

A MAGYAR SZÓALAKOK HANGIDŐTARTAM-TÉRKÉPEI

Olaszy Gábor

Bevezetés

A beszélő által kiejtett hangsorok időszerkezetének kialakítását az anyanyelvi szabályok alapján automatikusan végezzük az artikulációs mozgásokkal egyidejűleg. Ismeretes, hogy a beszédhangok időtartamának kialakítását sok tényező befolyásolja mind szegmentális, mind pedig szupraszegmentális szinten. Korábban már végeztünk olyan kísérleteket, hogy meghatározzuk a beszédhangok hangkapcsolati szintű specifikus időtartamait (Olaszy 2000), és ezzel közelebb kerülünk a beszéd időszerkezetének modell szintű leírásához. A specifikus időtartamok önmagukban jól jellemzik a vizsgált nyelv hangidőtartam-szerkezetét, azonban folyamatos beszédre való alkalmazásukkor a beszéd egysíkú, monoton hangzást ad, mivel nem valósulnak meg a magasabb szinten működő hangidőtartam-módosító mechanizmusok, vagyis a szó szerkezetéből adódó, és a mondat szintjén működő úgynevezett véglegesítő tényezők. Ebben a tanulmányban olyan kísérletekről számolunk be, amelyekben továbbléptünk, és meghatároztuk a magyar szóalakok szintjén működő hangidőtartam-módosító tényezőket. Külön szóalaknak nevezünk minden olyan szót, amelynek írott formája egy vagy több beszédhangnak megfelelő betűben különbözik. Tehát különböző szóalaknak tekintjük például a *fecskét* és a *fecskék* szót. Bevezettük a „szó hangidőtartam-térképe” fogalmat (HT), amelyikkel egységesített formában tudjuk a szó szintjén működő módosító tényezőket leírni, osztályozni. A HT és a specifikus időtartamok segítségével kiszámíthatók az adott szó hangjainak időtartamai. Feltételezésünk, hogy a magyar szóalakok véges számú HT segítségével jellemezhetők, így összeállítható egy olyan HT-tár, amelyikkel bármely szó hangidőtartamai mérés nélkül meghatározhatók.

A vizsgálat módszere

A kísérletek során iránymutatásként felhasználtuk a korábbi szakirodalmi adatokat, valamint hagyományos időtartam-méréseket is végeztünk. Arra, hogy a szó szintjén már működnek olyan befolyásoló tényezők, amelyek az időszerkezetre vonatkoznak, már Collier (1992, 207) is utal: „feltételezhetjük, hogy a szó szintén releváns elemként van jelen a beszéd temporális szervezésében”. Munkánk során arra törekedtünk, hogy a kísérletek megismételhetők legyenek. Ezért főleg az analízis szintézissel eljárásra alapoztuk a vizsgálódásainkat. Ennek lényege, hogy szintetizált mintákon, percepcióis ítéletekre alapozva végeztünk hangidőtartam-módosításokat, majd ezeket osztályoztuk, csoportosítottuk és szabályrendszerben összesítettük. A vizsgálat fontos szempontja volt, hogy a szavak időszerkezetét folyamatos beszédbe ágyazva vizsgáltuk, így a kapott eredmények a folyamatos beszédre (nem önállóan kiejtett szavakra) vonatkoznak. A szintézis alapját ugyanaz a kísérleti összeállítás képezte, amelyikkel a specifikus időtartamokat meghatároztuk (Olaszy 2000). E szerint a szintetikus beszéd férfihangon szólalt meg 12-13 hang/s-os artikulációs sebességgel. A mondatokat enyhén eső intonációval szólaltattuk meg. A mondat szavait hangsúlyozás nélkül valósítottuk meg. Az analízis szintézissel módszerrel először a szerző végzett méréseket olyan egyértelműnek látszó jelenségek megerősítésére, illetve elvetésére, amelyeket a korábbi kutatásokból lehetett származtatni. A mérések alapján megfogalmazott szabályokat beépítettük a szintetizáló rendszerbe és ezek után következtek a több személlyel végzett meghallgatásos tesztek. A mondatokat többször is meghallgathatta a négy tesztelő személy (részletesen lásd Olaszy 2000). Feladatuk az volt, hogy jelöljék be a szövegben, ha szerintük valamelyik hang túl hosszú, illetve túl rövid. A szövegtörzs első része számukra 45 kijelentő mondatból állt (mondatonként 5-10 szó), második része egy A4-es oldalnyi szabadon választott szöveget tartalmazott. A hangra adott „túl hosszú”, illetve „túl rövid” megjelölést akkor fogadtuk el, ha azokat legalább hárman ugyanazon hangsorrészre egyformán bejelölték. A tesztelésből kapott jóváhagyó, illetve nem helyeslő válaszok alapján módosítottuk, bővítettük a szabályokat, egészen addig, amíg elérkeztünk egy olyan állapothoz, hogy a szövegekben már csak olyan jelölések voltak, amelye-

ket egy, illetve két személy jelölt be ugyanarra a pontra. Ezekben az esetekben a szerző döntötte el, hogy kell-e módosítani a hangidőtartamon, avagy sem. A percepciók tesztjeink befejezése és a korrekciók elvégzése után szabályokat fogalmaztunk meg, majd a szabályok alkalmazásával szintetizáltunk 15 kijelentő mondatot. A kísérletnek ezen a pontján már számszerűsítve is összevetettük a szintetizált mondatok és a természetes ejtésű ugyanolyan mondatok időszerkezetét. Ennek során vizsgáltuk mind a tendenciákat, mind pedig az egyedi időtartamokat. Az eltérésekből következtettünk a szabályok működésének helyességére, illetve gyenge pontjaira. A munka önmagába visszatérő, több lépcsős elemzési, értékelési és javítási folyamat volt. A végleges szabályokat a szerző határozta meg. A szabályok szavakra történő alkalmazásából születtek meg a HT-k, vagyis a szóra jellemző hangidőtérképek.

Az előkészítő kísérleteket 1999-2000-ben végeztük, a HT-k meghatározására és működésük ellenőrzésére 2002-2003-ban került sor.

Szabálymeghatározás a korábbi vizsgálati eredmények alapján

A korábbi kutatásokból ismert, hogy a hangidőtartamokra ható legfontosabb módosító tényezők a következők: a szó hossza, a szó helyzete a hangsorban, és a hangsúly. A szó hossza tekintetében az a tendencia, hogy minél hosszabb a szó, annál rövidebbek benne a beszédhangok (Gombocz 1909; Tarnóczy 1974). Kevesebb szótagszám nagyobb időtartamátlagot eredményez (Kassai 1979). Ezt a kiegyenlítő-dés törvényével magyarázzák, amely szerint a produkció során az a törekvés, hogy a rövidebb és hosszabb hangsorokat nagyjából azonos idő alatt ejtsük ki. Fónagy (1959) kimutatta, hogy ez a csökkenés versmondás esetén 6 hangnál hosszabb hangsorokban már nem folytatódik. Ezt igazolták későbbi általánosított kísérletei is (Fónagy–Magdics 1960), amelyekben nem versmondást, hanem normál beszédet vizsgáltak. A beszédhang hangsorbeli helye szempontjából Magdics (1965) szerint a mondat elején rövidebbek a magánhangzók, mint mondat közepén, illetőleg a végén; az utóbbi helyzetben a leghosszabbak. Ugyanez a tendencia vonatkozik a mássalhangzókra is. Magdics tehát úgy találta, hogy a mondat kiejtése során a beszédtempó fokozatosan lassul. Kassai (1979) a következő eredményre jutott: a hangok a hang-

sor belsejében a legrövidebbek, hosszabbak hangsor elején és leg-hosszabbak a hangsor végén, tehát a tendencia: csökkenő a mondat belseje felé, majd növekvő a mondat vége felé. A hangsúly befolyását Kassai a következőképpen összegezte: a rövid magánhangzók megnyúlnak hangsúlyos helyzetben, a hosszúak változatlanok maradnak. Magdics is ilyen tendenciát állapított meg, az arány nála 1:0,8. Fónagy (1958, 16) ugyanakkor azt írta: „Semmivel sem nehezebb azonban olyan eseteket vagy esetek egész csoportját idézni, amikor a hangsúly nem esik egybe a nagyobb hangossággal, a magasabb hangfekvéssel, a nagyobb időtartammal.” A hangsúly hangidőtartamra vonatkoztatott hatása tehát nem tisztázott egyértelműen. A teszteredményeink szerint a magánhangzónak csak egy része nyúlik ebben a helyzetben a nagyobb része nem, sőt rövidül a specifikus időtartamhoz képest. Ez utóbbit erősítik meg Kovács (2002) legújabb mérései is.

Elsőként tehát megismételtük Gombocz Zoltán múlt század eleji kísérletét, és ennek alapján alkottunk szabályokat a hosszú magánhangzók viselkedésére különböző hosszúságú szavakban. A második lépésben az [l] és [r] hangok nyújtó hatását egyértelműen ki tudtuk mutatni az analízis szintézissel módszer segítségével, sőt az is kiderült, hogy ez a nyújtó hatás hangfüggő. A továbbiakban figyelembe vettük Olaszy (1993) korábbi mérési eredményeit is, az ezekből alkotott szabályokat is beépítettük a rendszerbe. Ezek után méréseket végeztünk a szerző által felolvasott mondatokban az első hangsúlyos szó első magánhangzójának viselkedésével kapcsolatban. A mérések alapján Magdicsnak azt a megállapítását fogadtuk el mérvadónak, hogy a mondat elején a hangok rövidebbek, mint a közepén, illetve a végén. Ezt az elvet alkalmaztuk a több szótagú szavakra is. Mindezek után végeztük a meghallgatásos kísérleteket több személy részvételével.

Hangidőtartam-módosító szabályok a szó szintjén

Az alábbi szabályleírásokban a következő hangokra, hangcsoportokra alkottunk hangidőtartam-módosító szabályokat:

– az [a:, ɔ, o, u, y, i, e:, ø, ɛ] magánhangzók mindegyikére szó eleji, szóvégi és szó belseji helyzetre a mássalhangzó-környezet, valamint a szó hosszúságának a függvényében;

- a rövid mássalhangzókra VCV, VCCV, VCCCCV, VCCCCCV helyzetre;
- a VV, VVV kapcsolódásokra;
- a hiányzó hosszú hangokat a rövid hang időtartamának kétszeres értékre való nyújtásával jellemeztük.

A fentiek alapján tehát szinte bármilyen szöveges állomány bármely pontja vizsgálhatóvá vált. Az időtartam-módosítást egy M szorzófaktorral fejezzük ki. A hang szó szintű végleges időtartamát tehát a következő képlet segítségével számítjuk ki a specifikus időtartamból:

$$t(\text{szóban}) = t(\text{spec.}) \times M(n)$$

Az n változóval azt fejezzük ki, hogy a szorzófaktor értéke különböző lehet (ahogy a szabályokból látható). Az eredmény ms-ban adja a hangidőtartamot. A szó minden hangjára meghatározzuk a szorzófaktorot. Az M -ek számsorozata alkotja a szó hangidőtérképét (HT). Ugyanazon értékű M hatása nem feltétlenül ugyanazt a fizikai módosítást jelenti, hiszen a specifikus hangidőtartamok is megfelelő eloszlás szerint alakulnak az őket megelőző, és követő hang függvényében.

A szó eleji rövid magánhangzók időtartam-módosulásai

A szó első rövid magánhangzója időtartamának kialakítására meghatározott M szorzófaktorokat az 1. táblázatban adjuk meg. Egyedi szabályokban írtuk le a rövid magánhangzók időtartamát az [l, r] hangok előtt. A percepciók teszt eredményei kimutatták, hogy VC helyzetben az [l, r] mássalhangzók időtartam-módosító hatása magánhangzó-specifikus. Külön szabályokat fogalmaztunk meg arra az esetre is, amikor szó előtt hangsúlytalan elem van (pl. névelő). Nézzük meg, hogy a kapott adatok milyen tendenciákra világítanak rá. A magánhangzók időtartam-módosulása $M=0,8-1,4$ -es tartományban valósul meg (1. ábra). A módosulás iránya és nagysága hangfüggő. A legkevésbé módosul az [o, i, ɔ, u], és szinte kizárólag rövidülnek. Az [l, r] nyújtó hatása a legnagyobb a [ɛ]-nél, kisebb az [ø]-nél, a többieknél nincsen ilyen hatás. Azokban az esetekben, amikor a magánhangzó után [l:] hang van (például *fellázadtak*) a nyújtás elérheti az 1,4-es faktort is. A mássalhangzó-környezet tehát az [ɛ] hangra gyakorolja a legnagyobb

hatást. Összegezve a rövidüléseket és a nyúlásokat azt mondhatjuk, hogy ezek az [ɛ]-ben valósulnak meg a legszélesebb skálán.

1. táblázat: A rövid magánhangzók időtartam-módosulása a szó első szótagjában

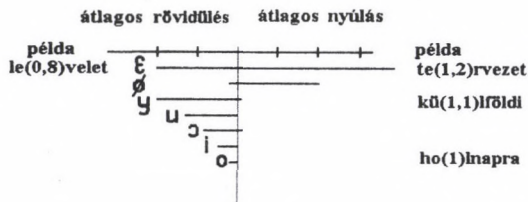
	Hangkapcsolat	Magánhangzók						
		[i]	[u]	[y]	[o]	[ɔ]	[ɛ]	[ø]
1	# C V C1	1	0,8	0,8	0,8	0,8	0,8	0,8
2	# C V C1 C	1	1	1	1	1	0,8	1
3	# C V C2	1	0,8	0,8	0,9	1	0,8	1
4	# □ C V C1	0,8	0,8	0,8	1	0,9	0,9	0,8
5	# □ C V C1 C	1	1	0,8	1	1	0,9	1
6	# □ C V C2	0,8	0,8	0,8	1	1	0,9	1
7	# V C	0,8	0,8	0,8	1	0,9	1	1
8	# V C1 C	0,8	1	1	1	1	1	1
9	# □ V C	1	1	0,8	1	0,9	1	1
10	# □ V C1 C	1	1	1	1	1	1	1,1
11	# V C2 C	1	1	0,8	1	1	1,3	1
12	# C V C2 C	1	1	1	1	1,1	1,3	1,1
13	# □ V C2 C	1	1	1	1	1	1,3	1
14	# □ C V C2 C	1	1	1,1	1	1	1,3	1
15	# V C3	1	1	0,8	1	1	1,4	1,2
16	# □ V C3	1	1	0,8	1	1	1,4	1,2
17	# C V C3	1	1	0,8	1	1	1,4	1,2
18	# □ C V C3	1	1	0,8	1	1	1,4	1,2

V = az adott magánhangzó; C = bármely mássalhangzó

C1 = bármely mássalhangzó, kivéve [r, l]; C2 = [r, l]; C3 = [l:]

□ = hangsúlytalan elem (főleg névelő) a szó előtt;

= szókezdő helyzet; szám = M szorzófaktor (például 0,8)



1. ábra

A szó eleji magánhangzók hangidőtartam-módosulásai

A szó belseji rövid magánhangzók időtartam-módosulásai

A szó belsejében a rövid magánhangzókat rövidíteni kellett, hogy a szó kiejtésében ne legyen érezhető bizonyos lomhaság, visszahúzás, erőltetett artikuláció. Ez korábbi kísérletek eredményeivel is egyezik (Olaszy 1993). A szabályok hangfüggősége itt nem mutat olyan széles skálát, mint a szó eleji magánhangzók esetében, amikor megindítjuk a szó artikulációját. A hangfüggőség csak annyit jelent, hogy az [i] hangot külön kell kezelni, mivel az önmagában is mindig rövid (a legrövidebb átlagértékkel rendelkezik a rövid magánhangzók között). A szabályokat a 2. táblázat mutatja.

2. táblázat: A rövid magánhangzók tartammódosulásai szó belsejében

	Hangkapcsolat	V1	V2	Megjegyzés
1	...CVC1...	0,9	1	
2	... CVC1C1...	1	1	
3	...CVC2C1	1	1	
4	...C3VC2C1	1,2	1,1	
5	...CVV...	1	1	
6	egyforma V-k a szóban a második szótagtól	0,9...0,8...0,8	..0,9...0,9...0,9.	ha nem CC, vagy V követi és nincs előtte hosszú magánhangzó

V1= bármely rövid magánhangzó; V2= [i];

C = bármely mássalhangzó

C1 = bármely mássalhangzó, csak nem C2 és nem C3;

C2 = [r, l]; C3 = [n, ʒ, c]

szám = M szorzófaktor (például 0,8)

A 2. táblázat szerint a CVC1C1 helyzetben lévő rövid magánhangzó időtartama nem rövidül, a specifikus értéken marad. Itt külön alkategóriát kell megkülönböztetnünk, amikor a magánhangzót megelőző mássalhangzó palatális, a magánhangzót követő CC kapcsolat első eleme pedig az [l, r] hang. Ebben az esetben az [l, r] nyújtó hatása szó belseji helyzetben is érvényesül. Például a *megnyergelte* szó magánhangzóinak időtartam-módosulási képe az 1. és 2. táblázat szabályai

alapján a következő lesz (a magánhangzók időtartam-korrigáló szorzó-faktorait zárójelben adtuk meg).

m e(0,8) g ny e(1,2) r g e(1) l t e(1)

Itt a második [ε] hang 1,2-es értékét a 2. táblázat 4. szabálya alapján írtuk elő. A CVV kapcsolatban a magánhangzó specifikus időtartama nem változik. Amennyiben a szóban ugyanazon rövid magánhangzó többször fordul elő (például: *alagutaknak, befeleledkezett*), akkor a szó második ilyen magánhangzója 0,9-es szorzófaktorral rövidül, a továbbiak 0,8-as szorzófaktorral, ha nem CC előtt állnak, illetve nem VV kapcsolatot valósítanak meg. A szó utolsó rövid magánhangzója a specifikus értékén marad, nem változik.

a(1) l a(0,9) g u t a(1) k n a(1) k

b e(0,8) l e(0,9) f e(0,8) l e(1) dk e(0,8) z e(1) tt

A hosszú magánhangzók időtartam-módosulásai

A hosszú magánhangzók fontos szerepet töltenek be a magyar beszéd időtartam-szerkezetének kialakításában. A rájuk kapott bonyolult időtartam-eloszlási térképek is ezt sugallják. A megállapított rövidítések, illetve nyújtások itt fizikailag nagyobb értéket jelentenek, mint a rövid magánhangzóknál, hiszen a hang eleve hosszabb. Például egy 160 ms-os [a:] hangnál 10%-os rövidítés férfihang esetén 2 periódus is lehet, női hangnál pedig ennek a kétszerese. A percepciós teszt eredményei azt mutatták, hogy a hosszú magánhangzók időtartam-kialakításának szabályai sokkal bonyolultabbak, mint gondoltuk, és másfajta szervezettséget mutatnak, mint amilyeneket a rövid magánhangzóknál láttunk. Az egyedüli közös vonás a hosszú és rövid magánhangzók között, hogy a hosszú magánhangzók is (főleg az [a:] és az [e:]) jobban nyúlnak [l] és [r] hangok előtt, mint más mássalhangzók előtt. A hosszú magánhangzó időtartam-módosulását a szó szótagszáma befolyásolja, amit Gombocz (1909) is megállapított. Eredményeink szerint további két esetet is figyelembe kellett venni, nevezetesen, hogy hány és milyen hosszúságú magánhangzó van a szóban és végül azt, hogy melyik szótagban helyezkednek el. E három szempont szerint összegeztük a végeredményeinket. Az eredmények alapján külön

kategóriába kellett sorolnunk az [a:] hangot és egy másikba az [e:]-t, továbbá egy harmadikba a többi hosszú magánhangzót. Ezen felül külön kategóriába kellett sorolni azokat az eseteket, amikor csak egy, illetve két hosszú magánhangzó van a szóban és egy másik továbbbiba, amikor ennél több. Ez utóbbit – eredményeink szerint – az előbbi két esetre kialakított szabályok alkalmazásával le lehetett vezetni.

Az [a:] hang időtartam-szerkezete

A tudományos folytonosság kedvéért elvégeztük Gombocz 20. század eleji kísérletét, amit már Tarnóczy (1974) is megismételt. Az [a:] hang időtartam-csökkenését a három kísérlet szerint a 3. táblázat mutatja. A három mérési eredmény tendenciájában hasonló.

3. táblázat: Az [a:] hang időtartamának csökkenése a szótagszám függvényében

Mintaszó	Gombocz	Tarnóczy	Olaszy (szintetizált minta, 2000-ben)
tát	272 ms	210 ms	217 ms
tátog	242 ms	180 ms	182 ms
tátogat	209 ms	140 ms	160 ms
tátogatók	190 ms	120 ms	145 ms
tátogatóknak	182 ms	110 ms	128 ms

Az [a:] hang időtartam-módosító faktoraira vonatkozó jelenlegi eredményeket a 4., 5. és 6. táblázatban foglaltuk össze. A 4. táblázatban azokat az időtartam-módosító szorzófaktorokat rendszereztük, amelyek az [a:] hang időtartam-módosulásait fejezik ki, ha a szóban csak egyszer szerepel, a többi magánhangzó pedig rövid. A főbb tendenciák a következők.

- Az [a:] hang időtartama az első szótagban a szóhosszúság növekedésével arányosan csökken. Négy szótagnál hosszabb szavakban már nem csökken tovább.
- A hang időtartama attól is függ, hogy hányadik szótagban található. Az előbbi csökkenő tendencia – a 4. szótagig – akkor is megtalálható,

ha az [a:] hang nem az első szótagban van, csak ekkor a rövidülés nem olyan markáns, mint az első szótag esetében.

c) Hosszú szavakban, ha az [a:] a 4. szótag utáni részben szerepel, akkor időtartama stabilan rövidebb, mint más esetekben.

d) Ha az [a:] az utolsó szótagban szerepel, akkor hossza fokozatosan csökken a szó hosszúságának a függvényében.

e) Az [l, r] mássalhangzók nyújtó hatással vannak az [a:] hang időtartamára minden helyzetben. Ez a nyújtó hatás csökken, ha a szó hosszabb.

4. táblázat: Az [a:] hang időtartam-módosulása, ha egy [a:] hang van a szóban és nincs benne több hosszú magánhangzó

Az [a:] helyzete a szóban	Környezet	A szó szótagszáma					
		1	2	3	4	5	6
1. szótag	VC1	–	1	0,9	0,85	0,8	0,8
	VC2	–	1,3	1,2	1,2	1,1	1
2. szótag	VC1	–	–	0,9	0,85	0,85	0,8
	VC2	–	–	1	1	1	1
3. szótag	VC1	–	–	–	0,9	0,8	0,8
	VC2	–	–	–	1	1	1
többi szótag	VC1	–	–	–	–	0,8	0,8
	VC2	–	–	–	–	1	1
utolsó szótag a szóban	VC1	1,2	1	0,95	0,95	0,95	0,9
	VC2	1,3	1,3	1,3	1,2	1,1	1,1

V = az [a:] hang

C1 = bármely mássalhangzó, csak nem [r, l] ; C2 = [r, l]

sám = M szorzófaktor (például 0,8)

Tehát az [a:] hang időtartama mind az első szótagban, mind a további háromban a szó hosszúságának függvényében folyamatosan csökken. Fontos megállapítás, hogy az utolsó szótagban is markáns időtartam-csökkenés van jelen a szóhosszúság függvényében. Ez az

artikulációs ökonómiával magyarázható. Ha az [a:] a szó elején van és hosszú a szó, akkor ezért rövidítjük meg a hangot, ha a szó utolsó szótagjában van, akkor pedig egyrésztől azért, mert az utolsó szótagra már nincs energia, másrésztől, mert ez a szótag már nem hordoz lényeges információt, egy befejező szakaszt képvisel. Ezt két példán illusztráljuk. A *láthatatlan* szóban a magánhangzók időtartam-korrigáló szorzófaktorait zárójelben adtuk meg. Az [a:] szorzófaktorát (0,85) a 4. táblázat 4 szótagos oszlopából vettük. A szó magánhangzóinak M faktorai azt mutatják, hogy az elején a tempó gyorsabb, majd lassul:

l á(0,85) t h a(0,9) t a(1) t l a(1) n

A *felebarát* szóban az utolsó szótagbeli [a:] szorzófaktorát a 4. táblázat utolsó szótag sorából vettük:

f e(1) l e(0,9) b a(0,9) r á(0,95) t

Az [a:] hang időtartama másképpen alakul, ha több hosszú magánhangzó van a szóban, és ezek közül az [a:] az első (5. táblázat). Ha az [a:] az első szótagban van, akkor időtartama hasonló képet mutat, mint a 4. táblázatban, ha a második szótagban van, akkor időtartama lényegesen lerövidül (5. táblázat). Az [l, r] mássalhangzók nemcsak az első, de a második [a:] hangban is érvényesítik nyújtó hatásukat. Ha a második [a:] az utolsó szótagban szerepel, akkor a rövidülése kissé erősebb, mint, amikor csak egy [a:] van a szóban. Ez valószínűleg annak tudható be, hogy az első [a:] -ra fordított artikulációs energia felemészti a másodikra fordítandó energia egy részét, így kevesebb marad a másodikra. Abban az esetben, ha több [a:] hang van a szóban akkor az 5. táblázat szabályait kell használni az első két [a:] időtartamának megállapítására, a további [a:] hangokra pedig azokat a szabályokat kell értelemszerűen alkalmazni, amelyek a második [a:] -ra voltak előírva. Vegyük például a *báránkjánál* szó magánhangzóinak időtartam-képét. Az 5. táblázatból az 5 szótagú szóra vonatkozó oszlopot kell kiválasztani. Az első [a:] -ra, mivel az [r] hang előtt van 1,1-es nyújtó szorzófaktorot kell alkalmazni. A másodikra, mivel az a második [a:] hang és a második szótagban található a 0,85-ös időtartam-csökkentő szorzófaktorot kell alkalmazni.

5. táblázat: Az [a:] hang időtartam-módosulása, ha több hosszú magánhangzó van a szóban és az [a:] az első

Az [a:] helyzete a szóban	Környezet	A szó szótagszáma					
		1	2	3	4	5	6
1. szótag	VC1	–	1	1	0,9	0,8	0,8
	VC2	–	1,1	1,1	1,1	1,1	1
2. szótag	VC1	–	–	0,95	0,9	0,9	0,85
	VC2	–	–	1	1	1	1
többi szótag	VC1	–	–	–	0,85	0,85	0,85
	VC2	–	–	–	1	1	1

V = az [a:] hang bármelyike

C1 = bármely mássalhangzó, csak nem [r, l]; C2 = [r, l]

szám = M szorzófaktor (például 0,85)

A harmadikra, negyedikre, mivel azok nem az első két szótagban szerepelnek, az 5. táblázat harmadik sorát kell alkalmazni, itt a szorzófaktor 0,85 lesz. Az ötödik [a:] az utolsó szótagban szerepel és [l] hang előtti, tehát 1-es szorzófaktorot fog kapni. A szó hosszú magánhangzóinak M faktoraival a kép a következő lesz:

b á(1,1) r á(0,9) ny k á(0,85) j á(0,85) n á(1) l

Az [a:] hang időtartama, másképpen alakul, ha több hosszú magánhangzó van a szóban és ezek közül az [a:] nem az első. Ekkor általánosságban 0,85-ös szorzófaktorot kap minden esetben, ha pedig [l, r] előtt van, akkor 0,95-öst.

Az [e:] hang időtartam-szerkezete

Az [e:] hangra vonatkozó hasonló mérési eredmények fő tendenciájukban hasonlóak, mint az [a:]-nál. Az [e:] hang időtartama is csökken a szóhosszúság növekedésével, azonban nem olyan markánsan, mint az [a:]-é. Ha az [e:] az utolsó szótagban szerepel, akkor hossza szintén kissé csökken a szó hosszúságának a függvényében. Hosszú szavakban az [e:] hang a 4. szótagtól számított szótagokban stabilan rövid, hasonlóan az [a:]-hoz. Az [l, r] mássalhangzók nyújtó hatással vannak

az [e:] időtartamára is. Ez minden helyzetben érvényes, foka azonban enyhébb, mint az [a:] hangnál.

A rövid magánhangzók hosszú párjainak időtartam-szerkezete

A rövid magánhangzók hosszú párjainak időtartam-módosulásait külön kategóriaként kezeltük. Az eredmények a következők. Az [i:, u:, y:, o:, ø:] hangok hasonlóan viselkednek, mint az [a:] és [e:] hangok, csak a nyúlások, illetve rövidülések nem olyan markánsak. Például a *kihívásáról* szó magánhangzóira vonatkozó M faktorok képe a következő lesz.

k i(1) h í(0,9) v á(0,85) s á(0,85) r ó(0,85) l

A mássalhangzók időtartam-módosulásai

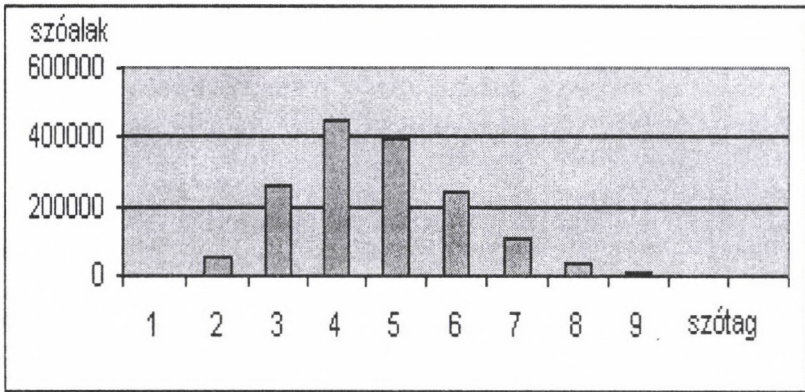
#CV, VCV, illetve VC# helyzetben a mássalhangzók nem módosulnak, a specifikus időtartamaikkal valósulnak meg. A VCCV helyzetű hangkapcsolatokban a mássalhangzók rövidítő szorzófaktort kapnak, a redukciós hatás eredményeképpen. Általánosságban 0,9-es faktorral szorozzuk meg a CC kapcsolat mindkét tagját. Kivételek: a rövidülés nem vonatkozik az [r] hangra, az [nk], [ng], [ln] és [lm] hangkapcsolatokra, ezekben az esetekben a specifikus időtartamot kell megtartani.

A hármas mássalhangzó-torlódások hangjai erőteljesebben rövidítendőek, mint a CC-k. Itt az általános szabályt a következőképpen határoztuk meg: 0,6-0,6-0,8-as szorzófaktor-sorozatot alkalmazunk a C1-C2-C3 hangokra a helyes időtartamok beállítására. Összesen 312 különböző CCC kapcsolatot vizsgáltunk meg, és az egyedi vizsgálatok eredményeképpen számos kivételt kellett tennünk az alapszabály alól. Néhány általános szabályt adunk itt meg a kivételekre. Amennyiben C1=[l, r, n, j], akkor azt nem rövidítjük, ha C1=[r] és C2=réshang, akkor egyiket sem rövidítjük. Ha két zárhang találkozik bármelyik kombinációban, akkor az elsőt nem rövidítjük.

A négyes mássalhangzó-torlódások ritkák a magyarban. Összesen 86-féle CCCC kapcsolódást vizsgáltunk meg, ezeknél az általános rövidítési szabály a következő: minden hangot 0,6-os szorzófaktorral rövidítünk. A kivételek hasonlóak, mint a CCC kapcsolatoknál.

A magyar szóalakok statisztikai eloszlásai

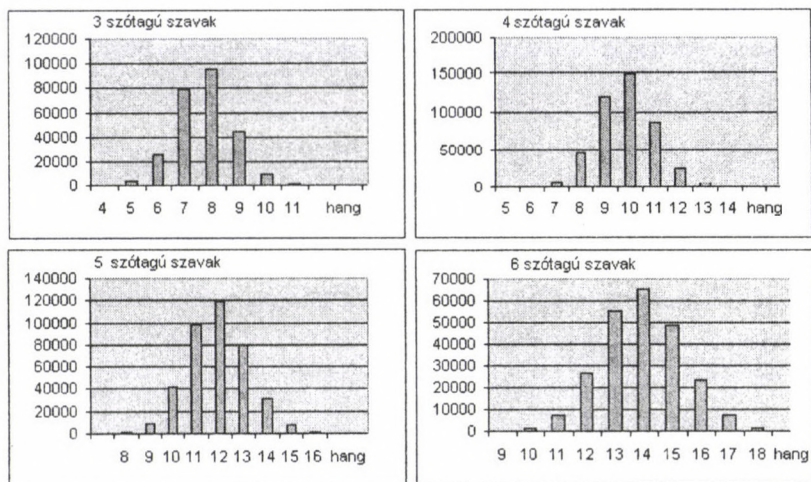
A célkitűzésünk a továbbiakban az volt, hogy a fenti szabályokkal megadjuk minden magyar szóalak hangjainak az M szorzófaktor-sorozatát, vagyis a szóra jellemző HT-t. A kiinduló adatállomány egy 50 millió szavas szövegtörzs volt (Németh–Zainkó 2002), ebből válogattuk ki a különböző szóalakokat, melynek eredményeképpen 1,5 millió elemű állományt kaptunk. Ebben az állományban tehát minden szó legalább egy hangnak megfelelő betűben különbözött egymástól. A célkitűzés eléréséhez a szavakat szótagszámuk, majd hangszámuk szerinti adathalmazokra kellett bontani. Első lépésben a szótagszámok szerint vizsgáltuk meg a korpusz szóanyagát. A kapott eloszlás képe azt mutatja, hogy a magyarban a 4 és 5 szótagú szavakból van a legtöbb, majd a 3 és 6 szótagúak következnek (2. ábra).



2. ábra

Magyar szóalakok szótagszám szerinti eloszlása az 1,5 millió szóalokot tartalmazó vizsgált korpuszban

Ahhoz, hogy a HT-eket meghatározhassuk minden szóalakra, tovább kellett bontani a szavakat a hangok szintjéig. Ennek eredményéből mutat be részleteket a 3. ábra.



3. ábra
Szóalakok hangsúly szerinti eloszlása

A szóalakok és hangidőtérképek

A több lépcsős felbontás után rendelkezésünkre állt az a szóállomány, amelyen a HT-k meghatározását elvégezhetjük. Erre speciális szoftvereket fejlesztettünk ki. Ennek lényege az volt, hogy a fenti időtartam-módosító szabályokat alkalmazva minden szóalakra meghatároztuk azokat az M szorzófaktorokat, amelyekkel kialakultak a szó hangjainak végleges időtartamai. Így a korpuszunkban szereplő 1,5 millió szó mindegyikére megkaptuk a HT-nek az adott szóra jellemző számsorozatát, tehát 1,5 millió HT-t. Ezek között természetesen vannak egyforma HT-k is, ezek csoportjait és a hozzájuk tartozó szavakat szeretnénk a vizsgálat célja szerint csoportokba szedni. Példaként bemutatunk néhány HT-t és a hozzájuk tartozó szavakból néhányat az 5. táblázatban.

A kutatásnak ezen a pontján jutottunk el ahhoz az állapothoz, hogy megvizsgálhattuk a hipotézist, mely szerint a szóalakok számánál lényegesen kisebb számú HT-vel leírhatók a magyar szóalakok hangjainak időtartamai.

5. táblázat: Különböző HT-k és a hozzájuk tartozó mintaszavak

HT és mintaszavak 10 hangú 3 szótagos szóra	HT és mintaszavak 8 hangú 4 szótagos szóra	HT és mintaszavak 13 hangú 5 szótagos szóra
1-1-0,9-0,9-1-0,6-0,6-0,8-1-1	1-1-0,8-1-0,9-1-1-1	1-0,8-1-1-0,9-0,9-0,95-1-1-0,9-0,9-0,9-1
célkönyvtár fagylaltban gázmaszkban háztömbhöz háztömbnek kontextus köznyelvben központban száznolcszor száznolcvan	adakozik alagutak amelyeken anyajuhot ecsetemet egyetemes idehozod igyekezel odajöhet ugyanazok	hazugságokból nehezményeznék
HT és mintaszavak 9 hangú 3 szótagos szóra	HT és mintaszavak 10 hangú 4 szótagos szóra	HT és mintaszavak 11 hangú 6 szótagos szóra
1-1-1-1-0,9-0,9-1,1-0,9-0,9	1-0,8-0,9-0,9-0,9-1-0,8-2-1-1	1-1-1-0,9-1-0,9-1-0,9-1-1-1
garantált gyerektánc harapdárd kicincált kiszolgálász kivizsgált leromlást maradványt sikoltást	becstelennek deszkalappal helynevekkel kedvesebben meglepetten megszerezzük rejtegettük szemgödörrel	aláírásokat automatákon kirohanásai kiszámítania várományosai

Szortírozó algoritmust készítettünk a HT-k összehasonlítására, amellyel az egyforma szótagszámú és hangszámú szóalakok HT-it gyűjtöttük csoportokba. Ez más szóval azt jelentette, hogy megkerestük mindazon szavakat, amelyek ugyanazzal a HT számsorozattal rendelkeztek. Ennek eredményéből mutat adatokat a 6. táblázat. E szerint tehát 1 332 936 szóalak időszerkezetének leírása 314 951-féle Ht-vel megoldható. Ez a szóállomány összegére vetítve 23,6%-ot jelent. Az adatok tehát igazolják a hipotézis gondolatát. Mint látható, az egyszer előforduló HT-k száma meglehetősen magas. Amennyiben megvizsgáljuk, hogy az egyszer előforduló HT-k miben különböznek a hoz-

zájuk legközelebb állótól, és azt is megnézzük, hogy besorolható-e abba, ha a különbözőséget adó hang M faktorát megfelelően módosítjuk, akkor ez a szám valószínűleg tovább csökkenthető. Ezt a módosítás utáni percepció vizsgálatával el lehet dönteni.

6. táblázat: A leggyakoribb szóalakok HT-inek száma

Szó	3 szótagú	4 szótagú	5 szótagú	6 szótagú
szószám a korpuszban	260 204	443 524	391 535	237 673
HT-k száma	12 148	58 190	122 064	122 549
egyszer előforduló HT-k száma	4749	28 045	65 990	65 942

A HT-k működésének percepció tesztelése

A szóidőterképek minőségét ellenőrző teszt összeállításakor a problémát – a HT-k nagy száma miatt – a tesztelésre kijelölt szavak számának meghatározása és a stimulusok kiválasztása jelentette. Úgy döntöttünk, hogy vizsgálatunkat a legtöbb szóalakot tartalmazó 4 szótagú korpuszon végezzük el. Véletlenszerűen választottuk ki a percepció teszt stimulusait az összes 4 szótagú szó (443 452 darab) halma-zából, és így az 58 190-féle időterképből is. Ez a módszer biztosította, hogy a percepció teszt anyagában az időterképek változatosságát megfelelően reprezentáljuk. A véletlen mintavételezésnek köszönhetően a stimulusok eloszlása hasonló volt az eredeti korpuszéhoz. A tesztben 40 személy vett részt három csoportban (20, 10, 10 fő). Az adatközlők egyetemi hallgatók voltak (átlagéletkoruk 22,3 év). Minden kísérleti személy 200 különböző szót hallott és értékelt. A tesztelés számítógépen történt, az adatközlők fejhallgatón hallották a szavakat és a következő instrukciót kapták:

„Döntse el, hogy elég természetesnek találja-e a szó ritmusát! Minősítse a hangsort az alábbi 3 osztályzat egyikével:

3-as osztályzat = a szó ritmusa elég természetes;

2-es osztályzat = a szó ritmusa furcsa, de nem nagyon zavaró;

1-es osztályzat = a szó ritmusa rossz, magyartalan.

A szavak között előfordulhatnak nagyon ritka vagy esetleg nem is létező alakok. Ha ilyennel találkozik, akkor ezt a 'Megjegyzések' címszó alatt szereplő a 'A szó nem érthető vagy nem magyar' felírra kattintva jelezze. Az osztályzat mindig csak a hangok időtartamára, a szó ritmusára vonatkozzék!"

A kísérleti személyek egy-egy szót kétszer hallgathattak meg, és módjuk volt jelezni, ha magát a szót találták furcsának. Ezen opció beépítését azért találtuk szükségesnek, hogy a hangsor esetleges idegensége a lehető legkevésbé befolyásolja az időviszonyok megítélését.

Mivel nehéz volt előre megjósolni, hogy mennyire fognak szórni az ítéletek, a lehallgatandó szómennyiséget dinamikusan határoztuk meg. Az alsó határt 4000 szóban állapítottuk meg. Ezt hallgatta meg a 20 személy. A következő lépésben 10 fő hallgatta meg a 200 szót, majd az utolsó 10 fő szintén. Ezzel a módszerrel nyomon kívántuk követni, hogy mennyire módosul az ítéletek átlaga, ha növeljük a meghallgatandó szavak számát. Mivel az átlagok a 4000, 6000 és 8000 válaszból számítva nem módosultak számottevően (7. táblázat), a lehallgatandó szavak számát nem növeltük tovább. Ennek megfelelően a továbbiakban szereplő számadatok és eredmények 8000 válaszra vonatkoznak.

7. táblázat: A lehallgatási teszt eredményei

Az ítéletek száma	Az ítéletek átlaga a kijelölt 3-as skálán
4000	2,594
4000+2000	2,5895
4000+2000+2000	2,608125

A tesztelés szempontjából fontosnak tartottuk, hogy a meghallgatandó szavak hossza (a hangok száma a szóban) ugyanolyan eloszlást mutasson, mint amilyen az egész 4 szótagos korpuszé (3. ábra). A 8. táblázat a hangszámokénti megoszlást mutatja. E szerint a tesztanyagban is a 10 hangból álló 4 szótagú szavakból volt a legtöbb. A percepciós teszt anyag 4928 különböző időtérképet ölelt fel.

8. táblázat: 4 szótagú szavak hangszámának eloszlása a tesztanyagban

Hangok száma a szóban	A szavak száma a teljes korpuszban (%)	A szavak száma a percepciós teszt anyagában (%)
6	0,14996	0,075
7	1,852963	1,8375
8	10,24981	10,3125
9	27,45957	27,3375
10	33,49652	32,7875
11	19,46502	20,3
12	5,96299	6,05
13	1,165402	1,075
14	0,173638	0,2
15	0,018717	0,025

Eredmények

A teszt eredményei azt mutatják, hogy a kísérleti személyek nagy arányban (68,4 %) elfogadták, megfelelően természetesnek ítélték a szavak ritmusát (9. táblázat), azaz a hangidőtérképekkel előállított szavak hangzása megfelel a magyar köznyelvi időtartam-struktúráknak.

9. táblázat: A percepciós teszt értékelése

Megítélés	Százalékos arány	Kumulatív százalékos arány
1-es osztályzat	7,9	7,9
2-es osztályzat	23,7	31,6
3-as osztályzat	68,4	100

A megjegyzéseket vizsgálva azt találtuk, hogy az idegennek tűnő szóalakok között voltak valóban furcsa szóösszetételek, például *asszonynéni*, *barhentszoknya*, *békepapi*, *bozótfutók*; ritka szavak például *ámbitusán*, *famulusnak*; idegen szavak és nevek, például *csehovian*, *beatlesért*, *bestselleri* stb. Összesen 604 ilyen jelölés volt a teszt végén, ami a teljes válaszok 7,5%-a.

Összegzés

A fent ismertetett kutatással egyrészt igazoltuk a hipotézist, másrészt elértük azt, hogy minden magyar szóalakra meg tudjuk adni a szót alkotó hangok várható időtartamát a HT és a specifikus időtartam segítségével, adott artikulációs sebességre. Ezzel közelebb kerültünk a magyar beszéd időszerkezeti képének teljes körű modellezéséhez. A fenti módszer előnyei: személyfüggetlen, bármikor megismételhető, hangzó formában is ellenőrizhető, minden hangra ad adatot minden hangkörnyezetben, különböző célzatú statisztikai vizsgálatokra is alkalmas. A hangidőtartam-módosító szabályok összességének vizsgálatából az látszik, hogy a szó szintjén a legtöbb esetben rövidíteni kell a specifikus időtartamon, a nyújtásra csak az [l, r] hangok előtt kerül sor. További részeredmény, hogy a szó hangidőtartam-szerkezetének kialakításában olyan tényezők is lényeges szerepet játszanak, hogy milyen hangok vannak a szóban, és azok hogyan csoportosulnak a szón belül. A kapott adatok számos más kutatás számára jelenthetnek jó kiindulási alapot (például: a magyar mint idegen nyelv tanítása, beszéd-sérültek oktatása, időmodell-tervezés, fonetikai kutatások). A továbbiakban tervezzük az egyszer előforduló HT-k szerkezeti vizsgálatát, valamint a percepció tesztek kiterjesztését más szótagszámú szócsoporthoz is.

Köszönetnyilvánítás

Köszönetemet fejezem ki mindazoknak, akik e kutatást munkájukkal segítették. Elsősorban a BME Távközlési és Médiainformaticai Tanszékén dolgozó Zainkó Csabának és Kiss Géának, akik a munka statisztikai és beszédtechnológiai programozási munkáit végezték, továbbá Kovács Magdolnának a Debreceni Egyetem Általános és Alkalmazott Nyelvészeti Tanszék munkatársának, aki a HT-k percepció teszzeit szervezte, végezte, adatolta. Külön köszönet jár Abari Kálmánnak, aki általános programozói munkájával segítette a kutatást és létrehozta az internetes lekérdező rendszert. Ezt a kutatást az OTKA 037288 támogatta.

Irodalom

- Collier, R. (1992): A comment on the prediction of prosody. In: *Talking Machines: Theories, Models and Designs*. Eds.: Bailly, G. – Benoit, C. – Sawallis, T. R. Elsevier Science Publishers, 205-208.

- Gombocz Zoltán (1909): A magyar beszédhangok időtartamáról. Nyelvtudomány II. Budapest, 93-100.
- Kassai Ilona (1979): Időtartam és kvantitás a magyar nyelvben. NytudÉrt. 102. Akadémiai Kiadó, Budapest.
- Fónagy Iván (1958): A hangsúlyról. NytudÉrt. 18. Akadémiai Kiadó, Budapest.
- Fónagy Iván (1959): A költői nyelv hangtanából. Budapest.
- Fónagy Iván – Magdics Klára (1960): Beszédsebesség, szólam ritmusérzék. MNy 56: 450-458.
- Kovács Magdolna (2002): Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában. Debreceni Egyetem, Kossuth Egyetemi Kiadó. Debrecen.
- Magdics Klára (1966): A magyar beszédhangok időtartama. NyK. 68: 125-139.
- Németh G. – Zainkó Cs. (2002): Multilingual statistical text analysis, zipf's law and Hungarian speech generation. Acta Linguistica Hungarica. 49/3-4: 385-405.
- Olaszy Gábor (1993): Hangidőtartamok számítógépes elemzése a beszéd ritmikai szerkezetének vizsgálatához. In: Beszédkutatás'93. Szerk.: Gósy Mária – Siptár Péter. MTA Nyelvtudományi Intézet. Budapest, 116-127.
- Olaszy Gábor (2000): Kísérlet a magyar beszédhangok specifikus időtartamainak meghatározására folyamatos beszédre. In: Beszédkutatás 2000. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 26-38.
- Tarnóczy Tamás (1974): A magánhangzók vizsgálatának akusztikai problémái. ÁNyT. X: 181-196.