

FIATAL KUTATÓI PÁLYÁZAT
MTA Nyelvtudományi Intézet

Kutatási terv
Neuberger Tilda

Zármássalhangzók temporális jellemzői:
fonológiai különbségek realizálódása

1. A kutatás célja és tárgya

A kutatás célja egy mássalhangzóosztály, a felpattanó zárhangok (explozívák) temporális sajátosságainak megismerése és rendszerszerű leírása. A beszédhangok ejtésekor az egyes artikulációs mozgások meghatározott időtartamban történnek, az artikulációs sajátosságok pedig tükröződnek az akusztikai vetületben. A felpattanó zárhangok képzésekor a levegőáram akadályba ütközik, amelyet nem tud áttörni, a feltorlódott levegő nyomása azonban felpattintja a zárat. Ebből adódik az explozívák összetett szerkezete: az akusztikai szerkezetben megváltoznak a paraméterek, az egyes részek időtartama külön-külön is meghatározható, így beszélhetünk a zárszakasz hosszáról, a felpattanás(ok) pillanatáról, továbbá a követő magánhangzóhoz vezető átmeneti szakasz időtartamáról, vagyis a zöngeskedési időről (VOT).

Ahogy a magyar mássalhangzókra általában jellemző, a zármássalhangzók is fonológiai oppozícióban állnak időtartamuk szerint, bár a mássalhangzók esetében az időtartam disztinktív szerepe kevésbé jelentős, mint a magánhangzórendszerben (vö. Siptár 1995). A korábbi, sporadikus fonetikai mérési tapasztalatok szerint a nyelvileg rövid és hosszú mássalhangzók objektív időértékei átfedéseket mutathatnak.

A zármássalhangzók időtartamának elemzési célja a fonológiai különbségek realizálódásának adatolása spontán beszédben. A jelen kutatás választ keres arra, hogy a nyelvi időtartam-különbség (hosszú–rövid) alátámasztható-e a felnőttek kiejtésében mért értékekkel, valamint hogy az artikuláció akusztikai vetületében hol és hogyan jelentkezik a fonológiai oppozíció megvalósulása: zárszakasz, zörejszakasz, zöngeskedési idő, lecsengés. Milyen mértékben van erre hatással az adott kontextus, illetve igazolható-e a spontán beszéd egyéb sajátosságainak befolyása az objektív tartamokra? Az elemzés tekintetbe veszi a zármássalhangzók időzítési viszonyait a beszédtervezési diszharmónia egyes jelenségeiben is, így teljes kép alakul ki a spontán beszéd artikulációs szerveződéséről a vizsgált mássalhangzók tekintetében.

2. A kutatás jelentősége

A kutatás eredményei részét képezhetik egy átfogó, rendszerezett hangidőtartam-vizsgálatnak, amely időszerű lenne, hiszen az utóbbi évtizedekben a beszélt nyelvben egyfajta gyorsulási tendencia figyelhető meg, melynek következtében a beszédhangok időviszonyaiban is változás tapasztalható (Kassai 1993, Adamikné 2000, Bóna 2009). A kapott adatok hozzájárulnak a mai magyar nyelv artikulációs és akusztikai sajátosságainak leírásához, továbbá felhasználhatók a foniátriai, logopédiai gyakorlatban vagy az afáziakutatásban. A leírás szükségességét beszédtechnológiai igények is sürgetik, hiszen a mérési eredmények alkalmazhatók a mesterséges beszéd-előállításban a mind természetesebb hangzás elérése érdekében (Collier et al. 1992, Olaszky 1993), valamint a beszédfelismerési algoritmusok létrehozásában.

3. A kutatás előzményei

A beszéd időszerkezetének vizsgálata az akusztikai fonetika egyik legkutatottabb területe, a beszédhangok időtartamával kapcsolatos kutatások nagy múltra tekintenek vissza mind a nemzetközi (pl. Meyer 1903, Metz 1914, Navarro Tomas 1916, Malmberg 1944, Umeda 1977, O'Shaughnessy 1981, Cho–Ladefoged 1999, van Alphen–Smith 2004, Beckman et al. 2011), mind a magyar szakirodalomban (pl. Gombocz 1909, Hegedűs 1941, Laziczius 1963, Magdics 1966, Tarnóczy 1965, Szende 1976, Kassai 1979, Bolla 1982, Kovács 1998, Olasz 2006, Gósy–Beke 2010). Az eszközfonetika megjelenésével már a múlt század elejétől kezdve lehetőség nyílt objektív adatokra épülő kvantitatív elemzésekre.

A szakirodalom a múlt század ötvenes éveinek végétől nagy teret szentel a zárhangok időszerkezeti vizsgálatának. Az eddigi eredményekből már sokat tudunk az explozívák szerkezeti felépítéséről: számos kutatás született a zöngelkedési időről (VOT), tudjuk például, hogy különbség mutatható ki az izolált szavakban és a spontán beszédben mért VOT-értékek között, valamint hogy a magyarban a mássalhangzók képzési helye hatással van a zöngelkedési idő értékére (Gósy–Ringen 2009). A fonológiai oppozíció szerint a rövid és hosszú párok zörejeleme nem mutat szignifikáns időtartam-változást, a zárszakasz jelentős megnyúlása számít a hosszúság legfontosabb jelölőjének (Kovács 2001). Ismeretes, hogy a mássalhangzók időtartamát befolyásolja a képzési hely, a fonetikai pozíció, valamint a közvetlen beszédhangkontextus, például a zárhangot követő magánhangzó minősége, időtartama hatással lehet a zöngelkedési idő értékére (Kassai 1979, Gósy 2000). Ugyanígy a zöngésségi oppozíció is befolyásoló tényező mind a beszédhang, mind a VOT tekintetében (Grácsi 2011).

A korábbi, kvantitás témakörében végzett kutatások nagyrészt olvasott vagy utánmondott beszédben, izolált szavak vagy mondatok ejtése alapján történtek. Ennek oka, hogy az előre meghatározott anyagban az adott vizsgálathoz szükséges specifikus adathalmaz létrehozására van lehetőség, ezzel szemben a spontán beszédben megjelenő (ugyanazon paraméterekkel rendelkező) jelenségek előfordulása esetleges, vagyis problémát okoz az előfordulási gyakoriság kiegyenlítetlensége. A megfelelő mennyiségű, statisztikailag elemezhető kutatási anyaghoz kiterjedt korpusz, nagy adatbázis szükséges.

4. A kutatás részterületei

4.1. A jelen kutatás anyagát a BEA adatbázis felvételeinek spontán és felolvasott részei képezik. A BEA modern technológián alapuló, a minőségi hangrögzítés kritériumait teljesítő adatbázis, amely nagy mennyiségű (spontán) beszédet tartalmaz, így különböző fonetikai vizsgálatok alapjául szolgálhat (Gósy 2006). Hangfelvételei alapján már számos fonetikai, pszicholingvisztikai és beszédtechnológiai tárgyú tanulmány született, amelyekhez a hanganyagok (kétféle típusú) átírata jelentős könnyebbséget nyújt a kutatók számára mind az ellenőrzésben, mind a (vissza)keresésben (Gyarmathy–Neuberger 2011). A vizsgálathoz 30 egynyelvű, budapesti beszélő (körülbelül 15 órányi) anyagát használjuk fel.

A vizsgált mássalhangzók a következők: [b, p, d, t, ʃ, c, g, k], valamint a nyelvileg hosszú párjuk. A [ʃ, c] hangok képzési módjával kapcsolatban megoszlanak a vélemények a szakirodalomban, a fonetikusok egy része palatális zárhangnak, másik részük palatális affrikátának tartja őket (vö. Siptár 1995). A jelen kutatásban az előbbi csoporthoz soroljuk őket. A tervezett vizsgálat alapkérdése, hogy milyen képet mutatnak az egyes zárhangok időtartam-értékei a felnőtt beszélők ejtésében spontán beszédben. További kutatási kérdések, hogy (i) milyen mértékben különülnek el a fonológiai oppozícióban álló explozívák az időtartam-paraméter mentén, illetve hogy (ii) egy-egy tényező miként befolyásolja az adott szegmentum időzítését. A temporális jellemzők ugyanis egyfelől függenek a mássalhangzó

akusztikai minőségétől, a hangsorban elfoglalt helyétől (szókezdő, szó belseji, szóvégi), a hangsor hosszától (hang-, illetve szótagszám), a kontextustól (CV, VC, VCV stb.), valamint a szupraszegmentumoktól (például artikulációs és beszédtempó), így az időtartamok meghatározása során ezeket a tényezőket is szem előtt tartom. A spontán és a felolvasott korpusz párhuzamos elemzése a két beszédhelyzetben produkált zárhangok összevetését teszi lehetővé.

A beszédsegmentumok időtartamainak kinyeréséhez a szövegeket három szinten annotálom a Praat (Boersma–Weenink 2010) szoftver segítségével: az első szinten a beszédszakaszokat, a másodikon a vizsgált beszédhangokat, a harmadikon pedig az összetett akusztikai szerkezetű mássalhangzók egyes részeit (zárszakasz, feloldás szakasza, VOT) címkézem fel. Az annotálás az első szakaszban automatikus módszerrel (speciális beszédfelismerő szoftver segítségével, MAUS) fog történni, majd manuálisan – vizuális és auditív ellenőrzés során – megtörténik a korrekció a mind pontosabb hanghatárok kijelölése érdekében. Az így annotált beszédanyag lehetővé teszi a vizsgált szegmensek időtartamértékeinek kinyerését. Az adatok statisztikai elemzéséhez az SPSS szoftvert használom.

Spontán beszédben a hanghatárok kijelölése több nehézséggel jár(hat) (vö. van Santen 1992, Olaszky 2006, Machač–Skarnitzl 2009, Gráczy 2012), azonban ez a hangidőtartamok meghatározásának alapja, így a szegmentálás módszerének kidolgozására nagy hangsúlyt kell fektetnünk. A természetes artikulációban ugyanis a szomszédos beszédhangok valamiféle átmenettel kapcsolódnak egymáshoz, és különféleképpen hatnak egymásra. A fonetikai koartikuláció során létrejött hangátmenetek az egyik hangból a másikba tartó artikulációs mozgások akusztikai vetületei, amelyek az adott beszédhangok észlelését és azonosítását segítik a percepció szempontjából, az egymástól való éles elkülönítésüket viszont nehezítik (Hardcastle–Hewlett 1999).

4.2. Fonológiai koartikulációnak azokat a folyamatokat tekintjük, amelyek során a beszédhangok úgy hatnak egymásra, hogy a módosult artikuláció egy másik fonéma alaprealizációját eredményezi (Gósy 2004). A fonológiai koartikuláció a beszédtervezés szintjén érvényesül, az artikulációs kivitelezése csak következmény: a beszélő a fonológiai szint aktiválása során hívja le a fonológiai szabályokat az adott lexikai egységgel együtt a mentális lexikonból (vö. Levelt 1989). A koartikuláció során a felszínen létrejöhetnek hosszú mássalhangzók. A zárhangok időviszonyainak elemzése során kitérek az ilyen típusú fonémarealizációk vizsgálatára, a koartikuláció eredményeképpen létrejött hangok időtartamát összehasonlítom az ugyanolyan artikulációs konfigurációval bíró, morfémán belüli hosszú mássalhangzók időtartamával (pl. *adta – vatta, megken – vekker*). A felszínen hosszú mássalhangzóként realizálódnak a hangsorban egymás mellé kerülő azonos rövid mássalhangzók is (pl. *szakkör, hat tál*). Az eredmények arra kívánnak rávilágítani, hogy a beszédprodukciónak fonetikai és fonológiai szintjeinek működése hogyan mutatkozik meg az időviszonyok tervezésében és az artikulációs kivitelezésben, a hasonlással és/vagy morfémák találkozásakor létrejövő hosszú mássalhangzók mutatnak-e különbséget a morfémán belüli hosszú mássalhangzókhoz szemben időtartamukban, illetve belső időszerkezetükben. A spontán beszéd variabilitása miatt a kutatás ezen részéhez más módszertan (pl. felolvasott szó-, illetve mondatlista) alkalmazására lesz szükség.

4.3. A beszédhangok fizikai időtartamára másfelől hatással lehet a beszédtervezés bizonytalansága is. A beszédprodukciónak folyamatában a lexikális hozzáférés és az artikulációs tervezés összehangolatlanságából adódnak a különféle hangnyújtások, amelyek megjelenhetnek szó elején, belsejében vagy a végén (Gósy 2004). A kezdő hang túlzott megnövekedése a lexikális hozzáférési problémát kívánja feloldani, a szóvég nyújtása

általában a gondolkodási szünetet helyettesíti. Korábbi kísérletek kimutatták, hogy a megakadások közül a nyújtásokra figyelnek fel legkevésbé a hallgatók, valamint hogy a nyújtások észlelésében az időtartam növekedésén kívül más tényezők is szerepet játszanak, például a hang minősége, a hangsorban elfoglalt helye vagy a megnyújtott beszédhangot tartalmazó szó szófaja (Bóna 2007). Ezen szempontok alapján (is) vizsgálom a zárhangokon történő nyújtások előfordulásait. Időtartamértékeiket összevetem az adott hang tipikus ejtésű realizációinak időtartamával, azzal a céllal, hogy megismerjük, hol húzódik a határ a nyújtott ejtésű beszédhang időtartama és a hangra jellemző átlagos időtartam között. További kérdésfelvetésem, hogy mi jellemzi a megnyújtott ejtésű zárhangok belső időszerkezetét.

4.4. Az ismétlések és az újraindítások háttérében szintén a beszélő tervezési bizonytalansága áll. Az előbbi esetben a beszélő egy teljes lexémát újból kiejt (pl. *én ezt ezt gondolom*), míg az utóbbi esetben egy lexéma kiejtése során az artikuláció abbamarad, majd ezt követi a teljes szó kiejtése (pl. *aka- akarom*). Az ismétlések és az újraindítások folyamán a percepció számára azonos fonémásorok hangzanak el, az első és a második kimondás akusztikai szerkezete azonban eltérő lehet: a második kimondás általában rövidebb időt vesz igénybe (Gyarmathy 2009). Ennek alapján feltételezhető, hogy a megismételt beszédhangok időtartamai is jellegzetes különbséget mutatnak. A tervezett kutatás összehasonlítja az első és a második kimondásban megjelenő zárhangok időtartamát az ismétlések és az újraindítások esetén.

4.5. A fonetikai elemzéseket beszédészlelési vizsgálatokkal egészítem ki, amelyek a mért értékek relevanciáját igazolhatják. Univerzális tény, amelyet számos szakirodalmi kísérlet alátámaszt több nyelvre, hogy a felpattanó zárhangok észlelése igen bizonytalan, gyakorlatilag kontextusfüggő, vagyis egyéb akusztikai kulcs vagy kulcsok is szükségesek a pontos azonosításhoz, mint például a követő magánhangzó első formánsának frekvenciája, az alaphangmagasság változása a zárfelpattanást követően, a jellegzetes zörej és a második formáns frekvenciája. A zöngékezdési időt az egyik elsődleges tényezőnek tekintik a zöngés és zöngétlen zármássalhangzók megkülönböztetésében. A tervezett kísérletemben arra keresem a választ, hogy a hallgatók milyen paraméterek vagy paraméteregyüttesek alapján azonosítják a vizsgált zárhangokat, illetőleg hogy az időtényező milyen szerepet tölt be a pontos észlelésükben. A kutatásnak ez a része is hiánypótló a magyar szakirodalomban.

Másfelől, mivel a nyújtások megítélése az észlelés függvénye, a percepció tesztek ezen a területen is alapvető fontosságúak. A hallgatói észlelésen alapuló ítéletek alapján kívánom meghatározni, hogy a zármássalhangzók esetén milyen időtartamérték tekinthető már biztosan nyújtásnak, és mi számít még az adott zárhang normál ejtésű realizációjának. Természetesen ezen a téren csak óvatos megállapításokat tehetünk, hiszen az egyéni artikulációs sajátosságok nagy szerepet játszanak a beszéd temporális szerveződésében, ugyanis a beszélők szabadon alakítják ki beszédük egyéni időtartamviszonyait bizonyos határértékeken belül.

5. A kutatás várható eredményei

A tervezett kutatás több szempont szerint kívánja szisztematikusan feltérképezni a különböző helyzetekben előforduló explozívák időtartamát. Ennek eredményeképpen lehetőség nyílik arra, hogy a nyelvileg rövid és hosszú zárhangok esetében számos beszélő egyéni artikulációs sajátosságai alapján meghatározzam az egyes beszédhangokra jellemző átlagos időtartamértéket, illetve a tipikus ejtésre jellemző határértékeket, s ezeket összevetsem a szakirodalom eddigi eredményeivel.

Az egyes fonémarealizációk artikulációs megvalósítása több tényező függvénye (pl. kontextus, szupraszegmentumok), az időtartamértékek meghatározása ezek függvényében fog

történni, így a részletes elemzés során kimutatható lesz az egyes paraméterek hatása a spontán beszédben produkált zárhangok időszerkezetére. A percepciós kísérlet eredményei megerősítik az objektív értékek hitelességét.

A koartikulációs folyamatokkal kapcsolatos adatok tovább árnyalják a mássalhangzók időtartamáról kialakult képet.

A jelen kutatás elsőként elemzi a zármássalhangzók időviszonyait a beszédtervezési diszharmónia egyes jelenségeiben, a várható eredmények újabb adalékot szolgáltatnak a megakadásjelenségek akusztikai fonetikai sajátosságainak megismeréséhez; valamint a beszédprodukció fonológiai és fonetikai szintjeinek pontosabb megismeréséhez.

Irodalom

- Adamikné Jászó Anna 2000. Változott-e húsz év alatt a főiskolások kiejtése és olvasása? *Beszédkutatás* 2000. 124–131.
- Alphen, P. M. van – Smith, R. 2004. Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: the role of prevoicing. *Journal of Phonetics* 32: 455–449.
- Beckman, J. – Helgason, P. – McMurray, B. – Ringen, C. 2011. Rate effects on Swedish VOT: Evidence for phonological overspecification. *Journal of Phonetics* 39: 39–49.
- Boersma, Paul – Weenink, David 2010. Praat: doing phonetics by computer (Version 5.1.13) [Computer program] <http://www.praat.org/>
- Bolla Kálmán 1982. *Fejezetek a magyar leíró hangtanból*. Akadémiai Kiadó, Budapest.
- Bóna Judit 2007. Magánhangzó-nyújtások akusztikai-fonetikai paraméterei a spontán beszédben. *Beszédkutatás* 2007. 99–107.
- Bóna Judit 2009. *A gyors beszéd. Produkciós és percepciós sajátosságok*. MTA Könyvtára– Lexica Kiadó, Budapest.
- Cho, Taehong – Ladefoged, Peter. 1999. Variations and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27: 207–229.
- Collier, R. – Leeuwen van, H. C. – Willems, L. F. 1992. Speech synthesis today and tomorrow. *Philips Journal of Research* 47: 15–34.
- Gombocz Zoltán 1909. A magyar beszédhangok időtartamáról. *Nyelvtudomány* 2: 93–100.
- Gósy Mária 2000. A [p, t, k] mássalhangzók zöngékezdesi ideje. *Magyar Nyelvőr* 124: 195–204.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2006. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás* 2006. 194–207.
- Gósy Mária – Beke András 2010. Magánhangzó-időtartamok a spontán beszédben. *Magyar Nyelvőr* 134: 140–165.
- Gósy, Mária – Ringen, Catherine O. 2009. Everything you always wanted to know about VOT in Hungarian. Elhangzott: *International Conference on the Structure of Hungarian* 2009, Budapest, 2009. szeptember 1. Elérhető: http://icsh9.unideb.hu/pph/handout/Ringen_Gosy_handout.pdf.
- Grácsi Tekla Etelka 2011. Intervokális explozívák a zöngésségi oppozíció függvényében. *Beszédkutatás* 2011. 46–60.
- Grácsi Tekla Etelka 2012. *Zörejhangok akusztikai fonetikai vizsgálata a zöngésségi oppozíció függvényében*. Doktori disszertáció. Budapest.
- Gyarmathy Dorottya 2009. A beszélő bizonytalanságának jelzése: ismétlések és újraindítások. *Beszédkutatás* 2009. 196–216.
- Gyarmathy Dorottya – Neuberger Tilda 2011. A BEA adatbázis alkalmazásfüggő lejegyzései. *Beszédkutatás* 2011. 109–120.
- Hardcastle, William J. – Hewlett, Nigel (eds.) 1999. *Coarticulation. Theory, Data and Techniques*. Cambridge University Press, Cambridge.
- Hegedűs Lajos 1941. Elektroakusztikai berendezések a beszéd- és nyelvjáráskutatás szolgálatában. *Magyar Nyelv* 27: 162–185.
- Kassai Ilona 1979. *Időtartam és kvantitás a magyar nyelvben*. Nyelvtudományi Értekezések 102. Akadémiai Kiadó, Budapest.
- Kassai Ilona 1993. Gyorsult-e a magyar beszéd tempója az elmúlt 100-120 évben? *Beszédkutatás '93*. 62–69.
- Kovács Magdolna 1998. A spektrális minőség és az időtartam szerepe a magánhangzók percepciójában. *Beszédkutatás '98*. 35–47.
- Kovács Magdolna 2001. A zöngétlen zárhangok és affrikáták időszerkezete. *Beszédkutatás* 2001. 46–60.
- Laziczus Gyula 1963. *Fonétika*. Tankönyvkiadó, Budapest
- Levelt, Willem J. M. 1989. *Speaking: From Intention of Articulation*. MIT Press, Cambridge, MA.
- Machač, P. – Skarnitzl, R. 2009. *Principles of Phonetic Segmentation*. Epona Publishing House, Prága.
- Malmberg, B. 1944. *Die Quantität als phonetisch-phonologischer Begriff. Eine allgemeinsprachliche Studie*. Lunds Universitets Arskr. Lund-Leipzig.
- Magdics Klára 1966. A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények* 68: 125–139.
- Metz, C. 1914. Ein experimentell-phonetischer Beitrag zur Untersuchung der italienischen Konsonanten-Gemination. *Vox* 23: 201–270.
- Meyer, E. A. 1903. Englische Lautdauer. In: *Nordiska Studier tillegnade A. Noren*. Uppsala.
- Navarro Tomas, T. 1916. Cantidad de las vocales acentuadas. *Revista de Filología Española* 3: 387–407.
- Olaszy Gábor 1993. Hangidőtartamok számítógépes elemzése a beszéd ritmikái szerkezetének vizsgálatához. *Beszédkutatás '93*. 116–127.

- Olaszy Gábor 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben*. Nyelvtudományi Értekezések 155. Akadémiai Kiadó, Budapest.
- Santen, J. P. H. van 1992. Contextual effects on vowel duration. *Speech Communication* 11: 513–546.
- O’Shaughnessy, D. 1981. A study of French vowel and consonant durations. *Journal of Phonetics* 9: 385–406.
- Siptár Péter 1995. *A magyar mássalhangzók fonológiája*. *Linguistica, Series A, Studia et Dissertationes* 18. MTA Nyelvtudományi Intézet, Budapest.
- Szende Tamás 1976. *A beszéd folyamat alaptényezői*. Akadémiai Kiadó, Budapest.
- Tarnóczy, Tamás 1965. Can the problem of automatic speech recognition be solved by analysis alone? *Reports of the Fifth International Congress of Acoustics II, Liege*. 371–387.
- Umeda, N. 1977. Consonant duration in American English. *Journal of the Acoustical Society of America* 61: 846–858.