

Az /i:/ artikulációs és akusztikai sajátosságai harmonikusan és antiharmonikusan toldalékolódó tövekben

Markó Alexandra^{1,2}, Bartók Márton^{2,3}, Csapó Tamás Gábor^{4,2},
Grácsi Tekla Etelka^{5,2}, Deme Andrea^{1,2}

¹ Eötvös Loránd Tudományegyetem

² MTA–ELTE „Lendület” Lingvális Artikuláció Kutatócsoport

³ Eötvös Loránd Tudományegyetem Nyelvtudományi Doktori Iskola

⁴ Budapesti Műszaki és Gazdaságtudományi Egyetem

⁵ Nyelvtudományi Intézet, Fonetikai Osztály

The aim of our study is to analyse the articulatory and acoustic characteristics of /i:/ in Hungarian monosyllabic harmonic and antiharmonic stems. The words were produced by 12 speakers (i) in isolation and (ii) in sentence-initial position, where they were followed by front and back vowels. The experiment was carried out using electromagnetic articulography, the acoustic signal was recorded using a head-mounted omnidirectional condenser microphone. The horizontal position of four receiver coils (tongue tip, tongue blade, tongue body 1, tongue body 2) and the first and second formant values were obtained at the temporal midpoint of the target vowels. The results showed that neither the horizontal positions of the receivers nor the formant values varied as a function of the harmonicity of the stem in either the isolated or the coarticulated setup.

Keywords: transparent vowel, harmonic stem, antiharmonic stem, electromagnetic articulometry, acoustic analysis

Kulcsszavak: áttetsző magánhangzó, harmonikus tö, antiharmonikus tö, elektromágneses artikulometria, akusztikai elemzés

Bevezetés

A magánhangzó-harmónia olyan fonológiai jelenség, amelyben egy adott tartományon (pl. tö vagy szóalak) belül a magánhangzók valamely sajátosságukban hasonlók egymáshoz. A jelenség eredetét sokan a magánhangzók között (a köztes mássalhangzón áthatolva) érvényesülő koartikulációs hatás grammatikalizálódásaként értelmezik (vö. pl. Beňuš–Gafos 2007).

A magyar előlségi harmónia az a jelenségekör, amelyben a tö és a szóalak magánhangzói előlségüket tekintve hasonlóak egymáshoz (azaz leegyszerűsítve vagy mind elől, vagy mind hátul képzettek). A magánhangzónak a tón belüli hasonlóságát *hangrend*nek nevezzük, és ennek típusai között szokás megkülönböztetni (vö. Nádasy–Siptár 1994/2006) hagyományos elnevezéssel 1. „magas” hangrendű töveket, ahol minden magánhangzó elülső vagy palatális, például *teher*, *bicikli*, *szekrény*, *ötvös*, *ünnep*; 2. úgynevezett „mély” hangrendű töveket, ahol minden magánhangzó hátulsó vagy veláris, például *ajtó*, *hullám*, *doboz*, *matrac*; végül 3. „vegyes” hangrendűeket, ahol a tö mind elülső/palatális (ma-

gas), mind hátsó/veláris (mély) hangzókat tartalmaz, például *ventilátor, hokedli, fiók, deszka*.

Fel kell hívnunk a figyelmet arra, hogy bár hagyományosan használatos a *magas*, illetve a *mély* elnevezés az elől, illetve hátul képzett magánhangzókra, ezek nem az artikulációra utaló jelzők (nincs közülük a nyelvemelkedés fokához vagy az állkapocs nyíltságához, még ha ezekkel a félrevezető „mély” és „magas” elnevezés miatt össze is téveszthetők). A magánhangzók elől vagy hátul képzett volta, azaz a vízszintes nyelvhelyzet akusztikai következménye a második formáns (F_2) frekvenciaértékében érhető tetten (ez kiegészül még az ajakműködés hatásával is). Az elől és hátul képzett magánhangzók perceptuális elkülönítése ugyanakkor nem pusztán az F_2 értéke alapján megy végbe. A hátul képzett magánhangzók esetében ugyanis ez első és a második formáns is relatíve alacsony frekvencián jelentkezik, ezért elbillen a spektrális egyensúly lefelé (és a felsőbb tartományokon arányaiban kevesebb lesz az energia), különösen mert az egymáshoz közelebbi rezonanciák (itt az F_1 és az F_2) erősítik is egymást, az F_3 éppen ezen okból kisebb amplitúdóerősítést okoz. Az egymáshoz közeli első két formáns integrálódik az észleletben. A palatálisoknál ezzel szemben kiegyenlítettebb a spektrum: ott az első két formáns távolabb van, nem erősítik egymást, ugyanakkor a harmadik és a második formáns erősíti egymást, tehát kiegyenlítettebb az amplitúdóeloszlás, a spektrális egyensúly nem billen el. Az elől képzetteknel az F_1 , illetve a kis különbség miatt integrálódó F_2 - F_3 góc két külön észlelt prominencia, nem egyetlen csúcs, mint a velárisoknál (Syrdal–Gopal 1986, Sussman 1990). Így tehát auditív alapon minősítjük „mély”-nek a hátul képzett magánhangzókat és „magas”-nak az elől képzetteket.

Visszatérve a harmónia jelenségköréhez, ha a tőhöz két- vagy többalakú toldalékok kapcsolódnak, ezek előlség tekintetében (háromalakúság esetén kerekességüket tekintve is) harmonizálnak a tő utolsó szótagjának magánhangzójával – ezt a jelenségcsoportot nevezzük *hangrendi illeszkedésnek*. Azaz: ha az utolsó magánhangzó elülső, a toldalékbeli magánhangzó is elülső lesz, például *kilincsen, szekrényben, ötvössel, ültet, ünnepelt*. Ha a tő utolsó magánhangzója hátsó, akkor a toldalékbeli magánhangzó is az lesz, például *piacon, dobozban, hullámos; fiókban, ventilátorral*. Mindez független attól, hogy a tő egészének hangrendje milyen, azaz attól, hogy az utolsó szótag előtti magánhangzók elülsők vagy hátsók.

A magyar előlségi harmóniát sokat vizsgálják a nemzetközi irodalomban is. Különösen nagy érdeklődésre tart számot az úgynevezett áttetsző / ϵ e: i i:/ magánhangzók viselkedése, amelyek fonetikai előlségük ellenére bizonyos esetekben nem elülsőként vesznek részt a harmóniában (vö. Nádasy–Siptár 1994/2006). Ezek az *áttetsző (átlátszó, semleges)* szegmentumok onnan kapták az elnevezésüket, hogy ők maguk a tő utolsó magánhangzójaként nem határozzák meg az illeszkedő toldalék magánhangzójának minőségét, hanem mintegy átengedik magukon a harmóniát (mintha ott se lennének), például *haver-kodik*,

kadét-hoz, kocsi-val, panír-oz (félkövérrel az áttetsző magánhangzókat jelöltük). Ezekben a szóalakokban rendre azt láthatjuk, hogy az /ε e: i i:/ utolsó szótagbeli magánhangzóként nem jelöli ki a toldalékbeli magánhangzó minőségét az elöltség tekintetében, hanem ezt a szerepet átengedi a megelőző magánhangzónak: *haverkodik, kadéthoz, kocsival, paníroz* (itt félkövérrel a harmóniában részt vevő magánhangzókat jelöltük). Annak ellenére van ez így, hogy az /ε e: i i:/ magánhangzókat fonetikailag elől képzettnek tartjuk, és sok esetben fonológiai viselkedésük is ennek megfelelő, az előlségi harmónia tekintetében is, például *kézbén, kerekség, hiszem, ívet*.

Az /ε e:/ az előlségi harmónia szempontjából az esetek egy részében elülsőként, azaz harmonikusan viselkedik, más esetekben azonban áttetszőként; az /i i:/ pedig mindig áttetszőként. (Különösen az /ε/ esetében találunk nagymértékű ingadozást is – lásd például Patay és munkatársai (megj.) vizsgálatát –, de erre a jelen tanulmányban nem térünk ki.)

Jó néhány olyan lexémával is találkozunk ugyanakkor, amelyben az áttetszőnek tekintett és aktuálisan áttetszően is viselkedő magánhangzók valamelyike a tő egyetlen magánhangzója (két szótagú példa is előfordul, bár igen ritka), azaz nincs előtte olyan nem elülső magánhangzó, amely irányítaná a harmóniát az /ε e: i i:/ áttetsző viselkedése esetén. Ezeket *antiharmonikus töveknek* szokás nevezni. Ilyen tövek a *cél* (vö. *céllal*), a *híd* (vö. *hídon*), a *sír* (vö. *sírok* főnévként és igeiként is) vagy a *derék* (vö. *derékba*). Mégsem minden tő tartozik ebbe a csoportba, amely hasonló magánhangzós szerkezetet mutat, hiszen harmonikusan toldalékolódnak a következők: *vél* (vö. *vélem*), *hír* (vö. *híres*), *kerék* (vö. *kerékbe*). Sőt, bár ritkán, de találunk olyan homofón töveket is, amelyek mindkét toldalékolási mintázatot mutatják, és ez attól függően váltakozik, hogy az adott szóalak igei vagy főnévi funkciót tölt be, mint a *nyír*_{ige}-ok, de *nyír*_{főnév}-ek; illetve a *szív*_{ige}-om, de *szív*_{főnév}-em. Az igei funkcióban ezek a szavak antiharmonikusan, a főnévi funkciójukban pedig harmonikusan toldalékolódnak. (A magyar szóalományban ezek tipikus mintázatnak számítanak, hiszen az /i:/-t tartalmazó egy szótagú igeik igen nagy arányban toldalékolódnak hátulsóként, míg az /i:/-t tartalmazó névszók az esetek döntő többségében harmonikusan toldalékolódnak, vö. pl. Blaho–Szeredi 2013.)

Mind korpuszelemzések, mind más (pl. kérdőíves) vizsgálatok készültek már a magyar előlségi harmóniának és általában a magánhangzó-harmóniának a nyelvhasználatban megvalósuló mintázatai feltérképezésére és magyarázatára, pl. Hayes et al. (2009), Zuraw–Hayes (2017), Patay et al. (megj.), Rebrus–Törkenczy (2019). Ugyanakkor a magyar elülső kerekítetlen magánhangzók áttetszősége különösen népszerű kutatási téma a nemzetközi fonológiában is, mert – ahogyan Rebrus és Törkenczy (2019) írja: „kifejezetten alkalmas fonológiai elméleti elképzelések ütköztetésére” (233). Fonetikai jellegű vizsgálatokból ugyanakkor viszonylag kevés van. Ezek közül feltehetően a leggyakrabban idé-

zett munka Beňuš és Gafos (2007) artikulációs kísérlete, amely egyébként szintén a fonológiai tanulságai és hozadéka miatt vált ismertté. Ennek a kísérletnek az eredményei a magyar áttetsző magánhangzókra vonatkozóan azt mutatták, hogy valamivel előrébb van a nyelv a szájüregben, ha a magánhangzó elülsőként (pl. *szél*) vesz részt a harmóniában, azaz a tő harmonikusan toldalékolódik, míg hátrébb helyezkedik el a nyelv a szájüregben, ha ugyanazon magánhangzó hátulsóként (pl. *cél*) vesz részt a harmóniában, azaz antiharmonikusan toldalékolódik. Ráadásul ezt izolált (nem toldalékolt, nem koartikulálódott) ejtésben is így találták (a kísérleteket alább részletesebben is bemutatjuk). Beňuš és Gafos (2007) ezt azzal magyarázta, hogy a toldalékolt alakokban és ezen keresztül áttételesen a toldalékolatlanokban is az esetlegesen köztes mássalhangzókon is átható magánhangzók közötti visszafelé ható koartikuláció, azaz a követő magánhangzó elülső vagy hátulsó volta van hatással az áttetsző magánhangzó ejtése közben a nyelv helyzetére a következők szerint. A követő magánhangzó előlsége esetén az áttetsző magánhangzó ejtésekor a nyelv előrébb van a szájüregben, mint akkor, ha a követő magánhangzó hátulsó. Az antiharmonikus tövek a többalakú toldalékok közül azonban csak nem elülső magánhangzót tartalmazó toldalékkal fordulnak elő, ami koartikulációsan hat a tőmagánhangzóra. Ezt a hatást örökli a toldalékolatlan alak is, és a beszélők ezt a mintázatot sajátítják el. A szerzők ezt a konklúziójukat a lexikális tárolás exempláralapú modelljeivel hozzák összefüggésbe: a magyar beszélők azért tárolnak és reprodukálnak finom fonetikai különbségeket a harmonikusan és antiharmonikusan toldalékolódó tövekben megjelenő, fonológiailag azonos magánhangzók között, mert az exemplárok (példányok) alapján tárolják a koartikulációs jellemzőket is, amelyek a különféle alakokban jelentkeznek.

A szerzők idézik egyebek közt Recasens (1999) munkáját, amely szerint az /i:/ viszonylag stabil mind akusztikai, mind perceptuális tekintetben, azaz a szomszédos hátul képzett magánhangzók hatására igen kismértékű eltérést mutat. Beňuš és Gafos (2007) ezzel a stabilitással magyarázzák, hogy az /i:/ nem realizálódik [u]-ként például a *papír* szóban a megelőző hátul képzett magánhangzó hatására, tehát bár az /i:/ ejtésekor a nyelv hátrébb húzódik az antiharmonikus tövekben a harmonikus tövekbeli helyzetéhez képest, akusztikailag és perceptuálisan mégis stabil marad, azaz megőrzi palatális minőségét. (Megjegyezzük, hogy Beňuš és Gafos 2007 a vizsgált anyagon sem formánselemzést, sem percepció tesztelést nem végzett, ezen érveik más szerzők korábbi, többnyire más nyelveken és különféle anyagokon végzett akusztikai és percepció vizsgálatainak eredményeire építenek.)

A szerzők szerint az /i:/ artikulációs variabilitása és akusztikai-perceptuális stabilitása a kvantális elméletből (Stevens 1989) ismert tények alapján magyarázható, amelyek szerint az artikulációbeli eltérések mértéke és az ezek nyomán az akusztikumban bekövetkező különbség nem áll egymással egyenes arányban,

azaz a beszédprodukción működések és az akusztikai szerkezet jellemzői közötti kapcsolat nem lineáris. Az artikulációban egy apró eltérés jelentős akusztikai különbséget idézhet elő, ha az egy ún. „kritikus régió”-ban történik; és fordítva, nagyobb artikulációs eltérések esetén sem feltétlenül tapasztalunk különbséget az akusztikai szerkezetben, ha ezek az eltérések kritikus régiókon kívül valósulnak meg. Az elmélet szerint ugyanígy nem lineáris a kapcsolat az akusztikai szerkezet eltérései és a percepcióbeli különbségészlelet között sem, azaz előfordulhat, hogy két hangjelenség nagymértékben különbözik egymástól akusztikai szempontból, és mégsem észlelhető közöttük különbség a hallgató számára; vagy fordítva, az is lehetséges, hogy egy kisebb akusztikai különbség nagy eltérést indukál az észlelt hangjelenségek tekintetében. Erre a nem lineáris összefüggésre utal a kvantális jelző: az eltérések ugrásszerűen képeződnek le (az artikuláció eltérései az akusztikumban, az akusztikai eltérések a percepcióban), nem pedig fokozatosan (Stevens 1989).

Beňuš és Gafos (2007) szerint az áttetsző /i/ tehát perceptuálisan stabilabb és artikulációsan variábilisabb, mint akár az ugyancsak felső nyelvállású, elől képzett, de kerekített /y/, ahol a kerekítettség miatt a nyelv hátrahúzásának lehetősége limitáltabb a perceptuális azonosság fennmaradása mellett (vö. pl. Stevens 1989; maguk a szerzők az /y/-t nem vizsgálták). Hasonlóképp, a nyelv hátrahúzásának a lehetősége az alsó nyelvállású, elől képzett, kerekítetlen /ε/ ejtésekor is limitáltabb, mint az /i/ esetében. Úgy vélik, ezzel magyarázható, hogy bár felső és elülső képzésű, az /y/ nem viselkedik áttetszően a harmónia szempontjából, illetve az, hogy az antiharmonikus tövekben a legnagyobb arányban /i:/-t találunk, /ε:/-t ritkábban, /ε/-t azonban nem.

Beňuš és Gafos (2007) kísérletében két nő és egy férfi vett részt, mindhárman fiatal budapesti beszélők. A szerzők kétféle artikulációs elemzést végeztek, az egyikben elektromágneses artikulometriát (azaz ún. elektromágneses artikulográfot, röviden EMA-t) alkalmaztak, amely a képzőszervek felületének egyes pontjaira rögzített szenzorok segítségével teszi követhetővé a szervek helyzetét. Ebben a kísérletben mindhárom beszélő részt vett, közülük kettő esetében két szenzort helyeztek (előlről, azaz az ajkaktól, hátrafelé haladva) a nyelvtestre (tongue body: TB1, TB2) és egyet a nyelvhátra (tongue dorsum: TD), a harmadik beszélő esetében pedig egy nyelv szenzort helyeztek a nyelvhegyre (tongue tip: TT), egyet a nyelvtestre és egyet a nyelvhátra. (Megjegyezzük, hogy a tanulmányból számunkra nem világos, hogy a nyelvtest – nyelvhat megkülönböztetés mit takar pontosan.) Az egyik beszélő esetében ultrahangos képalkotó eljárást is alkalmaztak, amely a nyelv midszagittális felszíni kontúráját (mint folytonos vonalat) teszi láthatóvá. A szerzők a nyelv (illetve a TB, TD szenzorok) helyzetét abban az időpillanatban vizsgálták, amikor a nyelv a célmagánhangzó ejtése közben a szájüregben legelőrébb helyezkedett el. Az adatokat a tő harmonikus, illetve antiharmonikus típusa szerint vetették össze.

A kísérletekben Beňuš és Gafos kétféle célszóstruktúrával dolgozott. Az egyik esetben a célmagánhangzó két szótagos tövek második szótagjában szerepelt, és egy szótagos (kétalakú) toldalék követte, pl. *bili-vel* és *buli-val*. A célszavak hordozó mondatban szerepeltek (lásd alább). Tekintettel arra, hogy ebben az anyagban a célmagánhangzók vagy csak elülső, vagy csak hátulsó magánhangzók között álltak, mint az várható is volt, a célmagánhangzó ejtése közben a nyelv helyzete a hangkörnyezet függvényében eltért: az elülső magánhangzók között a célmagánhangzóban mért nyelvhelyzet előrébb volt detektálható a szájüregben, mint a hátul képzett magánhangzók között.

A másik kísérleti anyagban a célszavak egy szótagosak voltak, és nem követték őket toldalék, de ugyancsak hordozó mondatba voltak ágyazva (lásd alább). A szerzők – ahogyan írják – megkíséreltek egymáshoz a lehető leghasonlóbb szópárokat összeállítani, amelyek lehetőleg csak az illeszkedésben való részvételüket tekintve (harmonikus vagy antiharmonikus) térnek el. A következő nyolc szópárt használták a kísérletben (a H és az AH az indexben a töveknek az idézett szerzők szerinti harmonikus vagy antiharmonikus viselkedésére utal): $vív_{AH} - ív_H$, $híd_{AH} - íz_H$, $ír_{AH} - hír_H$, $víg_{AH} - míg_H$, $síp_{AH} - cím_H$, $nyit_{AH} - hisz_H$, $cél_{AH} - szél_H$, $héj_{AH} - éj_H$. Első ránézésre is látszik, hogy a szótagstruktúra nem minden pár tagjai között azonos (CVC és VC, pl. $vív_{AH}$ és $ív_H$), aminek következtében a célmagánhangzók eleve eltérhetnek, hiszen a mássalhangzós szótagkezdet egyúttal azt is jelenti, hogy az azt követő célmagánhangzón érvényesül a mássalhangzónak a (mássalhangzó képzési helyének és zöngéességnek függvényében többé-kevésbé eltérő) koartikulációs hatása (vö. Hillenbrand et al. 2001), míg a magánhangzóval kezdődő szótagokban ilyen hatás nem jelentkezik. Az esetek többségében ugyanakkor a /h/ váltakozik a mássalhangzóhiánnyal ($híd_{AH} - íz_H$, $ír_{AH} - hír_H$, $héj_{AH} - éj_H$), ami bizonyos értelemben jó megfeleltetés, tekintve, hogy a /h/ ejtése nem igényel orális akadályt. Ugyanakkor az is feltehető, hogy – éppen mivel a /h/ ejtése nem kíván meg orális akadályt – a beszélők képzőszervei már a szókezdő /h/ létrehozása közben felvették a követő magánhangzó ejtéséhez szükséges artikulációs konfigurációt, ami akár oda is vezethetett, hogy a nyelv helyzete már a /h/ ejtése közben elérte a szájüregben a legelülső pozíciót. Éppen ezért, mivel a szerzők csak a célmagánhangzó időtartamán belül vizsgáloztak, ezekben az esetekben lehetséges, hogy szándékaikkal ellentétben nem a valóban legelülső helyzetet vették figyelembe.

Meg kell jegyeznünk azt is, hogy a *míg* legfeljebb metanyelvi használatban (pl. „*míg*”-ek vannak a mondatban) kaphat toldalékot.

A nazális-orális mássalhangzós szókezdet és szóvég összevetése az artikulációban szerepet játszó üregek eltérései miatt és ezáltal a nyelv működésének feltehető eltérései miatt ugyancsak kérdéseket vet fel a $víg_{AH} - míg_H$, $síp_{AH} - cím_H$, $nyit_{AH} - hisz_H$ szópárok esetében.

Azt is látjuk, hogy a célmagánhangzót körülvevő mássalhangzók szájüregi képzési helye is eltér néhány esetben, például posztalveoláris /ʃ/ váltakozik alveoláris /ʃ/ -vel, palatális /ɲ/ váltakozik laringális /h/-val, ami kifejezetten problematikus (különösen, ha ezek az adott magánhangzó-környezet bal oldalát képezik), hiszen a környező mássalhangzók képzési helye jelentősen befolyásolja a körülvevő magánhangzók ejtését, különösen az előség mentén (vö. Hillenbrand et al. 2001). Mi több, a kísérletben használt példákban rendre a szájüregben hátrébb képződő mássalhangzók álltak az antiharmonikus tövekben, és a szájüregben előrébb képződők a harmonikusokban, vagyis ezekben az esetekben a mássalhangzó-kontextus önmagában előidézhette a nyelv helyzetének a szerzők által várt különbségét.

Mindezek az észrevételek úgy summázhatók, hogy a nyolc pár közül valóban fonetikailag kontrollált és kiegyenlített kontextusú célszavak csak egyetlen esetben (*cél_{AH} – szél_H*) valósultak meg.

A célszavakat a szerzők olyan hordozó mondatba ágyazták, amelyben a célszó kétszer, de két eltérő helyzetben szerepelt: „Azt mondom, hogy ____, és elismétlem azt, hogy ____ még egyszer.” A két helyzetben megjelenő célszavakat együtt vizsgálták a szerzők. Összesen négy ismétlésben rögzítették a stimulusokat, mind a három szótagos, mind az egy szótagos célszavakat.

Az EMA mind térben, mind időben igen pontos mérést tesz lehetővé, ugyanakkor a nyelv alakjáról (vagyis arról, hogy a mérési pontok között, illetve azokon túl, elsősorban a nyelv hátulsó részén mi történik) nem ad információt. Erre a holisztikus megközelítésre a nyelvultrahang ad módot. Beňuš és Gafos (2007) egy beszélővel ultrahangfelvételt is készítettek. Az idézett szerzők az EMA-adatokat úgy tesztelték, hogy két beszélő esetében három szenzor (TD, TB2 és TB1), egy beszélő esetében pedig két szenzor (TD és TB) vízszintes helyzete szerepelt függő változóként. Az adatokat az egyes résztvevők esetén külön-külön elemezték, mivel a nyelv és a szájüreg mérete, alakja, valamint a szenzorok elhelyezése is egyénenként eltérő, illetőleg egyenként vetették össze ANOVA-val és Tukey-féle post hoc teszttel.

Tovább nem részletezve az idézett tanulmányt, az ismertetett anyag és módszer alapján (mint láttuk) Beňuš és Gafos (2007) azt az eredményt kapta, hogy hátrébb van a nyelv a szájüregben, ha a magánhangzó hátulsóként, azaz áttetsző magánhangzóként vesz részt a magánhangzó-harmóniában. Ezt mind a toldalékolt, így biztosan koartikuláció hatása alatt álló célszavakban, mind pedig az „izolált” – nem toldalékolt és a szerzők szerint koartikulációs hatást el nem szenvedő esetben (ami viszont a hordozó mondat jelenléte miatt nem feltétlenül volt így) – ejtésben is így találták.

Blaho és Szeredi (2013) akusztikai elemzést végzett annak ellenőrzésére, hogy a nyelv eltérő helyzete a harmonikus és az antiharmonikus tövekben, amit Beňuš és Gafos (2007) dokumentált, az akusztikumban is kimutatható eltérést

idéz-e elő. A kísérletben 12 beszélő vett részt, két eltérő nyelvjárásterületről. A kísérlet anyagaként hét szópárt használtak, túlnyomórészt a Beňuš és Gafos (2007) által is alkalmazottakat (három pár egyik tagját más szóra cserélték, ez egy esetben nagyobb mértékű fonetikai megfelelést idézett elő a párok tagjai között). Blaho és Szeredi (2013) is alkalmaztak hordozómondatot (bár a korábbitól némiképp eltérő: „Azt mondtam, ____, megismétlem, ____”), és ebben is kétszer jelent meg a célszó, itt azonban mindkét pozíció hasonlóbb prozódiai helyzetet reprezentált. Az akusztikai elemzést a magánhangzók időbeli közép-pontjában végezték el. Több paramétert is vizsgáltak, de itt csak a második formánshoz kapcsolódó eredményekre térünk ki. Ha volt is eltérés az akusztikai adatokban, az nem volt szignifikáns, illetve nem érte el az észlelhetőség küszöbét. Ráadásul egyes esetekben a formánsadatok ellenkező tendenciáról árulkodtak, mint a Beňuš és Gafos (2007) vizsgálatában kapott artikulációsak, azaz itt olykor a harmonikus (azaz elülső magánhangzóval toldalékolódó) töben találtak a szerzők alacsonyabb F_2 -értéket, ami a hátrébb képzettségre (és/vagy az intenzívebb ajakműködés következtében előrébb elhelyezkedő ajaknyílásra) utal.

Igaz ugyan, hogy Blaho és Szeredi (2013) nem találtak akusztikai eltérést az eltérő fonológiai viselkedésű magánhangzók között, és az ebben a tanulmányban idézett másik kísérlet (Szeredi 2012) alapján percepció eltérés sem volt felfedhető ezek között. Ugyanakkor ismerünk olyan elméleteket és kísérleti eredményeket, amelyek alapján akkor is feltételezhető, hogy az artikulációs eltérések tárolódnak a lexikonban, ha ezt akusztikai és/vagy percepció kísérleti adatok nem támasztják alá. Ennek kapcsán ide idézhetjük egyrészt a tükroneuronokkal kapcsolatos vizsgálatok eredményeiből levont azon következtetéseket (vö. pl. Théoret–Pascual-Leone 2002), amelyek szerint azért értjük meg mások, így a beszédpartnerek cselekvéseit, azaz például a nyelvi viselkedésüket, mert e cselekvések megfigyelése ugyanazokat az idegsejteket aktiválja, mint ezeknek a cselekvéseknek a kivitelezése. Másrészt az említett következtetés ugyancsak harmonizál a beszédészlelés motoros elméletében foglalt azon állítással, amely szerint a beszéd észlelése az akusztikai jel transformációján keresztül az artikulációs mozdulatsor reprezentációjának előhívásával megy végbe (Lieberman–Mattingly 1985). Feltételezhetjük tehát, hogy ha nincs is szignifikáns akusztikai eltérés a kérdéses magánhangzók között, maguk az artikulációbeli különbségek áttevődhetnek a koartikulált helyzetből az izolált helyzetben ejtett magánhangzókra, és eltárolódhatnak a mentális lexikonban.

Úgy véljük ugyanakkor, hogy Beňuš és Gafos (2007) kísérlete mégis felvet bizonyos metodológiai kérdéseket, például az alkalmazott nyelvi anyaggal kapcsolatosan, de említhetjük a résztvevők alacsony számát is. Saját kísérletünk anyagát és módszertanát ezért igyekeztünk úgy kialakítani, hogy kiküszöböljük ezeket a hibalehetőségeket.

Mindezek alapján a jelen kísérletben a fentiekől eltérő anyagon és részben eltérő módszerrel vizsgáltuk a következő kérdéseket:

1. A *nyír* és a *szív* a funkció szerint kétféleképpen toldalékolódó alakjai (főnév: elülső toldalék, ige: hátulsó toldalék) izolált ejtésben (azaz nem toldalékolt alakban és nem hordozó mondatban ejtve) mutatják-e akár artikulációsan, akár akusztikailag azt a Beňuš és Gafos (2007) által talált tendenciát, amely szerint az antiharmonikus tövek magánhangzója hátrébb képzett, mint a harmonikus töveké?
2. A kétféle homonim töben előforduló azonos /i:/ magánhangzó (azaz pl. a *szív_H* és a *szív_{AH}*) mutat-e eltérést abban a tekintetben, hogy milyen mértékben befolyásolja az ejtését és akusztikai szerkezetét a követő (egy köztes mássalhangzón túli) magánhangzó koartikulációs hatása? Azaz egy nem elülső követő magánhangzó „hátrébb húzza-e” az adott magánhangzót antiharmonikus esetben (pl. *szív_{AH(ige)} ugyan*), mint harmonikus esetben (pl. *szív_{H(főnév)} ugyan*), ahogyan egy követő elülső magánhangzó (pl. *szív_{AH(ige)} éppen*) „előrébb húzza-e” a kétféle töben ejtett magánhangzók közül a harmonikus töben megvalósultat?

Hipotéziseinket Beňuš és Gafos (2007) eredményeire alapozva az alábbiakban fogalmaztuk meg:

1. A harmonikus fő magánhangzója előrébb képződik, mint az antiharmonikus fő magánhangzója, izolált helyzetben is, és ennek mind artikulációs, mind akusztikai következménye mérhető.
2. A magánhangzók páronként a fentihez hasonló eltérést mutatnak a koartikulációs hatás jelentkezésekor is.

Kísérleti személyek, anyag, módszer

A kísérletben 12 nő vett részt, életkoruk 22 és 37 év között volt (átlag: 26,25 év, szórás: 4,77 év). Mindannyian magyar anyanyelvűek, nincs ismert beszéd- vagy hallásproblémájuk, és egészségesek voltak a felvétel készítésekor. Tájékoztatót kaptak a kísérlet menetéről és az adatkezelésről, majd beleegyező nyilatkozatot írtak alá.

A kísérletben az /i:/ magánhangzót vizsgáltuk egy szótagos, homofón szópárokban, amelyek igei funkciójukban antiharmonikusan, főnévi funkciójukban harmonikusan toldalékolódnak, azaz a célszavak a *nyír_{AH(ige)}*, *nyír_{H(főnév)}*, illetve a *szív_{AH(ige)}* és a *szív_{H(főnév)}* voltak. A célszavakat a számítógép monitorján jelenítettük meg a felvét elkészítés alkalmával.

Kétféle elrendezést teszteltünk. Az izolált elrendezésben nem alkalmaztunk sem toldalékolást, sem hordozó mondatot, a funkció (szófaj) azonosítását a résztvevőknek bemutatott képekkel segítettük elő: a képernyőn a célszó és felette a célszó denotátumát bemutató kép látszott. A résztvevők az egy szótagos szavakat mint önálló megnyilatkozásokat hangosították meg. Az izolált elrende-

zésben kilenc homofón párt alkalmaztunk disztraktorként: *fal_H*, *nyom_H*, *nő_H*, *nyúl_H*, *sír_{AH}*, *fűz_H*, *fej_H*, *vár_H*, *fél_H*, szintén képekkel együtt megjelenítve.

A másik elrendezésben a célszavak kontextusba ágyazottan jelentek meg. A célszavakat itt megnyilatkozás elején, főhangsúlyos helyzetben (ebben a tekintetben tehát az izolált elrendezéssel azonosan, kontrollált módon), de ezúttal követő magánhangzók koartikulációs hatásának kitéve helyeztük el. Bár ebben az elrendezésben a kontextus egyértelműsítette a funkciót, mégis, annak érdekében, hogy a feladat jellege ne változzon a kísérlet alatt, itt is szükségesnek tartottuk a képek közlését, így ebben az esetben ugyancsak látszott (az izolált kontextustól eltérő) egyértelműsítő kép. A célszavakat követő kontextusok a következők voltak: csak elülső magánhangzót tartalmazó *éppen*, illetve csak hátsó magánhangzót tartalmazó *ugyan*. A kontextusos elrendezés hordozó mondatai a következők voltak:

Nyír_{főnév} éppen/ugyan ez a csemete, csak még nem látszik rajta.

Nyír_{ige} éppen/ugyan a fűnyíró, de már régi.

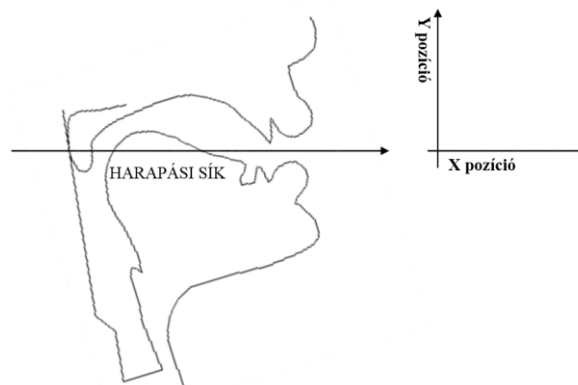
Szív_{főnév} éppen/ugyan, de nem emberé.

Szív_{ige} éppen/ugyan a porszívó, de már régi.

Mindkét elrendezés stimulusait (a szavakat és a mondatokat) ötszöri ismétléssel randomizáltuk, majd blokkokba rendeztük: 22 szó után 8 mondat, majd 22 szó, ismét 8 mondat következett, és így tovább, amíg minden stimulus 5-ször szerepelt. Három így előállított pszeudorandom sorrendet alkalmaztunk a kísérletben.

A vizsgálatot 16 csatornás AG501 típusú elektromágneses artikulográffal (Carstens Medizinelektronik GmbH) végeztük. A képzőszervek egyes pontjaira szenzorokat illesztettünk, amelyek elhelyezkedését és mozgását a transzmitter tekercsek által generált elektromágneses térben lehet követni 1250 minta/s felbontással.

A beszélő feje mozog, és ahhoz, hogy a szenzorok mozgását a beszélőhöz illesztett koordináta-rendszerben lehessen követni, szükség van referenciapontokra, amelyek meghatározzák ezt a koordináta-rendszert. Erre szolgálnak a referenciaszenzorok a fülek mögött és az ornyergén, valamint az úgynevezett „harapási lemez” (biteplate), amely segítségével a harapási vagy okklúziós síkot (az összezárt alsó és felső fogsor közti síkot) vesszük fel, és erre merőlegesen rögzítjük a keresztmetszeti síkot (1. ábra).



1. ábra: A harapási lemez helyzete és az ez alapján felvett koordináta-rendszer

A harapási lemezen négy szenzort helyezünk el, majd a beszélő ráharap a lemezre, így felvehető a harapási sík. További három szenzort alkalmazunk referenciaszenzorként, egyet-egyét a két fül mögött helyezünk el, egyet az ornyergén. Így alkotható meg a beszélőhöz illesztett háromdimenziós koordináta-rendszer, amely a fej elmozdulását is követi. Az artikulátorokra illesztett szenzorok mozgását az ekképpen rögzített háromdimenziós térben tudjuk vizsgálni.

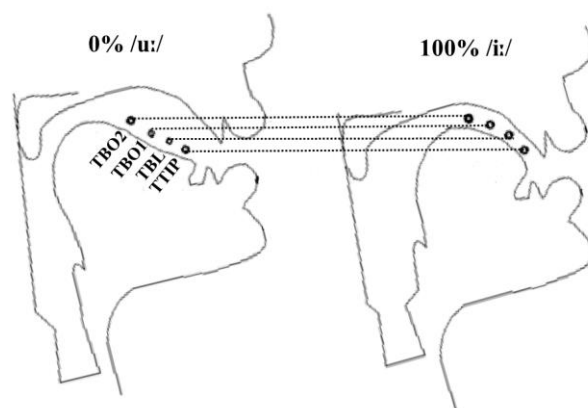
A jelen kísérletben a fej midszagittális vonalában az alsó és felső ajakra és az alsó fogínyre helyeztünk egy-egy szenzort, valamint a nyelv felső felszínére rögzítettünk további négyet (2. ábra). Egy szenzort rögzítettünk a nyelvhegyre (TT), egyet a nyelvperemre (TBL) és kettőt a nyelvhátra (TBO1, TBO2), a szenzorok közötti távolság mintegy 1 cm volt. Vizsgálatunkban a nyelvre helyezett négy szenzor helyzetét elemeztük.

A résztvevők bemondásait akusztikus csatornán is felvettük 44,1 kHz-es mintavételi frekvencián, a szájjughoz helyezett omnidirekcionális, kondenzátoros fejmikrofonnal.

Az artikulációs adatok utófeldolgozását részben a Carstens-szoftver erre szolgáló moduljaival végeztük. A mért adatokat olyan módon forgattuk el, hogy a fej és a harapási sík orientációja minden beszélő esetében egyezzen, a harapási sík a vízszintessel párhuzamos legyen, a keresztmetszeti sík mint a koordináta-rendszer (0,0,0) koordinátájú pontja pedig a metszőfogaknál pozícionálódjon. Ezután a háromdimenziós adatokat a midszagittális metszetet leképező kétdimenziós x - y síkra vetítettük, és az adatkinyeréshez használt Emu-adatbáziskezelővel (Winkelmann et al. 2018) kompatibilis kiterjesztésűvé konvertáltuk a Kölni Egyetem Fonetikai Intézetének (I/L Phonetik) saját készítésű konverterével.

A célmagánhangzókat az akusztikai jel alapján szegmentáltuk és címkéztük félautomatikusan a BAS webszolgáltatás graféma-fonéma konverterének (Reichel 2012) és a MAUS rendszernek (Schiel 1999) a segítségével, majd a

szükséges helyeken kézzel javítottuk a felismertetett hanghatárokat a Praat szoftverben (Boersma–Weenink 2018). Az így felcímkézett magánhangzók időbeli középpontjában nyertük ki a négy nyelv szenzor vízszintes helyzetének (x tengely) adatait (abszolút értékben, a fenti koordináta-rendszer origójához képest), majd az adatfeldolgozás során a szenzorok helyzetét a következőképpen normalizáltuk (Cho 2004 alapján). Az egyes beszélők által bemondott *nyúl* szó magánhangzójának időbeli középpontjában meghatároztuk a nyelv szenzorok vízszintes helyzetét, illetve a bemondott *sír* szó magánhangzójának időbeli középpontjában ugyancsak. Ez a két érték szolgált a továbbiakban a nyelv vízszintes helyzetének (x tengely) megragadására szolgáló skála „végpontjai”-ként úgy, hogy az *ú* esetében az x értéke a 0%-nak, az *í* értéke pedig a 100%-nak felelt meg minden beszélő esetében (2. ábra). Tekintettel arra, hogy a referenciaként szolgáló *sír* *tő* antiharmonikusan toldalékolódik, a hipotézisek alapján azt várhatjuk, hogy a harmonikusan toldalékolódó *nyír*, de különösen a *szív* esetében (ahol a célmagánhangzót megelőző mássalhangzó is előrébb képződik) a nyelv szenzorok helyzete az x tengelyen meghaladja a 100%-ot.



2. ábra: A szenzorok helyzetének normalizálásához használt referenciapontok: az *ú* (*nyúl*, 0%) és az *í* (*sír*, 100%) és az ezek által közrezárt skála az x tengelyen (Cho 2004 alapján), nyelv szenzorokként

Ezzel a módszerrel a beszélők közötti variabilitást kontrollálni tudtuk, és mivel egységesen normalizáltuk az egyes szenzorok helyzetére mint függő változóra kapott adatokat minden beszélő esetében, ezért a vizsgálatban az összes beszélő adatait egyben kezeltük. A módszertan nem korlátozta az elemzés lehetőségeit arra, hogy a beszélőkön belül készítsünk csupán összevetéseket, adataink ennél fogva jobban általánosíthatók.

A jelen kísérletben a változók hatását a szenzorok (normalizált) vízszintes pozíciójának négyelemű szekvenciájára generalizált additív modellekkel (GAM) elemeztük az R programban (R Core Team 2018, Wood 2001). A GAM használata lehetővé teszi több, valamely dimenzióban egymást követő adatpont értékeinek nonlinearis modellezését. A jelen vizsgálatban ezt a statisztikai eljárást az egyes, a nyelven egymást követő nyelv szenzorokon mért pozícióadatok együttes modellezésére alkalmaztuk. Míg Beňuš és Gafos (2007) az egyes szenzorok vízszintes helyzetét külön-külön modellezték, és vetették össze ezeket az értékeket a vizsgált kondíciók között, addig ezzel a módszerrel a négy szenzor együttes pozícióját modellezzük, és vetjük össze a harmóniabeli viselkedés szerint. Ennélfogva az elemzésünk holisztikusabb képet ad a nyelvhelyzet különbségeiről, mint az egyes szenzorok elhelyezkedésének összevetése.

Két modellt állítottunk, egyet az izolált ejtésű, egyet pedig a kontextusban ejtett szavak /i:/ magánhangzóiban mért pozícióértékekre. A modellek az izolált elrendezésben a harmóniabeli viselkedés (harmonikus/antiharmonikus) faktort, a mondatban ejtett megvalósulás esetén pedig a harmóniabeli viselkedés és a kontextus (elülső/hátulsó) faktorokat tartalmazták, valamint random hatásként a beszélő és a faktorok interakcióját. A GAM-ok illesztése az itsadug (van Rij et al. 2017), ábrázolásuk a TidyMV (Coretta 2018) R-csomaggal történt. A vizsgált hatások tesztelésére a két modellt a faktorokat nem tartalmazó minimális modellekkel vetettük össze a `compareML()` funkció segítségével, a két-két modell Akaike-féle információs kritérium-értékét (Akaike information criterion, AIC, Akaike 1974) összehasonlítva. Az Akaike-féle információs kritérium statisztikai modellek relatív jóságának mérésére alkalmas szám, és azt mutatja meg, hogy mennyi információt veszítünk, ha az adott modellel reprezentáljuk az adatokat. Az AIC relatív, csak ugyanazon kimeneti változót tartalmazó modellek összehasonlítására alkalmas mérőszám, melynek kisebb értékei jelentik a jobb modellillesszkedést (kevesebb információvesztést). Az AIC alapján történő modellválasztás során tehát a legkisebb értékű modellt tekintjük a legjobbnak, ugyanakkor Anderson és Burnham (2004) alapján azok a modellek, melyek a legkisebb AIC-értékkel (AICmin) rendelkező modelltől maximum 2 ponttal térnek el ($\Delta AIC_{min} \leq 2$), a legkisebb értékű modellhez hasonló illesszkedésűnek tekinthetők. Így tehát amennyiben a minimális modell AIC-értéke alacsonyabb volt, mint a tesztelendő faktorokat tartalmazóké, vagy kevesebb, mint 2 ponttal tért el azoktól, az azt jelentette, hogy a vizsgált hatások nem javítottak az adatok modellezésén.

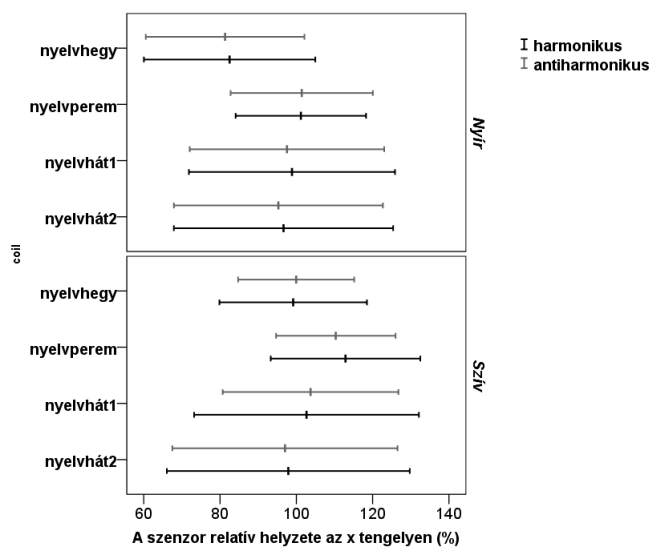
Az artikulációs adatok elemzése mellett szintén a magánhangzók időbeli középpontjában az akusztikai jelen megmértük a második formáns értékét is a Praat szoftverben (Boersma–Weenink 2018). A harmóniabeli viselkedés mint független változó hatását az így kapott formánsértékekre lineáris kevert modellekkel elemeztük az R-ben (Kuznetsova et al. 2015). A kevert modell a harmóni-

abeli viselkedés (harmonikus/antiharmonikus) és a kontextus (elülső/hátulsó) faktorok interakciója mellett az adatközlő változó random hatását tartalmazta.

Eredmények

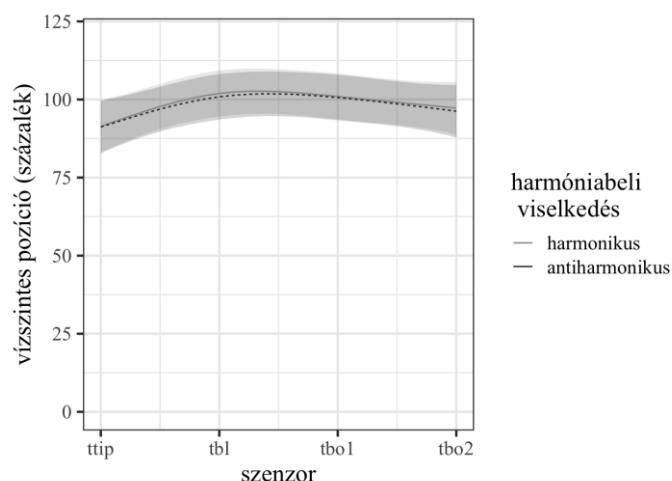
Az áttekintést az artikulációs vizsgálat eredményeivel kezdjük, majd az akusztikai elemzés eredményeit mutatjuk be. Ezekben belül előbb az izolált, majd a kontextusos elrendezés adatait ismertetjük.

Az izolált elrendezésben a két célszó között látszanak különbségek (3. ábra), amelyeket az eltérő mássalhangzós kontextus okozhat, hiszen míg a *nyír* esetében a célmagánhangzót megelőző mássalhangzó palatális, a követő pedig alveoláris, a *szív* esetében a megelőző alveoláris, míg a követő labiodentális képzésű. Feltehetően ebből adódik, hogy a *nyír* magánhangzójának ejtésekor a nyelvszensorok (de különösen a nyelvhegyen és a nyelvperemen lévő) hátrébb helyezkednek el a szájüregben, mint a *szív* esetében, amit az alacsonyabb százalékos értékek mutatnak. A homofón alakok harmóniabeli viselkedése mentén ugyanakkor nem láthatók jelentős eltérések, és különösen nem látjuk azt, hogy az antiharmonikus tő esetében a szensorok rendre hátrébb lennének (azaz kisebb százalékos értéket vennének fel), mint a harmonikus tő esetén.



3. ábra: A szensorok vízszintes helyzete (x tengely) %-ban kifejezve (0% = *nyúl*, 100% = *sír*) a célszavak harmóniabeli viselkedésének a függvényében az izolált elrendezésben, átlag és szórás

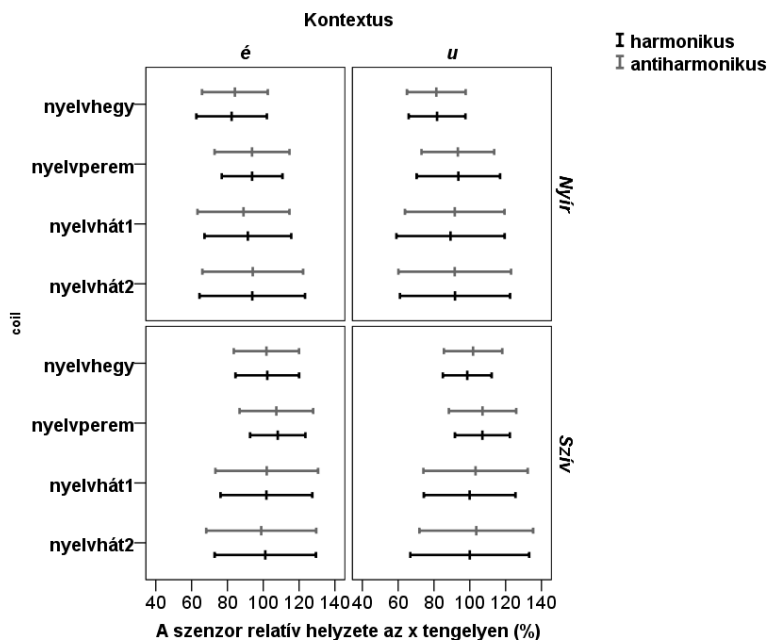
A nyelv szenzorok vízszintes helyzete a GAM szerint sem mutatott jelentős eltérést a harmóniabeli viselkedés függvényében az izolált elrendezésben (4. ábra). A harmóniabeli viselkedés főhatását tartalmazó modell AIC-értéke 8,44 ponttal magasabb volt a minimális modell AIC-értékéhez képest, tehát a főhatás nem javított a modellilleszkedésen.



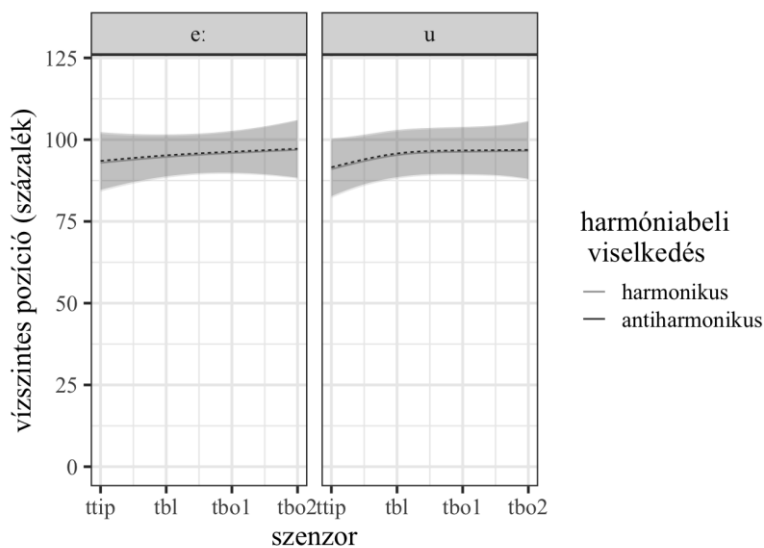
4. ábra: A nyelv szenzorok vízszintes pozícióértékeinek egymásutánjára illesztett görbék a harmóniabeli viselkedés szerint az izolált elrendezésben

A kontextusos elrendezésben a mondatkezdő helyzetű célszavakat vagy elülső (*éppen*), vagy hátsó (*ugyan*) magánhangzók követték. Azt vártuk, hogy a hátraható magánhangzó-magánhangzó koartikuláció miatt a kétféle kontextus eltérő mintázatokat hoz létre a nyelv helyzetben a harmóniabeli viselkedés függvényében. Az 5. ábrán látható, hogy nincsenek szisztematikus eltérések a nyelv szenzorok pozícióját tekintve. Ebben az elrendezésben a harmóniabeli viselkedés főhatását tartalmazó modell AIC-értéke volt az alacsonyabb, azonban csak 0,64 ponttal, mint a minimális modell AIC-értéke, tehát az eltérés az Anderson és Burnham (2004) által meghatározott küszöbérték alatt volt, így a főhatás ebben az esetben sem javított a modellilleszkedésen (6. ábra).

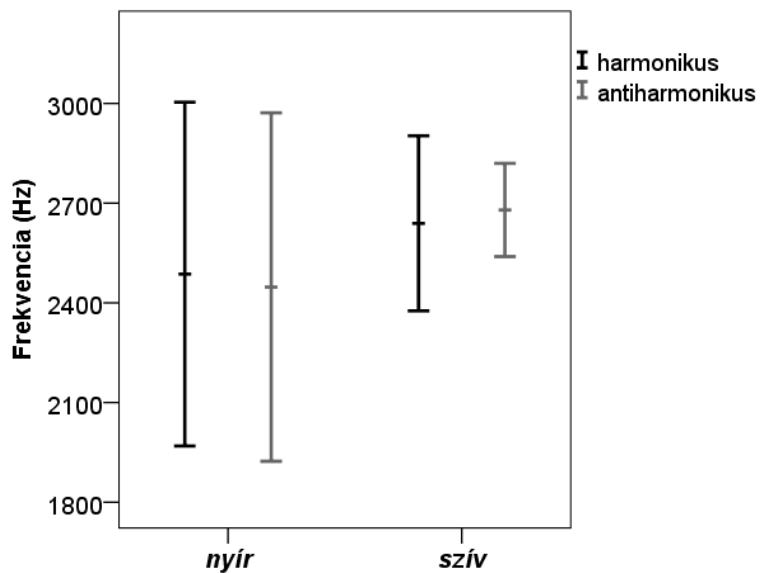
A célmagánhangzók időbeli középpontjában automatikusan nyertük ki a második formáns frekvenciaértékét. Az izolált elrendezésben az F_2 átlaga valamivel alacsonyabb volt a *nyír* esetében az antiharmonikus töben, mint a harmonikus töben, ami utalhat hátrébb húzott nyelvre, de akusztikailag mindenképpen centralizáltabb ejtésként értelmezhető, a *szív* előfordulásaiban ugyanakkor valamivel magasabb F_2 -t mértünk az antiharmonikus töben, mint a harmonikus töben, ami pedig ezzel éppen ellenkező tendencia (7. ábra). A statisztikai próba ugyanakkor egyik esetben sem mutatott ki matematikailag is alátámasztható eltérést.



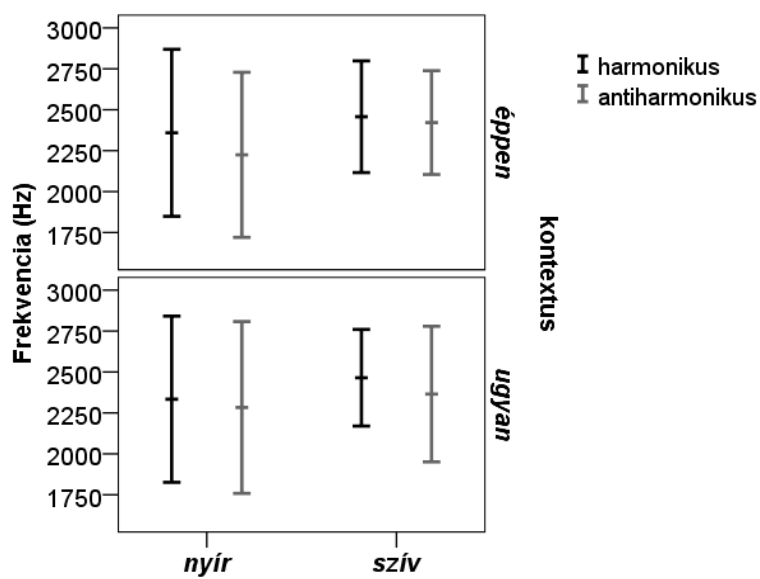
5. ábra: A szenzorok vízszintes helyzete (x tengely) %-ban kifejezve (0% = *nyíl*, 100% = *sír*) a célszavak harmóniabeli viselkedésének és a kontextusnak a függvényében a kontextusos elrendezésben, átlag és szórás



6. ábra: A nyelvszenzorok vízszintes pozícióértékeinek egymásutánjára illesztett görbék a harmóniabeli viselkedés és a kontextus szerint



7. ábra: A második formáns frekvenciaértéke a célszavak harmóniabeli viselkedésének a függvényében az izolált elrendezésben, átlag és szórás



8. ábra: A második formáns frekvenciaértéke a célszavak harmóniabeli viselkedésének a függvényében a kontextusos elrendezésben, átlag és szórás

A kontextusos elrendezésben (8. ábra) mind a *nyír*, mind a *szív* előfordulásaiban valamelyest hátrébb húzott nyelvre, illetve a harmonikus töben tapasztaltnál nagyobb fokú akusztikai centralizáltságra utalnak az antiharmonikus tövekben megjelenő magánhangzó formánsértékei, a kontextustól (*éppen* vagy *ugyan*) függetlenül. Ugyanakkor statisztikai eltérés itt sem volt kimutatható az egyes csoportok között.

4. Következtetések

Tanulmányunkban az áttetsző magánhangzók közül az /i:/ artikulációs és akusztikai jellemzőit, pontosabban az /i:/ ejtése közben a nyelv elejének szájüregbeli vízszintes helyzetét, illetve az F₂ formáns értékét vizsgáltuk elektromágneses artikulometriával és spektrális elemzéssel egy szótagos, homofón szópárokban, a *nyír* és a *szív* igei (antiharmonikusan toldalékolódó) és főnévi (harmonikusan toldalékolódó) funkcióban ejtett változataiban. Kétféle elrendezést alkalmaztunk annak tesztelésére, hogy a harmóniabeli viselkedés (harmonikusan, illetve antiharmonikusan toldalékoltság) függvényében eltérnek-e a nyelvre helyezett szenzorok pozíciói a szájüregben belül, vízszintes irányban, valamint az F₂ formáns értékei. Az egyik elrendezésben a célszavak izoláltan, toldalék és kontextus nélkül, önálló megnyilatkozásként szerepeltek, ekkor a funkciójukat a résztvevők számára képi ingerekkel elicitáltuk. A másik elrendezésben kontextusba ágyaztuk a célszavakat, mégpedig úgy, hogy azokat kiegyenlített módon vagy csak elülső (*éppen*), vagy csak hátulsó (*ugyan*) magánhangzó kövesse, mely magánhangzók tehát kontrollált módon hatottak (visszafelé) a magánhangzók közti koartikuláció révén a célhangra.

Azt vártuk, hogy az izolált elrendezésben a harmonikusan (elülső magánhangzóval) toldalékolódó célszavakban az /i:/ ejtésekor a nyelv előrébb helyezkedik el a szájüregben, mint az antiharmonikusan (nem elülső magánhangzóval) toldalékolódó célszavakban, és ennek mind artikulációs, mind akusztikai következménye mérhető. Az antiharmonikus tövek esetében ugyanis Beňuš és Gafos (2007) szerint a toldalékolatlan alak is örökli a toldalékolt szóalakban megvalósuló, hátraható magánhangzó-magánhangzó koartikulációs hatás artikulációs mintázatát, ami hátrébb képzett magánhangzó-megvalósulásokat eredményez az antiharmonikus toldalékolatlan tövekben. Ezt a feltételezést a vizsgálati eredmények nem erősítették meg, sem az artikuláció, sem az akusztikai szerkezet tekintetében.

A másik elrendezésben a kontextus megválasztása annak a feltételezésnek a tesztelését célozta, hogy (függetlenül attól, hogy az izolált ejtésben kimutatható-e artikulációs eltérés) a kétféle harmóniabeli viselkedést mutató tövek esetében az /i:/ ejtésekor mérhető vízszintes nyelv helyzet eltér a követő magánhangzók előlségének a függvényében. Ebben az elrendezésben a célszavak megnyilatkozás kezdetén, az izolált elrendezéshez hasonlóan szintén főhangsúlyos helyzet-

ben jelentkeztek. Tehát a két (izolált és kontextusos) elrendezés a követő magánhangzók koartikulációs hatásában tér(hetet) el, valamint abban, hogy az utóbbiban nem csak a képi, hanem a nyelvi anyag, azaz a teljes mondat is segíthette a célszavak aktuális – igei vagy főnévi – funkciójának előhívását a beszélőkben. A vizsgálati eredmények alapján a második hipotézisünk sem nyert megerősítést, tehát általánosságban azt találtuk, hogy a harmónia szempontjából áttetszőnek tekintett /i:/ magánhangzó nem valósult meg előrébb vagy hátrébb, illetve akusztikailag centralizáltabban vagy periférikusabban a magánhangzó-harmóniában való aktuális részvétele szerint. Sem az artikulációs, sem az akusztikai adatok nem utaltak a nyelv eltérő pozíciójára, illetve a magánhangzók eltérő mértékű akusztikai vagy artikulációs centralizáltságára a harmóniabeli viselkedés mentén a kontextus függvényében.

Beňuš és Gafos (2007) kísérletében – mint láttuk – a vizsgált és összehasonlított szópárok kapcsán több kérdés is felmerült, melyek az ott kapott eredményeket befolyásolhatták (a szótagszerkezet hatása, a mássalhangzók képzéshelyének a hatása, a hordozómondat hatása stb.). A jelen kísérletben éppen ezért maximálisan kontrollált, homofón szópárokat vetettünk össze, illetve valóban izolált és kontrollált kontextusos elrendezéseket vizsgáltunk. Mégis, a jelen kísérlettel kapcsolatosan is felmerülhet az a kérdés, hogy a képi ingerek mennyire segítettek a résztvevők számára az izolált elrendezésben az igei és a főnévi funkció megkülönböztetését. Nem lehetséges-e, hogy a résztvevők mechanikusan ejtették a szavakat, nem ügyelve a funkciók megkülönböztetésére? Ennek kiküszöbölésére is szolgálhatott ugyanakkor a kontextusos elrendezés, ahol a funkciók elkülönítésében a képi információn túl a mondat jelentése is a résztvevők segítségére volt.

További kérdéseket vethet fel a kísérletben részt vevők száma is, ugyanis ez a jelen kísérletnek alapjául szolgáló Beňuš és Gafos (2007) végezte kísérletben igen alacsony volt (három fő, ráadásul a kísérlet és az elemzés eltérően zajlott az egyik beszélőnél, mint a másik kettő esetében). Igaz, hogy az artikulációs kísérletekben a kísérleti személyek száma az eljárás komplexitása és az adatközlőtől igen nagymértékű rugalmasságot és igénybevételt kívánó volta miatt (valamint sokszor amiatt is, hogy nem minden jelentkező esetében lehet elemezhető kísérleti anyagot felvenni) általánosságban is igen alacsony. Összevetésül: a 2019-ben megrendezett ICPHS-kongresszus kötetében (Calhoun et al. eds. 2019) megjelent tanulmányok közül mintegy 25-ben alkalmaztak EMA-t, és ezek között mindössze kettő olyan volt, ahol a beszélőszám (egy adott vizsgálati csoportban) meghaladta a tízet, ugyanakkor a résztvevők száma nem érte el az ötöt sem a tanulmányok felében. Ez alapján tehát Beňuš és Gafos (2007) kísérletei tulajdonképpen még ma is megfelelnek az íratlan szakmai „sztenderdeknek”, és amikor a kutatás készült, ez még inkább így volt. Ugyanakkor az alacsony beszélőszámoknak nyilvánvalóan hatása van az eredmények általánosíthatóságára. Ez

önmagában is előidézhet például olyan eseteket, hogy az adatok mennyisége nem tesz lehetővé statisztikai elemzést, vagy pedig nem látszanak olyan tendenciák, amelyek nagyobb beszélőszám esetén kimutathatók lennének. Alacsony résztvevőszám esetén a beszélők közötti egyéni eltérések jelentősége ugyanis megnő.

Még ha Beňuš és Gafos (2007) három személy adatai alapján fonológiaelméleti megállapításokat tett is, a tizenkét beszélővel végzett vizsgálatunk eredményét nagyobb mértékben tartjuk általánosíthatónak. Ennek csak részben a nagyobb résztvevőszám az oka, lényeges eltérésnek tartjuk azt is, hogy a jelen elemzésben választott módszertan (az azonos skálára vetített, így normalizált nyelvhelyzet) segítségével a beszélők közötti variabilitást is kontrollálni tudtuk, így a különböző résztvevők adatait együttesen lehetett kezelni. Továbbá, a jelen kísérletben a változók hatását a szenzorok (normalizált) vízszintes pozíciójának négyelemű szekvenciájára generalizált additív modellekkel elemeztük, ami lehetővé tette a nyelven egymást követő nyelvszenzorokon mért pozícióadatok együttes modellezését. Ennélfogva az elemzésünk holisztikusabb képet ad a nyelvhelyzet különbségeiről, mint az idézett korábbi artikulációs vizsgálat.

Az akusztikai elemzést illetően (amelyet Beňuš és Gafos 2007 nem végzett), az eredményeink hasonlóak Blaho és Szeredi (2013) eredményeihez, azaz ahogy ők, mi sem találtunk eltérést a második formáns értékében a tő harmóniabeli viselkedése függvényében.

Habár az artikulációbeli eltérés ellenére is előfordulhat, hogy az akusztikai szerkezetben nincs statisztikailag alátámasztható különbség, mint ahogy az is lehetséges, hogy az artikulációban nem mutatkozik eltérés (egy adott módszertannal mérve), az akusztikumban mégis jelentősnek látszó különbség mérhető, a jelen esetben nem ez a helyzet. Eredményeink nem arra utalnak, amit Beňuš és Gafos (2007) sejtet, hogy az artikulációs eltérés nem kritikus régióban történik, hanem arra, hogy nincs artikulációs eltérés, vagy legalábbis nem éri el azt a mértéket, amely a generalizált additív modellekkel feltárható lehetne. Ez alapján pedig az is nehezen elképzelhető, hogy az artikulációbeli különbségek áttevedhetnek a koartikulált helyzetből az izolált helyzetben ejtett magánhangzókra, és eltárolódhatnak a mentális lexikonban, mivel az adatok arra utalnak, hogy nincs ilyen artikulációs különbség. Az eredmények inkább azt az állítást erősítik meg, mely szerint az /i:/ viszonylag stabil, azaz a szomszédos hátul képzett magánhangzók hatására igen kismértékű eltérést mutat (Recasens 1999) – bár egyes eredmények szerint ez az általános érvényűnek tartott elgondolás párhuzamos artikulációs és akusztikai eredmények tükrében árnyalásra szorul (Deme et al. 2019).

Természetesen további, percepció vizsgálatok adhatnak választ arra a kérdésre, hogy – függetlenül az itt kapott eredményektől – hatással van-e az észlelésre a homofón tövek harmóniabeli viselkedése. Amennyiben a percepció teszt alapján erre a kérdésre igen választ kapunk, további feladatunk az, hogy az arti-

kulációs mérés módszertant finomítsuk, és ha van artikulációs eltérés, azt ki tudjuk mutatni a műszeres elemzés segítségével. Mivel azonban a jelen eredmények nem támasztják alá a harmóniabeli viselkedés és az /i:/ megvalósulása közti korábban feltételezett összefüggéseket, a jelen adatok alapján le kell vonnunk azt a következtetést is, hogy az /i:/ viselkedése nem szolgál tudományos bizonyítékkal a beszédprodukción és beszédészlelés exemplárelméletének validitását illetően.

Irodalom

- Akaike, Hirotugu (1974), A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19/6: 716–723.
- Anderson, David R. – Burnham, Kenneth P. (2004), *Model selection and multi-model inference*. NY: Springer-Verlag.
- Beňuš, Stefan – Gafos, Adamantios I. (2007), Articulatory characteristics of Hungarian ‘transparent’ vowels. *Journal of Phonetics* 35: 271–300.
- Blaho, Sylvia – Szeredi, Dániel (2013), Hungarian neutral vowels: a microcomparison. *Nordlyd* 40(1): 20–40.
- Boersma, Paul – Weenink, David (2018), Praat: doing phonetics by computer [Computer program]. Version 6.0.43, <http://www.praat.org/>.
- Calhoun, Sasha – Escudero, Paola – Tabain, Marija – Warren, Paul (eds.) (2019), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*. Australasian Speech Science and Technology Association Inc., Canberra, Australia.
- Cho, Taehong (2004), Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics* 32: 141–176.
- Coretta, Stefano (2018), *tidymv: Tidy Model Visualisation*. R package version 1.5.4.9000. <https://github.com/stefanocoretta/tidymv>.
- Deme, Andrea – Bartók, Márton – Grácz, Tekla Etelka – Csapó, Tamás Gábor – Markó, Alexandra (2019), V-to-V Coarticulation Induced Acoustic and Articulatory Variability of Vowels: The Effect of Pitch-Accent. In: *Proc. Interspeech 2019*, 3317–3321.
- Hayes, Bruce – Zuraw, Kie – Siptár, Péter – Londe, Zsuzsa (2009), Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85/4: 822–863.
- Hillenbrand, James M. – Clark, Michael J. – Nearey, Terrance (2001), Effects of consonant environment on vowel formant patterns. *The Journal of the Acoustical Society of America* 109/2: 748–763.
- Kuznetsova, A. – Brockhoff, P. – H. B. Christensen, R. (2015), *LmerTest: Tests in linear mixed effects models*. R Package Version 2.
- Nádasdy Ádám – Siptár Péter (1994/2006), A magánhangzók. In: Kiefer Ferenc (szerk.) *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó, Budapest. 42–182.
- Patay, Fanni – Benkő, Ágnes – Lukács, Ágnes – Rebrus, Péter – Törkenczy, Miklós (megj.), Testing variability effects in Hungarian vowel harmony. In: Hegedűs, Veronika – Vogel, Irene (eds.), *Approaches to Hungarian 16. Papers from the 2017 Budapest conference*.

- R Core Team (2018), R: A Language and Environment of Statistical Computing, R Foundation for Computing, Vienna, <https://www.R-project.org>.
- Rebrus Péter – Törkenczy Miklós (2019), Magyar harmónia: a dolgok állása. Általános Nyelvészeti Tanulmányok XXXI. 233-333.
- Recasens, Daniel (1999), Lingual coarticulation. In: Hardcastle, Willam J. – Hewlett, Nigel (eds.), *Coarticulation: Theory, data and techniques*. Cambridge University Press, Cambridge, UK. 78–104.
- Reichel, Uwe D. (2012), PermA and Balloon: Tools for string alignment and text processing. In: *Proceedings of Interspeech 2012*. Paper no. 346.
- Schiel, Florian (1999), Automatic phonetic transcription of nonprompted speech. In: *Proceedings of the International Congress of Phonetic Sciences*. 607–610.
- Stevens, Kenneth N. (1989), On the quantal nature of speech. *Journal of Phonetics*, 17: 3–45.
- Sussman, Harvey M. (1990), Acoustic correlates of the front/back vowel distinction: A comparison of transition onset versus "steady state". *Journal of the Acoustical Society of America*, 88: 87–96.
- Syrdal, Ann K. – Gopal, H. S. (1986), A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America*, 79: 1086–1100.
- Szeredi, Dániel (2012), Acceptability of harmonic mismatch for neutral vowel stems in Hungarian. Qualifying Paper. New York University, New York.
- Thèoret, Hugo – Pascual-Leone, Alvaro 2002. Language acquisition: Do as you hear. *Current Biology* 12/21. pR736–R737. [http://www.cell.com/current-biology/pdf/S0960-9822\(02\)01251-4.pdf](http://www.cell.com/current-biology/pdf/S0960-9822(02)01251-4.pdf).
- van Rij, Jacolien – Wieling, Martijn – Baayen, R. Harald – van Rijn, Hedderik (2017), *itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs*, R package version 2.3.
- Winkelmann, Raphael – Jaensch, Klaus – Cassidy, Steve – Harrington, Jonathan (2018), *emuR: Main Package of the EMU Speech Database Management SystemR*, package version 1.1.1.
- Wood, Sara N. (2001), *mgcv: GAMs and generalized ridge regression for R*. *R news*, vol. 1, no.2, pp. 20–25.
- Zuraw, Kie – Hayes, Bruce (2017), Intersecting constraint families: an argument for Harmonic Grammar. *Language* 93(3): 497–548.

Köszönetnyilvánítás

Köszönjük Buza Ákos, Jankovics Julianna, Krepsz Valéria, Mrázik István, Puzder Zsófia és Weidl Zsófia segítségét a kutatás megvalósításában, valamint Rebrus Péternek és Törkenczy Miklósnak a konzultációt és a rendelkezésünkre bocsátott kéziratokat. A tanulmány a Tématerületi Kiválósági Program támogatásával készült.