodlagosan pedig az alaphangmagasság-változásra támaszkodtak, ugyanakkor „statisztikailag nem támasztható alá az, hogy az alacsonyabb frekvenciaérték mondat-határ-ítéletet idézne elő” (36). Következtetései között Gósy azt is megfogalmazza, hogy önmagában egy prozódiai sajátosság sem bír határjelző funkcióval, csak más eszközökkel együtt, és csak akkor, ha szemantikai megerősítést is kap, illetőleg hogy percepció rendszerünk nem objektív értékek, hanem viszonyok alapján dönt. Ennek a kutatásnak az eredményei alapján Váradi Viola (2008) különböző spontánbeszéd-műfajokat vetett össze percepció szempontból, és azt találta, hogy a narratív és leíró szövegekben is eltérő a határjelzés megvalósulása. A percepció megközelítés is azt támasztja tehát alá, hogy nem adhatunk kimerítő választ a szupraszegmentális határjelzés és az alapegység definícióját érintő kérdésekre.

Összefoglalva: a szerzők elsősorban a szünet, a beszéddallam és a szintaktikai-szemantikai szerkezet határjelző funkcióit tartják kulcsszerepűnek, a meghatározások azonban mind beleütköznek abba a problémába, amit ezeknek a jellemzőknek az interferenciájaként fogalmazhatunk meg – azaz a spontán beszéd (szemben például az igényes felolvasással) nem mutat egyértelmű megfeleléseket ezeknek az eszközöknek a használatában.

A jelen kutatásban azt vizsgáltuk, hogy a két legjelentősebbnek ítélt prozódiai kulcsnak, a szünetnek és a dallammenet változásainak milyen és mekkora szerepe van a spontán beszéd tagolásában, illetőleg hogy ezek milyen mértékű összefüggést mutatnak a szöveg szintaktikai-szemantikai-pragmatikai sajátosságaival. A korábbi szakirodalom alapján megnyilatkozás méretű

2. Ezzel egyébként Varga szemben áll azzal a felfogással, mely szerint a (magyar) intonációs frázis alulról, fonológiai frázisok kapcsolataként szerveződő egység, amelyben a logikai forma is releváns (vö. pl. Hunyadi 1999 – idézi Varga 2002: 60).

egységeket vizsgáltunk, háromféle teszttel. Hipotézisünk szerint a megnyilatkozathatárok többsége sem prozódiai, sem lexikai szempontból nem határozható meg teljesen egyértelműen, a két vetület együttese azonban javítja a felismerés hatékonyságát.

Anyag, módszer, kísérleti személyek

A vizsgálat anyagát az ún. BEA magyar beszélt nyelvi adatbázisból válogattuk. (Az adatbázis az MTA Nyelvtudományi Intézetének Fonetikai Osztályán készül, jellemzőiről lásd Gósy 2008.) 10 beszélő (5 nő és 5 férfi) átlagosan 4 perces (3'23"–4'42") spontánbeszéd-anyagát választottuk ki, olyan interjúrészt, amelyben főleg magukról (munkájukról, hobbijukról) beszélnek. A szövegek együttes időtartama 39'59". A beszélők életkora 20–64 év között szóródik, átlagéletkoruk 41,1 év.

A 10 spontánbeszéd-anyagot lejegyeztük, és háromféle tesztnek vetettük alá. Az első tesztben Gósy (2003) módszeréhez hasonlóan a résztvevők megkapták a helyesírásban lejegyzett, de központozás nélküli szöveget, és az volt a feladatuk, hogy egy hallgatás után függőleges vonallal jelöljék meg azokat a helyeket, ahol „mondatvég” van (a mondat mibenlétére további meghatározást nem adtunk). Minden szöveget 10-en hallgattak meg, egy-egy minősítőnek 1–5 szöveget játszottunk le, számítógépről, aktív hangfal segítségével. Csoportosan végeztük a tesztelést. Ebben a kísérletben 56 (elsősorban magyar szakos) bölcsészhallgató vett részt, életkoruk 18–36 év között szóródott, átlagosan 22,9 év.

A második tesztben részt vevők csak a központozás nélkül lejegyzett szöveget kapták meg, a hangfelvételt nem hallották. Ugyancsak azt kértük tőlük, hogy jelöljék a mondatathatárokat függőleges vonallal. Minden szöveget 10-en teszteltek, személyenként 1–4 szöveget osztottunk ki a 32 bölcsészhallgatónak. A résztvevők életkora 18–37 év (átlagosan 21,5 év). Szintén csoportokban zajlott a kísérlet, és nem adtunk meg időkorlátot a szöveg feldolgozására.

A harmadik tesztbe alul áteresztő szűrővel módosítottuk a hangfelvételeket. A szűrőt mindig az adott beszélő hangterjedelmének megfelelően állítottuk be, az így előállított hanganyag értelme nem volt kivehető, a tesztelők csak a prozódia alapján dönthettek a mondatathatárok felől. Ezt a tesztet a résztvevők egyenként, fejhallgatóval, számítógépen, a Praat programban végezték el, az volt a feladatuk, hogy címkézzék fel a hangfelvételt a mondatathatárok jelölésével. Minden szöveget 10-en annotáltak, a 13 adatközlő személyenként 5–10 szöveget hallgatott meg. A résztvevők életkora 21–36 év, átlagosan 25,8 év. Az adatközlők magyar vagy más nyelv szakos tanárok és egyetemi hallgatók voltak.

Összesen 300 tesztelést végeztünk (10 szövegen 3 kísérletben 10-10 személlyel). A három különböző kísérletben részt vevők között nem volt átfedés (tehát egy-egy személy csak egyféle tesztben vehetett részt).

A teszteredményeket összesítettük, majd Pearson-próbával elemeztük, hogy a különböző tesztek eredménye között van-e összefüggés. Azokat a „mondatvég”-eket, amelyeket legalább 8-an jelöltek, további elemzésnek vetettük alá. Megvizsgáltuk – a hallgatott tesztek esetében – a megnyilatkozásvéget követő szünet idejét, valamint a megnyilatkozást záró szakasz dallamát, illetőleg ezek viszonyát egymáshoz. Az írásbeli teszt esetében a jelölt mondatok szintaktikai, szemantikai és pragmatikai sajátosságait elemeztük, olyan (elsősorban lexikai) kulcsokat kerestünk, amelyek motiválhatták a mondathatárok egyöntetű jelölését. Külön elemeztük a teljes egyöntetűséget (10-ből 10) mutató jelöléseket is.

Az akusztikai elemzéseket (szünetidőtartam- és F_0 -mérés) a Praat 5.1 szoftverrel, a statisztikai számításokat (korrelációanalízis, egytényezős varianciaanalízis) az SPSS for Windows 16.0 verziójával végeztük el.

Eredmények

A jelölt megnyilatkozásathatárok számának tekintetében a háromféle teszt jellegzetesen eltérő eredményeket hozott (1. táblázat), ugyanakkor egyikben sem függ össze a szövegenkénti jelölések száma a szöveg hosszával.

1. táblázat: A jelölt megnyilatkozásathatárok száma szövegenként
(N = nő, F = férfi)

A beszélő azonosítója	Beszédidő	A megnyilatkozásathatárok száma (ebből 100%-ban egyöntetű jelölés)		
		Szűrt hanganyag (csak prozódia)	Teljes hanganyag	Írott szöveg (hang nélkül)
N1	4'42"	73 (2)	55 (5)	64 (0)
N2	3'23"	56 (0)	68 (1)	61 (0)
F1	3'55"	51 (1)	45 (3)	50 (0)
F2	4'17"	71 (2)	40 (5)	53 (1)
N3	4'09"	87 (2)	63 (6)	56 (2)
F3	4'03"	77 (1)	62 (8)	70 (2)
F4	3'42"	65 (2)	60 (4)	56 (2)
N4	4'21"	77 (0)	38 (3)	47 (3)
N5	3'39"	66 (2)	61 (4)	56 (3)
F5	3'48"	57 (0)	42 (4)	52 (0)
Összesen	39'59"	680 (12)	534 (43)	565 (13)

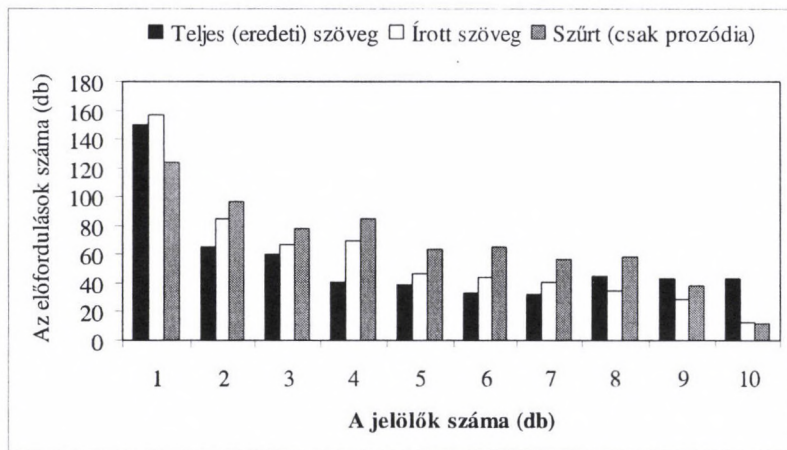
A tesztelők a legtöbb megnyilatkozásathatárt a szűrt hanganyagban jelölték: szövegenként 51–87 helyen jelölt legalább egy résztvevő „mondatvég”-et. Emellett ebben a kísérletben nagyon ritka volt az is, hogy mind a 10 tesztelő egyaránt megnyilatkozásvéget észlelt volna ugyanazon a ponton, ez mindössze az összes jelölt hely 1,8%-ában fordult elő. Ennek oka feltehetően a nagyfokú bizonytalanság, amelyet a nyelvi szerkezet hiánya okozott.

Arányaiban a legkevesebb és egyben legeggyöntetűbb jelölést a teljes hanganyag (a nyelvi és prozódiai szerkezet együttese) ismeretében kaptuk, itt szövegenként 38–68 helyet ítélték határnak a résztvevők, és minden szövegben volt legalább egy olyan határpont, amelyet egyöntetűen jelöltek. A szövegek közötti változatosság itt mutatkozik meg a legjobban, ami arra utal, hogy a nyelvi és prozódiai megvalósulás nagymértékben szöveg- és/vagy beszélőfüggő. Ebben a kísérletben összességében a határjelölések 8,1%-ában értett egyet minden tesztelő. Látható, hogy önmagában még ez az érték sem mutatja, hogy a magyar anyanyelvű beszélők a spontán beszéd „mondat”-ait azonos módon ítélnék meg.

Az írott szöveg láttán (a szöveg hangzásának ismerete nélkül) valamivel egyöntetűbb volt a jelölés, mint a szűrt hanganyag alapján, azaz a prozódia önmagában kevesebb támpontot adott, mint a szintaktikai, szemantikai és pragmatikai információk. A szövegenkénti 47–70 határjelből mindösszesen 2,3%-nyi volt olyan, amelyet minden tesztelő egyaránt mondatvégnak ítélt.

A háromféle teszt adatainak összevetése alátámasztja azt a megállapítást, hogy a prozódiai határok nem tükrözik ugyan pontosan a szintaktikai határokat, ugyanakkor a prozódia egyértelműsít(het)i a beszéd szintaktikai szerkezetét (Batliner et al. 1998).

A határjelölések relatív bizonytalanságát az is mutatja, hogy a legtöbb tagoló pontot csak egy-egy személy jelölte meg tesztenként (1. ábra).



1. ábra

Az ugyanazon a helyen megjelent határjelölések száma a három tesztben

Az írott szöveg és a szűrt hangfelvétel alapján történt jelöléseknél egyértelműen látszik az a tendencia, hogy a hallgatói ítéletek azonosságának növe-

kedésével csökken az általuk határnak ítélt szöveghelyek száma. Ugyanakkor az eredeti, teljes szöveg esetében ez a tendencia megtorpan, és 8 vagy több azonos jelölés egyenlő (és a korábbiaknál nagyobb) számban fordul elő.

Megvizsgáltuk, hogy az egyes határhelyek az egyes szövegtípusokban milyen mértékű egyezést mutatnak, azaz hogy ugyanott jelöltek-e tagolási pontot a résztvevők a szűrt és a teljes hanganyagban, illetve pusztán az írott szöveg ismeretében. A korrelációelemzés eredményeit a 2. táblázat foglalja össze – szövegenként és a teljes adathalmazra.

2. táblázat: A jelölések közötti összefüggések a három szövegtípusban

Beszélő		Teljes hanganyag	Írott szöveg
N1	Szűrt hanganyag (csak prozódia)	$r = 0,503; p < 0,001$	$r = 0,382; p < 0,001$
N2		$r = 0,465; p < 0,001$	$r = 0,419; p < 0,001$
F1		$r = 0,491; p < 0,001$	$r = 0,352; p = 0,002$
F2		$r = 0,575; p < 0,001$	$r = 0,411; p < 0,001$
N3		$r = 0,502; p < 0,001$	$r = 0,215; p = 0,022$
F3		$r = 0,398; p < 0,001$	$r = 0,031; p = 0,736$
F4		$r = 0,539; p < 0,001$	$r = -0,050; p = 0,624$
N4		$r = 0,562; p < 0,001$	$r = 0,430; p < 0,001$
N5		$r = 0,428; p < 0,001$	$r = 0,020; p = 0,842$
F5		$r = 0,482; p < 0,001$	$r = 0,310; p = 0,005$
Összesen		$r = 0,482; p < 0,001$	$r = 0,226; p < 0,001$

Beszélő		Írott szöveg
N1	Teljes hanganyag	$r = 0,872; p < 0,001$
N2		$r = 0,899; p < 0,001$
F1		$r = 0,869; p < 0,001$
F2		$r = 0,878; p < 0,001$
N3		$r = 0,726; p < 0,001$
F3		$r = 0,527; p < 0,001$
F4		$r = 0,392; p < 0,001$
N4		$r = 0,784; p < 0,001$
N5		$r = 0,547; p < 0,001$
F5		$r = 0,813; p < 0,001$
Összesen		$r = 0,714; p < 0,001$

A legerősebb összefüggést a teljes hanganyag és az írott szöveg alapján jelelt határok között találjuk: minél többen jelöltek határt egy adott ponton az eredeti hangfelvétel alapján, annál többen jelöltek mondatvéget ugyanott pusztán a szintaktikai, szemantikai, pragmatikai sajátosságok ismeretében. A Pearson-próba minden beszélő szövegénél szignifikáns eredményt adott,

bár vannak a korrelációs együttthatót tekintve kiugró szövegek (F3, F4 és N5 esetében csak gyenge vagy közepes összefüggés mutatható ki). Szignifikáns, közepes összefüggés mutatkozik a teljes hanganyagban és annak a szűrt (csak prozódiaát tartalmazó) változatában jelölt határok között (s bár itt nincsenek jelentős kiugrások, F3 és N5 esetében itt is kisebb korrelációs együttthatót találunk). Az írott szöveg és a szűrt hanganyag jelölései között többnyire szignifikáns, gyenge összefüggést találunk, illetőleg F3, F4 és N5 szövege esetében nincs korreláció.

A különböző szövegtípusok határjelölései közötti összefüggések vizsgálata azt mutatja, hogy a spontán beszéd megnyilatkozásáthatárainak megállapításában a legnagyobb mértékben a lexikális-grammatikai információra támaszkodik a percepció, és csak jóval kisebb mértékben veszi figyelembe a prozódiai struktúrát. Megjegyezzük azonban, hogy felolvasott szöveg esetén a prozódia és a szintaktikai-szemantikai struktúra illeszkedése akár száz százalékos is lehet. Egy ellenőrző kísérletben egy (nem hivatásos beszélő által) felolvasott szövegnek a fentiekhez hasonló módon szűrt változatát címkéztettük fel a résztvevőkkel, és minden jelölés tökéletesen egybeesett az eredeti írott szöveg mondathatáiraival. Váradi (2009) kísérletében is sokkal egyöntetűbbek voltak a jelölések a felolvasásban, mint a spontán narratívában. A spontán beszédben a prozódia tervezése a lexikális szintéhez képest csak másodlagos, megvalósulása ezért nagymértékben esetleges – szemben azzal, hogy felolvasáskor a produkciós folyamat nagyobb részét az artikuláció és benne a szupraszegmentális szerkezet tervezése teszi ki.

Elemztük, hogy a különböző szövegváltozatokban milyen (akusztikai és nyelvi) sajátosságok indokolhatták a határjelölést. A kutatás ezen részében azokat a szöveghelyeket vizsgáltuk meg (az 1. ábrán látható eredmények alapján), amelyeket a 10 tesztelő közül legalább 8-an azonosan ítélték meg.

Az **eredeti hangfelvétel** tesztadatai közül 129 felelt meg a fenti kritériumnak, 43-43 olyan tagoló pont van, amelyet 8, 9 vagy 10 személy jelölt. Ezeknek a megnyilatkozásoknak a végét négy esetben nem követi szünet, ekkor (egy esetben) ereszkedő a dallam, illetve (három esetben) glottalizált a szakaszvég (a glottalizáció is ereszkedő dallam/lelépés benyomását kelti a percepcióban, vö. Böhm–Ujváry 2008, Markó 2009). A szakirodalom szerint a szünet nagyon erős szereppel bír a határjelzésben (lásd a Bevezetésben idézett forrásokat). Mivel ezeken a helyeken nincs szünet, kérdés, hogy az ereszkedő jellegű dallam és/vagy a szöveg lexikális-grammatikai, illetve pragmatikai sajátosságai jelzik a megnyilatkozásáthárt, vagy más kulcs szolgál erre. A követő megnyilatkozás eső indítása jelzi a határt az első két példában (a példákat a tesztben követett módszernek megfelelően központosázás nélkül közöljük):

- (1) *úgyhogy én leginkább ezt kaptam egyelőre | meg hát persze rossz dolgokat is de azokról nem beszélünk*

- (2) *vitatkozzanak vele ha akarnak | hát mi még nem ezt a gyakorlatot követjük*

A (3) példában nem a határponton, hanem a következő szavak között jelennek meg rövidebb szünetek (szögletes zárójelben lásd az időtartamukat), amelyek alapján a hallgatók tagolhatták a szöveget, persze értelemszerűen a szintaktikai határra jelölték a közlésvéget:

- (3) *szal van egy kislány meg egy kisleány unokám | úgyhogy [147 ms] nagygyon [263 ms] örülök nekik*

Az iménti három esetben nem jelöltek a résztvevők egyértelmű határpusztán az írott szöveg alapján végzett tesztben, tehát a nyelvi struktúra önmagában nem volt elég erős jelzés. A (4) példában azonban a szövegtartalom lehetett a kiváltó kulcs (csak az írott szöveg láttán is jelölték ezt a helyet), egyrészt a *hát* diskurzusjelölő új témát indító használata, másrészt az alany- és a bizonyos mértékű témaváltás:

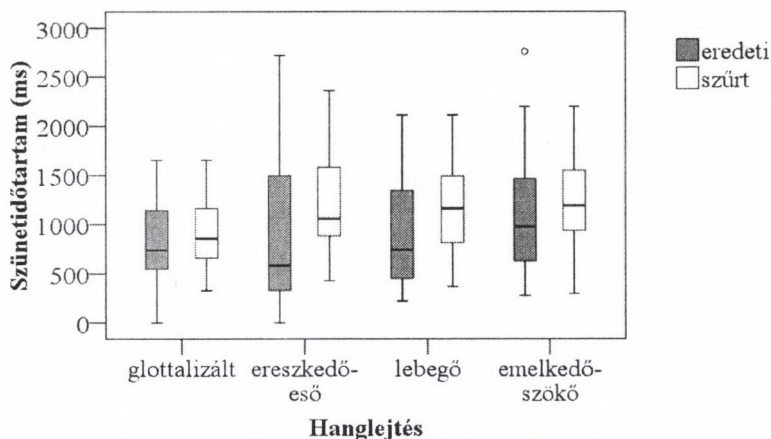
- (4) *ami azt jelenti hogy a föld feletti szerkezeteknek a statikájával kell foglalkozni | hát [1063 ms] a Lakótervben szereztem a legnagyobb gyakorlatot abból hogy hogyan kell egy épületszerkezetnek tényleg megállnia*

Ebben az esetben különösen érdekes, hogy a hosszú szünet nem a jelölt határon, hanem a *hát* után jelent meg, ennek megfelelően a szűrt hanganyagban végzett tesztben viszont természetesen ezt a szünethelyet jelölték tagoló pontnak a résztvevők.

A további 125 jelölt megnyilatkozásáért jelző szünet hossza 45–2759 ms közé esett, átlagosan 922,8 ms, az átlagos eltérés 414,1 ms. A szünetek előtti szövegrész 50,4%-ban glottalizáltan valósult meg (más kutatások is kimutatták, hogy a glottalizáció tipikus határjelző zöngeminőség a magyarban és más nyelvekben is, vö. pl. Redi–Shattuck–Hufnagel 2001, Markó 2009). A szakaszvégek hanglejtése (a hanglejtés minősítését Markó 2009-es kísérleti eredményei alapján végeztük el) az esetek további 16,8%-ában emelkedő, 13,6%-ban ereszkedő, 11,2%-ban szökő, 6,4%-ban lebegő és 1,6%-ban eső.

A szünethosszok összefüggése a szakaszvég hanglejtésével látható a 2. ábrán (eredeti hangfelvétel: szürke oszlopok; a szünet nélküli helyeket is ábrázoltuk, lásd 0 ms-os értékek).

A glottalizált szakaszvégek utáni szünetek 45 és 1654 ms között szóródnak, átlagosan 849,4 ms-osak, átlagos eltérésük 306,4 ms. Az ereszkedő és eső hanglejtést követően realizálódó szünetek időtartama 125–2720 ms, átlaguk 898,5 ms, átlagos eltérésük 569,7 ms. A lebegő dallamrealizációkhoz 219–2114 ms-os szünetek társulnak, átlag: 926,8 ms, átlagos eltérés: 506,4 ms. Az emelkedő és szökő hanglejtést követő szünetek 278–2759 ms-osak, átlagosan 1095,9 ms időtartamúak, átlagos eltérésük 464,0 ms. Jól kirajzolódik az a tendencia, hogy minél jobban eltér a szakaszzáró hanglejtésforma az ún. ponthanglejtéstől, átlagosan annál hosszabb szünet követi.



2. ábra

A szakaszrő hanglejtés és a szünetidőtartam összefüggése az eredeti és a szűrt (csak prozódiaát tartalmazó) hangfelvétel alapján végzett tesztek eredményeiben

A csak prozódiaát tartalmazó, **szűrt hangfelvételek** tesztjében 109 határt jelöltek legalább nyolcan, ebből 59-et nyolcan, 38-at kilencen és 12-t tízen. Ezek között egy olyan helyet sem találunk, ahol nem volt szünet, a legrövidebb szünet 298 ms, azaz mintegy szótagnyi időtartamú volt – a szöveg értelmében ennyi kellett ahhoz, hogy (relatív) egyöntetű határérték alakuljon ki a hallgatókban. Természetesen rövidebb szünetek is kaptak egy-egy jelölést (olyan is előfordult, hogy a zöngétlen mássalhangzó néma fázisát is jelölték határpontként), de ezek elszórtak voltak, a tesztelők nem vélekedtek róluk (kvázi) azonosan. A leghosszabb egyformán jelölt szünet 2361 ms volt, az átlagérték 1076,8 ms, az átlagos eltérés 374,2 ms.

Ezúttal is glottalizált szakaszvégek jelentek meg a határjelöléseknél a legnagyobb arányban, 51,4%-ban. A közlések további 14,7%-a emelkedő, 12,8%-a ereszkedő, 11,9%-a szökő, 8,3%-a lebegő és 0,9%-a eső hanglejtéssel realizálódott. A 2. ábra fehér oszlopai mutatják a szünethossz, valamint a zöngemínőség és a hanglejtés összefüggéseit a szűrt felvétellel készült teszt eredményeiben. A glottalizált szakaszvégeket követő szünetek 328–1654 ms-ban realizálódtak, átlagosan 931,9 ms-ban, átlagos eltérésük 306,5 ms. Az ereszkedő és eső hanglejtésű közlésvégek után 428 és 2361 ms közötti tartamú szüneteket tartottak a beszélők, átlagosan 1209,2 ms-osokat, amelyek átlagos eltérése 447,8 ms volt. A lebegő zárathoz tartozó szünetek adatai 370–2114 ms között szóródnak, átlaguk 1191,6 ms, átlagos eltérésük 409,5 ms. Végül emelkedő és szökő hanglejtésnél 298–2199 ms-os szünetek esetén jelöltek

határt az adatközlők, ezek átlaga 1252,7 ms, átlagos eltérése 397,9 ms. Itt is érvényesül az a tendencia, hogy az emelkedő és szökő hanglejtéshez tartozó szünetek átlagosan hosszabbak, mint az ereszkedő típusú dallamot követők. Emellett a szünetek minimumértékei minden hanglejtés esetén nagyobbak, mint a teljes hanganyagban jelölt határoknál, ami ismét arra utal, hogy a tartalom ismeretlen volta esetén csak a hosszabb szünet stimulálja a határérzékelést. Az egytényezős varianciaanalízis szerint a szűrt tesztben jelölt szünetek értékei szignifikánsan eltérnek az eredeti szöveg alapján jelöltekétől: $p = 0,018$, $F(1, 232) = 5,631$.

A pusztán az **irott szöveg** alapján végzett tesztben 76 helyet jelöltek hétnél többben, ebből 34-et nyolcan, 29-et kilencen és 13-at tízen. Megvizsgáltuk, hogy milyen lexikai-pragmatikai elemek utalhattak megnyilatkozásátharra. Többször a következő közlés elején megjelenő ún. pragmatikai kötőszók lehettek a határjelölők, például: *és, de, tehát, úgyhogy*. Gyakori témaindító a *hát*, illetve más diskurzusjelölővel együtt is előfordult: *hát egyébként*. A szakirodalom a *hát* szerepét elsősorban válaszelőzőnek tartja (Schiffrin 1987, Németh T. 1996) – ezekben a szövegekben (minthogy monologikus jellegűek) ez a funkció nem érvényesült. Németh T. (1996) a saját korpuszában az önjavítás vagy a magyarázkodás eszközeként találta a leggyakrabban. Ebben az anyagban azonban ez nem volt jellemző, lásd az alábbi példákat:

- (5) *tudtam hogy bölcsész szak és akkor olyan evidensnek tűnt hogy magyar latin | hát aztán sikerült*
- (6) *azt mondtam hogy ez egyszerűen fantasztikus | hát remélem hogy örökölnék tőle ilyen tehetséget*

A *mondjuk*, az *aztán* és a *na most* is szerepelt határjelzőként. További, komplex lexikai szerkezetű markerek: *úgy volt igazából hogy; meg az se utolsó szempont hogy*. Az esetek 55,3%-ában azonban nem találtunk a szövegben tagoló pontra utaló diskurzusjelölőt, illetve olyan is előfordult, hogy megjelent például a *hát* a szövegben, mégsem ítélték „mondathatár”-nak ezt a helyet. Témaváltás motivál(hat)ta a határjelölést – a diskurzusjelölő(k) ezt erősíti(k) –, például itt:

- (7) *és akkor kerestem egy munkahelyet magamnak ami a kilencedik kerületi önkormányzat építésügyi osztálya lett s végül is innen mentem nyugdíjba | hát egyébként ahhoz hogy az ember statikus legyen el kell végezni az egyetemnek megfelelő szakát*

Az összegző megállapításokat rendszerint új megnyilatkozásnak érezték a résztvevők, például:

- (8) *sőt hát volt olyan időszakom amikor volt öt emberem aki a műhelybe dolgozott egy közel kétszáz négyzetméteres műhelybe | hát voltak jobb időszakok*

- (9) *s aztán elmondtam neki hogy az óvodásoknak meg tudja magyarázni hogy az építész megálmodja az épületeket a statikus pedig számításokkal alátámasztja hogy ez az épület meg fog állni | nagyon érdekes dolog tulajdonképpen*

Az idiomaszerű kollokációk is motiválhatták a határjelölést:

- (10) *emíatt ez a fajta tevékenység egy picit hátrébb került | vanni van szóval mindig van munkám*
- (11) *főleg az elején mikor a nyers bőrök bekerülnek | van úgy hogy a hajóúton meg ugye Ausztráliából meg nagyon messzi földrészekről hoznak bőröket*

Egy esetben a modalitás megváltozása jelzett mondathatárt:

- (12) *az viszont kap belőle jutalmat | ne kapjon jutalmat*

Úgy tűnik, hogy a mutató és a személyes névmásként realizálódó alany ugyancsak mondatkezdő hatást kelt, például:

- (13) *meg gyerekkoromba is én mindig egyszerre többfélét csináltam | ez a munkámnál is így van*
- (14) *mer mindenki irodalom miatt megy oda | én viszont kifejezetten örültem annak hogy azért e ennyi nyelvtani dolog van*

Több esetben a megakadásjelenség felbukkanása miatt tagolhatták a szöveget, például:

- (15) *én kertésztechnikumot végeztem és van egy másik szakmám is az tímár | ééés sokáig négy gyerekem van és sokáig voltam otthon velük ilyen anyaságin*

Felmerül még az a lehetőség is, hogy a szövegrész hossza motiválja a határjelölést: az olvasó úgy érzi, hogy nem szerencsés, ha egy mondat másfél-két sornál hosszabb. Összegezve: nem nagyon találunk olyan egyértelmű jelzéseket, amelyek önmagukban a mondatot mint egységet az írott szövegben minden kétséget kizáróan behatárolják. Ezt az a mindennapi tapasztalat is alátámasztja, hogy újraolvasáskor a saját írott szövegeinket is könnyedén átszerkesztjük, például egy-egy többsoros mondat újratagolásával, egy-egy mondategységnek a szomszédoshoz való csatolásával stb.

25 határt mindhárom tesztben 8-an vagy többen jelöltek, vagyis ezeknél a prozódia és a nyelvtani-lexikai struktúra jól illeszkedett egymáshoz. A realizációk beszélőfüggő voltára utal azonban, hogy volt olyan szöveg, amelyben egy ilyen egybeesést sem találtunk, míg olyanok is, amelyben hatot. Például (szögletes zárójelben a szünet hossza és a tagoló pont előtti szövegrész hanglejtése):

(16) *amin dolgozom az a orvosi nukleáris műszerek tervezése* [428 ms, ereszkedő] *hát most már nem tervezek mert már nyugdíjas vagyok*

(17) *ennek a egyik fő objektuma az pontosan a Gamma volt* [1492 ms, lebegő] *aztán utána elvégeztem a egyetemet de előtte még katona is voltam*

A (16) példában a relatíve hosszú szünet és az ereszkedő hanglejtés a prozódiai kulcs, lexikai markerként pedig a témakezdő szerepű diskurzusjelölő *hát* szerepel. A (17) példában a kifejezetten hosszú szünet egyértelműsíti a határt (így a lebegő dallam sem zavarja meg a percepció döntést), lexikai szinten a markáns témaváltás, nyelvtani szempontból pedig az alanyváltás utal a megnyilatkozásvégre.

21 tagoló pontot jelöltek azonosan (7-nél többen) az írott szöveg és az eredeti hangfelvétel alapján, ezeknél a szövegrészeknél tehát a prozódia önmagában nem hordozott elegendő információt a megnyilatkozásvégre vonatkozóan. 33 határ pedig a szűrt és az eredeti hangfelvételt elemző tesztben esett egybe – ezekben az esetekben vélhetően a prozódia egyértelműsítette a szöveg alapján bizonytalan határt, például:

(18) *utána le kell ezt a bőrt meszel mesz meszes üzembe kerül* [765 ms, glottalizált] *és a meszes üzembe ezt leszórtelenítik*

(19) *igaz hogy csak a szigorlat után vagy amikor készültem arra* [1474 ms, szökő] *úgyhogy az abszolút jó volt*

Az utóbbi példában a szökő hanglejtés ellenére a prozódia megerősítette a határt, feltehetően a közel másfél másodperces szünet hatására. Végül mindössze két esetben fordult elő, hogy a csak prozódia és a csak az írott szöveg alapján alapuló tagolás egybeesett, miközben az eredeti hangfelvétel alapján csak néhányan jelöltek határt.

Következtetések

Magyar és külföldi kutatások egyaránt foglalkoztak már a tagolás kérdésével a produkciós és percepció szempontból – nemcsak az olvasott, hanem a spontán beszédben is. Olyan kísérletekről is olvashatunk, amelyekben a prozódia és a lexikális információt különválasztva kezelték, tesztelték, elemezték (pl. Lehiste–Wang 1977, de Pijper–Sanderman A. 1994, Gósy–Kovács 2008). A jelen kutatásban arra tettünk kísérletet, hogy megvizsgáljuk a prozódia szerepét ebben a folyamatban, illetve meghatározzuk azokat az összefüggéseket, amelyeket érdemes figyelembe venni a további kutatásokban.

Nemcsak a humán percepció megismerése és a hangzó nyelv leírása számára alapvető fontosságú a beszéd tagol(ód)ásának problémája, a mesterséges beszéd felismerés egyik legnehezebben megoldandó feladata ugyancsak az, hogy a beszéd nyelvi szerkezete hogyan dekódolható automatikusan, a prozódiai jegyek segítségével. Langlais és Méloni (1993) három típusban határozta meg a prozódia alapú automatikus felismerés nehézségeit:

a) a prozódiai tudás jelentős kontextusfüggő változatossága (a beszéd műfaja, a beszélő személye, a mondatok szerkezete és tartalma, a környezet sajátosságai – pl. zajossága – stb.);

b) a szupraszegmentális információ és az üzenet nyelvi szerkezete közötti kapcsolatok komplexitása;

c) a szupraszegmentális paraméterek pontos mérésének nehézségei, illetőleg a percepció szintjén történő integrálásuk.

Mindemellett további nehézséget jelent, ha a bemeneti szöveg spontán beszéd, amelyben indokolatlan intenzitáscsökkenések, néma és kitöltött tervezési szünetek, renyhe artikuláció ronthatja a beszédfelismerés pontosságát (Szászák–Vicsi 2007). Shriberg és munkatársai (2000) olvasott (rádióhírek) és spontánbeszéd-korpusz (Switchboard) összevetésekor azt találták, hogy míg az előbbiben a dallam megvalósulásai jó határjelzőnek bizonyultak, a spontán beszédben nem – itt az időtartam és a szóhatár volt megfelelő kulcs. Ugyanakkor mindkét típusú szövegekben nagy határjelző erővel bírtak a szünetek. Batliner és munkatársai (1998) szerint mindezek miatt az olvasott korpuszok automatikusan címkézhetők, míg a spontán szövegek esetében csak a manuális címkézés vezethet megfelelő eredményre.

E tekintetben figyelembe veendő még az emberi és a gépi feldolgozó mechanizmus működése közötti különbség is. Kutatások igazolják, hogy az emberi percepció hatékonysága és pontossága jelentősen felülmúlja a gépét, és ez a különbség többszöröse a zajos környezet, illetve spontán beszéd esetén. Míg a rosszabb körülmények ellenére az emberi percepció 5%-os hibaszázalékon belül teljesít, a számítógép teljesítménye akár 40%-os hibaszázalékig is leromolhat (Lippmann 1997).

A jelen kutatásban elvégzett három teszt eredményei jellegzetesen különböznek egymástól. Mint láttuk, a legkevesebb határt az eredeti, teljes hangfelvétel alapján jelölték a hallgatók, tehát ebben volt a legbiztosabb, leggyöntetűbb a döntésük. Itt a határok 24,2%-át jelölték legalább 8-an, és 8,1%-ban mindenki. Az írott szöveg alapján végzett tesztben, ahol nem hallották a hangfelvételt, valamivel több határhelyet találunk, ezek 13,5%-át jelölték legalább 80%-os egyöntetűséggel, 2,3%-ot pedig mind a tízen. Végül a szűrt hangfelvétel (csak prozódia) alapján jelentősen több határt jelöltek a résztvevők, vagyis itt volt a legnagyobb a bizonytalanság. Ugyanakkor 16,0%-ban jelölték ugyanazt legalább heten, bár összesen csak 1,8% volt teljesen egyöntetű. Érdekes továbbá, hogy a szűrt felvételen a legalább 8 tesztelő által jelölt határok közül csak 8 olyan volt (7,3%), amely nem szintaktikai határra esett. Azaz valamiféle összefüggés itt is tapasztalható volt a mondategység-határ és a prozódiai frázis határa között.

A korrelációanalízis azt az eredményt hozta, hogy a teljes hanganyag alapján történt jelölések nagyobb mértékben függenek össze az írott szöveg határjeleivel, mint a prozódiaalapú címkékkel. Az olvasott szövegen alapuló kontrolltesztben ugyanakkor a szintaktikai-prozódiai határok teljes egybeesése

látszott. Holland olvasott mondatokra de Pijper és Sanderman (1994) egy hasonló vizsgálatban 0,92-es korrelációs együtthatót kapott ($p < 0,01$), spontán beszédre azonban hasonló módszerrel végzett vizsgálatból származó korrelációs adatot nem találtunk az irodalomban.

A vizsgálat fontos tanulságának tartjuk, hogy egyes beszélőknél az általános tendenciák nem vagy kevésbé érvényesültek. Ez arra utal, hogy a tagolás beszélőfüggően realizálódik, véleményünk szerint összefügghet a beszélő koncentráltóságával/fáradtságával, a témához való viszonyával és más, a beszédtervezést befolyásoló tényezőkkel is.

A prozódiaalapú tesztekben (tehát azokban, amelyekben valamely hangfelvételt játszottuk le az adatközlőknek) jelölt határok mintegy felében glottalizált volt a szakaszvég. Egy korábbi vizsgálatban a spontán beszédben jelölt megnyilatkozásvégek 20-30%-a volt glottalizált, illetve a nők olvasott beszédében ez az arány meghaladta a 60%-ot (Markó 2009). Bár a glottalizáció egyéni beszédjellemző (vö. Böhm–Ujváry 2008), a jelen kutatásban minden beszélőnél előfordult, változó előfordulási gyakorisággal. Ez a gyakoriság felveti, hogy a glottalizációnak a magyarban is (más nyelvekhez hasonlóan, vö. Redi–Shattuck–Hufnagel 2001) határjelző funkciója (is) van. Különösen fontos lenne tehát, ha a beszédtechnológiában, illetve azon belül az automatikus beszéd- és beszélőfelismerésben kezelni tudnánk ezt a jelenséget. Bár már léteznek olyan alkalmazások (pl. automatikus osztályozó, vö. Böhm–Both–Németh 2009), amelyek bizonyos feltételek között képesek a glottalizáció azonosítására, a számítógépes felismerők ma még messze nem alkalmasak erre (vö. Böhm 2009).

Az eredmények azt mutatják, hogy a hallgatók részéről nem elvárás, hogy ereszkedő típusú hanglejtés jelenjen meg a tagoló ponton, hiszen a teljes hangfelvétel alapján 28,0%-ban, a szűrt hanganyag alapján pedig 26,6%-ban emelkedő/szökő dallamnál jelöltek határt. Ugyanakkor felismerhető volt az a tendencia, hogy a szünetek átlagos időtartama ezeknél a hanglejtésmintázatoknál hosszabb, mint az ereszkedő vagy a glottalizált szakaszvégeknél.

A jelen kutatás is igazolta, hogy a tagolásnak vannak akusztikai fonetikai korrelátumai (vö. Gósy 2003), az azonban, hogy a határjelző markerek egyértelműek és minden kétséget kizáróak lennének, az eredmények alapján nem bizonyítható. A prozódiai jelzések közül a kutatások a szünetnek tulajdonítják a legfontosabb szerepet (Carlson–Hirschberg–Swerts 2005), s bár az általunk vizsgált anyagban is jellemző volt a szünetek megjelenése a jelölt határokon, négy esetben nem tartott szünetet a beszélő. Az alaphangmagasság tulajdonságai közül egyet sem találtunk kizárólagosnak, bár kétségtelen, hogy a glottalizált szakaszvég a határjelölések túlnyomó többségét jellemezte. A prozódia és a lexikális jellemzők összefüggése sem egyértelmű, bár a szuprasegmentális sajátosságok pontosíthatják a szintaktikai-szemantikai szempontból kétértelmű szerkezeteket.

A mondat/megnyilatkozás azonosítása a spontán beszédben sem lexikai, sem szupraszegmentális szinten nem problémamentes, hiszen egyelőre nem találtunk olyan jegyeket, amelyek minden (kvázi) egyöntetű határjelölés esetében magyarázó erővel bírának. Felmerül tehát a kérdés, hogy akár a humán percepció, akár a gépi felismerés számára megfelelő alapegység-e a mondat/megnyilatkozás, illetőleg hogy milyen más szempontokat lenne érdemes bevonni a további vizsgálatokba, valamint a modellalkotásba.

Irodalom

- Batliner, Anton – Kompe, Ralf – Kießling, Andreas – Mast, Marion – Niemann, Heinrich – Nöth, Elmar 1998. M = Syntax + Prosody: A syntactic-prosodic labeling scheme for large spontaneous speech databases. *Speech Communication* 25. 193–222.
- Boersma, Paul – Weenink, David 2009. *Praat: doing phonetics by computer* (Version 5.1). http://www.fon.hum.uva.nl/praat/download_win.html
- Bolla Kálmán 1992. *Szupraszegmentális elemzések*. Egyetemi Fonetikai Füzetek 7. ELTE Fonetikai Tanszék, Budapest.
- Bóhm Tamás 2009. *Analysis and modeling of speech produced with irregular phonation*. PhD thesis. BME, Budapest.
- Bóhm Tamás – Both Zoltán – Németh Géza 2009. Automatic classification of regular vs. irregular phonation types. In: *An ISCA Tutorial and Research Workshop on Nonlinear Speech Processing (NOLISP) 2009*. Vic, Spain, 53–61.
- Bóhm Tamás – Ujváry István 2008. Az irreguláris fonáció mint egyéni hangjellemző a magyar beszédben. *Beszédkutatás* 2008. 108–120.
- Carlson, Rolf – Hirschberg, Julia – Swerts, Marc 2005. Cues to upcoming Swedish prosodic boundaries: Subjective judgment studies and acoustic correlates. *Speech Communication* 46. 326–333.
- Deme László 1971. *Mondatszerkezeti sajátosságok gyakorisági vizsgálata (Magyar szövegek alapján)*. Akadémiai Kiadó, Budapest.
- Dér, Csilla Ilona – Markó, Alexandra 2010. A pilot study of Hungarian discourse markers. *Language and Speech* 53/2. 135–180.
- Elekfi László 1962. *Vizsgálatok a hanglejtés megfigyelésének módjaihoz*. Nyelvtudományi Értekezések 34. Akadémiai Kiadó, Budapest.
- Gósy Mária 2003. Virtuális mondatok a spontán beszédben. *Beszédkutatás* 2003. 19–43.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás* 2008. 194–207.
- Gósy, Mária – Kovács, Magdolna 2008. Virtual sentences of spontaneous speech: Boundary effects of syntactic-semantic-prosodic properties. In Gardner-Bonneau, Daryle – Blanchard, Harry E. (eds.): *Human factors and voice interactive systems*. Springer, New York, 361–379.
- Langlais, Philippe – Meloni, Henri 1993. Integration of a prosodic component in an automatic speech recognition system. In: *Third European Conference on Speech Communication and Technology (EUROSPEECH '93)*. Berlin, 2007–2010.

- Lehiste, Ilse – Wang, William S-Y. 1977. Perception of sentence boundaries with and without semantic information. In Dressler, Wolfgang U. – Pfeiffer, Oskar E. (eds.): *Phonologica 1976*. Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck, 227–283.
- Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. A Bradford Book. The MIT Press, Cambridge (Massachusetts)–London (England).
- Lippmann, Richard P. 1997. Speech recognition by machines and humans. *Speech Communication* 22. 1–15.
- Markó Alexandra 2009. Stigmatizált hanglejtésforma a spontán beszédben. *Beszédkutatás* 2009. 88–106.
- Németh T. Enikő 1996. *A szóbeli diskurzusok megnyilatkozáspéldányokra tagolása*. Nyelvtudományi Értekezések 142. Akadémiai Kiadó, Budapest.
- de Pijper, Jan Roelof – Sanderman, Angelien A. 1994. On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *Journal of the Acoustic Society of America* 96/4. 2037–2047.
- Redi, Laura – Shattuck-Hufnagel, Stefanie 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29. 407–429.
- Schiffirin, Deborah 1987. *Discourse markers*. Cambridge University Press, Cambridge.
- Shriberg, Elizabeth – Stolcke, Andreas – Hakkani-Tür, Dilek – Tür, Gökhan 2000. Prosody-based automatic segmentation of speech into sentences and topics. *Speech Communication* 32. 127–154.
- Szszák, György–Vicsi, Klára 2007. Using prosody in fixed stress languages for improvement of speech recognition. In Esposito, Anna – Faundez-Zanuy, Marcos – Keller, Eric – Marinaro, Maria (eds.): *Verbal and nonverbal communication behaviours*. Springer-Verlag, Berlin–Heidelberg, 138–149.
- Szendé Tamás 1976. *A beszéd folyamat alaptényezői*. Akadémiai Kiadó, Budapest.
- Váradi Viola 2008. A virtuális mondatok műfaji meghatározottsága. *Beszédkutatás* 2008. 134–147.
- Váradi Viola 2009. Határjelzés a felolvasásban és a spontán beszédben. In Gecső Tamás – Sárdi Csilla (szerk.): *A kommunikáció nyelvészeti aspektusai*. Tinta Könyvkiadó, Budapest, 278–283.
- Varga László 1981. *A magyar intonáció – funkcionális szempontból*. Nyelvtudományi Közlemények 83. 313–339.
- Varga László 1999–2001. The unit of the Hungarian intonation. In Szathmári, István (red.) *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös nominatae*. Sectio Linguistica tomus XXIV. ELTE Eötvös Kiadó, Budapest, 5–13.
- Varga László 2002. *Intonation and stress. Evidence from Hungarian*. Palgrave Macmillan, Houndmills, Basingstoke.
- Wacha Imre 1988. Élő nyelvi (spontán) szövegek megnyilatkozásainak (szintaktikai) vizsgálati szempontjaihoz (a gazdagréti kábeltelevízió élő nyelvi felvételei alapján). In Kontra Miklós (szerk.): *Beszélt nyelvi tanulmányok*. Linguistica, Series A, Studia et Dissertationes I. MTA Nyelvtudományi Intézet, Budapest, 102–158.