

Magánhangzó-időtartamok a spontán beszédben*

Bevezetés

A világ nyelvei számos fonológiai és fonetikai tényezőben különböznek egymástól. Jellemző egy nyelvre, hogy a magánhangzói milyen időtartamúak, valamint hogy az időtartam-különbségeknek van-e funkciójuk. Az elemzett nyelvek egy részének magánhangzói úgy alkotnak párokat az időtartamuk szerint, hogy jelentéseket különítenek el, vagyis fonológiai oppozícióban állnak. Ez a jelenség a legkülönfélébb nyelvekben fellelhető, például a szingalézban, a japánban, a kikuyuban, a Rarotonga szigeten (Kelet-Polinézia), illetve Új-Zélandban beszélt maori nyelvben, avagy a lugandában (Uganda) (Laver 1994, Hermans 2006). A magánhangzó-időtartam oppozíciója különbözőképpen jelenhet meg egy nyelv magánhangzó-állományában. A Nyugat-Afrika több országában beszélt fula nyelvben például mind a tíz magánhangzó időtartam szerinti oppozícióban áll egymással. Sajátos nyelvi funkcióban érvényesül a fonológiai hosszúság egy holland nyelvjárásban, ahol a rövid magánhangzó a többes számú, a hosszú magánhangzó pedig az egyes számú főnevet jelöli. Néhány olyan nyelvet is ismerünk, amelyekben ugyanazon magánhangzó három különböző időtartama is fonológiai értékű, és azokat a szavak jelen-

* A kutatás az OTKA 78315 sz. pályázat keretében folyt.

tésének elkülönítésére használják fel, mint például a skót gaelikben vagy a lappban. A háromféle magánhangzó-hosszúság nyelvi funkcióját a Mexikó déli részén beszélt mixe nyelvben Hoogshagen (1959) közleménye után több mint három évtizeddel Ladefoged és Maddieson is igazolta (1996).

A magánhangzó-időtartam fonológiai funkciójának neurolingvisztikai vizsgálata igazolta a bal agyfélteke szerepét. Thai anyanyelvű afáziásokkal végzett kutatásban arra kerestek választ, hogy vajon a vizsgált betegek szóolvasásában milyen mértékben őrződik meg a rövid és a hosszú fonémák megkülönböztetése. Az afáziások beszédprodukciónak csaknem tökéletesen megmaradt az elkülönítés, és ez igazolta a bal agyféltekéhez köthető temporális szerveződést (Gandour–Dardarananda 1984). Későbbi kísérletek részben megerősítették a fonológiai tartam elkülönítésének megőrződését német anyanyelvű, cerebelláris betegeknek. A kísérletek szerint azonban a fonológiai időtartam megőrződése nem független az artikulációs gesztus összetettségétől (Ackerman et al. 1999).

A hosszúság vagy kvantitás mint fonológiai jegy és az időtartam mint fonetikai jegy nem azonosak. A fonológiailag hosszú magánhangzók¹ nem minden esetben realizálódnak hosszabban a kiejtésben, mint a fonológiailag rövidek. Ennek az az oka, hogy a magánhangzók objektív időtartamát számos tényező határozza meg, amelyek nemegyszer egyidejűleg hatnak, és alakítják ki az adott kontextusban a kiejtett hang fonetikai időtartamát. A fonológiai hosszúság fonetikai realizációi ennek megfelelően meglehetősen változatosak, az objektív időértékek jellegzetes átfedéseket mutatnak (vö. Laziczius 1944). Ezek az átfedések azt jelentik, hogy a fonológiailag rövid magánhangzó esetenként hosszabb időtartamban realizálódhat, mint a hosszú (és megfordítva). A rövid-hosszú magánhangzó párok időtartamának arányai nyelvenként jelentősen változók lehetnek; 28%-tól (pl. a thaiban) akár 90%-ig (pl. a németben) (vö. Lehiste 1970).

Az időtartam fonológiai kontrasztja sajátosan érvényesül a kinyarvanda nyelvben, amely egy Ruandában beszélt bantu nyelv. A rövid-hosszú szembenállás itt különböző fonetikai helyzetekben neutralizálódik (Myers 2005). Szó eleji és szóvégi helyzetben csak rövid magánhangzók fordulhatnak elő. Egy nazális és egy obstruens alkotta hangkapcsolatot csak hosszú magánhangzó előzhet meg; egy más-salhangzóból és egy simulóhangból álló hangkapcsolatot pedig csak hosszú magánhangzó követhet. Az akusztikai-fonetikai vizsgálatok azt mutatták, hogy a kinyarvandában fonetikailag háromféle magánhangzó-időtartam realizálódik, és ezek fonológiai kötöttséget mutatnak. A rövid magánhangzók időtartamát (mondatok felolvasása alapján) 96,4 ms volt, a hosszúaké pedig 169,4 ms. A fonetikai helyzet, illetve a hangkapcsolat-kontextus típusa szerinti hosszú magánhangzók átlaga 158,4 ms (nazális + obstruens előtt) és 168,6 ms (más-salhangzó + simulóhang után). Az adatok szignifikáns különbséget mutattak a négy beszélőnél, valamint ugyanazon beszélő ejtésében is. A fonológiai hosszúság tehát egyértelműen realizálódott a fonetikai időtartamokban.

¹ A tanulmányban a 'hosszú' és a 'rövid' magánhangzó kifejezést mindig fonológiai értelemben használjuk. A fonetikai időtartamokat objektív értékekkel adjuk meg.

Ugyancsak a bantu nyelvekhez tartozó civil nyelv (Afrikában, Gabonban és néhány szomszédos országban beszélnek) tíz magánhangzója fonológiai kontrasztot mutat a hosszúságot tekintve (Ndinga-Koumba-Binza 2008). Itt is jellemző a kontextuális, illetve pozicionális kötöttség. Mindig hosszú a magánhangzó, ha nazális + mássalhangzó kapcsolat előtt fordul elő, avagy ha mássalhangzó + simulóhang kapcsolatot követ. A fonológiai tartam a fizikai időtartamokban is realizálódik (szavakban és mondatokban vizsgálva). A rövid magánhangzók átlaga 91 ms, a hosszúaké 167 ms. A két afrikai nyelv magánhangzóira kapott értékek csaknem azonosak.

A magánhangzók időtartamát igen sok kutatásban vizsgálták; jellemzően szavak vagy mondatok (különböző tartalmú, illetve hordozó mondatok) alkották a beszédanyagot. A kutatások főként arra a kérdésre igyekeztek választ adni, hogy egy-egy tényező miként befolyásolja az adott szegmentum időzítését. A beszédhangok fizikai időtartama azonban – mint említettük – számos tényező kölcsönhatásának következményeként jön létre. Ez megnehezíti egy olyan hierarchia felállítását, amelyben a beszédhangok temporális sajátosságait kialakító tényezők szerepelnek. Nemegyszer a mérési módszerek különbözőségei gátolják az általánosítható következtetések megfogalmazását. Korábbi kutatási eredmények igazolták, hogy az időtartam kialakításában meghatározó maga a szegmentum, illetve annak spektrális jellemzői, a hangsúly, sőt az alaphangmagasság is (pl. Engstrand–Krull 1994, Kovács 2002, Greenberg et al. 2003), jellemző tényező a hangsor sajátossága, amelyben az adott beszédhang előfordul, valamint a hely, ahol az adott beszédhang megjelenik (fonetikai pozíció), a szótagszerkezet, a hangsor szótagszáma és a közvetlen beszédhangkontextus (a szomszédos hangok). A magánhangzót megelőző mássalhangzó időtartam-módosító szerepét nem minden esetben találták relevánsnak. Hatással lehet az időtartamra továbbá a közlés teljes hossza, a szintaktikai határ, a beszédstílus (pl. olvasás vagy spontán beszéd), a beszédtempó, a beszédhelyzet, a beszélő fiziológiai és pszichés állapota, avagy a szógyakoriság stb. (Lindblom–Rapp 1973, Klatt 1976, Crystal–House 1988, Bell et al. 2002 stb.). A kontextushatás minden beszélőnél megjelent spanyol, két szótagból álló értelmetlen hangsorok ejtésében; a magánhangzó időtartama megnövekedett a zöngétlen réshangok és a szonoránsok előtt. A szomszédos beszédhangok hatása ugyanakkor a magánhangzótól is függött, jelentősebb volt az /a/, mint az /i/ és az /u/ esetében (Mendoza et al. 2003). Az olasz hangsúlyos magánhangzók időtartamát vizsgálták különböző fonetikai pozíciókban. Az adatok nem igazolták egyértelműen azt az állítást, hogy az utolsó szótagi magánhangzó időtartama hosszabb, mint más helyzetekben (van Santen–D’Imperio 1999).

Nagy anyagon végzett elemzés alapján van Santen kimutatta (1992), hogy a magánhangzó időtartamára hét tényező hat (13 048 izolált mondat, férfi bemondó felolvasásában): a magánhangzó minősége, a szomszédos beszédhangok, a magánhangzó helyzete a szótagban, illetve a szótag típusa (VCV, CV, VC), a magánhangzót megelőző és követő szótagok száma a mondatban, a szótaghangsúly, a szó helyzete a mondatban (a frázishatár hatása és az, hogy a szón van-e mondathangsúly). A tempót és a szintaktikai szerkezetet külön nem elemezték.

A gyors beszédtempó a hosszú magánhangzók időviszonyait érintette jobban (rövidültek), a rövid magánhangzók azonban alig mutattak változást a koreai

rövid és hosszú magánhangzópárok vizsgálatában (Magen–Blumstein 1991). Hasonló eredménnyel zárult az a kutatás, amelyben rövid és hosszú japán magánhangzók időtartamait elemezték különböző beszédsebesség esetén. A tempóváltozás itt is jobban érintette a hosszú magánhangzókat, mint a rövideket; az arányok azonban azt igazolták, hogy valamilyen időzítési invariancia állandó maradt, függetlenül a beszélőtől és a beszédsebességtől (Hirata 2004).

Az anyanyelvi beszélők a beszédprodukcójukban igyekeznek a rövid és a hosszú magánhangzók elkülönítésére, noha az adott artikulációs megvalósítás számos tényező következménye. A különböző nyelvek vizsgálatai nyelvfüggetlen jelenségeket is igazoltak a fonetikai időtartamok alakulásában; ez pedig az időtartamviszonyok fiziológiai meghatározottságát támasztotta alá. Minden szegmentumnak feltételeztek egy specifikus átlagos időtartamot, amelyet magára a beszédhangra tekintenek jellemzőnek. Némely időzítési szerkezet a produkciós mechanizmus függvénye, például az artikulációs szervek univerzális biomechanikai működése (Perkell 1999). A beszédtervezés során azonban számos temporális paramétert a beszélőnek kell eldöntenie. Az időzítést rendszerint úgy tervezzük, hogy kielégítsük a beszédpartner kommunikációs igényeit (Wightman et al. 1992). Feltételezhető, hogy egy általános motoros program felügyeli azt, hogy a beszédírásban bekövetkező változások esetén az egyedi artikulációs mozgások kapcsolata lényegében ne változzon (Löfquist 1999). Ez biztosítja a temporális kohéziót, amely a spontán beszédben még fokozottabban érvényesül. Az artikulációs időkontroll kapcsán felmerül, hogy ez az ellenőrzés vajon belső vagy valamiféle külső mechanizmus. A külsőmechanizmus-felfogás szerint van egy központi óra vagy időszerkezet, amelyik irányítja és egyben ellenőrzi az artikulációs mozgásokat. A belső temporális ellenőrző mechanizmus képviselői azzal érvelnek, hogy az artikulációs mozgásokat csak a mechanizmuson belülről irányíthatja valami (Kelso–Tuller 1987). A hallgató beszédfeldolgozása egy bizonyos határon belül kevésbé érzékeny a szegmentális időviszonyok módosulásaira, mivel a mindenkori kontextus előjelzi a lehetséges szóvariációkat, és ily módon segíti a szelekciót. A szó mentális reprezentációja ennek megfelelően kellőképpen rugalmas (Gósy 2009).

A spontán beszédben tapasztalt időtartamviszonyok elemzése meglehetősen sporadikus, mivel ilyenkor nehéz kontrollált mintával dolgozni. A spontán beszéd és az olvasott beszéd sok tekintetben és szignifikánsan különbözik egymástól (Howell–Kadi-Hanifi 1991, Furui et al. 2005). Az úgynevezett laboratóriumi beszédben mért magánhangzók sajátos különbségeket mutatnak a spontán beszédprodukciónhoz képest spanyol nyelvi adatok alapján (Harmegnies–Poch-Olivé 1992).

A magyarban öt fonetikai magánhangzópár található, amelyek minőségüket tekintve nem vagy alig különböznek egymástól. A párok rövid és hosszú tagjainak szembenállása eltérő nyelvi funkciót mutat (pl. *szüret* ~ *szűret*, *irt* ~ *írt*, *bor* ~ *bór*, *örök* ~ *őrök*). A rövid és hosszú magánhangzóknak megfelelő betűket – az érvényes helyesírás szerint – eltérő jelekkel látjuk el (vagy jel nélküliek), ezek szemléltetik az időtartamok különbségét. A magyar magánhangzókat többen, többféle szempontból elemezték (Hegedűs 1941, Magdics 1966, 1969, Szende 1976, Kassai 1979, 1982, 1991, Bolla 1995, Kovács 2002, Olaszky 2006 stb.); az első vizsgálatok több mint száz évvel ezelőttiek (Gombocz–Meyer 1909). Ezek a kutatások

felolvasások alapján történtek, egy vagy néhány beszélő részvételével. Már több mint hat évtizeddel ezelőtt Laziczius (1944) felhívta a figyelmet arra, hogy a szegmentumok időtartamértékeit befolyásolják a mérési sajátosságok. A beszédakusztikai elemzési technikák fejlődése egyre pontosabb méréseket tesz lehetővé, a különböző kutatások adatai azonban – számos okból – csak óvatosan vehetők össze még ugyanazon nyelv esetében is. Mindezek tudatában a magyar magánhangzók időtartamviszonyairól korábban közölt adatok alapján megállapítható, hogy a fonológiaiailag rövid és hosszú magánhangzók időtartama jelentősen eltérni látszik valamennyi kutatásban; a határértékek (ahol vannak) jelzik az átfedéseket. Az objektív időtartamok függenek a magánhangzó minőségétől, a kontextustól, és különbségeket adatoltak a hangsúlyos és a hangsúlytalan helyzetű magánhangzók időtartamában is. A szakirodalomban található adatok (átlagok alapján) a rövid hangsúlyos *i, ü, ö, u* o^2 magánhangzókra: 60–140 ms, a hangsúlytalanokra: 50–105 ms; a hangsúlyos hosszú *í, ú, ő, ú, ó* magánhangzókra: 87–270 ms, a hangsúlytalanokra: 60–205 ms.

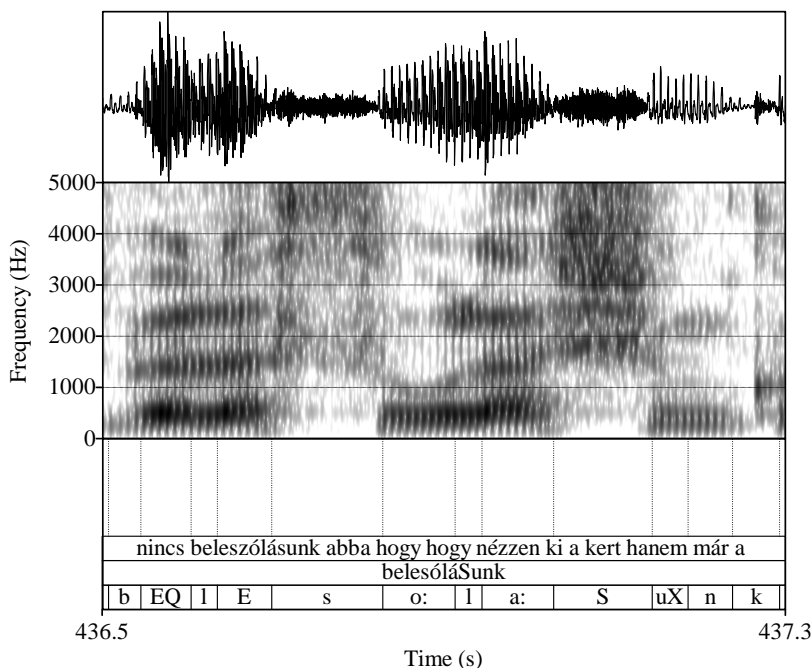
Spontán beszédanyagban eddig négy, különböző célú kutatás történt. Az első Hegedűs Lajos nevéhez fűződik 1941-ben; 2000 beszédhang időtartamát mérte le egy baranyai mesefelvételben. Közel négy évtizeddel később egy spontán rádióbeszélgetés anyagában történtek időtartammérések, ezek értékei a rövid hangsúlyos magánhangzókra: 60–100 ms, a hangsúlytalanokra: 40–100 ms; a hosszú hangsúlyosakra: 80–145 ms, a hangsúlytalanokra: 60–120 ms (Kassai 1979). Olasz (2006) szintén rádiós spontánbeszéd-felvételének megadott magánhangzó-időtartamai kissé rövidebbek (54 ms és 80 ms között szórnak). Bata (2007) szakdolgozatában spontán beszédben megjelenő magánhangzókat elemzett; a rövidek 65 ms és 83 ms közöttiek voltak, az *ó* átlaga 103 ms-nak, az *ú*-é pedig 90 ms-nak adódott. Noha említettük az összehasonlítás nehézségeit, az megállapítható, hogy a spontán beszédben ejtett magánhangzók rövidebbek, mint felolvasásban (ez persze számos belső és külső tényező következménye lehet); továbbá általánosan kimondható, hogy minél rövidebb a magánhangzók képzésére fordított idő, annál nagyobb a fonológiaiailag rövidek és hosszúak közötti átfedés.

A jelen kutatás spontánbeszéd-korpuszon vizsgálta a fonológiaiailag rövid és hosszú, e tekintetben párt alkotó, középső és felső nyelvvállású magyar magánhangzók időtartamviszonyait. A célunk az volt, hogy választ adjunk az alábbi fő kérdésekre: a) mekkora a fizikai tartam különbsége a rövid és a hosszú magánhangzók között, b) van-e eltérés a magánhangzó minőségének tekintetében, valamint c) a magánhangzók időtartamának alakulásában tényező-e a fonetikai helyzet. A hipotéziseink a következők: 1. napjaink spontán beszédében relatíve nagyok az időtartamok átfedései a kétféle fonológiai hosszúság esetén és 2. a felső nyelvvállású magánhangzók időtartamviszonyai nem mutatnak jelentős különbségeket, a középső nyelvvállásúaké azonban igen.

² A könnyebb olvashatóság érdekében a tanulmányban nem a fonetikai szakirodalomban szokásos IPA-jeleket használjuk, hanem a fonémákat (ill. adott esetben a beszédhangokat) a megfelelő dőlt betűvel jelenítjük meg.

Anyag, módszer, kísérleti személyek

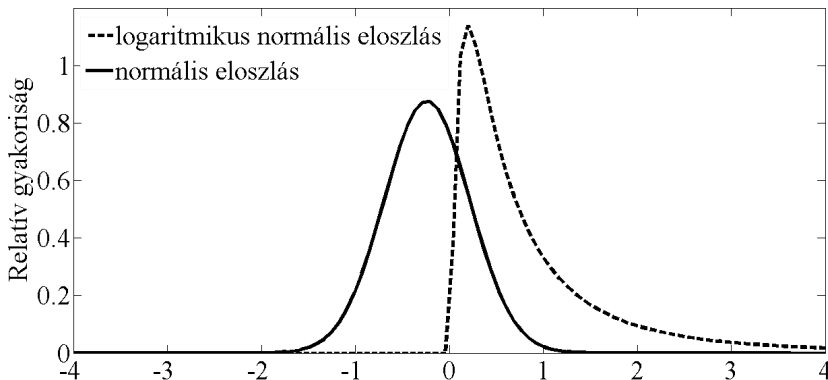
A BEA adatbázisából (Gósy 2008) 10 fiatal beszélő (5 férfi és 5 nő) spontán beszédét választottuk ki; életkoruk átlaga 25 év (a legfiatalabb 22, a legidősebb 30 éves). A hanganyagok spontán narratíváinak átlagos időtartama 8 perc volt, összesen több mint 110 percnyi anyagot elemeztünk (8214 szót). Néhány példa az anyagban előforduló szavak közül: *gyönyörű, túnt, színművészetire, összegyűlt, később, világrengető, tőlem, színész, korban, voltunk, külső, óra*. Az adatközlők beszédtempója hasonló volt, átlagosan 84 szó/perces sebességgel beszéltek. A hanganyagokat először szakaszszinten jegyeztük le a Praat szoftver alkalmazásával (Boersma–Weenink 2005); majd automatikus szegmentálóval (MAUS szoftver: ftp://ftp.bas.uni-muenchen.de/pub/BAS/SOFTW/MAUS) hangszinten annotáltuk. Ez azt jelenti, hogy a program a nemzetközi szakirodalomban szokásos hangelhatárolási kritériumok mentén jelöli be az egyes fonémák hangrealizációit (1. ábra). (Az eredetileg németre betanított szoftvert vizuális és auditív ellenőrzéssel használtuk.) Összesen 40 000 beszédhangot annotáltunk.



1. ábra. A *beleszólásunk* szó hangszínképe és annotációja (A Q az első szótagra, az X pedig az utolsó szótagra utal)

A gépi annotálást – ahol szükséges volt – manuálisan korrigáltuk. Az időtartamokat a Praatban írt gépi algoritmussal nyertük ki. Az időtartam kinyerésével egyidejűleg használt algoritmus biztosította a környezetfüggő címkézést is (pl. „B_k – o + J_r”, ez azt jelenti, hogy a rövid *o* magánhangzót a *k* mássalhangzó előzi meg, és

az r követi). Az automatikusan kinyert adatokon adatmodellezési vizsgálatokat végeztünk a MATLAB jelfeldolgozó és az SPSS statisztikai szoftver segítségével. Az adathalmaz vizsgálatok alapvetően fontos a tapasztalati eloszlást valamilyen nevezetes eloszlással közelíteni. Az időtartamértékek alapján kapott tapasztalati eloszlást minden esetben normalitás próbával közelítettük (Kolmogorov–Smirnov-teszt). Az eredmények az esetek döntő többségében logaritmikus normális eloszlást mutattak. Az adatok logaritmikus skálára történő transzformálásával normál eloszláshoz jutottunk (2. ábra).



2. ábra. Az eredetileg logaritmikus normál eloszlás (folytonos vonal) és az adatok megbízható elemzéséhez logaritmikus skálára transzformált adatok normál eloszlása (szaggatott vonal) az adatok relatív gyakoriságának (%) a függvényében

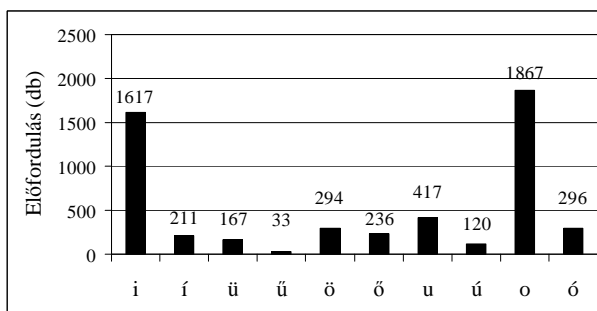
A statisztikai próbák eredményeiben nem találtunk jelentős különbséget az eredeti és a transzformált adatokban, ezért a végleges számolásokhoz az eredetileg kinyert értékeket használtuk.

Eredmények

Összesen 5258 magánhangzót elemeztünk, ebből 4362 volt fonológiaiailag rövid és 896 hosszú. Az időtartamok vizsgálata során az egyre részletezőbb elemzés módját választottuk. A következő tényezőket vettük figyelembe: a fonológiai hosszúságot, a magánhangzó gyakoriságát a szövegben, az artikulációs sajátosságokat, valamint a magánhangzónak a szóban betöltött helyét (a fonetikai pozíciót).

1. Gyakoriság

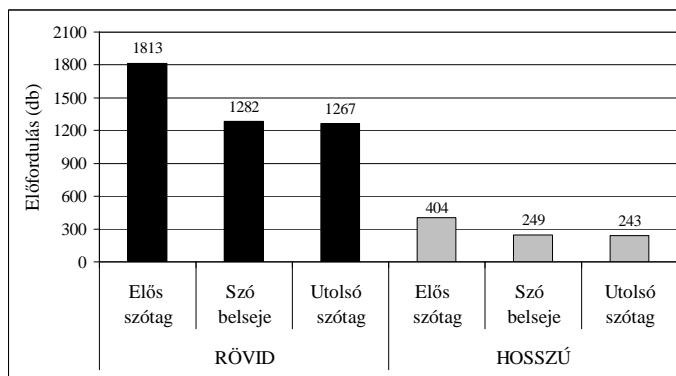
A spontán beszédben előforduló magánhangzók gyakorisága erősen változó (Gósy 2004); ezért nem valósítható meg az, hogy mindegyik elemzett magánhangzó azonos számú legyen (adataink tehát a valóságos eloszlást mutatják) (vö. 3. ábra).



3. ábra. A vizsgált magánhangzók előfordulása (darabszámban) a spontánbeszéd-anyagban

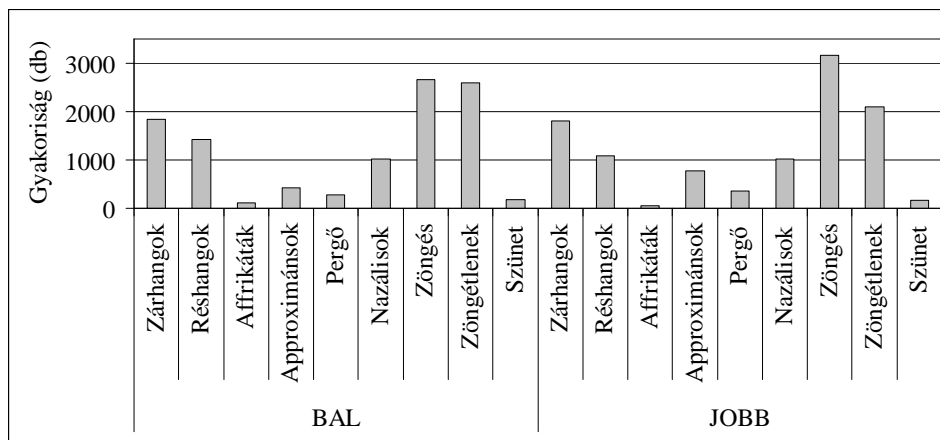
Korábbi elemzések alapján (pl. Gósy 2004) valószínűsítettük, hogy a spontán beszédben a rövid *o* és *i* realizációi fordulnak elő a leggyakrabban (vö. 1867 db *o*, 1617 db *i*), legritkábban pedig a felső nyelvállású, palatális, labiális magánhangzók. A magánhangzópárok rövid és hosszú tagjainak előfordulása különböző. A hosszú realizációk minden esetben ritkábbak, a legkevesebb előfordulása az *ú* és az *ű* realizációinak volt. A gyakori *i* magánhangzók mellett nagyon ritka a fonológiailag hosszú változat, ugyanakkor az *ö* és az *ő* realizációi csaknem azonos arányban fordultak elő (294 db és 236 db).

A magánhangzók gyakorisága aszerint is különbözhet, hogy a szó mely szótagjában helyezkedik el. Anyagunkban mindhárom pozícióban (első szótag, hangsor belseje, utolsó szótag) a rövid *i* és *o* realizációi voltak a leggyakoribbak; az első szótagban és a szó belsejében az *ű*, az utolsó szótagban pedig az *í* volt a legritkább (4. ábra). Az időtartamot befolyásolja a hangsorban elfoglalt helyzete; megállapították, hogy az adott nyelvtől és a hangsúlyhelytől függetlenül, a szóvégi magánhangzók megnyúlnak, illetve hogy a hangsúlyos helyzetű magánhangzók időtartama rendszerint hosszabb (vö. Magdics 1966, Kassai 1979, van Santen 1992, Kovács 2002 stb.).



4. ábra. A vizsgált magánhangzók előfordulása a fonológiai tartam és a fonetikai pozíció függvényében

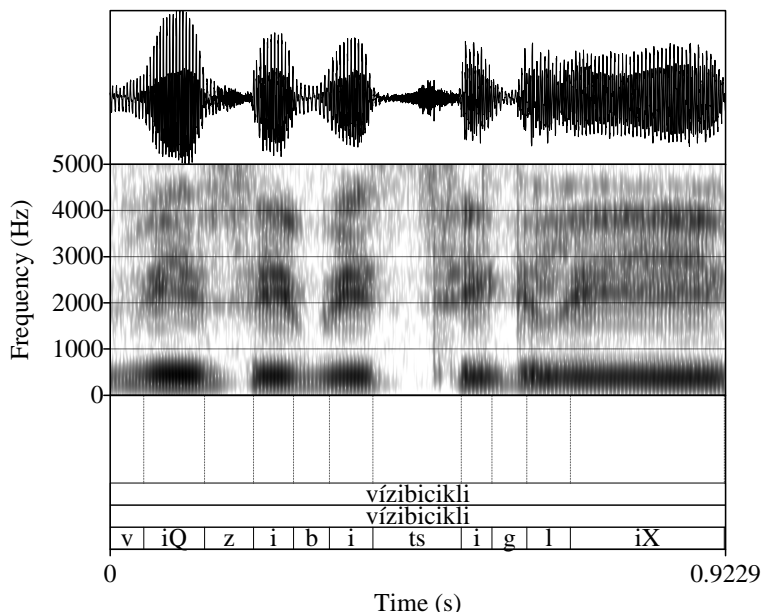
Anyagunkban különféle mássalhangzók és szünetek voltak adatolhatók a magánhangzókat megelőző és követő helyzetben (5. ábra), magánhangzó csupán néhányszor fordult elő (ezért ez utóbbiakkal a csekély szám miatt nem foglalkoztunk). A megelőző mássalhangzók hasonló arányban voltak zöngések vagy zöngétlenek, a magánhangzót követők között lényesen nagyobb mértékben fordultak elő a zöngések. Várhatóan a zárhangok voltak a leggyakoribbak mindkét helyzetben, ezeket a réshangok követték (gyakrabban a magánhangzót megelőzően). Gyakoriságban a következők a nazálisok voltak, elhanyagolható különbséggel a megelőző és a követő helyzetben, majd az approximánsok, amelyek gyakrabban fordultak elő a magánhangzót követő helyzetben, végül a pergőhangok (valamivel nagyobb mértékben a követő helyzetben). Az affrikáták megjelenési aránya rendkívül alacsony. Közel azonos arányban találtunk szünetet az elemzett magánhangzókat megelőzően és követően is (a bal kontextusban 172 db, a jobb kontextusban 168 előfordulás).



5. ábra. A magánhangzó jobb és bal oldali kontextusában előforduló hangok és a szünetek gyakorisága

2. Fizikai időtartamok: teljes anyag

Az elemzett magánhangzók átlagos időtartama 78 ms (szórás: 48 ms). A rövid magánhangzók átlaga 65 ms (szórás: 32 ms), a hosszúaké 90 ms (szórás: 47 ms). Valamennyi rövid és hosszú magánhangzó időtartama alapján megállapítható, hogy (a) tendenciájukban a rövidek rövidebb, a hosszúak pedig hosszabb fizikai időtartammal realizálódnak, ugyanakkor (b) a fonológiailag eltérő tartamú magánhangzók között nagy a fizikai időtartamok átfedése. A 6. ábra a *vízibicikli* szó ejtése alapján készült hangszínekben szemlélteti az *i*-realizációk változó időtartamát.



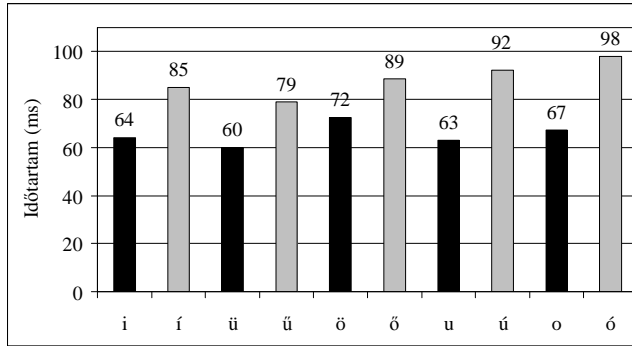
6. ábra. A vízibicikli szóban lévő *i/i* realizációk időtartamai:
80 ms, 73 ms, 75 ms, 39 ms, 240 ms

Az összes rövid és hosszú magánhangzó időtartamára elvégzett statisztikai elemzések szignifikáns különbséget igazoltak a fonológiai opozíció szerint, ezt az átlagok nagy különbsége is előjelzi (egytényezős ANOVA: $F(1, 5256) = 365,560$; $p = 0,001^{**3}$). Ha összevetjük adatainkat a korábbi kutatások adataival (ez az összevetés – mint már említettük – rendkívül bizonytalan részben az eltelt idő,⁴ részben számos eltérő módszertani tényező miatt), azt látjuk, hogy a rövid magánhangzókra az olvasott szövegek rövid magánhangzóihoz hasonló időtartamokat kaptunk; a hosszúak fizikai időtartama azonban anyagunkban jóval rövidebb. Az egyes magánhangzópárok eltérő arányban különböznek fizikai időtartamaikban (7. ábra). A legrövidebbtől a hosszabbak felé a rövid magánhangzók sorrendje a következő: $\ddot{u} < u < i < o < \ddot{o}$. Ugyanez a sorrend a hosszú magánhangzóknál némiképpen másként alakul: $\ddot{u} < \acute{i} < \ddot{o} < \acute{u} < \acute{o}$.

Minden magánhangzópár esetén szignifikáns eltérést találtunk a fizikai időtartamokban ($p = 0,001^{**}$). Az egyes magánhangzópárok időtartamának különbsége szoros összefüggést mutat a minőségükkel (1. táblázat).

³ A kettős csillag a 99%-os statisztikai megbízhatóságot jelzi.

⁴ Gombocz és Meyer 1908-ban például az *i/i* időtartamára 118 ms és 249 ms-os átlagértéket adott meg egy szótagú szavakban mérve, Magdics 1960-ban mondatokban mérve 102 ms-ot, illetve 180 ms-ot, Kassai (1979) spontán beszédben mért értékei 60 ms és 85 ms, Olaszky (2006) adatai saját felolvasásában 69 ms és 87 ms.

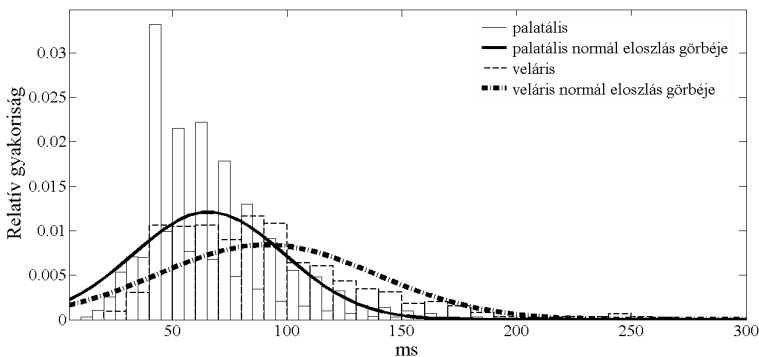


7. ábra. Az egyes magánhangzó-fonémák realizációinak átlagidőtartama

1. táblázat. A fonológiailag rövid és hosszú magánhangzók átlagos időtartama és szórása

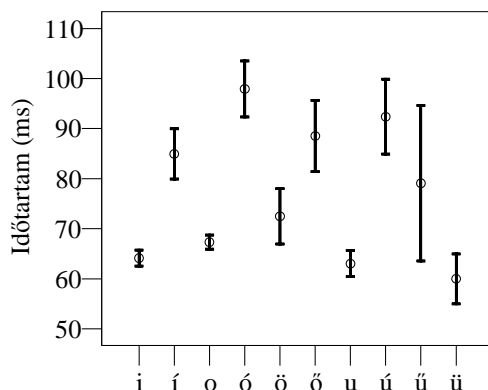
A magánhangzó minősége	A magánhangzók időtartama (ms)			
	Fonológiailag rövid		Fonológiailag hosszú	
	Átlag	Szórás	Átlag	Szórás
<i>i - í</i>	64	33	85	37
<i>ü - ű</i>	60	33	79	44
<i>ö - ő</i>	72	48	89	55
<i>u - ú</i>	63	27	92	41
<i>o - ó</i>	67	31	98	49

Az *o/ó* és az *u/ú* esetében volt a legjelentősebb a különbség (29 ms és 31 ms, mindkettő 46%), a legkisebb időeltérések pedig az *ö* és *ő*, az *ü* és *ű* között adódtak (17 ms: 23%; 19 ms: 32%). Ezek az adatok azt igazolják, hogy az elől képzett magánhangzók ejtésében kevésbé jelennek meg a realizációk időkülönbségei, míg a hátul képzettekénél ez nagyobb mértékű. A magánhangzó időtartamában statisztikai különbség van a nyelv vízszintes mozgása szerint ($F(1, 5256) = 4,382; p = 0,032$) (vö. 8. ábra).



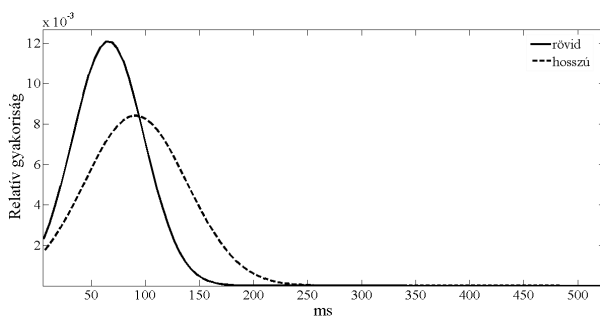
8. ábra. Az elemzett palatális és veláris magánhangzók időtartamainak eloszlása az időtartamértékek és a relatív gyakoriság (%) függvényében

A felső nyelvállású rövid magánhangzók időtartamátlagja 59 ms, a középső nyelvállású rövideké 64 ms. A rövidek esetében a nyelvállásfok meghatározó tényező, a felső nyelvállásúak időtartamátlagja szignifikánsan rövidebb, mint a középső nyelvállásúaké ($F(1, 1812) = 10,675; p = 0,001^{**}$). A hosszú, felső nyelvállásúak átlagos időértéke 85 ms, a hosszú, középső nyelvállásúaké 91 ms (ez a különbség matematikailag nem igazolható). A hosszú magánhangzókánál a nyelv vízszintes mozgása tűnik jellemzőbbnek; az elöl képzettek rövidebbek, mint a hátul képzettek, de ez a különbség sem szignifikáns. Az egyes magánhangzó párok időtartamai nagy fizikai értékkülönbségekkel realizálódnak; az átlagokat és a szórástartományt a 9. ábra szemlélteti.



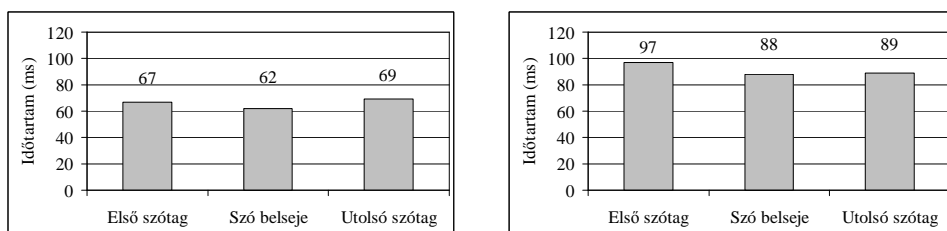
9. ábra. Az elemzett magánhangzók időtartamainak átlaga (medián) és szórása

Az ábráról leolvasható, hogy a beszélők melyik magánhangzót milyen változatossággal ejtik az időtengely mentén. Legkevésbé az *o* és az *i* realizációi variálódnak, a legnagyobb mértékben pedig a hosszú, labiális, palatálisok, valamint az *ö*. Megállapítható, hogy a hosszú magánhangzók artikulációs tartama kevésbé állandó, mint a rövideké. A legrövidebb mért magánhangzók 25 ms-osak, a leghosszabbak a 280 ms-ot is meghaladják. A nagyfokú átfedéseket a rövid és a hosszú magánhangzók között a 10. ábra szemlélteti.



10. ábra. Az összes elemzett rövid és hosszú magánhangzó időtartama az érték és a relatív gyakoriság (%) tükrében

Elemeztük az időtartamok alakulását a magánhangzónak a szóban elfoglalt helyétől függően. A rövid magánhangzók első szótagi előfordulásban átlagosan 67 ms-osak, valamivel hosszabbak a szó végén, átlagosan 69 ms-osak, és rövidebbek a szó közepén, az átlagérték itt 62 ms. (A szórás 30–35 ms közötti.) A szóban elfoglalt helyzet szignifikáns különbséget mutat ($F(1, 5256) = 17,327$; $p = 0,000^{**}$; a Tukey post-hoc teszt szerint az első és az utolsó szótagi magánhangzó időtartama különbözik szignifikánsan a szó belseji szótagban lévő időtartamától). A hosszú magánhangzók fizikai időtartamában nagyobbak az időkülönbségek a fonetikai pozíció függvényében. Az első szótagban átlagosan 97 ms-osak, a szóvégen 89 ms az átlag, míg szó belsejében 88 ms (11. ábra), ezek az eltérések azonban nem szignifikánsak. (A szórás 41–56 ms közötti.)



11. ábra. A magánhangzók átlagos időtartama a hangsorban elfoglalt helyzet szerint (rövid magánhangzók = bal oldalon, hosszú magánhangzók = jobb oldalon)

A rövid magánhangzók a szavak első és többedik, de nem utolsó szótagjában hasonlóan rövidebb időtartammal realizálódnak, mint a szavak utolsó szótagjában. Ezzel szemben a hosszú magánhangzók relatíve hosszabbak a szó első szótagjában, míg az időtartamaik hasonlóbbak a szó belsejében és az utolsó szótagban. Adataink azt mutatják, hogy az első szótagi magánhangzók időtartam-növekedése a hosszú magánhangzóknál kifejezettebb, mint a rövideknél (vö. pl. Greenberg et al. 2003). Az utolsó szótagban lévő magánhangzók időtartamának jellegzetes növekedése inkább a rövid magánhangzóknál tapasztalható. A spontán beszédben a fonetikai pozíció sajátosan hat az egyes magánhangzók időtartamára (2. táblázat). A hosszú, felső nyelvvállásúaknál nem mutatható ki semmilyen hatás, hasonlóképpen a rövid *u* és *ü* esetében sem. A többi magánhangzónál egy-egy pozíció esetében volt szignifikáns különbség (pl. az *i* esetében a szó belsejében és az utolsó szótagban lévő magánhangzó időtartamában vagy az *ó* esetében az első és az utolsó szótagban megjelenő magánhangzók időtartamában).

Ha a magánhangzók az első szótagban fordulnak elő, akkor a rövid és hosszú hangok fizikai időtartamainak különbsége átlagosan mintegy 31 ms (37%). Ha a szó belsejében vannak, akkor a rövid/hosszú oppozíció átlagosan mintegy 25 ms-os (36%) különbségben realizálódik. Az utolsó szótagban a rövidek és a hosszúak között átlagosan mintegy 20 ms (35%) az eltérés; a tendencia tehát az időkülönbségek csökkenését mutatja, az arányok azonban jelentősen nem változnak. Meg kell jegyeznünk, hogy spontánbeszéd-anyagunkban a szavak első és utolsó szó-

2. táblázat. A magánhangzók időtartama a fonetikai pozíció, a hangminőség és a fonológiai tartam függvényében

Magánhangzók	A magánhangzók időtartamának átlaga (ms)					
	Fonológiailag rövid			Fonológiailag hosszú		
	Első szótag	Szó belseje	Utolsó szótag	Első szótag	Szó belseje	Utolsó szótag
<i>i – í</i>	62	60	71	87	83	102
<i>ü – ű</i>	59	47	70	69	66	91
<i>ö – ő</i>	71	77	64	87	91	87
<i>u – ú</i>	62	64	63	98	91	87
<i>o – ó</i>	72	63	70	108	92	91

tagja az esetek döntő többségében nem jelent virtuális mondathatárt is egyben; továbbá elemzésünk nem terjedt ki a szótagszerkezetre és arra, hogy a szó belsejében hányadik szótagi magánhangzóról van szó.

3. A magánhangzó-időtartamok korrigált értéktartománya

A spontán beszéd fonetikai jellemzőinek egyike, hogy a beszédhangok esetenként megnyúlnak, néha indokolatlanul, illetve szándéktalanul. Ezeknek a nyújtásoknak az oka gyakran a beszédtervezés diszharmoníájára vezethető vissza. A megnyújtott magánhangzóknek rendszerint extrém fizikai időtartamértékük van (anyagunkban pl. előfordult 320 ms-os magánhangzó is). A nyújtások előfordulásai és időviszonyai beszélőnként változók; időtartamuk behatárolása még egyetlen beszélő esetében is csak percepciók kísérlet alapján lehetséges. A nyújtás észlelése azonban multifaktoriális; egy magyar kutatás eredményei szerint (Bóna 2007) nem csupán a hangidőtartamok növekedése játszik szerepet a nyújtások azonosításában, hanem meghatározó a magánhangzó minősége is, továbbá befolyásoló tényező a beszédszakaszban elfoglalt pozíció, a beszéddallam, sőt az intenzitás. Ugyancsak kísérleti eredmény erősítette meg, hogy a hallgatók a megakadások közül a legkevesbé a nyújtás jelenségére figyelnek fel. A fizikai időtartamok elemzésekor azonban figyelemmel kell lennünk a nyújtás tényére, hiszen az ekkor kialakuló, esetleg szélsőséges időértékek (statisztikailag) meghamisíthatják a spontán beszédben realizálódó magánhangzók időtartamarányait.

A spontán beszéd a nyújtások mellett rendkívül rövid (például 25 ms-os) magánhangzókat is tartalmaz, amelyek izoláltan nem vagy igen bizonytalanul azonosíthatók. A folyamatos beszédben a hangsor, illetve a szó felismerése és megértése biztosítja a hallgató számára az igen rövid beszédhangok „helyreállítását” (mindehhez a közvetlen kontextus szemantikai és grammatikai sajátosságai is hozzájárulnak). A beszédhangok észlelésében fontos a megkülönböztetési küszöb értéke, ezt azonban számos tényező befolyásolja; ahhoz, hogy az időtartam-különbség fonológiailag releváns információ lehessen, meg kell haladnia a 25 ms-ot (Klatt–Cooper 1975, Nooteboom 1999).

A fentiek miatt adatainkat „normalizáltuk”, tekintetbe véve az adott magánhangzóra jellemző szórást, valamint az egyes értékekhez rendelt előfordulási számot. A perceptíósan irreleváns és az extrém időtartamokat – minden esetben az adott magánhangzónál meghatározva – a további elemzésekből kizártuk. Az így kapott adathalmazra a korrigált értéktartomány a jellemző, amelynek határértékei 30 ms és 130 ms (az adott magánhangzó minőségétől függően). A korrigált értéktartomány kijelölése a szerzőknek a megadott kritériumok szerinti „önkényes” döntése, amely természetesen nem független a kísérleti adathalmaztól. A szükséges statisztikai eljárások eredményeinek ismeretében kijelenthető, hogy a korrigált értéktartomány reális, és alkalmas a további elemzésekre.

3.1. A magánhangzók időtartamadatai

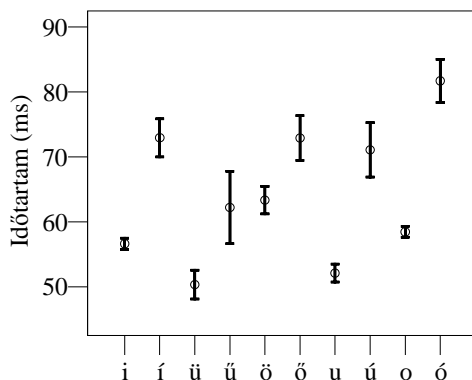
A korrigált adathalmazban az összes elemzett magánhangzó átlaga 64 ms; a rövid magánhangzók átlaga 57 ms, a hosszúaké pedig 75 ms; a közöttük lévő különbség szignifikáns ($F(1, 4111) = 613$; $p = 0,001^{**}$). A rövid magánhangzók minimálisan 30 ms-osak; a maximális értékek a magánhangzó minőségétől függően változók (80 ms-os az *u* és az *ü*, 90 ms-os az *i* és az *o*, 100 ms-os az *ö*). A hosszú magánhangzók minimálisan 40 ms-osak; maximális értékük szintén minőségspecifikus (az *ű* realizációi 80 ms-osak az *ú* magánhangzók 100 ms-osak, az *í*-k 110 ms-osak, az *ó*-k 120 ms-os, az *ó*-k pedig 130 ms-osak). Az extrém időtartamú előfordulások főként az elől képzett, középső nyelvállású labiális magánhangzókat érintették. Az egyes magánhangzókra kapott adatokat a 3. táblázat tartalmazza.

3. táblázat. A magánhangzók időtartamai a korrigált értéktartományban

A magánhangzók minősége	A magánhangzók időtartama (ms)			
	Fonológiailag rövid		Fonológiailag hosszú	
	Átlag	Szórás	Átlag	Szórás
<i>i – í</i>	57	16	73	19
<i>ü – ű</i>	50	12	62	13
<i>ö – ő</i>	63	16	73	24
<i>u – ú</i>	52	11	71	18
<i>o – ó</i>	58	16	82	25

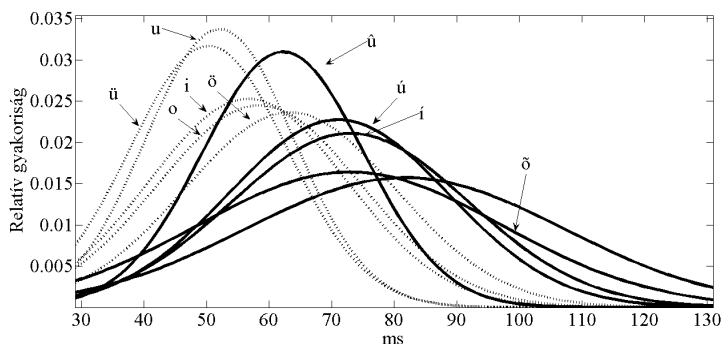
A korrigált értéktartomány adatai hasonló tendenciákat mutatnak az időtartamok alakulásában, mint a teljes adathalmazéi (vö. 12. ábra).

A felső nyelvállású rövid és hosszú magánhangzók rövidebbek, mint a középső nyelvállásúak. Az *i* és az *o* átlaga között alig van eltérés. A rövid, felső nyelvállású, labiálisok rövidebbek, mint az illabiálisok. Az *í* és az *ú* fonémák realizációinak időtartamátlagában azonban alig van különbség. A rövid magánhangzóknál az *ö*, a hosszúaknál az *ó* átlagidőtartama a legnagyobb. Az *ó* átlagosan



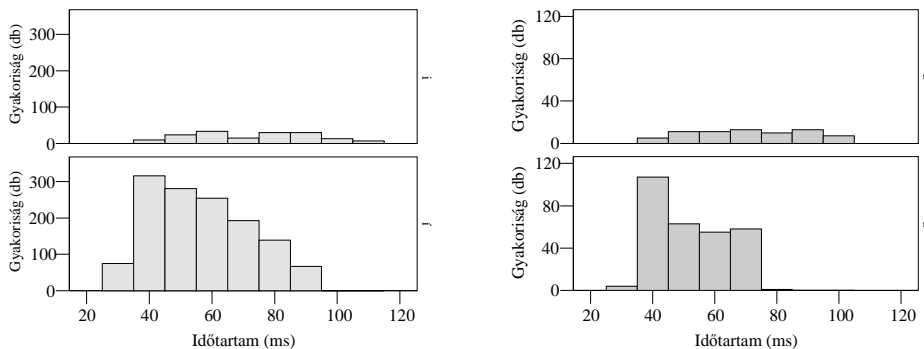
12. ábra. A korrigált adathalmaz alapján megállapított időtartam-átlagok (medián) és szórások

hosszabb időtartamban realizálódik, mint az \ddot{o} , ez éppen ellentétes azzal, amit a megfelelő rövid magánhangzónál látunk. A felső nyelvállású magánhangzók kerekítettségével kapcsolatosan annyi állapítható meg, hogy legrövidebb időtartamúak az \ddot{u}/\acute{u} realizációi. A 13. ábra az elemzett magánhangzók összes időtartamának relatív gyakoriságát szemlélteti az előfordulás függvényében (a tapasztalati eloszlásra illesztett normális eloszlású görbékkel szemléltetve).



13. ábra. A magánhangzók időtartamának relatív gyakorisága (%) az időtartam függvényében (az \acute{u} jel az \ddot{u} , az \ddot{o} pedig az \ddot{o} helyett áll)

A görbék a következő főbb eredményeket szemléltetik: a) a rövid és a hosszú magánhangzók időtartamai egyértelműen elkülönülnek, b) a fonológiai oppozícióban álló fonémák realizációi eltérő mértékben különböznek az időtartamértékeik tekintetében, c) a fonológiai tartamtól függően egyes magánhangzók időtartamai szűkebb sávban (pl. u vagy \acute{u}), mások szélesebb sávban jelentkeznek (pl. \ddot{o} vagy \acute{o}) és d) valamennyi fonológiai oppozícióban álló magánhangzópár jellegzetes átfedést mutat az időértékekben. A 14. ábra az u/\acute{u} és az i/\acute{i} példájával mutatja a konkrét értékeket.



14. ábra. Az időtartam és a gyakoriság összefüggései az *ú/u* (bal oldalon) és az *í/i* párok (jobb oldalon) esetében

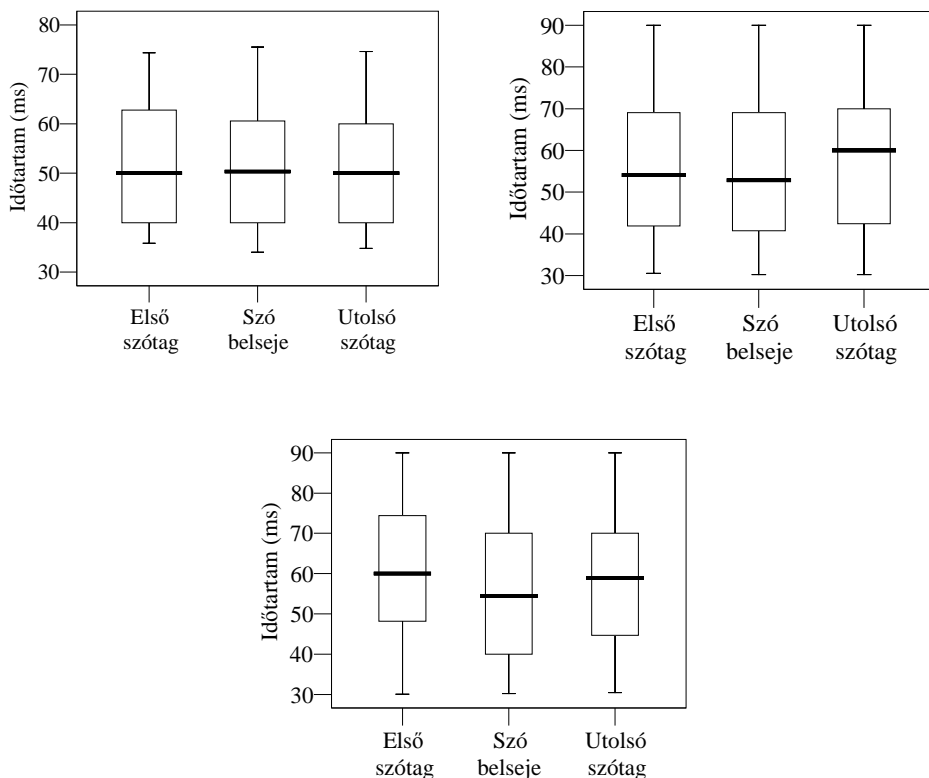
A korábbi elemzések igazolták azt, hogy az időtartamértékek a magánhangzó képzésének függvényében is jellegzetesen változók. A statisztikai elemzések anyagunkban több tekintetben is igazolható eltéréseket mutattak. A középső nyelvállású rövid magánhangzók átlaga 59 ms, a hosszúaké 78 ms, a különbség szignifikáns ($F(1, 2117) = 334,401; p = 0,001$). A felső nyelvállású rövidek átlaga 55 ms, a hosszúaké 71 ms, az eltérés itt is szignifikáns ($F(1, 1993) = 236,505; p = 0,001$). Az elől, illetve hátul képzett rövid magánhangzók átlagidőtartama azonos, 57 ms. A hosszúaknál az elől képzettek átlagos időtartama 72 ms, a hátul képzetteké 79 ms, ez a különbség szignifikáns ($F(1, 672) = 15,238; p = 0,001$). A kerekítettség tekintetében nem találtunk matematikailag igazolható különbséget a fonológiai oppozíció függvényében (a rövid kerekített magánhangzók átlaga 57 ms, a kerekítetleneké 56 ms; a hosszú kerekítetteké 76 ms, a rövideké 72 ms).

3.2. A fonetikai pozíció

Az első szótagban előforduló magánhangzók időtartama – fonológiai hosszúságtól függetlenül – 61 ms (szórás: 19 ms); a szó belsejében megjelenőké 59 ms (szórás: 18 ms); az utolsó szótag magánhangzóinak átlagos időtartama szintén 59 ms (szórás: 17 ms). Az első szótagban előforduló rövidek átlaga: 58 ms, a hosszúaké 79 ms, a szó belsejében adatolt rövideké 56 ms, a hosszúaké 75 ms. Az utolsó szótagban található rövid magánhangzók átlagos időtartama 57 ms, a hosszúaké 70 ms. Noha az időtartamértékek különbsége nagyon kicsi, a rövid magánhangzók hosszabbak az első szótagban, mint a másik két fonetikai helyzetben; a hosszúak időtartama ugyancsak az első szótagi előfordulásakor a leghosszabb, de az utolsó szótagban a legrövidebbek. A kiugró értékek kizárása – jelen esetben a nyújtásoké is – főként az utolsó szótagban eredményezett időtartam-módosulásokat (vö. 11. ábra).

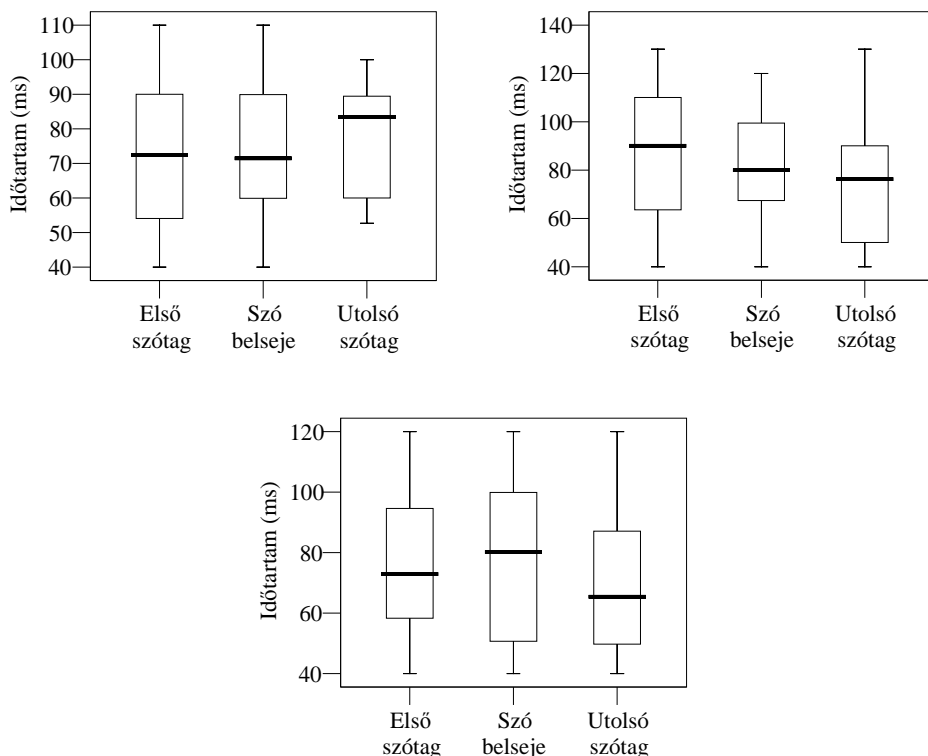
A rövid magánhangzók közül az *u* és az *ü* realizációi nem mutatnak statisztikailag értékelhető különbséget a fonetikai pozíció függvényében. Matematikailag releváns eltéréseket látunk ugyanakkor az *i* esetében a szó belseji (átlagérték: 55 ms)

és a szóvégi helyzet (átlagérték: 58 ms) között ($F(2, 1323) = 3,523$; $p = 0,030$), továbbá az *o* esetében (az átlagérték az első szótagban 61 ms, szó belsejében 57 ms; $F(2, 1453) = 7,948$; $p = 0,001$) és az *ö*-realizációknál (az átlagérték az első szótagban 67 ms, szóvégen 59 ms; $F(2, 243) = 5,778$; $p = 0,004$). Ez azt sugallja, hogy az első szótagi helyzet öt magánhangzóból háromnál gyakorol hatást az időtartamok alakulására (vö. 15. ábra).



15. ábra. Rövid magánhangzók időtartamai (medián és szórás) a fonetikai pozíció függvényében (*u* = bal oldalt, felül; *i* = jobb oldalt, felül; *o* = alul)

A hosszú, felső nyelvállású magánhangzók esetében az első szótagi pozíció irreleváns az időtartamokat tekintve. Az *ó* esetében az első és az utolsó szótagban realizálódók (átlag: 87 ms, ill. 75 ms; $F(2, 222) = 4,430$; $p = 0,013$) mutatnak matematikailag igazolható eltérést. Az *ő*-nél a szó belseji és a szóvégi magánhangzók közötti időtartam-különbség volt szignifikáns (az átlagok: 77 ms, ill. 68 ms; $F(2, 190) = 3,178$; $p = 0,044$) (vö. 16. ábra).

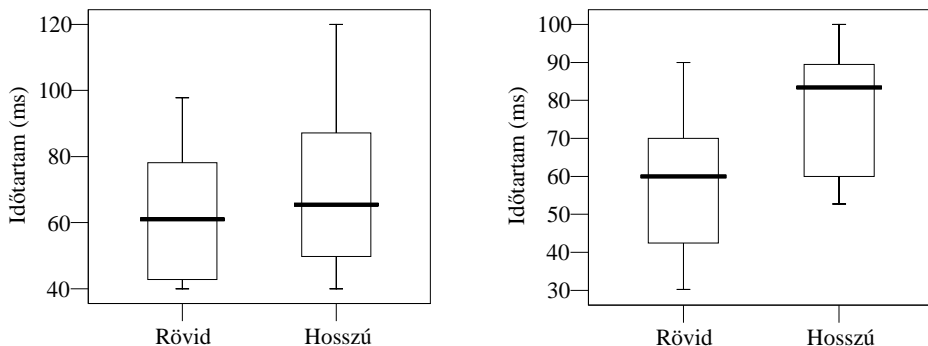


16. ábra. Hosszú magánhangzók időtartamai (medián és szórás) a fonetikai pozíció függvényében (í = bal oldalt, felül; ő = jobb oldalt, felül; ö = alul)

Elemeztük a hangsor belseji helyzetben megjelenő magánhangzók időtartamát az első és az utolsó szótagban megjelenőkéhez képest. Mindössze három esetben találtunk figyelemre méltó különbséget. A rövidek közül az *o* (szó belsejében: 48 ms, egyéb helyzetekben: 64 ms; $F(1,1454) = 412,833$; $p = 0,001$) és az *ö* (szó belsejében: 59 ms, egyéb helyzetben: 66 ms; $F(1,242) = 10,225$; $p = 0,002$); a hosszúak közül pedig csak az *ő* (a hangsor belseji átlag: 77 ms, a többi helyzetben: 70 ms; $F(1, 191) = 4,488$; $p = 0,035$) időtartamértékei különböznek szignifikánsan.

A fonológiai oppozícióban álló magánhangzópárok matematikailag igazolható különbséget mutatnak – két kivétellel – minden fonetikai helyzetben (nincs eltérés szó belsejében az *ü/ű* és az utolsó szótagban az *ö/ő* között). A 17. ábra az *ö/ő* adatai mellett példaként mutatja az *i/i*-re kapott értékeket ugyancsak szóvégi pozícióban.

Megvizsgáltuk, hogy az egyes változók miként befolyásolják a magánhangzók időtartamának alakulását (egyváltozós varianciaanalízis alkalmazásával). A független változók a fonológiai tartam, az ajakműködés, a nyelv vízszintes és függőleges



17. ábra. Az időtartamok nem szignifikáns (az *ö/ő* realizációjánál, bal oldalon) és szignifikáns (az *i/í* realizációjánál, jobb oldalon) különbsége a fonológiai tartam függvényében az utolsó szótagban

mozgása, valamint a fonetikai pozíció voltak. Az eredmények azt mutatják, hogy a legnagyobb hatása az időtartamra a fonológiai tartamnak van, vagyis annak, hogy az adott magánhangzó rövid vagy hosszú. Ezt követi a nyelv függőleges mozgása, majd jóval kisebb mértékben az ajakműködés, a nyelv vízszintes mozgása, legvégül pedig az, hogy első szótagban fordul-e elő a magánhangzó.

3.3. A magánhangzók automatikus osztályozása a kontextus és az időtartam alapján

A rövid és a hosszú magánhangzókat döntési fával modelleztük. A döntési fa alapötlete az, hogy az összetett összefüggések egyszerű döntések sorozatára vezethetők vissza. Az algoritmus képes felismerni, hogy mely paraméterek jelentősek a csoportosításban. Noha a szakirodalomban nem modellezik mindig külön a hangok időtartamát a beszédfelismerésben – mivel az általában alkalmazott rejtett Markov-modell időben igen rugalmas mintaillesztést tesz lehetővé –, ez a nézet korántsem egységes. Számos tanulmány fontosnak tartja a magánhangzók pontos hangidőtartam-modelljeinek kialakítását (Pylkkönen–Kurimo 2004, Kocsor–Tóth 2005). A jelen kutatás elemzéseiből látható, hogy jelentős különbség van a rövid és hosszú magánhangzók között, jóllehet – számos oknál fogva – azt várnánk, hogy a spontán beszédben a magánhangzó-oppozíció nem jelenik meg mindenütt egyértelműen a fizikai értékekben, illetve egyfajta kiegyenlítődéss történik. Általában elmondható, hogy azok a jellemzők, amelyek fontosak az egy csoportba soroláshoz, azok a fa tövéhez vannak közel, attól távolodva a szerepük csökken (18. ábra).

```

időtartam <= 90
| B_APPROXIMÁNS = IGEN
|| K_KÖZÉPSŐNYELVÁLL = IGEN
||| J_PALATÁLIS = IGEN: RÖVID
||| J_PALATÁLIS = NEM
|||| J_APPROXIMÁNS = IGEN: HOSSZÚ
|||| J_APPROXIMÁNS = NEM
||||| időtartam <= 88.546207
||||| J_RÉS = IGEN
||||| J_ALVEOLÁRIS = IGEN: RÖVID
||||| J_ALVEOLÁRIS = NEM
||||| K_PALATÁLIS = IGEN
||||| B_ALVEOLÁRIS = IGEN: HOSSZÚ
||||| B_ALVEOLÁRIS = NEM: RÖVID
||||| K_PALATÁLIS = NEM: HOSSZÚ
||||| J_RÉS = NEM
||||| K_HANGSÚLYOS = IGEN
||||| K_PALATÁLIS = IGEN: RÖVID
||||| K_PALATÁLIS = NEM
||||| J_NAZÁLIS = IGEN: HOSSZÚ
||||| J_NAZÁLIS = NEM
||||| J_ALVEOLÁRIS = IGEN
||||| B_ALVEOLÁRIS = IGEN: RÖVID
||||| B_ALVEOLÁRIS = NEM: HOSSZÚ
||||| J_ALVEOLÁRIS = NEM: RÖVID
||||| K_HANGSÚLYOS = NEM

```

18. ábra. A döntési fa modelljének részlete

A döntési fa kialakításához az adatok 60%-át használtuk (2404 magánhangzó-időtartam), míg a teszteléséhez a fennmaradó részt (1707 db-ot). A magánhangzókat a fonológiai oppozíciónak megfelelően rövid és hosszú osztályokra bontottuk. A döntési fa a teszteléskor a rövid magánhangzók 97%-át, a hosszú magánhangzók 52%-át osztályozta helyesen. Ezek az eredmények alátámasztják, hogy a fizikai időtartam meghatározó akusztikai jegy az elemzett magánhangzók elkülönítésében, jóllehet a hosszú magánhangzók azonosítása jóval bizonytalanabb.

Következtetések

A jelen kutatásban azokra a kérdésekre igyekeztünk választ találni, amelyek a fonológiai időtartampárt alkotó magánhangzók fizikai időtartamainak alakulására vonatkoznak a spontán beszédben. A fizikai tartam különbsége a rövid és a hosszú magánhangzók között jelentős, szignifikánsan igazolható. Ez azt jelenti, hogy mindegyik magánhangzó pár objektív időértékeiben egyértelműen tükröződik a fonológiai hosszúság különbsége. A spontán beszédben természetesen igen nagyok

az átfedések a rövid és a hosszú hangok időtartamai között, ez azonban csak jellemzi az értéktartományokat, de mégsem teszi azokat azonossá vagy nagyon hasonlóvá. Választ kerestünk arra, hogy a magánhangzó minősége milyen mértékben játszik szerepet a rövid és a hosszú magánhangzó párok fizikai időtartamaiban. Az artikulációs gesztusok bizonyos mértékig befolyásolják az időtartamok alakulását, de a hatásuk a spontán beszédben nem független a magánhangzó fonológiai hosszától. A nyelv függőleges mozgása meghatározó a fizikai időtartamok alakulásában ($F(1, 672) = 11,560, p = 0,001$); a nyelv vízszintes mozgása csak a hosszú magánhangzók esetében időtartam-módosító tényező ($F(1, 672) = 15,238, p = 0,001$). Az ajakműködés nem eredményez szignifikáns különbségeket sem a rövid, sem a hosszú magánhangzóknál.

A fonetikai helyzet sajátosan gyakorol hatást a magánhangzók időtartamára, amely részben az artikulációs sajátosságokkal, részben pedig a fonológiai hosszúság tényével magyarázható. A spontán beszédben az első szótagi pozíció nem jár együtt feltétlenül időtartamnövekedéssel. Ez azzal is összefügg, hogy a spontán beszéd hangsúlyviszonyait feltehetően nagyobb mértékben befolyásolják a tételhangsúlyok, esetleg az érzelmi hangsúlyok, illetőleg az aktuális beszédtervezési sajátosságok.

Hipotézisünk volt, hogy napjaink spontán beszédében relatíve nagyok lesznek az időtartamértékek átfedései a kétféle fonológiai hosszúság esetén. Az adatok ezt megerősítették, ám az átfedések időtartománya nem volt független az adott magánhangzó párt artikulációs sajátosságaitól. Feltételeztük továbbá, hogy a felső nyelvállású magánhangzók időtartamviszonyai nem mutatnak jelentős különbségeket, a középső nyelvállásúaké azonban igen. A mért adatok ezt csak részben támasztották alá; a fonológiai oppozícióban álló felső nyelvállású magánhangzók időtartamértékei is szignifikáns különbséget igazoltak.

A beszédhangok időtartamának változatosságát és változékonyságát a már említettek túl befolyásolja a beszéd- és artikulációs tempó, a beszélők egyéni temporális különbségei, sőt a rejtett beszédtervezési folyamatok jellemzőinek az időviszonyokra gyakorolt hatása. Laziczius azt írta (1944), hogy a tempókülönbségek zavaró hatását a mérések nagy számával lehet ellensúlyozni. Ennek figyelembevétele mellett a kutatásunk kezdetén igyekeztünk úgy kiválasztani az adatközlőket, hogy a beszédtempójuk nagyon hasonló legyen. Az egyéni különbségek a beszédsebességtől függetlenül is érinthetik a beszéd temporális viszonyait. Hertrich és Ackermann azonban úgy vélekedik (1997), hogy a magánhangzó időtartam-realizációinak kontrollja nem egyetlen mechanizmus függvénye, hanem azt úgy kell felfognunk, mint célzottan irányított viselkedést, amely bizonyos fokig magában foglalja az egyéni stratégiákat (és ezáltal a nagy különbségek csökkenésének irányába mutat).

Ismeretes az a feltételezés, hogy a köznyelvi ejtésben jelentősek az időtartambeli ingadozások, amelyek a rövid és a hosszú magánhangzók különbségének csökkenésére utalhatnak, főként a felső nyelvállású párok esetében. A hallgatói beszédészlelés alapján ez a benyomás jogos, és számos okkal magyarázható. Ezek az okok egyrészt azokban a tényezőkben keresendők, amelyek a magánhangzók időtartamát befolyásolják, másrészt pedig azokban, amelyek a hallgató időtartam-

észlelését meghatározzák. Láttuk, hogy a hosszú magánhangzók fizikai időértékei kevésbé állandók, mint a rövideké. A hosszú magánhangzók, különösen a felső nyelvvállásuk ritka előfordulása a spontán beszédben odavezet, hogy a hallgatónak nincsenek megbízható stratégiái az időzítésük azonosítására. Továbbá a hosszú magánhangzók fizikai időtartamának nagymértékű variálódása bizonyos értelemben megzavarja a mentális reprezentációt, és ennek következtében a temporális viszonyok percepcióját. Egy-egy hosszú magánhangzó időtartamát például annak a szónak a jelentése is befolyásolhatja, amelyben előfordul. Kísérleti adatok igazolták, hogy – felolvasásban – a gyakran előforduló, illetve a jobban valószínűsíthető szavak időtartama rövidebb (Baker–Bradlow 2009). Feltehető, hogy a spontán beszédben ez még fokozottabban jelentkezik.

A magánhangzók időtartamának elkülönülése 5–6 éves gyermekek beszédében még nem stabil, de 9 éves korban a fizikai időtartamokkal jól elkülöníthetők a rövid és a hosszú magánhangzók (Bóna–Imre 2010). Ez a kísérletileg igazolt tény is arra utal, hogy a hallott mintában a fonológiai oppozíciónak tükröződnie kell a fizikai időtartamokban. Az sem hagyható figyelmen kívül, hogy a kilencévesek a rövid és a hosszú magánhangzók (és mássalhangzók) különbségének írásos megerősítésével szembesülnek tanulmányaik során.

Kutatásunkat megelőzően az adatoktól annak a megerősítését vártuk, hogy a fonológiai különbség bizonyos mértékig „semlegessé” válik a spontán beszédben. A most elemzett anyag, a tíz adatközlő objektív adatai és a statisztikai vizsgálatok azonban azt mutatják, hogy a beszélők – még ha nem is tudatosan – tendenciaszerűen ügyelnek a hosszú magánhangzók időtartamának hosszabb ejtésére.

SZAKIRODALOM

- Ackermann, Hermann – Gräber, Susanne – Hertrich, Ingo – Daum, Irene 1999. Phonemic vowel length contrasts in cerebellar disorders. *Brain and Language* 67: 95–109.
- Baker, Rachel E. – Bradlow, Ann R. 2009. Variability in word duration as a function of probability, speech style, and prosody. *Language and Speech* 52: 391–413.
- Bata Sárolta 2007. *Időszerkezeti mintázatok a spontán beszédben*. Szakdolgozat. ELTE, Budapest.
- Bell, Alan – Gregory, Michelle L. – Brenier, Jason M. – Jurafsky, Daniel – Ikeno, Ayako – Griand, Cynthia 2002. Which predictability measures affect content word durations? *Proceedings of the ISCA workshop on pronunciation modeling and lexicon adaptation for spoken language technology*. Estes Park, Colorado, 1–5.
- Boersma, Paul – David Weenink 2005. Praat: Doing phonetics by computer, <http://www.praat.org/>. [Letöltés: 2005. március 12.]
- Bolla Kálmán 1995. *Magyar fonetikai atlasz*. Nemzeti Tankönyvkiadó, Budapest.
- Bóna Judit 2007. Magánhangzó-nyújtások akusztikai-fonetikai paraméterei a spontán beszédben. *Beszédkutatás* 2007: 9–107.
- Bóna Judit – Imre Angéla 2010. A rövid/hosszú magánhangzók óvodás és kisiskolás gyermekek beszédprodukcójában. In: Navracsics Judit (szerk.): *Nyelv – beszéd – írás. Pszicholingvisztikai tanulmányok* I. Tinta Kiadó, Budapest, 49–56.
- Crystal, Thomas H. – House, Arthur S. 1988. Segmental durations in connected-speech signals: Preliminary results. *Journal of the Acoustical Society of America* 72: 705–16.
- Engstrand, Olle – Krull, Diana 1994. Durational correlates of quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of adaptive dispersion. *Phonetica* 51: 80–91.

- Furui, Sadaoki – Nakamura, Masanobu – Ichiba, Tomohisa – Iwano, Koji 2005. Analysis and recognition of spontaneous speech using Corpus of Spontaneous Japanese. *Speech Communication* 47: 208–19.
- Gandour, Jack – Dardarananda, Rochana 1992. Prosodic disturbance in aphasia: Vowel length in Thai. *Brain and Language* 23: 206–24.
- Gombocz Zoltán 1908. A magyar hangok időtartamáról. *Nyelvtudomány* 2: 93–100.
- Gombocz Zoltán – Meyer, E. A. 1909. *Zur Phonetik der ungarischen Sprache*. Edv. Berlings Buchdruckerei, Uppsala.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutató 2008*: 194–207.
- Gósy Mária 2009. Szóejtés és szóészlelés: változatosság és adaptálódás. *Beszédkutató 2009*: 46–75.
- Greenberg, Steven – Carvey, Hannah – Hitchcock, Leah – Chang, Shuangyu 2003. Temporal properties of spontaneous speech – a syllable-centric perspective. *Journal of Phonetics* 31: 465–85.
- Harmegnies, Bernard – Poch-Olivé, Dolores 1992. A study of style-induced vowel variability: Laboratory versus spontaneous speech in Spanish. *Speech Communication* 11: 429–37.
- Hegedűs Lajos 1941. Elektro-akusztikai berendezések a beszéd- és nyelvjáráskutatás szolgálatában. *Magyar Nyelv* 27: 162–85.
- Hermans, Ben 2006. Quantity. In: Brown, Keith (ed.): *Encyclopedia of language and linguistics*. Elsevier, Cambridge, UK, 308–14.
- Hertrich, Ingo – Ackermann, Hermann 1997. Articulatory control of phonological vowel length contrasts: Kinematic analysis of labial gestures. *Journal of the Acoustic Society of America* 102: 523–36.
- Hirata, Yukari 2004. Effects of speaking rate on the vowel length distinction in Japanese. *Journal of Phonetics* 32: 565–89.
- Hoogshagne, Searle 1959. Three contrastive vowel lengths in Mixe. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 12: 111–5.
- Howell, Peter – Kadi-Hanifi, Karima 1991. Comparison of prosodic properties between read and spontaneous speech material. *Speech Communication* 10: 163–9.
- Kassai Ilona 1979. *Időtartam és kvantitás a magyar nyelvben*. NyttudÉrt. 102. Akadémiai Kiadó, Budapest.
- Kassai Ilona 1982. A magyar beszéd időtartamviszonyai. In: Bolla Kálmán (szerk.): *Fejezetek a magyar leíró hangtanból*. Akadémiai Kiadó, Budapest, 115–54.
- Kassai Ilona 1991. Hol tart a magyar magánhangzók időtartamának ingadozása? Szociolingvisztikai próbainterjúk tanulságai. *Magyar Nyelvőr* 115: 69–80.
- Kelso, Scott J. A. – Tuller, Betty 1987. Intrinsic time in speech production: Theory, methodology, and preliminary observations. In: Keller, Eric – Gopnik, Myrna (eds.): *Motor and sensory processes of language*. Lawrence Erlbaum, Hillsdale, New Jersey, 137–53.
- Klatt, Dennis H. 1976. Linguistic uses of segment duration in English. Acoustic and perceptual evidence. *Journal of the Acoustic Society of America* 59: 1208–21.
- Klatt, Dennis H. – Cooper, William E. 1975. Perception of segment duration in sentence context. In: Cohen, Antonie – Nooteboom, Sieb (eds.): *Structure and process in speech perception*. Springer Verlag, Heidelberg, 135–88.
- Kovács Magdolna 2002. *Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen.
- Ladefoged, Peter 1982. *A course in phonetics*. Harcourt Brace Jovanovich, San Diego.
- Ladefoged, Peter – Maddieson, Jan 1996. *The sounds of the world's languages*. Blackwell Publishers, Oxford.
- Laver, John 1994. *Principles of phonetics*. Cambridge University Press, Cambridge.

- Lindblom, Björn – Rapp, K. 1973. Some temporal regularities of spoken Swedish. *PILUS* (Papers from the Institute of Linguistics) 21: 1–58.
- Laziczius Gyula 1944. *Fonétika*. Királyi Magyar Egyetemi Nyomda, Budapest.
- Lehiste, Ilse 1970. *Suprasegmentals*. Cambridge University Press, Cambridge, Massachusetts.
- Löfquist, Anders 1991. Proportional timing in speech motor control. *Journal of Phonetics* 19: 343–50.
- Magen, Harriet S. – Blumstein, Sheila E. 1991. Effects of speaking rate on the vowel length distinction in Korean. *Journal of the Acoustic Society of America* 89: 1918–28.
- Magdics Klára 1966. *A magyar beszédhangok időtartama*. Nyelvtudományi Közlemények 68. 125–39.
- Magdics, Klára 1969. *Studies in the acoustic characteristics of Hungarian speech sounds*. Indiana University Publications 97. Indiana University, Bloomington, Mouton, The Hague.
- Mendoza, E. – Carballo, G. – Cruz, A. – Fresneda, M. D. – Munoz, J. – Marrero, V. 2003. Temporal variability in speech segments of Spanish context and speaker related differences. *Speech Communication* 40: 431–47.
- Myers, Scott 2005. Vowel duration and neutralization of vowel length contrasts in Kinyarwanda. *Journal of Phonetics* 33: 427–46.
- Ndinga-Koumbo-Binza, Hugues Steve 2008. *Phonetic and phonological aspects of Civilian vowel duration: an experimental approach*. PhD dissertation. Stellenbosch University, Stellenbosch (Dél-Afrika).
- Nooteboom, Sieb 1999. The prosody of speech: Melody and rhythm. In: Hardcastle, William J. – Laver, John (eds.): *The handbook of phonetic sciences*. Blackwell Publishers, Oxford, 640–73.
- Olaszy Gábor 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben*. NyudÉrt. 155. Akadémiai Kiadó, Budapest.
- Perkell, Joseph S. 1999. Articulatory processes. In: Hardcastle, William J. – Laver, John (eds.): *The handbook of phonetic sciences*. Blackwell Publishers, Oxford, 333–70.
- Pylkkönen, Janne – Kurimo, Mikko 2004. Duration modeling techniques for continuous speech recognition. In: *Proceedings of the 8th International Conference on Spoken Language Processing* (Interspeech 2004), October 4–8, 2004, Jeju Island, Korea, 385–8.
- Santen van, Jan P. H. 1992. Contextual effects on vowel duration. *Speech Communication* 11: 513–46.
- Santen van, Jan – D’Imperio, Mariapaola 1999. Positional effects on stressed vowel duration in standard Italian. In: Ohala, John J. – Hasegawa, Yoko – Ohala, Manjari – Granville, Daniel – Bailey, Ashlee C. (eds.): *Proc. of the 14th International Congress of Phonetic Sciences*. International Phonetic Association, San Francisco, 241–4.
- Szende Tamás 1976. *A beszédfolyamat alaptényezői*. Akadémiai Kiadó, Budapest.
- Tóth László – Kocsor András 2005. Explicit duration modelling in HMM/ANN hybrids. *TSD*, 310–7.
- Wightman, Colin W. – Schattuck-Hufnagel, Stefanie – Ostendorf, Mari – Price, Patti J. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustic Society of America* 91: 1707–14.

Gósy Mária – Beke András
 MTA Nyelvtudományi Intézet
 – ELTE Fonetikai Tanszék

SUMMARY

*Gósy, Mária and Beke, András***Vowel durations in spontaneous Hungarian speech**

Vowel durations have been investigated by many researchers; the issue they have especially wanted to resolve is how a given factor influences the timing of a segment. Hungarian has five pairs of vowels whose short and long members do not differ (or only minimally differ) in their quality. The present paper reports on an investigation of a spontaneous speech corpus with respect to temporal relationships of high and mid Hungarian vowels phonologically constituting five short/long pairs. Differences in physical duration between short and long vowels were found to be significant and were mathematically confirmed. In spontaneous speech, overlaps between the durations of short and long vowels are nevertheless large; the exact temporal data are also influenced by vowel quality and phonetic position. Vowel durations attested in 110 minutes of recorded spontaneous speech by ten subjects suggest that speakers, albeit without being aware of the fact, tend to make sure that long vowels are pronounced with a longer duration than short ones.

Keywords: spontaneous speech, duration, phonological opposition, phonetic position, automatic classification