

A mássalhangzókra ható prozódiai erősítés vizsgálata a magyarban

Markó Alexandra¹, Deme Andrea^{2,1}, Juhász Kornélia^{2,1,3}, Bartók Márton¹,
Csapó Tamás Gábor^{4,1}, Grácsi Tekla Etelka^{1,3}

¹MTA–ELTE Lendület Lingvális Artikuláció Kutatócsoport

²Eötvös Loránd Tudományegyetem

³Nyelvtudományi Kutatóközpont

⁴Budapesti Műszaki és Gazdaságtudományi Egyetem
Távközlési és Médiainformatikai Tanszék

Abstract

Prosodic strengthening is the spatial and/or temporal extension of articulatory gestures that occurs at prosodically salient locations, such as the edges of larger prosodic units or stressed syllables. Hungarian consonants have not been analysed with reference to traits of prosodic strengthening so far. In a material of pseudowords with the structure $pV_1pV_1pV_1pV_1$ forming independent utterances (recorded previously for the examination of prosodic strengthening in vowels), bilabial voiceless plosive consonants were analyzed in terms of the number of bursts, duration, VOT, spectral moments and intensity, as a function of syllable position (first and stressed, internal and final) and adjacent vowels' quality (/i/, /u/, /ɒ/ and /a:/). The presence and the number of bursts indicated prosodic strengthening (especially) at the domain-initial edge of the phrase. A lack of burst was not detected in the first, stressed syllable. Realizations containing one burst were also most typical for the first syllable, and the occurrence of more than one burst was the least typical here. As for VOT, the interaction of syllable position and vowel quality had a significant effect, and the model explained the data to a large extent. In the first syllable, shorter VOT was measured; however, the difference between the syllable positions (first vs. final) reached the significance level only in close vowels. CoG and SD were lower, while skewness and kurtosis were higher in the first syllable than in the final syllable, reaching the significance level only in open vowels. The intensity of bursts was also higher in the first syllable with a gradual and significant decrease towards the end of the phrase, independently of vowel quality. In general, prosodic strengthening was observed in the first and also stressed syllables' consonant; however, in most parameters, this effect was dependent on the height of the adjacent vowel. Furthermore, in the analysed parameters the final syllable differed significantly from the first one, which may serve as a clue for boundary perception. Results support the theory that different languages use different combinations of acoustic keys to indicate stress and prosodic boundaries.

Keywords: prosodic strengthening, consonants, Hungarian

Email addresses: marko.alexandra.phd@gmail.com (Markó Alexandra),
deme.andrea@btk.elte.hu (Deme Andrea), juhasz.kornelia@nytud.hu (Juhász Kornélia),
bartokmarton@gmail.com (Bartók Márton), csapot@tmit.bme.hu (Csapó Tamás Gábor),
graczi.tekla.etelka@nytud.hu (Grácsi Tekla Etelka)

1. Bevezetés

A prozódiai szerkezet két, a beszédprodukciónak szempontjából lényeges funkciót lát el: az egyik a határjelölés, azaz a prozódiai alkotóelemek hierarchikus csoportosítása, a másik pedig a prominencia, azaz a prozódiai alkotóelemek közötti viszony jelölése (Pierrehumbert, 1999: 679). Egy adott megnyilatkozás prozódiai szerkezete természetesen a szegmentális szerkezet sajátosságaira is hatással van. *Prozódiai erősítésnek* (*prosodic strengthening*, vö. Cho, 2005, 2015) nevezik azt a jelenséget, hogy a prozódiai jellemzők kifejeződnek a szegmentális fonetikai sajátosságokban. A terminus az artikulációs gesztusok olyan térbeli és/vagy időbeli expanziójára utal, amely feltűnő helyeken jelentkezik, mint amilyenek a prozódiai egységek szélei vagy a hangsúlyos szótagok. (Más megközelítésben a szerzők – vö. pl. White et al., 2020 – szétválasztják a gesztusok térbeli expanzióját, ezt nevezve *erősítésnek* (*strengthening*), valamint az időbeli expanziót (*nyújtás/lengthening*). A jelen tanulmányban azonban a fenti definíciót alkalmazzuk, amely mind a térbeli, mind az időbeli expanziót, attól függetlenül, hogy ezek együtt járnak-e, erősítésként értelmezi a prozódiailag kitüntetett helyeken.)

A jelenség kapcsán meg kell említenünk Lindblom (1990) H&H elméletét, amelynek az elnevezése a *hyperspeech* és a *hypospeech* kifejezések rövidítéséből állt elő (418). Az előbbi a fonetikai információban gazdag (*overarticulate/d*), *túl/hiperartikulált*), az utóbbi pedig az információhiányos (*underarticulate/d*), *alulartikulált*) jelre utal, amelyeket egy kontinuum két végpontjaként képzelhetünk el. Lindblom szerint a beszélő az információgazdagság és az információs alulszabottság végpontjai közötti megvalósítási formák közül választ az alapján, hogy mit feltételez a hallgató számára (a jeltől függetlenül) elérhető információról.

Mind a hangsúlyos szótagokban tapasztalható, mind a nagyobb prozódiai egységek határán megvalósuló prozódiai erősítés vezethet lokális hiperartikulációhoz, amelyet az motivál, hogy a beszélő a hallgatói feldolgozáshoz kíván

több kulccsal szolgálni (ami persze nem feltétlenül tudatos). Lindblom (1990) terminusát továbbvive de Jong (1995) a hangsúlyos szótagok elemzése kapcsán az általa *lokális hiperartikulációnak* (*localized hyper-articulation*) nevezett jelenségre alkalmazta, azaz a gesztusoknak a prominencia hatására megvalósuló expansziójára. A *lokális hiperartikuláció/túlargikulálás* terminust kiterjeszthetjük a prominens szótagok mellett más kitüntetett fonetikai helyzetekre is, és Cho és munkatársai (2011: 345) alapján mindezekre alkalmazhatjuk a *prozodikusán motivált (lokális) túlargikulálás* (*prosodically driven (local) hyper-articulation*) terminust. A jelenséget már 1997-ben dokumentálta Fougeron & Keating (ők így nevezték: *articulatory strengthening at prosodic-domain edges*), és mind a térbeli (extrém artikulációs gesztusok), mind az időbeli (tartamtöbblet) dimenzióban dokumentált eltéréseket ide értették.

A frázisvégi nyújtás/nyúlás (az angol nyelvű szakirodalomban *preboundary* vagy *domain/phrase-final lengthening*) az egyik legszembetűnőbb és leggyakoribb példája annak, hogy a prozódia hatással van a szegmentális szerkezet megvalósítására Cho (2015). Ez a hatás abban jelentkezik, hogy a prozódiai határhoz közeledve a szegmentumok időzítése megváltozik, azaz a szegmentumok időtartama hosszabb az intonációs frázis végén, mint a frázis belsejében (pl. Edwards et al., 1991; Byrd, 2000). A frázis eleji nyújtás/nyúlást (az angol nyelvű szakirodalomban *postboundary* vagy *domain-initial lengthening*, vö. Cho, 2015: 519) elsősorban mássalhangzók esetében dokumentálták, és bár jellemzően *nyúlásként* említik, itt is a gesztusok nagyobb (és nem feltétlenül időbeli) expansziójáról van szó. Artikulációs vizsgálatokban, elektropalatográfias módszerrel elemelve a koronális mássalhangzók esetében nagyobb kiterjedésű lingvoplatális érintkezést találtak magasabb szintű prozódiai egységek (intonációs frázis) elején, mint alacsonyabb szintű határ (pl. szóhatár) esetében (összefoglalóan lásd Cho et al., 2007; Ip & Cutler, 2022), tehát az erősítés valóban bármely dimenzióban (térben, azaz a gesztusok térbeli megvalósításában, illetve időben, azaz a gesztusok időtartamában) megvalósulhat. Akusztikai vizsgálatok alapján pedig intenzitás-erősítésre és a spektrális momentumok (azaz a spektrum

frekvenciaátlag, csúcossága és ferdesége, vö. Forrest et al., 1988) eltérésére is vannak adatok (lásd a vonatkozó szakirodalom összefoglalását alább).

Az a tény, hogy mind a frázis végén, mind a fráziskezdő helyzetben mutatnak időbeli és/vagy térbeli erősítést a különféle nyelvek, akár a mássalhangzókra, akár a magánhangzókra vagy mindkét szegmentumtípusra vonatkozóan, arra utalhat, hogy mindkét esetben univerzális tendenciáról van szó. A nyelvek közötti eltérések alapján ugyanakkor arra lehet következtetni, hogy a jelenség fonetikai implementációja a beszélő által kontrolláltan, nyelvenként sajátos módon megy végbe, amely további tanulmányozást igényel. A szakirodalom áttekintése nyomán azt is láthatjuk, hogy a különféle prozódiai szerkezettel rendelkező nyelvek eltérő módon és mértékben jelenítik meg a prozódiai erősítés mintázatait ezeken a kitüntetett helyeken. Mivel a frázisvégi nyúlásnak és a hangsúly indukálta prozódiai erősítésnek magyar nyelven is viszonylag nagy irodalma van (lásd az alább hivatkozott forrásokat), az alábbiakban a frázis kezdetén mérhető prozódiai erősítésre koncentrálnak, és azon belül is a mássalhangzókra, illetve az explozívákra, hiszen a jelen vizsgálatunk a /p/ különféle fonetikai helyzetekben mutatózó realizációira irányul.

Akusztikai vizsgálatok alapján az angol, a német és a koreai nyelvben a zöngétlen hehezetes (aspirált) felpattanó zárhangok zöngékezdési ideje (azaz VOT-ja) hosszabbnak bizonyult fráziskezdő helyzetben, mint frázis belsejében (lásd Ip & Cutler, 2022 összefoglalását); másrészt a holland zöngétlen (aspirátlan) explozíva /t/ esetében VOT-rövidülést figyeltek meg a hangsúlyos szótagokban és a fráziskezdetben a frázis belsejéhez képest (Cho & McQueen, 2005). A VOT-t az artikulációs erőfeszítéssel szokás összefüggésbe hozni, amit a fortisz-lénisz skálával képeznek le (vö. pl. Lisker & Abramson, 1964). Itt a fortisz az erőteljesebb ejtésre, vagyis nagyobb izomfeszítettségre utal, és a zöngétlenekre jellemző, tehát nagyobb pozitív VOT társul hozzá, szemben a lénisz (zöngétlen, negatív VOT) ejtéssel. Ezen a skálán a hehezetes ejtésű explozíva a fortisz végponthoz közelebb helyezkedik el (a nem hehezetes zöngétlenhez képest is), ami lényegében azt jelenti, hogy a hangsúly hatására az angolban a nagyobb erőfeszítéssel

képzett felé tolódik el az ejtés (nagyobb pozitív VOT), míg a hollandban ellenkezőleg.

Vicenic (2010) a grúz nyelv három explozívátípusát (zöngétlen hehezetes, ejektív és zöngés) vizsgálta a bilabiális, alveoláris és veláris képzéshelyeken, tekintettel többek között a prozódiai erősítésre is. Akusztikai méréseket végzett egyebek mellett a zárfeloldás intenzitására és a spektrális momentumokra (azaz a spektrum frekvenciaátlagára, csúcosságára és ferdeségére) vonatkozóan. A magasabb rendű prozódiai helyzetnek (intonációs frázis eleji, szó eleji) kismértékű hatása volt a zárfeloldás intenzitására a hehezetes és a zöngés felpattanó zárhangokban, ugyanakkor az ejektívákban éppen az alacsonyabb szintű (szó belseji) pozíció esetén volt intenzívebb a zárfeloldás. A grúz alveoláris explozívák nem mutattak eltérést a zárfeloldást kísérő turbulens zörej átlagos frekvenciaértékében, de kisebb csúcosság- és ferdeségértékeket kapott a szerző a magasabb rendű (intonációs frázis eleji, szó eleji) prozódiai helyzetekben. A veláris zárhangok magasabb átlagos frekvenciaértékeket mutattak, ami alacsonyabb csúcosság- és ferdeségértékekkel párosult a magasabb rendű prozódiai helyzetekben. A bilabiális zöngés zárhangoknál hasonlóképpen magasabb átlagos frekvenciaértékek, valamint alacsonyabb csúcosság- és ferdeségértékek adódtak a magasabb rendű prozódiai helyzetekben, míg a bilabiális hehezetes zárhang és a bilabiális ejektíva alacsonyabb frekvenciaátlagot mutatott ugyanabban a helyzetben, de nem volt eltérés a csúcosságban és a ferdeségben az alacsonyabb rendű prozódiai helyzethez (szó belseje) képest.

Mindezek alapján tehát azt mondhatjuk, hogy a prozódiai erősítés megjelenése az adott nyelv mássalhangzóképzésétől és az abban található fonemikus kontrasztoktól függően alakul, ugyanakkor hatással vannak rá az adott nyelv prozódiai sajátosságai is. Cho és munkatársai (2011) az angol és a koreai nyelv között talált eltéréseket a nyelvspecifikus prozódiai mintázatokra vezették vissza, mégpedig arra, hogy a lexikai hangsúly (lexical stress) és dallamhangsúly (pitch accent) nélküli nyelvekben (pl. koreai) nagyobb az artikulációs gesztusok térbeli és időbeli expanziójának a mozgásteret, mint az olyan nyelvekben (pl. az angolban), ahol a lexikai prominencia korlátozza ennek a lehetőségeit. Korábban

a frázis eleji erősödést négy nyelvben, köztük szintén az angolban és a koreai-
ban is vizsgálva Keating és munkatársai (2004) arra a következtetésre jutottak,
hogy a koreaiiban robusztusabb a fráziskezdet erősítő hatása, mint az angolban,
mivel a koreaiiban a prozódiai szerkezet elsősorban a tagolást szolgálja, míg az
angolban a tagolás mellett a prominenciát is jelöli. A magyar nyelv sajátosnak
mondható ebből a szempontból, mert ugyan lexikai szinten nincs a hangsúlynak
megkülönböztető szerepe, és a fonémarealizációk jellegzetességeit tekintve sem
térnek el a hangsúlyos és a hangsúlytalan szótagbeli magánhangzók (nincs szisz-
tematikus magánhangzó-redukció a hangsúlytalan szótagban, vö. Gósy, 1997),
egyed tanulmányok eredményei mégis a szóhangsúlyos magánhangzó akusztikai
száliciájára utalnak (pl. Szalontai et al., 2016; Mády et al., 2017). Ugyan-
akkor az első szótagi kötött hangsúly miatt a magyarban igen gyakori, hogy
egybeesik a fráziskezdet és a (szó)hangsúly, így nem is mindig választható szét
a két prozódiai erősítő hatás.

Tabain és munkatársai (2016) öt, különböző mássalhangzókészlettel rendel-
kező és eltérő prozódiai mintázatokat mutató nyelvben vizsgálták a szóhangsúly
hatását a felpattanó zárhangok megvalósulására, zöngétlen zárhangokat elemez-
ve (az egyes nyelvekben 3, 4, 5 vagy 6 képzéshely alkotott fonológiai kontrasztot),
referenciaként az angol nyelv 3 zöngétlen felpattanó zárhangját alkalmazva. A
magyarhoz a vizsgált nyelvek közül az Indonéziában beszélt makassar nyelv ha-
sonlít a leginkább az ott található négy képzéshellyel, melyek a bilabiális (/p/),
az alveoláris (/t/), a palatális (/c/) és a veláris (/k/). Ugyanakkor ebben a
nyelvben a szóhangsúly az utolsó előtti (vagy elég gyakran az azt megelőző)
szótagra esik – amely (két, illetve három szótagos szó esetén) akár egybe is es-
hetne az első szótaggal, de a tanulmány a vizsgált szavak szótagszámáról nem
szolgál információval. A vizsgálat tárgyát két olyan nyelv is képezte, amelyben
a magyarhoz hasonlóan első szótagi lexikális hangsúly van, ezek a pitjantjatjara
és a warlpiri, ezekben azonban öt-öt zöngétlen zárhang alkot kontrasztot. Habár
egyik nyelv sem vethető össze egy az egyben a magyarral, az elemzések tanul-
ságosak a jelen vizsgálatunk szempontjából is. A kutatásban az időtartamot,
a spektrális lejtést és a négy spektrális momentumot (spektrális súlypont, azaz

CoG; szórás, ferdeség, csúcosság) elemezték, az utóbbiak közül csak az első kettőt mutatták be a tanulmányban. A zárfeloldás időtartama az angolban a hangsúlyos szótagok mássalhangzóiban hosszabb volt, mint a hangsúlytalanokban; a makassarban nem találtak ilyen eltérést. Ez a magyar szempontjából azért releváns, mert a magyarhoz hasonlóan az angolban és a makassarban is van zöngésségi kontraszt az explozívákban, de míg az angol aspiráló nyelv, ahol a zöngétlen aspirált (magas pozitív VOT) és a zöngétlen nem aspirált (alacsony pozitív VOT) áll szemben egymással (azaz pozitív VOT-kontraszt van), addig a makassar és a magyar zöngésítő nyelv, ahol a zöngés (negatív VOT) és a zöngétlen (alacsony pozitív VOT) áll szemben (azaz negatív VOT-kontraszt van). A szerzők más kutatások alapján is úgy találták, hogy az angolban a zöngétlen (és aspirált) explozívák kevésbé aspiráltak (értsd: kisebb a VOT-értékük) akkor, ha hangsúlytalan szótagban jelennek meg; ugyanakkor úgy vélték, hogy ha a makassar mássalhangzókészletből figyelembe vették volna az előzöngés hangzókat (negatív VOT), abban a körben eltérést találtak volna a hangsúlyos és a hangsúlytalan szótagbéli előfordulások VOT-értékei között. A pitjantjatjara esetében kismértékű eltérést találtak a hangsúlyos és a hangsúlytalan szótagbéli /p/-k VOT-je között (kisebbit, mint a /c/ és a /k/ esetében), nem találtak ugyanakkor ilyen hatást a warlpiri nyelvben. A CoG az angolban csak a /k/ esetében mutatott eltérést a hangsúlyos és a hangsúlytalan szótagok között (alacsonyabb a CoG a hangsúlyosokban), a makassarban viszont nem volt eltérés egyik mássalhangzóban sem. A pitjantjatjara az angollal egyezően a /k/-ban alacsonyabb CoG-t mutatott a hangsúlyos szótagban, a warlpiriban viszont semmilyen hatás nem volt mérhető. A spektrális szórásban kizárólag az angol /k/ esetében volt eltérés. Összességében tehát a mássalhangzókészletét tekintve a magyarhoz leginkább hasonló makassarban semmilyen (a szignifikanciát is elérő) hatást nem találtak a szóhangsúlyosság függvényében a VOT-t és a spektrális momentumokat illetően, sem a /p/, sem a többi zöngétlen zárhang esetében, ugyanakkor a szerzők utalnak a nagymértékű egyéni variációra. Azt is említik, hogy ebben a nyelvben nagy időtartambeli különbség van a hangsúlyosság függvényében a magánhangzók között, és úgy vélik, ez a tény is támogatja azt

az elméleti elgondolást, amely szerint a különféle nyelvek az akusztikai kulcsok eltérő kombinációit alkalmazzák a hangsúly jelzésére (Cutler, 2012).

A magyarra vonatkozóan a prozódiai erősítés hatását (még ha nem is feltétlenül nevezték így a szerzők) több tanulmány vizsgálta a hangsúlyos helyzetben lévő magánhangzók körében (pl. Magdics, 1965, 1966; Kassai, 1979; Gósy, 1997; Kovács, 2001; Mády et al., 2008; Gósy & Beke, 2010; Szalontai et al., 2016; Mády et al., 2017; Markó et al., 2018, 2019; Deme et al., 2019a,b, 2021, 2022, megj., kézirat), illetve a legújabban hangsúlyos és frázisszállan lévő magánhangzókban is (Markó et al., 2022). Bár az eredmények nem vágnak egybe, ami vélhetően a számos jellegzetességet tekintve eltérő módszertanokból adódik, többségükben mégis azt igazolják ezek a kísérletek, hogy a hangsúlyos szótagi magánhangzók akusztikai és artikulációs sajátosságai esetenként (a magánhangzó-minőségtől nem függetlenül) eltérnek a hangsúlytalan szótagbeli megvalósulásuktól. Céltan mássalhangzókra vonatkozó vizsgálatokról a jelen tanulmány szerzőinek nincs tudomása, de mivel a magyarban gyakran fráziskezdő is az, ami hangsúlyos, az alábbiakról adhatunk számot a szakirodalom alapján.

A magyar fráziskezdő mássalhangzókra vonatkozó adatok ellentmondásosak: míg Magdics (1966) rövidebbnek találta a mássalhangzókat a felolvasott mondatok elején (ugyanazon beszédhangoknak a hangsor belsejében található előfordulásaihoz képest), addig Kassai (1979) éppen ellentétes tendenciát talált. Az eltérések mögött módszertani eltérések is húzódnak (vö. Kohári, 2018), ugyanakkor más nyelvek esetén is dokumentáltak hasonlóképpen ellentmondó tendenciákat. White (2002) angol nyelvű kísérleteiben például hangsúlyos és hangsúlytalan helyzetben is vizsgált különböző szavakat mind frázis elején, mind frázis belsejében, és kevert mintázatokat talált a különféle szócsoportok esetén az egyes pozíciók tekintetében.

A különféle prozódiai határokhoz közeledve megjelenő nyújtás jelenségét a fentebb idézett szerzők univerzálisnak tartják, illetve a nyúlás – bár eltérő mértékben és módon – érintheti mind a magánhangzókat, mind a mássalhangzókat. A nyúlás mértéke függ a frázis végéhez képest elfoglalt helyzettől: a frázis utolsó szótagjában a legnagyobb mértékű a nyúlás, de már a frázis végéhez közeledve is

megfigyelhető az időtartamtöbblet (ha nem is feltétlenül fokozatosan/lineárisan nő a szegmentumok tartama). A nyúlás mértéke a szótagsajátosságoktól (zárt vagy nyílt szótag), továbbá az érintett szegmentumok különféle jellemzőitől is függ, ilyen például a fonológiai kvantitás, a magánhangzó-minőség, illetve a mássalhangzók képzésmódja (összefoglalóan vö. Kohári, 2018).

A magyarban a frázisvégi nyúlást több tanulmány is adatolta (pl. Magdics, 1966; Kassai, 1979; Kovács, 2001; Kohári & Markó, 2015; Gósy & Krepesz, 2016), mind magánhangzók, mind mássalhangzók időtartamában. Habár a módszertanok eltérnek, és az eredmények ezért nehezen összevethetők, a magyar beszéd is egyértelműen mutatja ezt a globális tendenciát. A komplex időszerkezetű beszédhangok (például a felpattanó zárhangok) esetében ugyanakkor a belső időszerkezetre (például a zöngeskedési időre) vonatkozóan nincsenek még magyar nyelvű kutatások.

A jelen tanulmány célja a prozódiai erősítés vizsgálata, azaz a prozódiai szempontból szálens fonetikai helyzeteknek (a fráziskezdő és egyben hangsúlyos, illetve frázisvégi helyzetnek) a szegmentális szerkezetben kifejeződő akusztikai mintázatait elemezzük, ezúttal a szótagkezdő (CV szótagbeli) mássalhangzóban. A vizsgálat jelentőségét főként abban látjuk, hogy a magyarban – az eddig legtöbbet vizsgált nyelvektől több szempontból eltérően – nem elsősorban prozódiai, hanem szintaktikai fókuszjelölés jellemző, ritmikai szempontból szótag-időzítésű nyelv, kötött szóhangsúllyal rendelkezik, és hangsúlytalan szótagban nem mutat fonológiai magánhangzó-neutralizációt.

A vizsgálat tárgyát négy szótagos álszavak önálló megnyilatkozásként való meghangosítása képezte, amelyben a kezdő szótag viseli a frázishangsúlyt, azaz egyszerre jelentkezhet a fráziskezdethez és a prominenciához kötődően prozódiai erősítés. Vizsgáltuk a frázisvégi nyújtás jelenségét is. E két szélső pozícióban jelentkező mássalhangzók szegmentális jellemzőit viszonyítjuk a közbülső mássalhangzókéhoz a következő paraméterek tekintetében:

- a felpattanások száma (ideértve a felpattanás esetleges elmaradását is);

- a zárhang időtartama (zöngétlen felpattanóról lévén szó ezt csak a 2–4. szótagban tudtuk elemezni);
- a zöngeskedési idő (VOT);
- a négy spektrális momentum, azaz a spektrális súlypont (CoG), a szórás, a ferdeség és a csúcsosság;
- a zárfeloldás intenzitásának mediánja és annak a követő magánhangzó legnagyobb intenzitásától való eltérése.

A vizsgálat központi kérdése tehát az, hogy a fenti paraméterekben eltérnek-e a feltételezett prozódiai erősítést mutató első és utolsó szótagokban álló mássalhangzók a frázis belseji szótagokban ejtett mássalhangzóktól. Azt, hogy a prozódiai erősítés hatása legalábbis bizonyos esetekben megmutatkozik a frázisok kezdetén és végén, egy korábbi, ugyanezen az anyagon végzett mérésorozatunk (Markó et al., 2022) a magánhangzók tekintetében igazolta: magánhangzó-minőségtől függő mértékű és hatókörű expansziót láttunk az időtartamban a frázis végén; a frázis kezdetén (és egyben a hangsúlyos szótagban) a hosszú és legalsó nyelvállású /a:/ mutatott egyértelmű artikulációs gesztusbeli expansziót (nagyobb ajaknyílást). Mivel a prozódiai erősítés a különféle nyelvekben és esetekben eltérő kombinációkban jelenhet meg, elképzelhető, hogy a jelen vizsgálatban a zármássalhangzók elemzése kapcsán a magánhangzóknak látottaktól eltérő (vagy éppen ellentétes) tendenciákat kapunk.

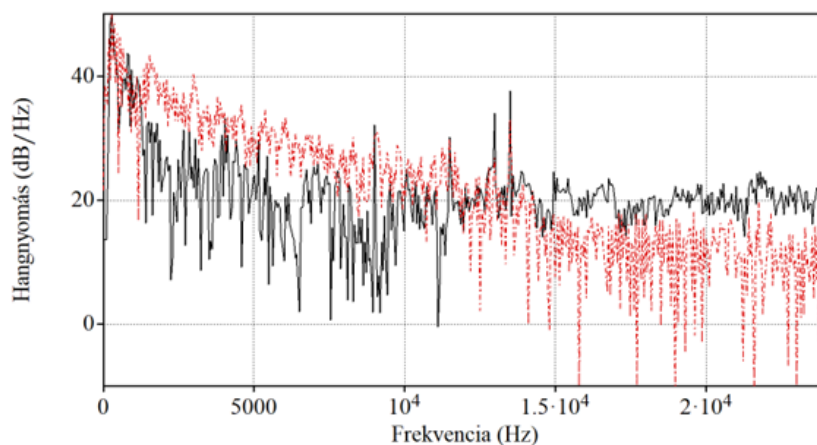
Meg kell jegyeznünk, hogy jelentősége van annak, hogy a vizsgált mássalhangzók CV szótagok onszetjei, hiszen így az utolsó mássalhangzó nem a hangsort záró szegmentum. Ennélfogva nem számítottunk erőteljes gyengülésre az utolsó szótagi mássalhangzóban (vö. Keating et al., 2004).

Mivel a követő magánhangzó minőségének szerepe van a VOT alakulásában (vö. pl. Gósy, 2000), a zárfelpattanások számának alakulásában (vö. Grácsi & Kohári, 2012), valamint különféle szótagpozíciókban megjelenő magánhangzók artikulációs és akusztikai jellemzői is eltérnek a magánhangzó-minőség függvényében (vö. Markó et al., 2022), ezért a szótagpozíció mellett a magánhangzó-

minőség hatását is vizsgáltuk. A magyarban az első szótag egyben hangsúlyos is, ezért itt a prozódiai erősítés hatását vártuk. Grácsi & Kohári (2012) szerint a zárfelepattanás elmaradásának hátterében az áll, hogy az izmok feszítettsége és/vagy a zár mögött kialakult nyomás mértéke nem elegendő, ilyen jelenséget tehát az első szótagban nem feltételeztünk, a további szótagokban azonban számítottunk rá. A magyar zöngétlen zárhangok közül a bilabiálisokra a legkevésbé jellemző a több felepattanás Neuberger & Beke (2017) szerint (a rövid /p/-k 27,6%-a volt ilyen a frázisbeli helyzetre nem kontrollált anyagokban). Más irodalmi előzményről nem lévén tudomásunk, nem volt hipotézisünk a zárfeleoldás többszöröződésére vonatkozóan a szótaghelyzet függvényében, de azt (Grácsi & Kohári, 2012 alapján) feltételeztük, hogy ennek gyakorisága függ a követő magánhangzó minőségétől. Gósy & Ringen (2009) magyar (izoláltan, szólistaként) felolvasott szavakban hosszabb VOT-értékeket mért szó belsejében álló zöngétlen zárhangokban, mint szó elején.

A spektrális momentumoknak nevezett (Forrest et al., 1988) paraméterek az energiaeloszlást jellemzik a spektrumban. A spektrális középpontot vagy súlypontot (CoG) a különböző frekvenciájú energiakomponensek súlyozott átlagaként határozzuk meg: ha magasabb frekvenciákon koncentrálódik az energia, akkor magasabb a CoG értéke, és fordítva. Ehhez viszonyítva határozható meg a spektrális szórás, azaz a spektrum frekvenciaátlagától mért átlagos eltérés. Az eloszlás szimmetrikus voltát ragadja meg a ferdeség paramétere; illetve a csúcosság jellemzi az eloszlás formáját. Ha normális az eloszlás, így például a spektrum energiájának, illetve a felepattanó zárhang felepattanási zöreijének spektrális eloszlása, akkor a ferdeség és a csúcosság értéke is 0. Attól függően, hogy balra vagy jobbra tolódik el a spektrum energiaeloszlása, a ferdeségi érték negatív vagy pozitív. A pozitív csúcosságérték azt jelzi, ha az eloszlás a CoG-érték környezetében csúcsosabb, mint a normál eloszlás, a negatív pedig azt, ha lapultabb, azaz, hogy mennyire koncentrálódik az energiaeloszlás egy területre. Az 1. ábrán mindezt a jelen kutatás anyagából származó két /p/ mássalhangzó (ugyanazon ejtésből a frázis első = fekete és összefüggő vonallal jelölt, valamint negyedik = piros és szaggatott vonallal jelölt zárhangja) ese-

tében mutatjuk be a 2. fejezetben ismertetett mérési beállítások mentén. A folytonos (fekete) görbe olyan mássalhangzót reprezentál, amelynek a spektrális középpontja 2178 Hz-nél van, a szórás pedig 4657 Hz. A ferdesége 2,97, a csúcossága 8,076. A szaggatott (piros) görbe által reprezentált /p/ zárfelpattanási zörejének spektrális súlypontja alacsonyabban, 1744 Hz-nél mérhető, hiszen az alsóbb frekvenciákon magasabb intenzitásértékeket látunk, a magasabb frekvenciatartományban pedig alacsonyabb intenzitásértékeket. Ennek a spektrális szórása (vagyis a CoG-től mért átlagos eltérés) a folytonossal jelölt mássalhangzóénál kisebb, 3345 Hz. A ferdeség a másik mássalhangzóhoz hasonlóan pozitív (2,82), hiszen mindkét esetben az alsóbb frekvenciatartomány felé tolódik el az eloszlás. A csúcosság 8,071, hasonlóképpen a folytonos görbével reprezentált /p/-éhez, azaz mindkét esetben pozitívak, ami a CoG-érték körüli frekvenciatartományban mindkét irányban alacsonyabb intenzitást jelent.



1. ábra. Két /p/ felpattanási zörejének energiaeloszlása a frekvencia függvényében (200 Hz alatt szűrés 100 Hz-es simítással, majd 1 kHz-től oktávonként 6 dB-es előerősítéssel).

Habár általában a képzési hellyel összefüggésben vizsgálják a spektrális momentumokat, a korábbi kutatások alapján ezek a zöngesség-zöngétlenség tekintetében is mutatnak eltéréseket. Magasabb frekvenciaértékek adódtak több vizsgált nyelvben is a zöngétlen zárhangok esetében (vö. van Alphen & Smits,

2004; Sundara, 2005; Chodroff & Wilson, 2014), ami percepció kulcsként is szolgál a zöngésségi kategória felismerésében. A grúz bilabiális zöngés zárhangok alacsonyabb átlagos frekvenciaértéket, illetve magasabb ferdeség- és csúcsosságértékeket mutattak az azonos képzéshelyű más explozívákhoz/ejektívákhoz képest (Vicenik, 2010). Ha a zöngés-zöngétlen közötti eltérést a lénsz-fortisz skálára vetítjük, ez azt jelenti, hogy a magasabb frekvencia (CoG) és az alacsonyabb ferdeség és csúcsosság fortiszabb képzésre utal.

A magyar /p t k/ zárhangokat a spektrális momentumokat illetően vizsgálták már ugyan, de annak a kutatásnak a középpontjában nem a mássalhangzó szó- vagy frázisbeli helyzete, hanem a fonológiai hosszúsága állt. Neuberger & Beke (2017) 10 férfi beszélő spontán beszédéből származó 221 darab /p/ átlagos CoG-értékét 1243 Hz-nek mérte. Konkrét számadatokat nem ismertetnek a szerzők, de az kiderül, hogy a ferdeség pozitív értékeket adott, azaz az eloszlás aszimmetrikus, és jobbra nyúlik el; valamint a csúcsosság értékei is pozitívak voltak, vagyis az eloszlás csúcsosabb volt, mint a normál eloszlás. Az intenzitás átlagértékéként a /p/ zárhang esetében 62 dB-t mértek. Meg kell jegyeznünk, hogy az idézett kutatás és a jelen mérések anyaga és módszertana több tekintetben eltért, ezért az ismertetett adatok nem szolgálnak közvetlen referenciaként a jelen elemzéshez.

A nemzetközi és a magyar szakirodalom ismeretében nem tudtunk a vizsgálati kérdésekre és paraméterekre hipotéziseket felállítani, vizsgálatunk exploratív jellegű.

2. Kísérleti személyek, anyag, módszer

A kutatásban kilenc magyar anyanyelvű női beszélő vett részt, átlagéletkoruk $25,2 \pm 5,9$ év volt, ismert hallás- vagy beszédproblémájuk nem volt.

A kutatás anyagát négy szótagos álszavak képezték, amelyeket önálló megnyilatkozásként hangosítottak meg a résztvevők. Az álszavak CV szerkezetű szótagokból álltak, amelyek mindegyikében a /p/ bilabiális zöngétlen explozív volt a mássalhangzó. Azért a bilabiális /p/ kontextust használtuk a kísérletben,

mert ez a hang csak kismértékben lép lingvális koartikulációba a szomszédos magánhangzókkal. Egy álszón belül a vizsgált anyagban a magánhangzó minősége sem változott – ezzel kívántuk elérni azt, hogy ne érvényesülhessen a magánhangzók közötti koartikulációs hatás (vö. Deme et al., 2019a,b, 2022, megj., kézirat), és (a lehetőségek mértékében) állandó legyen a szomszédos mássalhangzó és magánhangzó közötti koartikulációs hatás is. Négy, a magyar magánhangzókészletben legszélsőnek számító magánhangzót vontunk be a vizsgálatba: a hátul képzett és felső nyelvállású /u/-t, az elől képzett és felső nyelvállású /i/-t, a hátul képzett és alsó nyelvállású /ɒ/-t és a nyelv vízszintes helyzetét tekintve centrálisnak tartott és legalsó nyelvállású /a:/-t. Így végül a következő álszavak (négyelemű szótagsorok) szerepeltek a vizsgálati anyagban: *pupupupu*, *pipipipi*, *papapapa*, *pápápápá*.

A célszavakból típusonként és beszélőnként hat-hat előfordulást vizsgáltunk (összesen 216 hangsort, azaz $864 /p/-előfordulást = 9 \text{ beszélő} \times 4 \text{ magánhangzóminőség} \times 4 \text{ szótag} \times 6 \text{ ismétlés}$), melyeket random sorrendben, disztraktorok között rögzítettünk. Az álszavakat képernyőn jelenítettük meg a résztvevők számára, egyszerre egyet, és arra kértük őket, hogy „kijelentő mondat”-ként olvassák fel őket. Az itt bemutatott elemzés az akusztikus csatornán (44,1 kHz-es mintavételi frekvencián, a szájjughoz helyezett omnidirekcionális, kondenzátoros fejmikrofonnal) rögzített felvételen készült, de megjegyezzük, hogy ezzel párhuzamosan artikulációs vizsgálatra alkalmas felvétel is készült elektromágneses artikulográfiával (Markó et al., 2022). (Az eredeti felvételi módszertan indokolja a résztvevők viszonylag alacsony számát.) A jelen tanulmányban csak akusztikai méréseket mutatunk be.

A szegmentumokat az akusztikai jel alapján szegmentáltuk és címkéztük fél-automatikusan a *BAS* webszolgáltatás graféma-fonéma konverterének (Reichel, 2012) és a *MAUS* rendszernek (Schiel, 1999) a segítségével, majd a szükséges helyeken kézzel javítottuk a felismertett hanghatárokat a *Praat* szoftverben (Boersma & Weenink, 2022). Ezt követően kézzel felcímkéztük az adott mássalhangzó zárfeloldásának kezdetét (az első felpattanás kezdetét), valamint dokumentáltuk a zárfeloldások számát is. Az így felcímkézett mássalhangzók

időtartamát (a 2. szótagtól a 4. szótagig) a megelőző magánhangzó második formánsának végétől a követő magánhangzó második formánsának kezdetéig mértük, a zöngelkedési idejét pedig az első felpattanás kezdetétől a követő magánhangzó zöngéjének megjelenéséig. A zöngelkedési idő és a spektrális mérések esetében is (értelemszerűen) kizártuk azokat a megvalósulásokat, amelyekben nem volt detektálható a zárfeloldás. Ez összesen 30 előfordulást érintett (a realizációk 3,5%-a). A spektrális jellemzőket az alábbi módon mértük a *Praat* szoftver segítségével. A hangfájl intenzitásmenetét 100 Hz-es minimális alaphfrekvencia-beállítással 0,002 s-os ablakolással, átlagolással nyertük ki. A zárfeloldás spektrális jellemzőit annak kezdetétől a zöngé kezdetéig mértük.

A mássalhangzóban a spektrális momentumokat az alábbi módon határoztuk meg. A hangot 200 Hz alatt szűrtük (100 Hz-es simítással), majd 1 kHz-től oktávonként 6 dB-es előerősítést alkalmaztunk. Ezután a gyors Fourier-transzformált spektrumon nyertük ki a spektrális súlypontot, szórást, ferdeséget és csúcosságot a *Praat*-ban (Vicenik, 2010 módszerével egyezően).

A magánhangzóban 5 ms-os ablakkal, 0,5 ms-os eltolással megkerestük a legnagyobb intenzitású időszakaszt, és ennek középpontja körül ± 10 ms-mal mértünk intenzitásmediánt. A zárfeloldásban egyrészt elemeztük a feloldás intenzitását (mediánérték), másrészt a magánhangzóhoz képest mért intenzitását, azaz a magánhangzó fentebb leírt módon megkeresett legintenzívebb pontja körül 20 ms-os intervallumban kimértük az intenzitás mediánját, és ebből kivontuk a /p/ zárfeloldásában mért intenzitásmediánt.

A statisztikai elemzéseket az *R* szoftverben (R Core Team, 2021) végeztük, az ábrák a *ggplot2* csomagban készültek (Wickham, 2016). A kiugró értékeket (a mediántól több mint az interkvartilis tartomány 1,5-szeresére lévő értékeket) kizártuk az elemzésből.

A magánhangzó-kontextus és a szótagsorszám hatását a VOT-re, az időtartamra, a spektrális jellemzőkre, valamint az intenzitásjellemzőkre lineáris kevert modellekkel elemeztük (*lme4* csomag: Bates et al., 2015; *lmerTest* csomag: Kuznetsova et al., 2017). A vizsgált paraméterből kiindulva az időtartam, a VOT, a spektrális jellemzők vagy az intenzitásjellemzők egyike volt a függő

változó, a magánhangzó-kontextus és a szótag sorszáma a fix hatások, közöttük interakciót megengedve. A random hatások a két fix hatásra illesztett beszélőnkénti random meredekségek voltak, a beszélőre és az ismétlésre illesztett random konstansokkal. (Ismétlések alatt az egyes logatompéldányokat értjük.) A random meredekségek több esetben szinguláris illesztést okoztak, így azokat ezekben az esetekben elhagytuk, és csak a random konstansokat tartottuk meg. Amennyiben szingularitási probléma nem, hanem konvergációs hiba lépett fel, a *BOBYQA*-optimalizációt (*Bound Optimization BY Quadratic Approximation*, Powell, 2009) alkalmaztunk a maximális függvényértékelés értékének 20000-re emelésével (alapérték: 10000) (Singmann & Kellen, 2019 ajánlásaiból kiindulva). A random, majd a fix hatások alapján történt a modellszelekció az *anova()* paranccsal. A legszűkebb olyan modellt tartottuk meg, amely a legbővebb (fent leírt) modelltől még nem tért el szignifikánsan. Amennyiben mind a magánhangzó, mind a szótagsorszám elhagyása a kritériumoknak megfelelő modellt eredményezett, az alacsonyabb AIC-számút választottuk (Akaike, 1974). A p -értékeket Satterthwaite-approximációval nyertük ki az *anova()* parancs segítségével. A hatásnagyságokat a *MuMIn* (Bartoń, 2022) csomaggal elemeztük. A modellen Tukey-féle post hoc tesztet is futtattunk az interakcióra (*emmeans*: Lenth, 2020). Akkor is az interakcióra készítettünk post hoc próbát, ha az interakció nem bizonyult szignifikáns hatásnak.

A zárfelpattanások számát külön elemeztük szótagonként, magánhangzó-minőségenként és beszélőnként. Az adott faktor mentén összegeztük a zárfelpattanások számát, majd a gyakori cellánkénti alacsony előfordulások és a magas általános elemszám miatt Fisher-féle egzakt próbát alkalmaztunk a p -érték 2000-es szimulációjával (Harrison & Pius, 2021). A magánhangzók, illetve a szótagok páronkénti összevetését is elvégeztük ugyanilyen módon. A szignifikanciaszint korrekcióját a 4 csoportra alkalmaztuk mindkét faktor esetében, a 0,05-ös határértéket ($4 * 3/2 =$) 6-tal osztva, a 0,008-as p -értéket tekintettük szignifikanciaküszöbnek ezekben a post hoc tesztekben.

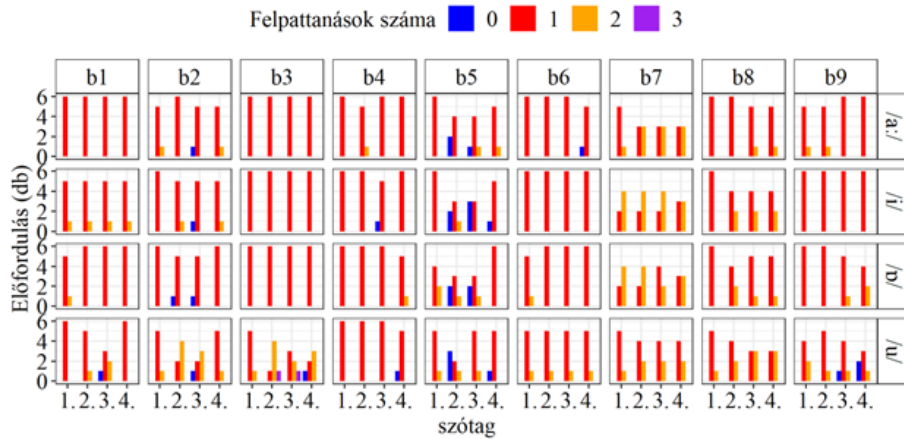
3. Eredmények

3.1. A zárfelpattanások száma

A zárfelpattanás nélkül realizálódott mássalhangzók száma (mint említettük) 30 volt, ami az összes példány 3,5%-át teszi ki. Egy felpattanással 703 előfordulás, azaz a vizsgált beszédhangok 81,4%-a realizálódott, egynél többel pedig 131, azaz 15,2%. A 2. ábra a felpattanások számának eloszlását mutatja beszélőnként, szótagonként és magánhangzó-minőségenként. Látható, hogy a zár feloldásának elmaradása és a többszöri feloldás egy, illetve két beszélőre jellemző főként, de mindkét jelenség jelentősen gyakoribb az /u/ magánhangzó környezetében. A szótagpozíció kapcsán megállapítható, hogy a felpattanás elmaradása az első szótagban nem fordult elő, a többszöri felpattanás pedig inkább beszélőnként változóan alakult a további szótagok között. A második szótagban mindösszesen 10 (4,6%), a harmadikban 14 (6,5%), a negyedikben pedig 6 (2,8%) esetben maradt el a zárfelpattanás. Egy felpattanással realizálódott az első szótagi mássalhangzók 89,4%-a (193 db), a második és a harmadik szótagbeliek 77,3%-a (167 db), valamint a negyedik szótagbeliek 81,5%-a (176 db). Egynél több felpattanással az első szótagban 23 /p/ realizálódott (az első szótagiak 10,6%-a), a másodikban 39 (18,1%), a harmadikban 35 (16,2%), a negyedikben pedig 34 (15,7%). Az eredményeket Fisher-féle egzakt próbával elemeztük a három faktorra külön-külön. Mindhárom faktor esetében szignifikáns hatást találtunk (szótagpozíció: $p = 0,002$; magánhangzó-minőség: $p = 0,001$, beszélő: $p < 0,001$). A szótagpozícióra készített post hoc tesztben azt az eredményt kaptuk, hogy az első szótag szignifikáns eltérést mutat az összes többi szótagtól, míg a többiedik szótagok között nincs szignifikáns eltérés a felpattanások számának tekintetében. A magánhangzó-környezet esetében az /u/ mutatott szignifikáns eltérést az összes többi magánhangzótól, míg a többi vizsgált magánhangzó-minőség között nem kaptunk szignifikáns eltérést.

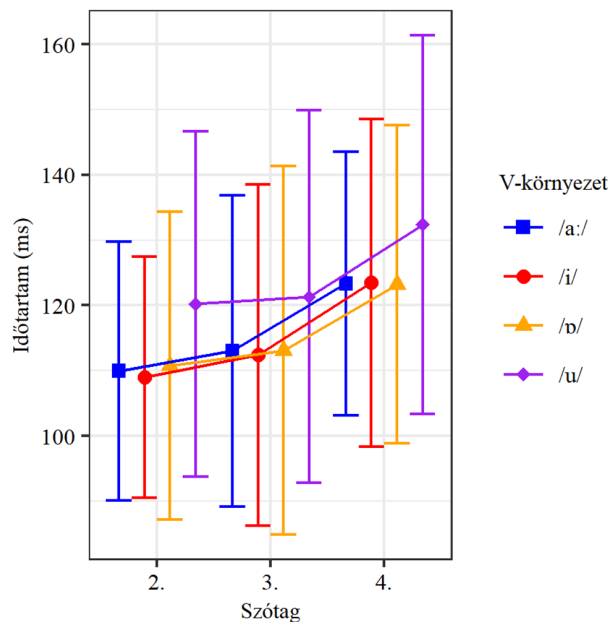
3.2. Időtartam

Az időtartamot a mássalhangzó kezdetén lévő néma fázis miatt a magánhangzókat követő kontextusban, azaz a 2., a 3. és a 4. szótagban tudtuk mérni,



2. ábra. A felpattanások számának eloszlása beszélőnként (b1–b9), szótagpozícióként (1–4) és magánhangzó-minőségenként.

ennélfogva nincs adatunk arról, hogy a mássalhangzó teljes időtartamára hatásal lehetett-e a frázis kezdeti és egyben hangsúlyos helyzet. A csak a szótagok sorszámát mint fix hatást tartalmazó modell bizonyult az adatokat legjobban leíró modellnek. A szótag sorszám (2–4.) szignifikáns hatással volt a mássalhangzó időtartamára ($F(2, 410, 49) = 92,986, p < 0,001$). Míg a második szótagban mért átlagos időtartamoktól (/a/: 110 ms, /i/: 109 ms, /ɔ/: 111 ms, /u/: 120 ms) a harmadik szótagban mért értékek (113 ms, 112 ms, 113 ms, 121 ms – azonos sorrendben) csak kis mértékben tértek el, a negyedik szótagban az elsőhöz képest legalább 10%-kal hosszabb átlagos időtartamértékeket kaptunk (/a:, i, ɔ/: 123 ms, /u/: 132 ms). A Tukey-féle post hoc teszt alapján a 4. szótagban mért időtartamok bizonyultak szignifikánsan hosszabbnak a korábbi szótagokban mérttől (3. ábra), azaz a frázis végi nyújtás hatását találtuk az adatokban – hasonlóan a magánhangzókhoz (vö. Markó et al., 2022). A modellre kapott hatásnagyságok alapján ($r_m^2 = 0,057, r_c^2 = 0,813$) azonban fontos kiemelni, hogy a szótagpozíció csak kis mértékben magyarázza az adatokat, míg a beszélők közötti variabilitás jelentős.

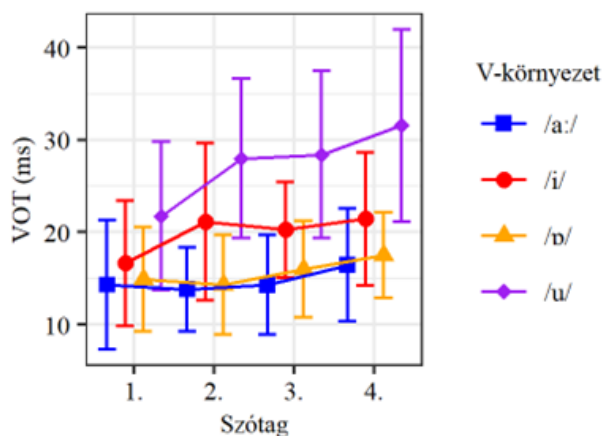


3. ábra. Az időtartam alakulása a szótagpozíció (2–4.) és a magánhangzó-környezet függvényében (átlag és szórás).

3.3. Zöngelkedési idő (VOT)

Az adatok leírására a legjobb modellnek az bizonyult, amely a két faktor (szótagpozíció és magánhangzó-minőség) interakcióját, a beszélőnkénti és ismétlésekénti random konstanst és a magánhangzó-minőségekénti random meredekséget tartalmazta. A szótagpozíció és a magánhangzó-minőség interakciója szignifikáns hatású volt ($F(9, 571, 31) = 2,681, p = 0,005$). A marginális hatásnagyság ($r_m^2 =$) 0,344, a kondicionális ($r_c^2 =$) 0,574, tehát a modell viszonylag nagymértékben magyarázza az adatokat. A post hoc tesztek alapján szignifikánsan eltért a /p/ zöngelkedési ideje a következő kombinációk közötti párokban (csak a szélső és a belső szótagok viszonylatában ismertetve az azonos magánhangzó-minőségek esetében): az /i/-s kontextusban az első (16,14 ms) és a második (19,47 ms) szótag között, valamint az első és a negyedik (20,11 ms) szótag között; az /u/-s kontextusban pedig az első szótagbeli /p/ VOT-értéke (20,48 ms) tért el az összes többi (második: 23,88 ms, harmadik: 24,34 ms, ne-

gyedik: 25,43 ms) szótagbeli értéktől. Habár a szignifikanciaküszöböt nem éri el a különbség minden magánhangzó-minőség (mint kontextus) esetében, a 4. ábrán az látszik, hogy a frázis végén a zöngékezdesi idő általánosságban (valamennyivel) hosszabb, mint a frázis kezdetén, ami a frázisvégi prozódiai erősítés (egyik) jele lehet. A 2–4. szótagokban a felső és az alsó nyelvállású magánhangzós kontextusok szinte minden esetben szignifikáns eltérést mutattak (/a:/ vs. /i/ és /a:/ vs. /u/ a 2., 3., 4. szótagban; /ɒ/ vs. /i/ a 2.; /ɒ/ vs. /u/ a 2. és 3. szótagban).

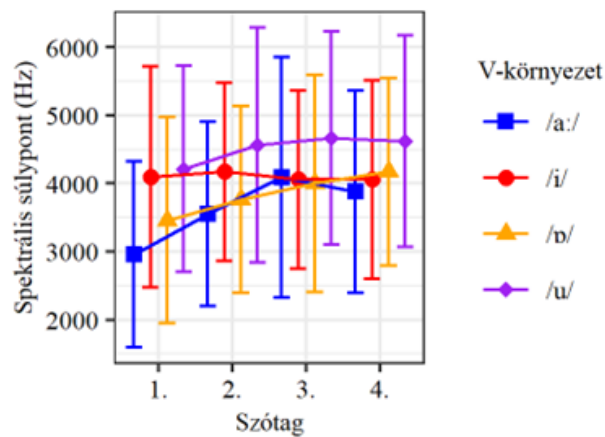


4. ábra. A zöngékezdesi idő (VOT) alakulása a szótagpozíció és a magánhangzó-környezet függvényében (átlag és szórás).

3.4. Spektrális súlypont (CoG)

A spektrális súlypont esetében a legjobb modell tartalmazta a két faktor interakcióját, beszélőnkénti és ismétlésekénti random konstanssal, random meredekség nélkül. A szótagpozíció és a magánhangzó-minőség interakciójának hatása szignifikáns ($F(9, 551,37) = 2,569; p = 0,007$), a marginális hatásnagyság azonban alacsony ($r_m^2 = 0,103$), ugyanakkor a kondicionális hatásnagyság közepes ($r_c^2 = 0,585$), azaz a beszélőnkénti és ismétlésekénti variabilitás bír valamelyes magyarázó erővel. Az 5. ábrán az látszik, hogy a CoG-átlagértékek magánhangzó-minőségenként eltérő mintázatokat mutatnak a különböző szótag-

pozíciókra vonatkozóan. Az /i/ esetét tekintve lényegében nincsenek különbségek a szótagpozíciók között, az /u/ előtti mássalhangzóban vannak ugyan kis mértékű eltérések, de nem érik el a szignifikanciaküszöböt a post hoc tesztek alapján. Az /v/-t megelőző /p/ spektrális súlypontja az első (3460 Hz) és az utolsó szótag (4172 Hz) között mutat szignifikáns eltérést; az /a:/-környezetben pedig az első (2959 Hz) és a harmadik (4093 Hz), valamint az első és a negyedik (3880 Hz) szótag között. Ezekben az esetekben a magasabb CoG a frázis vége felé fortisabb képzésre utal. Az egyes magánhangzó-környezetek között azonos szótagokban az /a:/ és az /u/ esetében minden szótagban szignifikáns eltérést találtunk a mássalhangzók spektrális súlypontjában, míg az egyéb párok között nem.

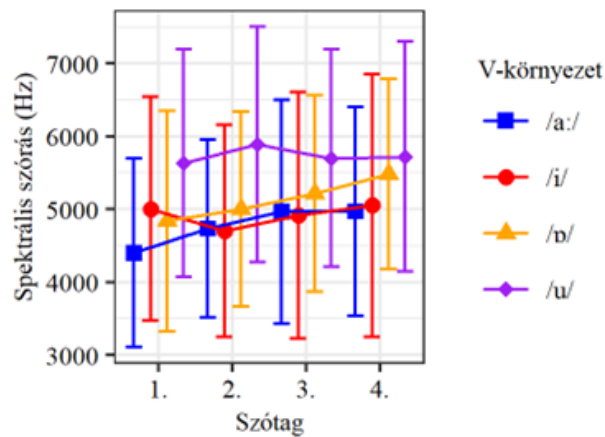


5. ábra. A spektrális súlypont (CoG) alakulása a szótagpozíció és a magánhangzó-környezet függvényében (átlag és szórás).

3.5. Spektrális szórás (SD)

A spektrális szórást leíró legjobb modellnek az bizonyult, ahol mindkét faktor szerepel, a beszélőnkénti és ismétlésenkénti random konstanssal. A két faktor interakciója szignifikáns ($F(9, 585,75) = 2,938, p = 0,020$), de kicsi a marginális hatásnagyság ($r_m^2 = 0,076$). A kondicionális hatásnagyság alapján ($r_c^2 = 0,752$) ugyanakkor a beszélő és az ismétlés jelentős mértékben magyarázza a variabili-

tást. A 6. ábráról leolvasható tendenciák ismét meglehetősen változatos képet mutatnak magánhangzó-minőségenként, és a CoG-hez hasonlóan szintén a nyílt magánhangzók esetében mutattak eltérést a post hoc tesztek. Az /v/-t megelőző /p/ spektrális szórása az első (4837 Hz) és az utolsó szótag (5482 Hz) között mutat szignifikáns eltérést; az /a:/-környezetben pedig az első (4399 Hz) és a harmadik (4968 Hz), valamint az első és a negyedik (4970 Hz) szótag között. Az /u/ környezetében álló /p/ CoG-értéke minden más magánhangzó-kontextushoz képest szignifikánsan magasabb volt (az /a:/-s környezettől minden szótagban, az /i/-stől a 2–4. szótagokban, az /v/-stől az első két szótagban mutatott eltérést).

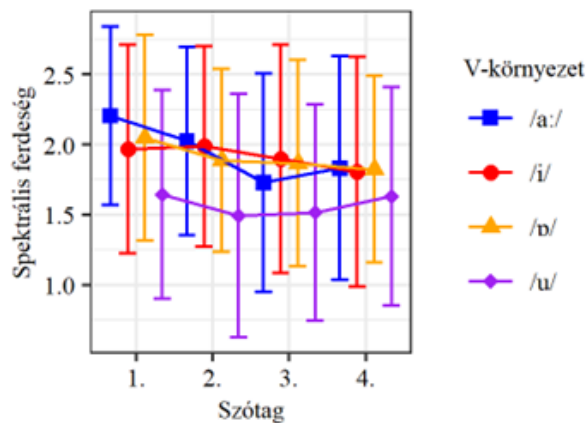


6. ábra. A spektrális szórás (SD) alakulása a szótagpozíció és a magánhangzó-környezet függvényében (átlag és szórás).

3.6. Spektrális ferdeség

Mint említettük, a spektrum eloszlásának szimmetriáját ragadja meg a ferdeség paramétere, és a /p/ esetében pozitív értéket várunk, hiszen a középponttól jobbra nyúlik el az eloszlás. A kérdés az, hogy az eltolódás mértéke összefüggést mutat-e a szótagpozícióval és a magánhangzó-minőséggel. A legjobb modellnek az bizonyult, ahol mindkét faktor szerepel, a beszélőnkénti és ismétlésenkénti random konstanssal. A két faktor interakciója nem volt szigni-

fikáns, míg egyesével mind a szótagpozíció ($F(3, 554,40) = 6,700, p < 0,001$), mind a magánhangzó-minőség ($F(3, 200,42) = 12,410, p < 0,001$) főhatása igen. A marginális hatásnagyság alacsony ($r_m^2 = 0,079$), a kondicionális pedig közepes ($r_c^2 = 0,542$), vagyis bizonyos mértékig magyarázzák az adatokat a beszélők és az ismétlések egyéni jellegzetességei, de valamilyen további, itt nem vizsgált tényezőnek is hatása van. A magánhangzó-minőségek ezúttal is változatos mintázatokat adnak (7. ábra). A post hoc tesztek alapján csak az /a:/-környezetben volt eltérés, mégpedig az első (2,20) és a harmadik (1,73) szótagbeli, valamint az első és a negyedik (1,83) szótagi ferdeségi értékek között. Ebben az esetben az alacsonyabb ferdeségérték a frázis vége felé fortiszabb mássalhangzóképzésre utal. A magánhangzó-környezet tekintetében az /u/ esetében itt is több szótagban előfordul a többitől való szignifikáns eltérés (/a:/ vs. /u/ és /v/ vs. /u/ az első két szótagban, /i/ vs. /u/ a 2. és 3. szótagban).

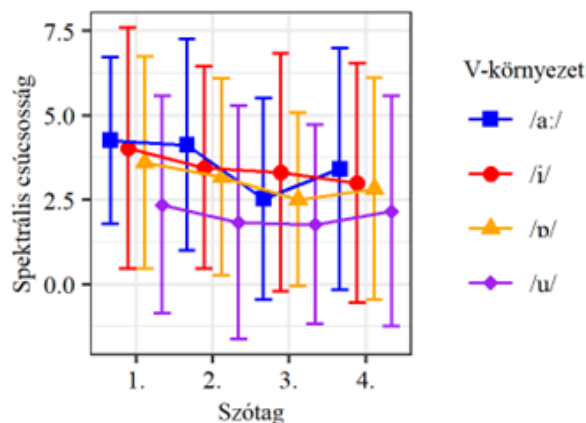


7. ábra. A spektrális ferdeség alakulása a szótagpozíció és a magánhangzó-környezet függvényében (átlag és szórás).

3.7. Spektrális csúcsosság

A csúcsosság értéke a spektrális eloszlás formáját jellemzi, mégpedig oly módon, hogy a pozitív csúcsosságérték jelzi azt, ha az eloszlás csúcsosabb, mint a normál eloszlás – a /p/ esetében erre számítottunk. A kérdés itt is az, hogy van-e

hatása a szótagpozíciónak és/vagy a kontextust képező magánhangzó minőségének a csúcosság értékére. A legjobb modellnek az bizonyult, ahol mindkét faktor szerepelt, beszélőnkénti és ismétlésenkénti random konstanssal, random meredekség nélkül. A két faktor interakciója nem bírt szignifikáns hatással, a két főhatást így külön elemeztük. A szótagpozíciónak ($F(3, 534,86) = 6,596$, $p < 0,001$) és a magánhangzó-minőségnek ($F(3, 194,02) = 9,104$, $p < 0,001$) a hatása is szignifikánsnak adódott. A marginális hatásnagyság ezúttal is elhanyagolható ($r_m^2 = 0,064$), a kondicionális pedig közepes ($r_c^2 = 0,575$), vagyis a vizsgált tényezők nem magyarázzák teljes körűen az adatokat. A post hoc tesztek közül ismét csak az /a:/-környezetben volt szignifikáns eltérés, mégpedig az első (4,26) és a harmadik (2,53) szótagban mért értékek között (8. ábra). Ebben az esetben az alacsonyabb csúcosságérték a frázis vége felé fortiszabb mássalhangzóképzésre utal. A spektrális ferdeséghez hasonlóan ismét kisebb az /u/-s környezet többtől való eltérése, de az /a:/-stól az 1. és 2., az /i/-stól a 3. szótagban szignifikáns eltérést mutat.



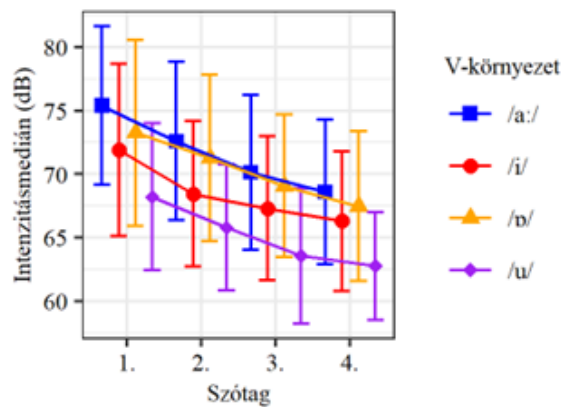
8. ábra. A spektrális csúcosság alakulása a szótagpozíció és a magánhangzó-környezet függvényében (átlag és szórás).

3.8. A zárfeloldás intenzitása

A zárfeloldás intenzitásának mediánjára vonatkozóan a legjobb modellnek az bizonyult, ahol mindkét faktor szerepel, a beszélőnkénti és ismétlésekénti random konstanssal, random meredekség nélkül. A két faktor interakciója ismét nem bizonyult szignifikánsan meghatározónak az eredményekben, míg a két faktor hatása külön-külön szignifikáns volt (a szótagpozícióra $F(3, 604,86) = 213,582$, $p < 0,001$; a magánhangzó-minőségre $F(3, 201,87) = 91,410$, $p < 0,001$). A marginális hatásnagyság alacsony ($r_m^2 = 0,256$), de elmondható, hogy az adatok mintegy negyedét magyarázzák az elemzett főhatások. A kondicionális hatásnagyság emellett jelentősen magasabb, azaz a vizsgált tényezők (beszélő, ismétlés) hatásai nagymértékben magyarázzák az adatokat ($r_c^2 = 0,860$). A post hoc tesztek alapján minden magánhangzó-minőség esetén szignifikáns eltérés van az első és a második, az első és a harmadik, az első és a negyedik szótagbeli /p/-k intenzitása között. Az /a:/-s és az /v/-s környezetben ezen túlmenően szignifikáns a különbség a második és a harmadik, valamint a második és negyedik szótag között is; az /i/ és az /u/ esetében pedig a második és a negyedik szótag között. Mindez azt jelenti, hogy az első szótag legintenzívebb zárfeloldásától kezdődően fokozatos intenzitáscsökkenés figyelhető meg a frázis utolsó (két) szótagjáig (9. ábra). Az azonos szótagban a magánhangzó-környezetek alapján ismét nagyobb különbségek mutatkoztak: az /a:/ minden, az /v/ több szótagban is szignifikáns eltérést mutatott az /i/-től és az /u/-től. Az /i/-/u/ ugyancsak minden szótagban szignifikáns különbséget mutatott. Az /a:/-/v/ környezetek között csak az első szótagban kaptunk szignifikáns eltérést.

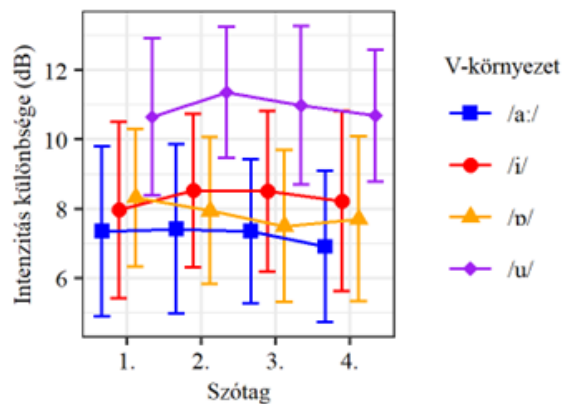
3.9. A mássalhangzó intenzitásának a követő magánhangzó legnagyobb intenzitásától való eltérése

A mássalhangzó intenzitásának a követő magánhangzó legnagyobb intenzitásától való eltérését a legjobban az a modell jellemezte, amelyben a magánhangzó-minőség és a szótagpozíció interakciója szerepelt, random konstanssal a beszélőre és az ismétlésre. Csak a magánhangzó-minőségre lett szignifikáns az eredmény ($F(3, 196,12) = 63,787$, $p < 0,001$), a marginális hatásnagyság alacsony



9. ábra. A zárfeloldás intenzitásmédiánjának alakulása a szótagpozíció és a magánhangzó-környezet függvényében (átlag és szórás).

($r_m^2 = 0,264$), a kondicionális pedig közepes ($r_c^2 = 0,491$). A post hoc tesztek nem mutattak eltérést egyazon magánhangzós környezetben a szótagpozíciók között; a magánhangzó-környezetek között azonban az /u/ az összes szótagban magasabb értékeket adott a többi magánhangzóhoz képest. A többi magánhangzó környezetében mért értékek között nem volt eltérés (10. ábra).



10. ábra. A zárfeloldás intenzitásmédiánjának eltérése a követő magánhangzó intenzitásától a szótagpozíció és a magánhangzó-környezet függvényében (átlag és szórás).

4. Következtetések

A prozódiai erősítés az artikulációs gesztusok térbeli és/vagy időbeli expansziója, amely prozódiailag kiemelkedő helyeken, például nagyobb prozódiai egységek szélein vagy hangsúlyos szótagokon történik. Ezeknek az artikulációs módosulásoknak akusztikai következményei is vannak/lehetnek, még ha nem is jellemezhető egyenes arányossággal az artikulációs gesztusbeli eltérés és az akusztikai szerkezetben jelentkező különbség közötti viszony (Stevens, 1989). A magyar mássalhangzókat tudomásunk szerint ez idáig nem elemezték a prozódiai erősítés szempontjából, így csak más nyelvekre vonatkozó vizsgálatok eredményei alapján feltételezhetjük, hogy a magyarban is találhatunk eltéréseket a frázisbeli pozíció (prozódiai tartomány eleje, vége; hangsúlyos helyzet vs. tartomány belseji, hangsúlytalan helyzet) függvényében a mássalhangzók valamely jellemzőiben. A szakirodalomban olvasható elemzések olyan nyelvek mássalhangzóit vizsgálták, amelyek a magyartól többek között a prozódiai szerkezet jellemzőiben (például a szóhangsúly helye, a szónál tágabb tartományhoz kötődő hangsúly funkciója és akusztikai manifesztációja) és/vagy a mássalhangzókészletüket tekintve is eltérnek, így előzetesen nem tudtunk hipotéziseket megfogalmazni.

A vizsgálatban önálló megnyilatkozást alkotó $pV_1pV_1pV_1pV_1$ szerkezetű álszavak anyagában (amelyet korábban a magánhangzók prozódiai erősítésének elemzése céljából rögzítettük) elemeztük a bilabiális zöngétlen felpattanó mássalhangzókat. A független változók a szótagpozíció (a szótag sorszáma 1–4.) és (korábbi eredmények alapján) a szomszédos magánhangzók minősége (/i/, /u/, /ɒ/ és /a:/) voltak. Ezek függvényében elemeztük a /p/-ket a zárfelpattanás megléte/hiánya és száma, az időtartam, a VOT, a spektrális momentumok és az intenzitás (a zárfeloldás intenzitása, valamint ennek a követő magánhangzó legintenzívebb szakaszának intenzitásához viszonyított eltérése) szempontjából.

Általános, nyelvfüggetlen aerodinamikai elvek alapján nem vártuk a zárfelpattanás elmaradását az első szótagban, a további szótagokban azonban igen. (Megjegyezzük, hogy a frázisbeli helyzet függvényében a zárfelpattanás elmara-

dására vagy többszöröződésére vonatkozó elemzést a nemzetközi szakirodalomban sem találtunk.) Adataink alapján a zárfelpattanások megléte, illetve száma utal a prozódiai erősítésre az első és hangsúlyos szótagban.

A VOT esetében a szótagpozíció és a magánhangzó-minőség interakciója szignifikánsan meghatározó volt, és a modell nagymértékben magyarázta az adatokat. Az első szótagi helyzetben rövidebb VOT-t mértünk a szó belseji előfordulásokhoz képest. A post hoc tesztek alapján az /i/-s kontextusban az első és a második, valamint az első és a negyedik szótag között; az /u/-s kontextusban pedig az első szótagbeli /p/ VOT-értéke tért el az összes többi szótagbeli értéktől. Feltehető, hogy az /i/ és az /u/ a nyelvvállásfoka miatt alkalmasabb a magasabb nyomás hosszabb megtartására (másfelől közelítve: kevésbé alkalmas a gyorsabb nyomáskiegyenlítésre), ezért a felső nyelvvállású magánhangzók környezetében jobban hat a prominencia. A holland zöngés explozívák esetében ugyancsak VOT-rövidülést figyeltek meg a hangsúlyos szótagokban és a fráziskezdetben a frázis belsejéhez képest (Cho & McQueen, 2005).

A spektrum frekvenciaeloszlása kapcsán Tabain et al. (2016) az angol /k/-t vizsgálva a hangsúlyos szótagban alacsonyabb, Vicenik (2010) pedig a grúz bilabiális zárhangokban szintén alacsonyabb átlagos frekvenciát mért a magasabb prozódiai szinteken megvalósuló prominens pozíciókban, így a prozódiai erősítés jele ezekben feltehetően az alacsonyabb CoG. Megjegyzendő ugyanakkor, hogy ilyen jellegű tendenciákat csak az alsó nyelvvállású magánhangzók esetében találtunk a jelen anyagban, amelyek előtt a koartikulációs hatás (a zárfeloldás tágabb ajaknyílás felé tart, mint a felső nyelvvállású magánhangzók előtt) magyarázza az alacsonyabb CoG-értékeket is. Az alsó nyelvvállású magánhangzók esetében ráadásul az ajaknyílás mértéke kevésbé korlátozott (nagyobb a mozgástér), mint a zártak esetén, amelyeknek az ajaknyílását (és az ebből adódó, a megelőző más-salhangzóban mérhető CoG-értéket) korlátok közé szorítja a középső nyelvvállású magánhangzó-minőségekbe való átcsapás lehetősége.

A fenti VOT-adatokkal összevetve a CoG-értékeket feltehetjük, hogy a prozódiai erősítés a felső nyelvvállású magánhangzók szomszédságában inkább az időzítés paraméterében, az alsó nyelvvállásúakéban pedig inkább a spektrális

mintázatban fejeződik ki, ami természetesen összefügg a magánhangzók artikulációs jellemzőiből adódó aerodinamikai hatásokkal. Másrészt ugyanakkor a prozódiai határ közeledtének kulcsolására is szolgálhat az alsó nyelvállású magánhangzók környezetében a magasabb spektrális súlypont, amit a frázisvégen (illetve az /a:/ esetben a frázis vége felé) láttunk; valamint a felső nyelvállású magánhangzós kontextusban a hosszabb VOT. Az itt vizsgált egységek mindössze négy szótagosak voltak, de lehetséges, hogy hosszabb frázisok esetében nagyobb mértékben tér(né)nek el a frázis szélén és a frázis belsejében jelentkező mássalhangzók akusztikai jellemzői.

A zárfeloldás intenzitásadatait a korábbi, grúz nyelvre végzett vizsgálat adataihoz (Vicenik, 2010) nem tudjuk egy az egyben hasonlítani, mivel ott bilabiális zöngétlen zárhangból csak aspirált ejtésűt tartalmaz a mássalhangzókészlet. Mégis, a vizsgált tendenciák bizonyos fókig egybecsengenek a magyaréval, hiszen a magasabb rendű prozódiai helyzetnek (a magyar esetében az első szótagi és egyben hangsúlyos helyzetnek) erősítő hatása volt a zárfeloldás intenzitására.

A szakirodalom alapján feltehető, hogy a VOT értéke és a felpattanási zörej intenzitása összefügg (Sundara (2005) idézi Pickett (1999) felvetését), mivel a szájüregi zár időtartamának aerodinamikai következményéről van szó. A magasabb VOT-értékkel jellemezhető zárhangok hosszabb záridőtartamot is mutatnak, ennél fogva a szájüregi nyomás felépülésének intenzívebb felpattanási zöreje a következménye. Mindez természetesen ismét a nagyobb VOT-nak a feszebb ejtéssel való összefüggésére is utal. A mi adataink ugyanakkor ezt az összefüggést nem mutatják, mert bár az intenzitásmedián a frázisban folyamatosan csökken, a VOT (magánhangzónként változó mértékben és léptékben) inkább növekszik. Az egyes magánhangzók környezetében a frázis vége felé a CoG ugyancsak magasabb, a ferdeség és a csúcsosság pedig alacsonyabb, ami szintén fortiszabb ejtésre utal. Így tehát az intenzitás csökkenése mellett más akusztikai paraméterek inkább erősítést mutatnak a frázis vége felé.

Az intenzitásadatok csökkenési tendenciája lehet annak a sokak által általánosnak tartott és fízológiai alapokra visszavezetett folyamatnak a része, hogy a megnyilatkozás kezdetétől a végéig fokozatos gyengülés (ereszkedés) történik,

amely megnyilvánul az artikulációs gesztusokban, a frekvenciában, az intenzitásban és az időzítésben egyaránt, mind szegmentális, mind szupraszegmentális tekintetben (vö. pl. Vayra & Fowler, 1992). Mindez a megnyilatkozásban előre haladván a szubglottális nyomás fokozatos csökkenésével függhet össze, amely a (beszéd)légzés fiziológiai sajátosságaiból ered (Cohen et al., 1982). Ez a fiziológiai magyarázat ugyanakkor azt sugallja, hogy a beszédműködés olyan következményéről van szó, amely nem áll a beszélő kontrollja alatt. Ezzel szemben számos példát ismerünk arra, hogy a légzési és a gégeizmok a beszélő kontrollálásával működnek (ilyen egyebek között a magyar kérdő intonáció, azaz a frázisvégi frekvenciaemelkedés és -csökkenés). Ennélfogva több szerző is azt a magyarázatot támogatja, hogy a frázis végéhez közeledve jelentkező artikulációs-akusztikai módosulások (például a frázisvégi nyúlás) aktív folyamat eredményei, és az a szerepük, hogy ráirányítsák a hallgató figyelmét a következő prozódiai határra, ezzel is kulcsot adva a beszédfeldolgozáshoz (vö. pl. Byrd, 2000; Cho, 2006; Cho et al., 2011). Ezt az elképzelést támogatja az a tény is, hogy például az említett frázisvégi nyújtás esetében – habár univerzális jelenségnek tűnik – eltérések találhatók a nyelvek között a mértékét és a kiterjedését tekintve (vö. Kohári, 2018 összefoglalását).

Amint láttuk, az alkalmazott magánhangzó-környezetek függvényében is kirajzolódtak mintázatok, amelyek további kérdéseket is felvetnek. Az /u/ környezetében majdnem minden paraméter, az /i/-kontextusban a VOT és az intenzitás tekintetében tértek el a /p/ akusztikai tulajdonságai a többi környezetben mértéktől. Mivel a /p/ kisebb koartikulációs hatással bír a környező magánhangzókra, ezért e magánhangzók szomszédságában a megemelt nyelvhelyzetnek nagyobb mértékű hatása lehet a szájüreg méretére, így a szupraglottális nyomás magasabb lehet, mint az alsóbb nyelvállásfokok esetében. Ez a hatás az /u/ hátul képzett volta következtében (az /i/-éhez képest is) kisebb üregméret miatt tovább fokozódhat. Ennek a nyomásviszonynak köszönhető feltehetőleg a Bernoulli-hatásnak a felpattanások számát megnövelő hatása. Ugyancsak emiatt lassabban egyenlítődik ki a zárfeloldás után a külső légnyomás és a belső (intraorális) nyomás, ezért a VOT is hosszabb lehet. A zár mögött közvetlenül

felgyűlő levegő mennyisége és nyomása pedig az intenzitásértékekre volt hatással. A spektrális momentumokban kisebb, de (az első kettőben erősebben) meglévő eltérés háttérében ezen okok együttesen játszhatnak szerepet. Mindennek vizsgálata további, ezt a kérdést célzó kutatások feladata, okai között szerepelhet a magánhangzó célkonfigurációjához közeli állapotban megtartandó nyelvforma.

Elemzésünk egy más céllal felvett hanganyagon természetesen korlátozottan adhatott választ a magyar mássalhangzók prozódiai erősítésével kapcsolatban felmerülő tudományos kérdésekre. A rendelkezésre álló szekvenciák viszonylag rövidek voltak, és mindösszesen egy mássalhangzó elemzését tették lehetővé, ráadásul e mássalhangzó zöngétlensége miatt bizonyos szempontokat (például a teljes mássalhangzó-időtartamot az első szótagban) nem tudtunk vizsgálni. Mindemellett a szakirodalom (Tabain et al., 2016) alapján azt is számításba kell vennünk, hogy a /p/ az ilyen szempontú vizsgálatra az artikulációs sajátosságaiból adódóan nem ideális választás (még ha az anyag eredeti vizsgálati célja szerint az volt is), mivel az akadály mögötti teljes üreg gerjesztődik a zárfelpattanás során létrejövő zörej által. Másrészt a mássalhangzó-észlelés szempontjából szerepet játszhat az a tény is, hogy a bilabiális zár vizuálisan egyértelműsíti a képzési jellemzőket (az ajakpozícióból és az izomtevékenység látható jeleiből a képzés helye, módja, a zöngéesség és az időtartam is egyértelműen azonosítható), így nincs szükség akusztikai erősítésre. Mindezen érvényességi korlátok és kérdések tudatában mégis arra utalnak az elemzési eredmények, hogy a prozódiai erősítés jelenségét a magyar mássalhangzók tekintetében is érdemes (tovább) vizsgálni, erre a célra összeállított kísérleti anyagon, hiszen bizonyos, más nyelvekre is jellemző tendenciák itt is megmutatkoztak.

Köszönetnyilvánítás

Köszönjük az adatfelvételben és -feldolgozásban nyújtott technikai segítséget Krepsz Valériának. A tanulmány a Bolyai János Kutatási Ösztöndíj támogatásával (D. A., Cs. T. G.), az Innovációs és Technológiai Minisztérium ÚNKP-21-3 (J. K.) és ÚNKP-21-5 (D. A., Cs. T. G.) kódszámú Új Nemzeti

Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott támogatásával, valamint az Innovációs és Technológiai Minisztérium Nemzeti Kutatási, Fejlesztési és Innovációs Alapból nyújtott támogatásával (ELTE TKP2020-IKA-06 sz. támogatói okirat) és az FK128814-es pályázat számára nyújtott támogatásával készült. Köszönjük az anonim lektorok hasznos megjegyzéseit, javaslatait.

Hivatkozások

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723. doi:10.1109/TAC.1974.1100705.
- van Alphen, P. M., & Smits, R. (2004). Acoustical and perceptual analysis of the voicing distinction in dutch initial plosives: the role of prevoicing. *Journal of Phonetics*, *32*, 455–491.
- Bartoń, K. (2022). Mumin: Multi-model inference. r package version 1.46.0. <https://cran.r-project.org/web/packages/MuMIn/index.html>.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.
- Boersma, P., & Weenink, D. (2022). Praat: doing phonetics by computer. version 6.1.14. <http://www.praat.org/>.
- Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica*, *57*, 3–16.
- Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /ɑ, i/ in English. *Journal of the Acoustical Society of America*, *117*, 3867–3878.
- Cho, T. (2006). Manifestation of prosodic structure in articulation: Evidence from lip kinematics in English. *Laboratory Phonology*, *8*, 519–548.

- Cho, T. (2015). Language effects on timing at the segmental and suprasegmental levels. In M. Redford (Ed.), *The handbook of speech production* (pp. 505–529). Hoboken, NJ: Wiley-Blackwell.
- Cho, T., Lee, Y., & Kim, S. (2011). Communicatively driven versus prosodically driven hyper-articulation in Korean. *Journal of Phonetics*, *39*, 344–361.
- Cho, T., & McQueen, J. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, *33*, 121–157.
- Cho, T., McQueen, J., & Cox, A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, *35*, 210–243.
- Chodroff, E., & Wilson, C. (2014). Burst spectrum as a cue for the stop voicing contrast in American English. *Journal of the Acoustical Society of America*, *136*, 2762–2772.
- Cohen, A., Collier, R., & t' Hart, J. (1982). Declination: Construct or intrinsic feature of speech pitch? *Phonetica*, *39*, 254–273.
- Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press.
- Deme, A., Bartók, M., Csapó, T. G., Grácsi, T. E., Juhász, K., & Markó, A. (2021). Acoustic and articulatory vowel variation as quality shift and increased variance in anticipatory and carryover vowel-to-vowel coarticulation. In M. Tiede, D. H. Whalen, & V. Gracco (Eds.), *Proceedings of the 12th International Seminar on Speech Production* (pp. 32–35). New Haven, CT, Egyesült Államok: Haskins Press.
- Deme, A., Bartók, M., Csapó, T. G., Grácsi, T. E., Juhász, K., & Markó, A. (2022). A magánhangzók centralizációja és produkciós homogenitása az előrefelé és hátrafelé ható magánhangzók közti koartikulációban – artikulációs

- és akusztikai adatok. *Általános Nyelvészeti Tanulmányok XXXIV*, (pp. 15–49).
- Deme, A., Bartók, M., Csapó, T. G., Grácsi, T. E., Juhász, K., & Markó, A. (kézirat). The effect of pitch-accent on the acoustic and articulatory variability of vowels: a real-word EMA study.
- Deme, A., Bartók, M., Grácsi, T. E., Csapó, T. G., Juhász, K., & Markó, A. (megj.). A koartikulációs ellenállás és agresszió hangsúlyos helyzetű magánhangzókbán: artikulációs és akusztikai adatok. *Nyelvtudományi Közlemények*, 118. doi:10.15776/NyK/2022.118.
- Deme, A., Bartók, M., Grácsi, T. E., Csapó, T. G., & Markó, A. (2019a). A mondathangsúly hatása a magánhangzók megvalósulásának változatosságára. *Nyelvtudományi Közlemények*, 115, 199–232.
- Deme, A., Bartók, M., Grácsi, T. E., Csapó, T. G., & Markó, A. (2019b). V-to-V coarticulation induced acoustic and articulatory variability of vowels: The effect of pitch-accent. In G. Kubin, & Z. Kačič (Eds.), *The 20th Annual Conference of the International Speech Communication Association (Interspeech 2019)* (pp. 3317–3321). Graz, Austria: International Speech Communication Association (ISCA).
- Edwards, J. E., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America*, 89, 369–382.
- Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America*, 84, 115–123.
- Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, 1018, 3728–3740.

- Gósy, M. (1997). Semleges magánhangzók a magyar beszédben. *Magyar Nyelvőr*, 121, 9–19.
- Gósy, M. (2000). A [p, t, k] mássalhangzók zöngeskedési ideje. *Magyar Nyelvőr*, 124, 195–204.
- Gósy, M., & Beke, A. (2010). Magánhangzó-időtartamok a spontán beszédben. *Magyar Nyelvőr*, 134, 140–165.
- Gósy, M., & Krepsz, V. (2016). A magánhangzó nyúlása a szótagszám és a mondatpozíció függvényében. *Beszédkiutató*, 2016, 59–88.
- Gósy, M., & Ringen, C. O. (2009). Everything you always wanted to know about VOT in Hungarian. (előadás). In *International Conference on the Structure of Hungarian 2009*. Budapest. URL: https://www.researchgate.net/publication/253974237_Everything_You_Always_Wanted_to_Know_About_VOT_in_Hungarian.
- Grácsi, T. E., & Kohári, A. (2012). A zöngeskedési idő egy módszertani kérdés függvényében. In A. Markó (Ed.), *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngeskedési időig* (pp. 228–248). Budapest: ELTE Bölcsészettudományi Kar és MTA Nyelvtudományi Intézet.
- Harrison, E., & Pius, R. (2021). R for health data science. (electronic version of the healthyr book published by chapman & hall/crc. https://argoshare.is.ed.ac.uk/healthyr_book/.
- Ip, M. H. K., & Cutler, A. (2022). Juncture prosody across languages: Similar production but dissimilar perception. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 13, 1–49.
- de Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, 97, 491–504.

- Kassai, I. (1979). *Időtartam és kvantitás a magyar nyelvben. Nyelvtudományi Értekezések 112*. Budapest: Akadémiai Kiadó.
- Keating, P. A., Cho, T., Fougeron, C., & Hsu, C. (2004). Domain-initial strengthening in four languages. In J. Local, R. Ogden, & R. Temple (Eds.), *Laboratory phonology, Vol. IV: Phonetic interpretation* (pp. 145–163). Cambridge, UK: Cambridge University Press.
- Kohári, A. (2018). *Időztési mintázatok a magyar beszédben*. Budapest: ELTE Eötvös Kiadó.
- Kohári, A., & Markó, A. (2015). A megnyilatkozás végének jelzése: temporális szerkezet és zöngeminőség felolvasásban. *Beszédkutatás, 23*, 35–52.
- Kovács, M. (2001). *Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában*. Ph.D. thesis Debrecen. URL: <https://dea.lib.unideb.hu/dea/bitstream/handle/2437/79756/ertekezes.pdf?sequence=6&isAllowed=y>.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*, 1–26. doi:10.18637/jss.v082.i13.
- Lenth, R. (2020). Emmeans: Estimated marginal means, aka least-squares means. <https://CRAN.R-project.org/package=emmeans>. R package version 1.4.4.
- Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H & H theory. In W. Hardcastle, & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 403–439). Dordrecht: Kluwer Academic Publishers.
- Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word, 20*, 384–422.
- Mády, K., Bombien, L., & Reichel, U. (2008). Is Hungarian losing the vowel quantity distinction? In R. Sock, S. Fuchs, & Y. Laprie (Eds.), *Proceedings of 8th*

- International Seminar on Speech Production* (pp. 449–453). Strasbourg: INRIA. URL: https://phon.nytud.hu/mady/pub/mady_bombien_reichel.pdf.
- Mády, K., Reichel, U., & Szalontai, Á. (2017). A prozódiai prominencia (nem-)jelölése a németben és a magyarban. In *Általános Nyelvészeti Tanulmányok XXIX* (pp. 77–98).
- Magdics, K. (1965). *A magyar beszédhangok akusztikai szerkezete. Nyelvtudományi Értekezések 49*. Budapest: Akadémiai Kiadó.
- Magdics, K. (1966). A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények*, 68, 125–139.
- Markó, A., Bartók, M., Csapó, T. G., Deme, A., & Grácsi, T. E. (2019). The effect of focal accent on vowels in Hungarian: articulatory and acoustic data. In *Proceedings of the 19th ICPHS* (pp. 2715–2719).
- Markó, A., Bartók, M., Grácsi, T. E., Deme, A., & Csapó, T. G. (2018). Prominence effects on Hungarian vowels: A pilot study. In *9th International Conference on Speech Prosody* (pp. 868–872).
- Markó, A., Juhász, K., Bartók, M., Csapó, T. G., Grácsi, T. E., & Deme, A. (2022). Magyar magánhangzók artikulációs és akusztikai jellemzői a fonetikai pozíció függvényében álszavakban. In *Általános Nyelvészeti Tanulmányok XXXIV* (pp. 51–79).
- Neuberger, T., & Beke, A. (2017). A zárfelpattanás spektrális jegyei a hosszúsági oppozíció függvényében. *Beszédkutató*, 25, 7–23.
- Pickett, J. M. (1999). *The Acoustics of Speech Communication*. Boston: Allyn and Bacon.
- Pierrehumbert, J. (1999). Prosody and intonation. In R. A. Wilson, & F. Keil (Eds.), *The MIT encyclopedia of cognitive sciences* (pp. 479–482). Cambridge, MA: MIT Press.

- Powell, M. J. D. (2009). *The BOBYQA algorithm for bound constrained optimization without derivatives*. *Cambridge NA Report NA2009/06* volume 2009. Cambridge: University of Cambridge.
- R Core Team (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: <https://www.R-project.org/>.
- Reichel, U. D. (2012). Perma and balloon: Tools for string alignment and text processing. In *Proceedings of Interspeech 2012* (p. 346).
- Schiel, F. (1999). Automatic phonetic transcription of nonprompted speech. In *Proceedings of the International Congress of Phonetic Sciences* (pp. 607–610).
- Singmann, H., & Kellen, D. (2019). An introduction to mixed models for experimental psychology. In D. H. Spieler, & E. Schumacher (Eds.), *New Methods in Cognitive Psychology* (pp. 4–31). PsychologyPress.
- Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, *17*, 3–45.
- Sundara, M. (2005). Acoustic-phonetics of coronal stops: A cross-language study of Canadian English and Canadian French. *Journal of the Acoustical Society of America*, *118*, 1026–1037.
- Szalontai, Á., Wagner, P., Mády, K., & Windmann, A. (2016). Teasing apart lexical stress and sentence accent in Hungarian and German. In C. Draxler, & F. Kleber (Eds.), *Tagungsband 12. Tagung Phonetik und Phonologie im deutschsprachigen Raum (P&P 12)* (pp. 216–219). München: Ludwig-Maximilians-Universität.
- Tabain, M., Breen, G., Butcher, A., Jukes, A., & Beare, R. (2016). Stress effects on stop bursts in five languages. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *7*, 1–23.

- Vayra, M., & Fowler, C. (1992). Declination of supralaryngeal gestures in spoken Italian. *Phonetica*, *49*, 48–60.
- Vicenik, C. (2010). An acoustic study of Georgian stop consonants. *Journal of the International Phonetic Association*, *40*, 59–92.
- White, L. (2002). *English Speech Timing: A Domain and Locus Approach*. Edinburgh: The University of Edinburgh. URL: <https://era.ed.ac.uk/handle/1842/23256> PhD Thesis.
- White, L., Benavides-Valera, S., & Mády, K. (2020). Are initial-consonant lengthening and final-vowel lengthening both universal word segmentation cues? *Journal of Phonetics*, *81*. URL: <https://doi.org/10.1016/j.wocn.2020.100982>.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.