

VIRTUÁLIS MONDATOK A SPONTÁN BESZÉDBEN

Gósy Mária

Bevezetés

A beszélők és hallgatók a kommunikáció során mintegy ösztönösen, ritkábban tudatosan betartják az ún. együttműködési elvet (Gósy 1999). Ennek az a lényege, hogy a beszélő igyekszik úgy beszélni, hogy a lehető legnagyobb mértékben biztosítsa a hallgató számára az elhangzottak feldolgozhatóságát; a hallgató pedig úgy igyekszik működtetni a beszédészlelési és beszédmegértési folyamatait, hogy az elhangzottakat mind pontosabban, a közlő eredeti szándéka szerint dekodeolja. Az együttműködési elv azonban a mindennapi kommunikációban gyakran sérül, sokszor anélkül, hogy akár a beszélő, akár a hallgató tudatában lenne ennek. A beszédtapasztalat során kialakult korrekciós stratégiák – mindkét mechanizmus esetében – lehetővé teszi a kommunikációban résztvevők számára, hogy az együttműködési elv tökéletlen megvalósulása esetén is megtörténjen az információcsere. Különösen nagy a jelentősége mindennek a spontán beszédben, illetőleg a spontán narratívák feldolgozása során, mivel a spontán beszéd nem olyan tökéletesen szerkesztett, mint az előre megfogalmazott szövegek meghangosítása. Továbbá a tervezés és kivitelezés nagymértékű egyidejűsége, illetőleg átfedése következtében, a produktum nyelvileg nem feltétlenül mintaszzerű.

Az elhangzó hosszabb közlések megértésének és értelmezésének egyik alapfeltétele a megfelelő tagolás. A gondolatfolyam nyelvi formába öntésekor – a beszédprodukciónál ún. mikrotervezési szakaszában (Levelt 1989) – a beszélő már igyekszik megtervezni a tagolásokat egyfelől a saját gondolatainak szemantikai, logikai váza mentén, másfelől a majdan elhangzó közlések fonetikai interpretációjának tervezésében. A spontán beszéd eddigi elemzési eredményei azt mutatták, hogy az artikulációs megvalósítás túlnyomórészt ösztönös, a beszélő nincs feltétlenül tudatában annak, hogy mely szegmentális vagy szuprasegmentális tényezőt alkalmazza tagoló funkcióban, illetőleg meg-

lehetősen nagyok az egyéni különbségek (vö. Kohler 1983). A hierarchia ún. alsóbb nyelvi szintjein a beszélőnek nincs nagy választási lehetősége a prozódiai megvalósítást illetően (vö. a szegmentális szerkezet fonológiai struktúrája). A magasabb szinteken azonban opcionális lehetőségei vannak, s ezeknek az aktuális együttes kiválasztása számos tényező függvénye. A korábbi elemzések kimutatták, hogy a mondat szintaktikai struktúrája hatással van ugyan a prozódiai megvalósításra, de nem határozza meg azt (Frazier et al. 2003).

A spontán beszéd tagolása jelentősen eltér az interpretatív megnyilatkozások, avagy a felolvasás kiejtési sajátosságaitól; hiszen ez utóbbi esetekben a beszédprodukciós mechanizmus kezdeti folyamatainak tervezésére nincsen szükség, a beszélő „készen kapott” szöveget hangosítja meg. Ennélfogva sokkal nagyobb mértékben képes a hangosítás jellemzőire figyelni, illetőleg azokat tudatosan megvalósítani. Megfelelő központoszással ellátott, felolvasott szövegekben a mondathatárok percepciósi jelölése 90%-osnál valamivel jobb egyöntetű eredményt hozott (Batliner et al. 1998).

Spontán beszédkor a beszélő a beszédprodukciós folyamatnak csak a legvégső szakaszában – az artikulációs tervezés szintjén – jut el a közlés számos artikulációs sajátosságának az összerendezéséhez, így például a tagoláshoz. Gyakorlatlan beszélők a verbális nyelv tagolásainak tudatos megvalósítására nemigen képesek, azt fiziológiai tényezők, például a légzés, sokkal inkább szervezik, mint a szöveg tartalmi és formai egységeinek tudatos jelzései. A hallgató ugyanakkor mindvégig arra kényszerül, hogy valamiképpen tagolja, bizonyos egységekre bontsa az elhangzottakat, s ezeknek az egységeknek mentén ismerje fel az összefüggéseket. A tagolási egység lehet mondatnyi hosszúságú közlés, avagy annál hosszabb, összefüggő szövegrész. Lehiste ez utóbbit ’bekezdésnek’ (vö. paragraph) nevezi (1979), annak ellenére, hogy tudatában van, a szakkifejezés elsősorban az írott szövegek bizonyos egységeinek (tipográfiai) elkülönítésére szolgál. Úgy véli, a spontán beszédben is léteznek ezek a nagyobb, összefüggő egységek, amelyeket magasabb szintű egységeknek nevez.

A mondatok, illetőleg a bekezdések mint egységek szintaktikailag és szemantikailag meghatározható összefüggésrendszerrel mutatnak, és a szuprszegmentumok jellegzetes érvényesülésén keresztül különül-

nek el (pl. Schafer et al. 2000). Ilyen szupraszegmentum a szünet (mint jelkimaradás), az alaphangmagasság változása, az intenzitás csökkenése, valamint a szünetet megelőző szó időtartamának növekedése. Az intenzitás csökkenését és a frekvencia változását mondathatárokon spontán diskurzusokban igazolták (Hird–Kirsner 2002). A mondatvég temporális változása, az utolsó szó (szavak) lassabb artikulációját más kísérletek is megerősítették (pl. Lindblom 1978). Lehiste és Wang eredményei szerint a bekezdéseket nagy arányban azonosan ismerték fel angol anyanyelvű hallgatók (1977).

A határjelzés kérdése a beszédben nemcsak a mondathatárokkal vagy az annál nagyobb egységekkel kapcsolatos, hanem például a szóhatárokkal is. A szóhatár jelzése univerzális, még inkább talán fiziológiai jellegű sajátosság, amely azonban aktuális megvalósulásában nyelvspecifikus sajátosságokat mutat és mind a beszélőknél, mind a hallgatóknál különbözőképpen érvényesül. A szóhatárjelzés megjelenhet a szegmentum, illetőleg a szótag időtartamának változásában, az alaphangmagasság módosulásában, sőt bizonyos koartikulációs hatásokban, marginálisan még jelkimaradásokban is. A szóhatár felismerése az anyanyelv-elsajátítás során alakul ki, elsősorban a disztribúciós szabályok, másrészt a fonotaktikai sajátosságok megtanulása révén (Brent–Cartwright 1996). Ez a kiindulás a szöveg nagyobb egységeinek szegmentálásához.

Német beszéd elemzése alapján arra a következtetésre jutottak, hogy a közlések szövegkohéziója különböző szintek egységeinek többé-kevésbé hierarchikus egymásra épülésén alapszik, s ezen a módon jönnek létre a nagyobb egységek (Kohler 1983, 129). A szövegkohézió nem a szintaktikai struktúrában vagy a szintaxis és a prozódia együttjárásában keresendő, hanem a szemantikai meghatározottságban, illetőleg a szemantika és a prozódia összefüggéseiben. Két mondat következhet egymás után anélkül, hogy bármilyen információs kapcsolat lenne közöttük, ugyanakkor a prozódiai megvalósításban nem különböznek a kohéziót mutató, összetartozó mondatkapcsolatoktól. A szemantikai kohézió önmagában nem eredményez szöveg-egységet, azt a beszélő szándéka szerinti szövegkohézió fogja létrehozni, amely a szemantikai és a prozódiai sajátosságok együttes megvalósulása.

Hipotézisünk szerint a spontán közlések egysége valamiképpen kódolva van a beszélőben, és a beszélő mint hallgató ezt az intuitív tudást felhasználja a beszédfeldolgozás során tagolásra. Hosszabb szöveg esetén ennek a legkisebb egységnek a mondatnak kell lennie, függetlenül attól, hogy a 'mondat' nyelvészeti definíciójával az átlagos beszélő/hallgató nyilvánvalóan nem rendelkezik. A 'mondat' szót tehát jobb híján használjuk a spontán beszéd alapegységeinek a meghatározására. Valójában virtuális mondatokról kell beszélnünk, a 'virtuális' jelzőt a 'benne rejltő' értelmében használva. Ezek a virtuális mondatok – feltételezésünk szerint – többé-kevésbé azonos paraméterekkel jellemezhetők, s ezeket a paramétereket a hallgatók képesek tagoló funkcióban használni spontán közlések hallgatásakor. A fonetikai és a pszicholingvisztikai szakirodalomban nagyobb szintaktikai/szemantikai egységek elkülönítésére a spontán beszédben a 'közlés' szót használják, amely állhat egyetlen szóból vagy – a tradicionális értelemben vett mondatot használva – akár 3-4 mondatból. A jelen kutatáshoz azonban a 'közlés' terminus használata túlzottan tág, nem definiálható egyértelműen; lényegében tehát ugyanolyan problémát jelent a használata, mint a mondaté. Másfelől pedig sokkal nehezebb egy adott kísérletben a résztvevőknek az instrukciót megfogalmazni. Azt kellene ugyanis mondanunk, hogy egy hosszabb közlést fog hallani, amelyben a kisebb (?) közléseket keressük... (vö. a 'megnyilatkozás' magyarázatát is: Wacha 1988). Nem véletlen, hogy a spontán beszéd elemzésével foglalkozók felvetik egy egészen másfajta grammatika, illetőleg grammatikai leírás lehetőségét a spontán beszéd elemzésére (Fábricz 1988). A továbbiakban tehát – jobb híján – maradunk a virtuális mondat kifejezés használatánál.

A kérdésünk az, hogy a beszélő vajon hogyan tagol az artikulációs kivitelezésben mondatszinten, továbbá a hallgatók feldolgozó mechanizmusa mennyire egyöntetűen működik az elhangzottak észlelésében? A hipotézisünk ugyanis az, hogy ha a mondattagolás relatíve nagymértékű egyöntetűséget mutat, akkor a virtuális mondatok kognitív szinten egyértelműen léteznek – függetlenül a terminus problémás voltától, avagy a grammatikai definíció nehézségeitől. Kimondható-e tehát, hogy a 'virtuális mondat' valóban létezik, akusztikai-fonetikai korrelátumai vannak, s a hallgatók ezek individuális stratégiai felhasznál-

nálását végzik a spontán beszéd szegmentálásakor? Kísérletünk célja annak a feltételezésnek az elemzése (igazolása vagy elvetése) volt, hogy léteznek virtuális mondatok, amelyeket a szünetek és az alaphangmagasság aktuális értékétől függően észlelnek a kísérleti személyek. Felmerülő kérdés az is, hogy van-e mindennek bármilyen kapcsolata a spontán szöveg megértésével.

Anyag, módszer, kísérleti személyek

A tesztanyag egy magnetofonszalagra rögzített spontán narratív szöveg volt. A beszélő 35 éves nő (beszédhibája nincs), a felvétel készítésének tudatában volt. A CSL 4300B típusú jelfeldolgozó segítségével végeztük a temporális sajátosságokra vonatkozó méréseket, továbbá az F0 változásainak objektív adatolását a szüneteket megelőző és követő szóban, valamint a kísérleti személyek által mondathatárnak jelölt valamennyi további helyen. Szemantikai és szintaktikai kritériumok alapján meghatároztuk a szöveg mondatait, a szerkezethatárokat, a grammatikai anomáliákat, a megakadásjelenségeket.

A szöveget lejegyeztük (helyesírásban), de központosítás nélkül, és ezt a két oldalas tesztlapot kapták meg a kísérleti személyek. A résztvevők feladata az volt, hogy kétszeri meghallgatás alapján jelöljék meg a tesztlapon egy függőleges vonallal azokat a helyeket, ahol véleményük szerint a beszélő szándéka szerint mondatvég volt. Semmilyen egyéb magyarázatot vagy definíciót a mondat fogalmára nem adtunk, s a kísérleti személyeknek lehetőségük volt akár az első elhangzás, akár a második elhangzás alkalmával jelölniük. Hangsúlyoztuk azonban, hogy spontán beszédet fognak hallani.

A kísérleti személyek – célzottan – két csoportba kerültek. Az egyik csoportot magyar szakos bölcsészhallgatók (átlagéletkoruk 21 év) alkották (ELTE, Budapest), valamennyien nők. A másik csoportot pedig mérnökhallgatók (átlagéletkoruk 21 év) a budapesti Műszaki és Gazdaságtudományi Egyetemről, valamennyien férfiak. Mindkét csoportban 28-an voltak. Feltételezésünk szerint – éppen tanulmányaik különbözősége miatt – a bölcsészhallgatóknak mások a kódolt ideáik a mondatról, mint a mérnökhallgatóknak. A feladat egyik csoportnak sem okozott nehézséget, s a hallgatók többsége mindkét csoportban már az első elhangzáskor jelölte a vélt mondathatárokat.

A mondatjelölést követően még egy feladatot kaptak a kísérleti személyek, amelyre azonban korábban nem készítettük fel őket. Öt, megértést ellenőrző kérdésre kellett válaszolniuk az elhangzottakkal kapcsolatosan. A kérdések a szöveg jelentősebb részleteire vonatkoztak, amelyeket az elhangzottak megértéséhez nélkülözhetetlenek ítéltünk. Négy kérdés megválaszolása egy-két szóval megoldható volt; egyetlen kérdés igényelt egy mondatnyi kifejtést (ez a kérdés: „*Hogyan zajlott az esküvő Egyiptomban?*”). A kérdéseket a kísérletvezető tette fel, és a résztvevőknek egy külön lapra kellett a válaszokat leírniuk.

A spontán szöveg fonetikai jellemzése

A rögzített hanganyag terjedelme 5 perc 10 másodperc, a kiejtett szavak száma 550. A beszélő artikulációs tempója: 107,84 szó/perc, beszédtempója 141,02 szó/perc. A szövegben előforduló 100 ms-nál hosszabb szünetek száma 135, időtartamuk szórása: 100-1350 ms, az átlag 532,44 ms. A csendes szüneteken túl hezitálások is találhatók a szövegben, ezek elsősorban az [ə] magánhangzó különböző időtartamú változatai, néhány esetben a [m] és [hm] is előfordul. A hezitálás és a néma szünet különféle kombinációja 15 alkalommal fordult elő.

1. táblázat: A spontán tesztszöveg szünetei (sz.=szerkezethatáron, n.=nem szerkezethatáron)

Szünet idő-tartama (ms)	Típusa (db)					
	néma		kitöltött		kombinált	
	sz.	n.	sz.	n.	sz.	n.
100-300	13	13	1	7		
310-500	12	8		7		1
510-800	24	5		9	3	1
810-1000	6	1			2	3
1010-1200	3	1		2	2	2
1210-1300	5	1		2	1	
Összesen	63	29	1	27	8	7

Az összes nem tartalmas jel (különféle szünetek) időtartama: 71,88 s. Szerkezethatáron összesen 72 alkalommal, míg nem szerkezethatáron 63 esetben tartott szünetet a beszélő. Leggyakoribbak a néma szünetek voltak, majd a hezitálások, ezekhez képest ritkábban fordultak elő kombinált szünetek. Az 1. táblázat a szünetek időtartamának és számának összefüggéseit szemlélteti.

A beszélő dallamformái megfelelnek az elvártaknak, átlagos alaphangmagassága 198,5 Hz, az F0 terjedelme: 150-400 Hz. Az intenzitás szint stabilnak mondható, sehol nincsenek feltűnő erősödések vagy elhalkulások. A hangsúlyok megvalósítása nem mond ellent a nyelvi szabályoknak. A szöveg a néma és kitöltött szüneteken kívül tartalmaz megakadásjelenségeket, amelyek közül jellegzetesek a nyújtások, különösen a névelőkön, valamint a szóban tartott szünetek. A névelőnyújtásra 10, egyéb nyújtásra további négy adatunk van. Az [ɔ] és [ɛ] magánhangzókon tapasztalt nyújtások időtartama megközelíti a 300 ms-ot. A beszélőre nem volt jellemző a szünet előtti szó vagy szavak időtartamának megnyújtása; jellegzetesen az utolsó szótag magánhangzója szenvedett nyúlást, de ez nem érvényesült következetesen. Sok esetben éppen a szünet előtt tapasztaltuk a szavak vagy az utolsó szó tempójának növekedését. A szöveg a következő megakadásjelenségeket tartalmazza (példákkal). Ismétlés: *hogy jön ide, hogy jön ide*; egyszerű nyelvbtlás (artikulációs kivitelezési probléma): *találkozóhoz találkozóhoz* vagy *márs és más nevük* (javítás nélkül); ismétlés morfológiai változtatással: *a neve* (452 ms néma szünet) *nevén keresztül* vagy *testrésznek testrésszel*; agrammatizmus: *sokkal magas szociális* (ti. sokkal magasabb szociális); szünet a szóban: *kultusz* (187 ms) *ból* vagy *nép* (354 ms) *zenében*; téves kezdés: *ter teherbeesés*; újrazdés: *pro prostituáltak*; változtatás: *rendszeres öö általános*.

Eredmények

A kapott adatokat a következő szempontok szerint elemeztük: a virtuális mondatok jelöléseinek sajátosságai (kvantitatív és kvalitatív eredmények); a jelölések és a szünetek összefüggésének elemzése; a jelölések és az alaphangmagasság összefüggésének elemzése; a szövegértés elemzése.

Virtuális mondatok jelölése

A két csoport összes virtuális mondatvég jelölésének száma 58. Ebből egyöntetű jelölés 49, ez azt jelenti, hogy az adott helyen mindkét csoportból jelöltek mondatvéget. Összesen 9 esetben tapasztaltunk különbséget a két kísérleti csoport között. A bölcsészek összesen 52 helyen, a mérnökhallgatók összesen 55 helyen észleltek virtuális mondatvéget a szövegben; a fiúk és a lányok között e tekintetben gyakorlatilag nincs különbség. Ezt a statisztikai elemzések is alátámasztották. A bölcsészhallgatók átlagosan 23,75 mondatot, a mérnökhallgatók átlagosan 26,11 mondatot jelöltek, a lányok csoportjában a szórás: 14-34, a fiúknál pedig 8-43. Nem választva szét a kísérleti személyek két csoportját, a virtuális mondatok átlaga 24,92, az átlagos eltérés pedig 6,14. Az adatokban feltűnő az, hogy a kísérleti személyek igen nagy arányban ítélték ugyanazon a helyen mondatvéget. Az egyöntetűen jelölt 49 lehetőségből 16 helyen teljesen azonos az ítéletek száma (ez a jelölések 32,7%-át jelenti)

Egy személy eltéréssel 28,57%-ban ismét egyöntetű a jelölés, két személy eltéréssel további 8,16%-ban. Ez összesen 69,38%, vagyis azt mondhatjuk, hogy az anyanyelvűek virtuális mondatészlelése csaknem 70%-ban azonosnak tekinthető (ezt azért is állíthatjuk, mert a két hallgató az összes résztvevőnek csupán 7,14%-a). Adataink azt mutatják, hogy 3-4 hallgató különbséget 8 helyen tapasztaltunk (ez 16,32%); s mindössze 14,3%-ban tekinthetők nagy különbségűnek a két csoport közötti eltérések (ez a szöveg hét helyén fordult elő). A számadatok arra engednek következtetni, hogy egyfelől a hallgatók ugyanazon paramétereket használhatták döntéseikben, illetőleg, hogy az azonos anyanyelvűek (s azonos életkorúak) nemtől függetlenül virtuálisan nagy hasonlósággal tagolják az elhangzó szöveget. A lányok és a fiúk jelölése között szignifikáns különbség nincs. Anyagunkban a lányok 3 helyen jelöltek mondatvéget, ahol a fiúk nem, a fiúk pedig 6 olyan helyen, ahol a lányok nem (ezek elemzésére visszatérünk).

A tesztanyagban 9 olyan helyet találunk, ahol a kísérleti személyek nem jelöltek virtuális mondatot, pedig szintaktikailag, illetőleg szemantikailag szerkezethatárra estek a beszélő által tartott szünetek. Ebből három eset valódi szerkezethatár volt, két esetben az *és* kötőszóval elválasztva (a szünetidőtartamok: 360-650 ms). Hat esetben pedig alá-

rendelő összetett mondat határán fordultak elő a szünetek, vonatkozó névmási kötőszók, illetőleg a *hogy* szomszédságában. A relatíve hosszú szünetek ellenére – 187-860 ms, átlag 553 ms – a kísérleti személyek nem észleltek mondathatárokat.

A kísérleti tesztanyagot szemantikai és szintaktikai kritériumok alapján négy csoportba soroltuk azokon a helyeken, amelyeket a kísérleti személyek megjelöltek: a) valódi mondatvég, b) kötőszó következik, c) szerkezethatár, de nem tekinthető valódi mondatvégnak d) szerkezet belseje. (A továbbiakban a 'valódi mondat' kifejezést szemben a 'virtuális mondattal' akkor használjuk, ha egyébként félreérthető a szövekkörnyezetben, hogy a percepció ítéletről van szó, avagy nem.) Példákkal illusztráljuk az egyes kategóriákat, a központozással segítve a szövegértést. Valódi mondatnak tekintettük a következőket: „*A másik jellegzetesség a mozgás, ami eltér az európai kultúrától.*” vagy „*Eredetileg termékenységi kultusból indult ki.*” Példák olyan közlésrészletekre, amikor kötőszó következik, és a kötőszó előtt mondatvéget jelöltek a kísérleti személyek (az egyértelműség kedvéért a kötőszót félkövérítettük): „*Az ókorban ezeknek szent helyük volt, tehát így sokkal magas szociális státuszban voltak.*” vagy „*Az nem elsősorban dallamon, hanem ritmuson alapul, és a ritmus is egészen más, mint mint az európai zenében.*” Szerkezetek, amelyek határán mondatvéget jelöltek a kísérleti személyek: „*Inkább nézzük a következő produkciót, /mondatvég jelölése/ nem felejtve a gyökereket, most elsősorban a szórakoztató célú hastáncot fogják látni.*” vagy „*Annyiban tér el, és ezt majd a következőkben tapasztalni fogják, /mondatvég jelölése/ a zene, a keleti zene az európai zenétől...*”. A szerkezetben belül ítélt mondatvégre példák: „*Egy másik aspektusával, de szintén a szüléssel kapcsolatos /mondatvég jelölése/ az úgynevezett berber törzsi születési ceremónia.*” vagy „*A francia gyomortáncról kezdve a görög, úgynevezett ciftetelli, ami egy török ritmusnak a neve /mondatvég jelölése/ nevén keresztül...*”. A kötőszók, amelyek előtt mindkét csoportban mondatvéget jelöltek a hallgatók, mellérendelő kötőszók, leggyakrabban az *és* és a *tehát*. A *tehát* szó esetében feltételezhetjük azt is, hogy nem kötőszó funkcióban, hanem töltelékszóként használta a beszélő. A percepció eredmény azonban – az adatok szerint – független az esetleges funkciókülönbségtől.

A szünetek elemzése

A spontán beszéd szünetei kapcsolatban vannak a szintaktikai szerkezettel, ennek az összefüggésnek a pontos meghatározása azonban nem egyszerű (vö. Duez 1993). A pszicholingvisztikai kutatások kezdeti vizsgálataiban azt találták, hogy a mondatokat elkülönítő szünetek közel 80%-ának az időtartama 500 ms-nél hosszabb volt, míg a szerkezeteken belüli szünetek mintegy 66%-ánál rövidebb, mint 500 ms (Goldman-Eisler 1972). Francia és angol interjúk összehasonlítása azt mutatta, hogy a szünetek mintegy 70%-a nagyobb frázisegységek határára esett, s ezek hosszabbak voltak, mint a frázisokon belüliek (Grosjean-Deschamps 1975). A szünetészlelési eredmények azonban nem csupán a szintaktikai szerkezettel, hanem más prozódiai paraméterekkel is mutatnak összefüggést (Duez 1985; Gósy 2000).

A 2. táblázatban összegeztük az időtartamok és a virtuális mondatok jelölésének összefüggéseit (a néma szünetet két esetben hezitálás követte, ezt plusz jellel kapcsoltuk a néma szünet értékéhez), azokban az esetekben, amikor a szemantikai és szintaktikai elemzések alátámasztották, hogy „valódi” mondat hangzott el.

A szünetek szórása 205-1350 ms közötti, az átlag 793,458 ms. Az összes kísérleti személyt figyelembe véve a 24 lehetséges helyen átlagosan 79,97%-ban észleltek mondatot. Két esetben volt teljesen egyöntetű a vélemény, ekkor a szünetek időtartama 1212 ms (néma), illetőleg 1127 ms (néma szünet, amit egy 224 ms-os hezitálás követett) volt. További négy helyen 98%-on felüli egyöntetű jelölést kaptunk, a szünetidőtartamok itt is relative hosszúnak mondhatók: 680 ms-tól 1062 ms-ig, az átlag 946,09 ms. A 69% és 87% közötti egyöntetű ítéleteknél a szünetek átlaga 735,25 ms; az 50% és 59% közötti ítéletek esetén a szünetátlag 612 ms. Az 50%-osnál kisebb egyöntetű ítéleteket percpiciósan már egyáltalán nem tekinthetjük relevánsnak, ekkor a szünetek 291 ms-tól 423 ms-ig szórnak, az átlag: 323 ms. Az adatok alátámasztják a szünet időtartamának meghatározó szerepét a mondatvég észlelésében. Minél hosszabb szünetet tart a beszélő, annál inkább az a hallgató benyomása, hogy gondolati egységet zárt le, vagyis megjelenik a virtuális mondat vége.

2. táblázat: A szünetidőtartam és a mondatvég jelölésének összefüggése (az előfordulás sorrendjében)

Szünet (ms)	Mondatvég- észlelés (%)	Szünet (ms)	Mondatvég- észlelés (%)
397	53,5	205	41,07
812	98,21	1040	94,64
996	87,5	1263	98,21
1038	91,07	710	98,21
812+250	96,42	1212	100
860	58,92	590	55,35
680	98,21	900	91,07
600	57,14	423	46,42
700	69,64	710	80,35
530	94,64	1150	91,07
290	39,28	1127+224	100
1224	92,85	300	85,71

A 3. táblázat a szünetek és a szerkezethatárok közötti összefüggést szemlélteti. A beszélő ezeken a helyeken már jóval rövidebb szüneteket tart, s két esetben megjelenik a hezitálás is (mindkettőnél a néma szünetet követően). A mondat-határ-jelölések nagyobb szórást mutatnak, de a két csoport között itt sincs lényeges eltérés. A szünetidőtartamok szórása 150-830 ms, az átlag 465,87 ms.

3. táblázat: Szünetek és szerkezethatárok

Szünet (ms)	Mondatvég- észlelés (%)	Szünet (ms)	Mondatvég- észlelés (%)
632+198	67,85	540+240	12,5
491	1,78	282	8,92
229	46,42	150	1,78
665	17,85	300	1,78

A kötőszó előtti szünetek is nagymértékben keltették a virtuális mondat végének benyomását a hallgatókban. Ezeknek a szüneteknek a szórása hasonló a szerkezethatáron tartottakéhoz, 150-800 ms, az át-

lagérték azonban hosszabb, ami azt jelenti, hogy a tartományon belül valamivel nagyobbak az időtartamok, mint a szerkezethatáron mértéké (vö. 4. táblázat). Az átlag: 509,87 ms.

4. táblázat: A mondatvég észlelése kötőszó előtt

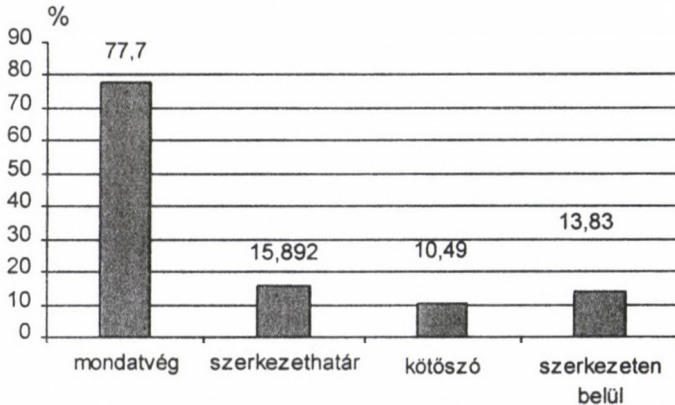
Szünet (ms)	Mondatvég-észlelés (%)	Szünet (ms)	Mondatvég-észlelés (%)
515	14,28	nincs szünet	3,57
386	7,14	587	16,07
473	1,78	603	3,57
521+284	1,78	791	37,5
665	35,7	347	1,78
334	23,21	150	7,14
597	12,5	540	16,07
644	5,35	286	1,78
435	1,78		

Sem a mondathatároknál, sem szerkezethatároknál nem tapasztaltuk, hogy a hallgatók virtuális mondatot jelöltek volna és nem volt szünet. Az ilyen típusú szüneteket Duez 'szubjektív szüneteknek' nevezi (1993, 23). A hallgató valamilyen belső percepció tagolás következtében jelkimaradást érzékel annak valóságos akusztikai manifesztálódása nélkül. A kötőszók esetében azonban egy ilyen eset előfordult, ahol két hallgató is jelölt mondatvéget. A kötőszó itt a *tehát* volt. Összegeztük az egyértelműen szerkezeten belül megjelenő mondatjelöléseket (5. táblázat).

5. táblázat: Virtuális mondatok jelölése szerkezeten belül

Szünet (ms)	Mondatvég-észlelés (%)	Szünet (ms)	Mondatvég-észlelés (%)
452-498	8,92	nincs szünet	3,57
nincs szünet	3,57	584-450	64,28
679	8,92	633	3,57
257	3,57	705	14,28

A kísérleti személyek jelöltek mondathatárt szintaktikai szerkezetben belül is, mindössze nyolc esetben, ez az összes határjelölés mintegy 14%-a, tehát elenyésző arány. Érdekes mégis, hogy mi eredményezheti néhány hallgatóban a közlésegyesség végének benyomását olyan helyzetben, amikor mind a szemantikai, mind a szintaktikai sajátosságok ellentmondanak ennek. A táblázatban látható szünetek átlaga 709,66 ms, vagyis relatíve hosszúnak mondhatók. A határértékek: 257-1034 ms. Két esetben a néma szünetet hezitálás is követi, ezáltal a szegmentális jelkimaradás időtartama jelentősen megnövekszik. A hallgatók több, mint 60%-a ott jelöl mondatvéget, ahol egy toldalékválasztási hiba miatt tart a beszélő hosszabb szünetet, és nem egyértelmű, hogy a folytatás önálló egység kezdete, avagy a megelőző szerkezet bizonytalan folytatása. Itt a beszélő alaphangmagassága leesik 170 Hz-re, s ez ismét a virtuális mondat befejezésének benyomását eredményezi (szemben a másik hosszú szünet esetén, ahol az alaphangmagasság kismértékű emelkedését tapasztaljuk). Két helyen jelöltek a kísérleti személyek mondatvéget, ahol objektíve nem volt szünet, vagyis ismét előállt a 'szubjektív szünet' helyzete. Mindkét esetben a beszélő által „elkövetett” szintaktikai hiba okozhatta a zavart. Az 1. ábra a virtuális mondatvég jelölésének átlagait szemlélteti.



1. ábra

A szemantikai, szintaktikai kritériumok és a virtuális mondatok átlaga

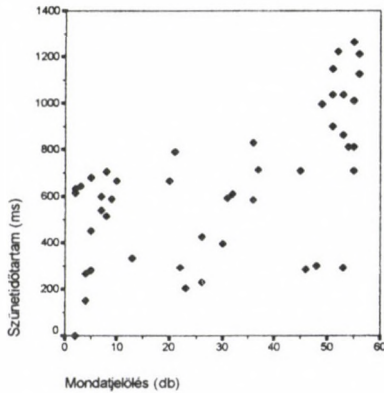
Ha a négy – szemantikai/szintaktikai kritériumok szerint elkülönített – helyzetet összehasonlítjuk, akkor közöttük jelentős különbséget tapasztalunk, ezt szemlélteti az 1. ábra. A statisztikai elemzéseket az ANOVA ún. Welch-próbájával végeztük, eszerint a szemantikai/szintaktikai csoportok szerint a különbség 1%-os szinten szignifikáns (72,867, $df_1=3$, $df_2=19,086$, $p < 0,0001$). A 4 kategória post hoc tesztekkel történt összevetése is megerősítette, hogy a „mondatvég” a döntő tényező. A hallgatók tehát igen nagy mértékben támaszkodnak a tartalomra és a formára együttesen a virtuális mondatathárók felismerésében.

Az adatok alapján elmondható, hogy a közlések tartalma meghatározó a virtuális mondat megítélésében, vagyis adott esetben mintegy felülbírálja a formát. A szünetek hossza ugyanakkor valamilyen mértékben befolyásolja a döntést. Kimondható, hogy minél hosszabb a szünet, annál jobban kelti a hallgatóban a mondatvég benyomását. Ez magyarázza azt a tényt, hogy a nyelvi és jelentésbeli egyértelműség ellenére a valóságos mondatvégek jelölésében is csak akkor kaptunk 90%-os vagy 100%-os egyöntetű ítéletet, ha a beszélő által tartott szünet tartama relatíve hosszú volt. Ez az érték átlagosan 900 ms. Korábbi szünetészlelési kísérletek – amelyekben funkciótól függetlenül kellett a beszélő által tartott szüneteket jelölni – azt igazolják, hogy már 500 ms tartamú néma szüneteket is képesek a hallgatók 80-100%-ban felismerni (Gósy 2000). Ez azt támasztja alá, hogy a szünet fontos, de nem egyedüli tényező a virtuális mondatok azonosításában. A szünet-időtartamok és a mondatthár-ítéletek közötti összefüggést korrelációszámítással ellenőriztük. A Spearman-féle korrelációs együttható pozitív és közepesen erős ($r=0,663$, $p=0,01$). Ez azt jelenti, hogy az esetek nagy hányadára igaz az, hogy a nagyobb szünetidőtartamhoz a virtuális mondatthárók nagyobb száma társul (vö. a 2. ábra pontdiagramját).

A szünetek időtartama különbözik a megállapított szemantikai, szintaktikai kategóriák szerint. Kimondható, hogy a szünetek időtartama azokon a virtuális mondat helyeken a leghosszabb, amelyek egyben szemantikai és szintaktikai kritériumok szerint is valóságos mondatvégek; az eredmény szignifikáns (Welch-próba: 5,305, $df_1=3$, $df_2=8,398$, $p=0,025$).

Az alaphangmagasság elemzése

A következő elemzett paraméter, az alaphangmagasság változása volt, illetőleg annak a mondatvég észlelésére gyakorolt hatása. Az alapkérdés itt az, hogy vajon a beszélő jelzi-e valamiképpen az alaphangjával az egység végét. Ha igen, a következő kérdés az, hogy vajon a hallgató képes-e ezt az információt feldolgozni. Holland anyanyelvűekkel végzett kísérletek azt igazolták, hogy mind a beszélő, mind a hallgató használja a befejezettség érzékeltetésére az alaphangváltozást. Ezek a változások a dallamcsúcsban, valamint az átlagos alaphangmagasságban és annak szórásában voltak megfigyelhetők (Swerts–Collier 1991). Német nyelvű kísérletek alapján az látható, hogy az alaphang frekvenciacsúcsának elhelyezkedése meghatározó a határjelzés szempontjából. Ha a hangsúlyos helyzetű szótagban nem a szótag vége felé jelent meg a frekvenciacsúcs (hanem az elején vagy a szótag közepén), akkor annak kisebb határjelző hatása volt az egyidejűleg megjelenő szünet ellenére, s így ekkor gyakoribbak voltak a megszakítások (Moosmüller 1997).



2. ábra

A szünetek tartamának és a virtuális mondatok jelölésének összefüggései

Saját anyagunkban a korábban felállított négy kategóriát tekintve azt látjuk, hogy a szemantikai és a szintaktikai kritériumok szerinti mondatvégek jelentős részénél a beszélő alaphangja csökken, vagyis a frekvencia egy magasabb értékről egy alacsonyabb értékre jut az utolsó szóban vagy az utolsó szótagban. Az anyagunkban található 25 mondatvég esetében ez 14 esetben valósul meg. Jellemzően esik az F0 190-200 Hz-ről 170-150 Hz-re. Más esetekben a beszélő lebegteti az alaphangját, vagyis sem csökkenés, sem emelkedés nem következik be, a frekvenciaérték vagy 160 Hz körüli, vagy jóval magasabb, 300 Hz körüli. Az utóbbi a tesztszövegben 4 esetben fordult elő, az F0 tartománya ekkor: 290-325 Hz. A fennmaradó 7 esetben a beszélő alaphangmagasságában bekövetkező változás emelkedés, amelynek a mértéke eltérő (de szökésnek, tehát igen rövid időtartam alatt nagy frekvenciaemelkedésnek nem tekinthető). A frekvenciaváltozás különbsége relatíve nagy tartományban szór: 80 Hz és 220 Hz között; változó a kiinduló frekvenciaérték, valamint az az időtartam is, ami alatt a változás bekövetkezik. A frekvenciatartomány: 140-220 Hz, az időtartam: 150-300 ms. A 6. táblázat a frekvenciaváltozás függvényében szemlélteti a virtuális mondatokra kapott ítéleteket.

6. táblázat: Az F0 és a virtuális mondatok észlelése valódi mondatvég esetén

F0 (Hz)	Virtuális mondatvég észlelése (%)	
	átlag	szórás
ereszkedő	85,16	39,28-100
emelkedő	65,77	23,21-91,07
lebegő	72,22	46,42-94,64

A kísérleti személyek tehát felhasználják döntéseikben az alaphangmagasság változását vagy stagnálását. Legnagyobb mértékű a mondatvég jelölése az ereszkedő dallam esetében, majd a lebegő dallamnál, és ezekhez képest legkevésbé ítélnék mondatvéget, ha a dallammenet emelkedik. Az alaphangmagasság változását a másik három kategória esetében is elemeztük.

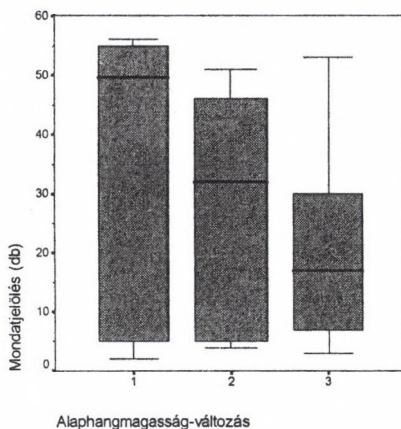
A szerkezetvégeket a beszélő döntően az alaphangmagasság növelésével vagy lebegtetésével jelzi, ez következik be az esetek 75%-ában, ebből 50%-ban történik a dallam alacsonyabb értékről magasabb értékre növelése. A különbségek itt 30-140 Hz közöttiek, s a változás kiindulása 170-310 Hz-es tartományban szór (a frekvenciaváltozás átlagosan 400 ms alatt következik be). Ha összevetjük ezeket az adatokat a valóságos mondatvégek akusztikai jellemzőivel, az látható, hogy szerkezethatáron a frekvenciaváltozás különbségei kisebbek, s a kiinduló érték tendenciaszerűen magasabb szerkezethatáron. A beszélő tehát kétféleképpen is jelzi, hogy másféle egység határán van (mint mondatok befejezésekor). A szerkezetvégek azonosítása virtuális mondatként meglehetősen alacsony, az ítéletek és az alaphangváltozás tekintetében nem mutatható ki semmiféle összefüggés.

Kötőszó előtt is az alaphangmagasság változásának mindhárom említett típusát megfigyelhetjük. Adataink szerint 50%-ban tapasztaljuk a frekvenciaemelkedést, közel 19%-ban a lebegést (magas frekvenciaértéken), a többi esetben pedig alacsony frekvenciaértéken enyhe csökkenést regisztráltunk. A dallammenet frekvenciaemelkedésének különbsége 90-160 Hz, tehát a tartomány szűkebb, mint a mondatoknál, de nagyobb, mint a szerkezethatárnál. A kiinduló frekvenciaérték szórása 150-260 Hz, vagyis nagyobb, mint a mondatok esetében, és szűkebb, mint a szerkezethatárok esetében tapasztaltak. A beszélő tehát sajáttságosan, de tendenciaszerűen jelzi, hogy másféle szerkezeti határon van, mint akár valódi mondatvég, avagy szerkezethatár esetében. Itt sem találtunk egyértelmű összefüggést a dallammenet és a virtuális mondatvég között. Kötőszó előtt például 37,5%-ban ítétek mondatvéget a hallgatók, amikor a szünetet (és a kötőszót) megelőző dallammenet 130 Hz-et emelkedett. Csaknem ugyanilyen arányban jelölték mondatvéget (35,7%-ban) ott, ahol az F_0 értéke 100 Hz-nyit csökkent a szünetet, illetőleg a kötőszót megelőzően. Nyilvánvaló, hogy a dallamváltozás mint paraméter nem elsődrendű fontosságú a szövegegység elkülönítésében.

Szerkezetben belül három esetben fordult elő hangmagasság-emelkedés 60 Hz, 80 Hz, illetőleg 150 Hz különbséggel; a kiinduló értékek 150-250 Hz között mozogtak. Az összes többi esetben (ötben) az F_0 alacsony frekvencián volt vagy kismértékben csökkent átlagosan 170

150-250 Hz között mozogtak. Az összes többi esetben (ötben) az F0 alacsony frekvencián volt vagy kismértékben csökkent átlagosan 170 Hz-re. Az egyetlen nagyobb arányú, egységes mondatészlelés egy ilyen frekvenciaérték esetén következett be (64,28%-os egyöntetű ítélet).

A statisztikai adatok ezúttal nem mutatnak olyan összefüggést a hallgatók ítéleteivel, mint a szünetek és az előzetesen meghatározott kategóriák esetében. Az F0 értéke és a jelölt mondatatharok közötti korreláció negatív, gyenge és nem szignifikáns. Ez azt jelenti, hogy statisztikailag nem támasztható alá az, hogy az alacsonyabb frekvenciaérték mondatathar-ítéletet idézne elő. A középértékek ugyan eltérő tendenciát jeleznek, azonban az adatok erőteljes szórása miatt összefüggés nem mutatható ki. A tendencia arra utal, hogy a csökkenő F0 a mondatvég benyomását kelti (vö. 3. ábra).



3. ábra

Az alaphangmagasság (F0) és a virtuális mondatok észlelésének tendenciaszerű összefüggése (1=alacsony F0, 2=magasan lebegő F0, 3=emelkedő F0)

Az elemzett adatok arra utalnak, hogy a mondatvég összefüggésben van az alaphangmagasság csökkenésével, illetőleg az alacsony

F0-értékkel, míg a dallamemelkedések azt jelzik a hallgatónak, hogy az adott egységnek nincs vége, tehát a kísérleti személyek általában nem fognak mondatvéget jelölni. A beszélő azonban nem feltétlenül konzekvens a dallammegformálásban, a szemantikai, szintaktikai tervezés nem mindig jár együtt ugyanazzal a prozódiai megformáltsággal. Nem egy esetben azonban nemcsak a szünet előtti F0-változást, hanem a szünetet követő frekvenciaértéket is vizsgálunk kell, hiszen a hallgató a közlés folyamatosságában dönt, tehát minden lehetséges paraméter befolyásolhatja az ítéletét. Ha a szünetet követő szó első szótagjának frekvenciaértéke magasabb, mint a szünet előtti szó utolsó szótagjában mért érték, akkor előfordul, hogy mondatvéget jelöl a hallgató. Az ilyen jelölések anyagunkban azonban ritkák és csak néhány kísérleti személynél jelentkeztek.

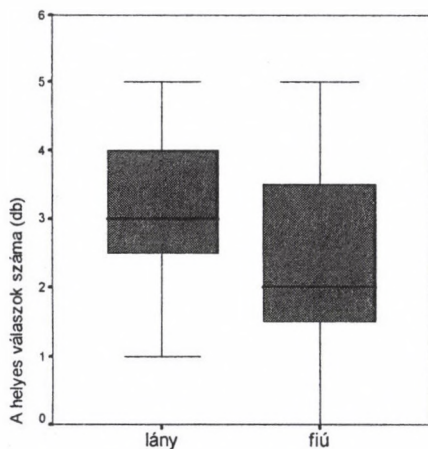
Az eddigi elemzésekből kihagytuk azokat a mondatjelöléseket, amelyek vagy csak az egyik, vagy csak a másik kísérleti csoportból kerültek ki. A következőkben ezt a 9 virtuálismondat-jelölést vizsgáljuk meg. A lányoknál 3 ilyen eset adódott, érdekességképpen egyik helyen sem tartott szünetet a beszélő, sőt jelentős alaphangmagasság-változást sem mértünk. A három esetből kettő megfelel a valódi mondat szemantikai és szintaktikai kritériumainak, míg egy esetben szerkezethatárra esett a jelölés. A fiúk 6 olyan helyen jelöltek, ahol a lányok nem; közülük 4 kötőszó előtti helyzet, két esetben pedig szerkezeten belüli a mondatvég észlelése. A kötőszók egy kivétellel alárendelő tagmondatokat kapcsolnak össze (az egy kivétel az *és így tovább* közléség, amely egyben valódi mondatvég is; itt nyilvánvaló jelölési hibáról van szó). Mind a 6 helyen szünetet tartott a beszélő, ezek szórása 257-521 ms, az átlag pedig 420,66 ms.

Elemeztük a nemek szerinti esetleges különbséget. Hangsúlyozzuk, hogy tekintettel a kísérleti személyek kiválasztására, a két csoport nemcsak abban különbözik, hogy lányok és fiúk alkotják, hanem abban is, hogy – feltételezhetően – az ún. nyelvi tudatosság a magyar szakos bölcsészhallgatókban nagyobb, mint a mérnökhallgatókban. Ezért nagy különbséget feltételeztünk a két csoport között. A kapott adatok azonban ellentmondanak a feltételezésnek. A lányok virtuálismondat-jelöléseinek átlaga 23,75 (átlagos eltérés 10,16), a fiúké 24,81 (átlagos eltérés, 10,86). A két csoport között nincs szignifikáns kü-

lönbség. Hipotézisünk volt az is, hogy a lányok jelentősen több mondatot fogtak bejelölni, mint a fiúk; továbbá hogy a lányok és a fiúk között a szerkezetben belüli és a kötőszó előtti kategóriákban jelentős lesz a különbség. Adataink egyiket sem támasztották alá. A két csoport között statisztikailag semmilyen különbség nem volt igazolható.

Szövegértés

Az elhangzott szöveg értését a megértést ellenőrző kérdésekre adott válaszok elemzésével vizsgáltuk. A lányok általában több kérdésre válaszoltak helyesen (vö. 4. ábra), mint a fiúk; a virtuális mondatok jelölése és az értés között azonban statisztikai összefüggés nincsen. Az egytényezős ANOVA szignifikáns különbséget mutat a két csoport összevetésekor ($F(1, 54)=4,799; p=0,033$).



4. ábra

A spontán szöveg értése a két kísérleti csoportban (átlag és szórás)

Elemeztük az egyes kérdésre adott helyes válaszokat. A két csoport hasonlóan válaszolt helyesen az ötből három kérdésre; s ugyanaz a kérdés okozta valamennyi kísérleti személynek a legtöbb problémát.

Ez utóbbit nem egészen 20%-uk tudta csak jól megválaszolni. Nagy különbséget a bölcsészek és a mérnökhallgatók között két kérdésben tapasztaltunk. Az elsőre a lányok 75%-ban, a fiúk csak 32,24%-ban adtak korrekt választ; a negyedik kérdésben pedig 50%, illetőleg 85,7% a helyes válaszok aránya, ismételten a lányok javára. Mindkét csoportban akadt természetesen olyan személy, aki egyetlen kérdésre sem tudott válaszolni, de olyan is, aki az összes kérdésre hibátlanul felelt. Összevetettük a kérdésekre adandó válasz helyét a szövegben a kapott válaszokkal. Az első kérdés a spontán szöveg legelején, a második, a harmadik és a negyedik a szöveg első felében elhangzottakra kérdez rá, míg az utolsó kérdésre elvárt válasz a szöveg második felében jelenik meg. A 7. táblázat a kérdésre várt válasz szintaktikai/prozódiai/észlelési jellemzőit veti össze a helyes válaszok arányával (a „mondajelölés” a válasz elhangzását követő jelölést mutatja). Az első „válasz” a mondat közepén, az összes többi a mondat végén hangzott el.

7. táblázat: A megértést ellenőrző tesztben a kérdésekre várt válaszok szöveghegyeinek szünetjellemezői, a virtuális mondatok ítéletei és a helyes válaszok (K.=kérdés)

K.	A kérdésre várt válasz jellemzői		Helyes válaszok aránya (%) b.=bölcsészek, m.=mérnökhallgatók	
	szünet (ms) előtte/utána	mondatje- lölés (%)	b. – m.	összesen
1.	812/996	87,5	75 – 32,14	53,57
2.	601/703	66,07	64,28 – 60,7	62,5
3.	597/205	41,07	14,28 – 25	19,64
4.	1263/710	98,21	85,7 – 50	67,85
5.	590/791	37,5	67,85 – 71,4	69,64

Az adatok nem mutatnak egyértelmű összefüggést, vagyis a prozódiai tényezők nincsenek hatással a megértésre. Ugyanakkor az sem igazolható, hogy a virtuális mondatok jelölése miatt nem tudtak a kísérleti személyek a kérdésekre jól válaszolni, hiszen a mondatjelölés-

től függetlenül a megértést ellenőrző kérdésekre adott helyes válaszok.

Következtetések

A kutatások eddigi megállapításai szerint a spontán beszédben szoros (bár nem minden esetben tökéletesen megvalósított) a kapcsolat a prozódia és a szintaktikai szerkezet között (vö. pl. Cutler et al. 1997). Ezt a tényt német anyanyelvűekkel végzett agyi kiváltott potenciál vizsgálatok (ERP) is alátámasztották, amennyiben a prozódiai határok azonnal szintaktikai szerkezeteket aktiváltak (Steinhauser–Friederici 2001; Steinhauer 2003). Valahányszor a kísérleti személy intonációs határt észlelt, pozitív hullám jelent meg az ERP-ben (amit a szerző(k) CPS-nek neveztek, vö. closure positive shift). A kísérletben azonban nem spontán beszéd, hanem összetett mondatok képezték az ingeranyagot, amelyben a szemantikai, szintaktikai és prozódiai határok egyértelműen egybeestek. A kutatók alapkérdése az volt, hogy a prozódiai jellemzők vajon a szintaktikai struktúrával vagy a szemantikai sajátosságokkal mutatnak-e szorosabb kapcsolatot.

A szemantikai koherenciát igazolják angol nyelvű kísérletek a szintaktikai szerkezet és a prozódiai megvalósítás egységével szemben (Frazier et al. 2003). A szerzők kimondják, hogy adataik szerint nem létezik olyan prozódiai szerkezet, amely képes biztosítani a szintaktikai szerkezeti egységet a hallgató számára. A hallgatók a szintaktikai szerkezetek prozódiai megtörését nem is érzik természetellenesnek. A lényeg tehát az, hogy nem a szünet jelenléte vagy hiánya az, ami döntő, hanem az, hogy a létrejött prozódiai struktúra által kijelölt közlésszakasz szemantikailag koherens-e vagy sem. A hallgatók 'nem természetesnek' ítélték olyan több mondatból álló szintaktikai egységeket, amelyek prozódiailag ugyan tökéletesek voltak, de szemantikailag kevésbé koherensek.

A spontán beszéd egységeinek jelölése, felismerése, illetőleg elkülönítése az automatikus beszédfelismerés egyik alapvető problémája (pl. Batliner et al. 1998; Gallwitz et al. 2002). Korpuszokon végzett vizsgálatok eredményei azt mutatták, hogy a prozódiai (és egyéb) kulcsok különböznek a szöveg típusától függően. A szünet és az alaphang sajátosságai meghatározóak például a hírek szegmentálásakor, ugyan-

akkor a párbeszédekben a szünetnek, a beszédhangok időtartamának és a szavak szegmentálásának van nagyobb jelentősége (Shriberg et al. 2000). A mesterséges beszédfelismerés eredményei szerint a következő paraméterek meghatározóak a határjelzésekben (digitális jelfeldolgozás, majd lineáris regressziószámítás, illetőleg a rejtett Markov-móddel alkalmazásával): F0 és változásai, szótag/szó intenzitása, temporális viszonyok (elsősorban a szótag/szó időtartama), a szót megelőző és/vagy követő szünet tartama, a beszédtempó normalizálása, a szóhangsúly jelenléte egy adott szón (Batliner et al. 1998). A humán beszédfeldolgozó mechanizmus részben más folyamatok működtetése, részben más feldolgozás révén dekódolja a spontán beszédben lévő egységeket, de természetesen ugyanazokat az akusztikai-fonetikai paramétereket használja fel.

A humán mechanizmus rendkívül rugalmas, nem konkrét értékek alapján, hanem viszonyok alapján dönt, amelyek egy bizonyos egységben belül, avagy egységek sorozatában realizálódnak. A bekezdés végét általában hosszabb szüneteknél jelzik a hallgatók, szemben a mondatvégekkel. A konkrét adatok mégis azt mutatják, hogy bekezdésvéget észleltek 1380 ms-os szünet esetén és mondatvéget 1390 ms-os szünetnél (Lehiste 1979, 196). A spontán beszéd bármely prozódiai sajátossága, amely a virtuális egység befejezését jelezheti, önmagában nem, csak bizonyos összefüggésekben bír határjelző funkcióval, s csak akkor, ha a szemantikai sajátosságok ezt a határt megerősítik.

Kísérletünkben a virtuális mondatok percepció határait a szünetek (néma vagy kitöltött) megjelenése, az alaphangmagasság csökkenése vagy lebegése és a szemantikai strukturáltság eredményezte. Megállapítható, hogy a 'virtuális mondat' valóban létezik (mind a beszélő produkciójában, mind a hallgató percepciójában), akusztikai-fonetikai korrelátumai vannak, s a hallgatók nagymértékben egyöntetűen ismerik fel ezeket a spontán beszédben. Az elhangzott közlések megértése azonban független a szegmentálási stratégiáktól és működésektől.

Irodalom

- Batliner, A. – Kompe, R. – Kiessling, A. – Mast, M. – H. Niemann – Nöth, E. (1998): M = Syntax + Prosody: A syntactic-prosodic labelling scheme for large spontaneous speech databases. *Speech Communication* 25: 193-222.
- Brent, M. R. – Cartwright, T. A. (1996): Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition* 61: 93-125.
- Cutler, A. – Dahan, D. – van Donselaar, W. (1997): Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 40: 141-201.
- Duez, D. (1993): Acoustic correlates of subjective pauses. *Journal of Psycholinguistic Research* 22: 21-39.
- Fábricz Károly: A beszélt nyelvi szövegalkotás kérdéséhez. In: *Beszélt nyelvi tanulmányok*. Szerk.: Kontra Miklós. MTA Nyelvtudományi Intézet. Budapest, 76-90.
- Frazier, L. – Clifton, Ch. Jr. – Carlson, K. (2003): Don't break, or do: Prosodic boundary preferences. *Lingua* 1: 1-25.
- Gallwitz, F. – Niemann, H. – Nöth, E. – Warnke, V. (2002): Integrated recognition of words and prosodic phrase boundaries. *Speech Communication* 36: 81-95.
- Goldman-Eisler, F. (1972): Pauses, clauses and sentences. *Language and Speech* 15: 103-147.
- Gósy Mária (1999) *Pszicholingvisztika*. Corvina Kiadó. Budapest.
- Gósy Mária (2000): A beszédszünetek kettős funkciója. In: *Beszédkutatás 2000*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 1-15.
- Grosjean, F. – Deschamps, A. (1975): Analyse des variables temporelles du français spontané. *Phonetica* 28: 191-226.
- Hird, K. – Kirsner, K. (2002): The relationship between prosody and breathing in spontaneous discourse. *Brain and Language* 80: 536-555.
- Kohler, K. J. (1983): Prosodic boundary signals in German. *Phonetica* 40: 89-134.
- Lehiste, I. (1979): Perception of sentence and paragraph boundaries. In: *Frontiers of Speech Communication Research*. Eds.: Lindblom, B. – Öhman, S. Academic Press. London, New York, San Francisco, 191-201.
- Lehiste, I. – Wang, W. S.-Y. (1977): Perception of sentence boundaries with and without semantic information. In: *Phonologica 1976: Innsbrucker Beiträge zur Sprachwissenschaft*. Eds.: Dressler, W. U. – Pfeiffer, O. E.

- Institut für Sprachwissenschaft der Universität Innsbruck. Innsbruck, 227-283.
- Levelt, W. (1989): *Speaking*. MIT Press. Cambridge, Mass.
- Moosmüller, S. (1997): The relevance of fundamental frequency contour for interruptions: A case study of political discussions in Austria. In: *Communicating Gender in Context*. Eds.: Kotthoff, H. – Wodak, R. John Benjamins. Amsterdam, Philadelphia, 401-420.
- Schafer, A. J. – Speer, S. R. – Waren, P. – White, S. D. 2000. Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research* 29: 169-182.
- Shriberg, E. – Stolcke, A. – Hakkani-Tür, D. – Tür, G. (2000): Prosody-based automatic segmentation of speech into sentences and topics. *Speech Communication* 32: 127-154.
- Steinhauser, K. (2003): Electrophysiological correlates of prosody and punctuation. *Brain and Language* 86: 142-164.
- Steinhauser, K. – Friederici, A. D. (2001): Prosodic boundaries, comma rules, and brain responses: The Closure Positive Shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research* 30: 267-295.
- Swerts, M. – Collier, R. (1991): On the controlled elicitation of spontaneous speech. *Speech Communication* 12: 35-48.
- Wacha Imre (1988): Élő nyelvi (spontán) szövegek megnyilatkozásainak (szintaktikai) vizsgálati szempontjaihoz. In: *Beszélt nyelvi tanulmányok*. Szerk.: Kontra Miklós. MTA Nyelvtudományi Intézet. Budapest, 10-159.