

AZ [i] + V ÉS V + [i] HANGKAPCSOLÓDÁSOK AKUSZTIKAI ELEMZÉSE A HIÁTUSTÖLTÉS MAGYARÁZATÁHOZ

Olaszy Gábor

Bevezetés

A hiátustöltés (HT) két magánhangzó közé beékelődött palatális zöngés approximánsként van jelen a magyarban (Siptár 2002). Fonológiai szempontból ez a definíció teljességgel elegendő. Beszédtechnológiai vonatkozásban azonban pontosabb adatokra van szükség. Fonetikai szempontú elemzéssel mindeddig egyetlen korábbi magyar vizsgálat foglalkozott a magánhangzókapcsolódások elemzésének keretein belül (Menyhárt 2006). A kutatás szerint a leggyakoribb hiátustöltést generáló hang az [i]. Menyhárt vizsgálatában az [i] a vizsgált hiátustöltéses kapcsolatok 64,2%-ában fordult elő (például *dió*), ezt követte az [ɛ] 19,6%-kal (például *tea*), majd az [y] következett 13,6%-kal (például *menüért*). Ebben a tanulmányban célzott összehasonlító elemzések eredményeit adjuk közre. A fő kérdés, hogy mennyiben tekinthető a magyar intervokális helyzetű /j/ fonéma hangrealizációja megegyezőnek a hiátustöltéskor létrejövő hangelemmel, amely az [i] + V, illetve a V + [i] hangkapcsolatokban jön létre. A további vizsgálandó témakör, hogy a hiátustöltések között milyen akusztikai fonetikai különbségek mutathatók ki. Az elemzésben a következő kérdésekre igyekeztünk választ kapni. Van-e különbség a /j/ fonéma hangrealizációja és az ugyanolyan pozíciójú HT, azaz a két magánhangzó közötti palatális zöngés approximáns között, és ha van, akkor miben különbözik a kettő. Független-e a HT akusztikai minősége a hanghelyzettől a hangsoron belül? Hogyan befolyásolja a hiátustöltést, ha előtte a /j/ fonéma realizációja fordul elő? Hogyan valósul meg a HT három- és négyelemű magánhangzókapcsolatokban?

Hipotézisünk egyrészt az, hogy a hiátustöltés más-más időtartammal változik meg az egyes hanghelyzetekben, másrészt pedig az, hogy időtartama rövidebb, mint a hasonló helyzetben lévő [j] hangé.

Anyag, módszer, kísérleti személyek

A vizsgálatot szólista felolvasásából kapott beszédanyagon végeztük. Az átlagos artikulációs sebesség 9,8 hang/s a teljes hanganyagra vonatkoztatva. A vizsgálat során a hangidőtartamokat mértük meg és hasonlítottuk össze a kérdéses hangokra.

A szólista szavainak belső hangsorszerkezetét célzottan állítottuk össze, igazodva a vizsgálatban megfogalmazott kérdésekhez. Az [i] + V és a V + [i] kapcsolatokat úgy tekintettük, hogy mindenképpen létrejön bennük a hiátustöltés, csak különböző mértékben (amennyiben a hanghullámban a folytonosság megvan a teljes hangátmenetben). A /j/ fonémát intervokális helyzetben vizsgáltuk egyrészt az [i]-hez kapcsolódva, másrésztől más helyzetekben is. A magánhangzó-kapcsolatoknál a hangkapcsolódás helyzeti pozícióját is figyelembe vettük (hangsor elején, közepén, végén). A VVV kapcsolatokban a középső hang mindig az [i] volt. A fenti kritériumoknak megfelelően szócsoportokat határoztunk meg.

Az első csoport szavai egyetlen /j/ fonémát tartalmaztak intervokális helyzetben (*kijárat, ijedős, erdélyi* stb.). Az egyik kapcsolódó V minden esetben az [i] hang. A szóválasztás célja az volt, hogy össze tudjuk hasonlítani a palatális zöngés approximánst a hiátustöltéskor létrejövő hasonló hanggal.

A második szócsoportban olyan szó belseji VV kapcsolatokat szerepeltetünk, amelyeknek egyik eleme [i], a másik pedig más magánhangzó volt, és feltételeztük, hogy minimális a HT megvalósulása (*katonáinduló, június, július, jóindulat* stb.). Ezt úgy ellenőriztük, hogy megpróbáltuk a szó kiejtése során betoldani a hiátustöltő palatális zöngés approximánst a magánhangzók közé. Amennyiben az ilyen kiejtés erőltettnak tűnt, akkor minimális hiátustöltésként kezeltük a hangot, ellenkező esetben pedig erősnek (a harmadik hangcsoport). Úgy találtuk, hogy a minimális hiátustöltést az jellemezte, hogy nem lehet betoldani egy palatális zöngés approximánst a magánhangzók közé, mert erőltettnak éri a beszélő az artikuláció szempontjából. Az ilyen eseteket úgy tekintjük, hogy nincs hiátustöltés. A harmadik csoportban olyan szavak szerepeltek, amelyekben erős HT valósul meg (*fiatalok, kiabál, fiával, sietnek* stb.). Ezeknél a kiejtés során is hallani lehet a hiátustöltés palatális zöngés approximánsát.

A negyedik csoportba olyan szavakat vettünk fel, amelyekben a HT a szó végén valósul meg, és a hiátustöltés az egyik csoportban az [i] után (*stúdió, vakáció* stb.), a másokban az [i] előtt (*támadói, katonái* stb.) valósult meg. A cél itt az volt, hogy a hiátustöltéseken belül vizsgáljuk meg, hogy van-e időtartambeli különbség más megvalósulási formák között.

Az ötödik csoportban olyan szavak szerepeltek, amelyekben a hiátustöltést megelőző mássalhangzó a [j] hang (*bójátok, szabályainak, fejeiben* stb.). Itt az egymás utáni palatális hangelemek viszonyát elemeztük. Az volt a cél, hogy megvizsgáljuk a /j/ fonéma hatását hangsor belseji helyzetben a HT létrejöttének fokozataira.

A hatodik csoportba tartoztak azok a szavak, amelyekben a hiátustöltést megelőző mássalhangzó a [j] hang, és a HT az utolsó szótagban jön létre (*osztályai, vejei, folyói* stb.). A vizsgálati cél hasonló volt, mint az ötödik szócsoportban, csak a szó végén.

A hetedik csoport szavaiban a hiátustöltést megelőző mássalhangzó palatális, és a HT az utolsó két magánhangzó között jön létre (például *szempontjai*, *sebesültjei*, *tornyai*). Itt azt vizsgáltuk, hogy egyéb palatális mássalhangzók hatása hasonló-e, mint a /j/ fonémáé a HT megvalósulására.

A nyolcadik szócsoporthoz szavaiban V + [i] + V kapcsolatok voltak szó belseji helyzetben (például *lemezeiért*, *röplapjaiért*). Ezekben a szavakban a hármas hangkapcsolatokban létrejövő hiátustöltés időtartamait mértük meg.

Az [i] + V + [i] + V négyelemű magánhangzó-kapcsolatban létrejövő hiátustöltéseket egyetlen szóban (*misszióiért*) vizsgáltuk. Mindkét esetben azt akartuk megtudni, hogy van-e szignifikáns eltérés a HT-k időtartamai között.

A teljes szólista mindösszesen 66 szót tartalmazott. A szólistát 13 személyllyel olvastattuk fel (6 nő és 7 férfi, 30–60 év közötti magyar anyanyelvű köznapis beszélők) professzionális hangstúdióban, a digitalizálás 22 kHz-es mintavételezéssel és 16 bites kvantálással történt. Minden szót külön fájlként kezeltünk. A hullámformán bejelöltük a hanghatárokat, valamint a HT határait is (gépi feldolgozással és kézi korrekcióval).

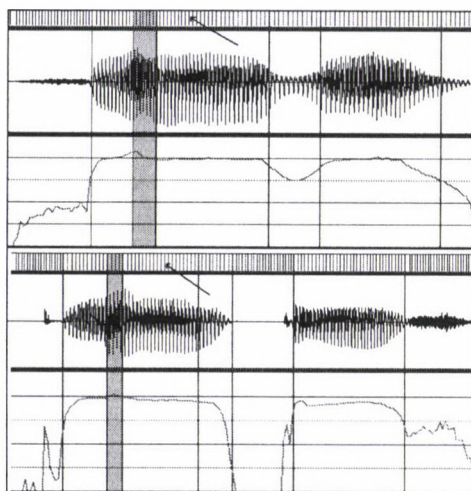
A HT-t úgy tekintettük, hogy azt a hangátmenet bizonyos része képviseli a fizikai jelben. A HT frekvenciaszerkezete hasonló a [j], illetve az [i] hang szerkezetéhez is, ezért a két hang közötti spektrális különbségek kimutatása nehéz. Ennek tükrében úgy döntöttünk, hogy a méréseket a hangidőtartamok összehasonlítására alapozzuk. Ez utóbbi meghatározásában auditív-vizuális ítélettel döntöttük el a HT határpontjait, periódusonkénti lépcsőzetes bővítéssel, szűkítéssel mindkét irányból elvégezve a meghallgatást.

Erre a Profivox szövegfelolvasó fejlesztői rendszerét vettük igénybe (Olaszy et al. 2002). A hanghatárok kijelölésénél támpontot jelentett a precíz intenzitásgörbe párhuzamos vizsgálata a hanghullámmal (1. ábra). A szócsoporthoz mért hiátustöltések időtartamait összehasonlítottuk egymással is, és a mért [j] hangokéival is. A szignifikanciavizsgálatokhoz az R statisztikai programcsomagot használtuk.

A mérés hanganyagát közreadjuk a <http://magyarbeszed.tmit.bme.hu> honlapon.

Eredmények

Mindenekelőtt a hipotézisünk második gondolatára vonatkozó vizsgálati eredményeket, azaz a palatális approximánsok időtartamait hasonlítottuk össze. Az első szócsoporthoz vizsgálatából a mért /j/ fonémák hangrealizációinak átlagos hossza 36,5 ms, a szórás 7,41 ms. Ez feltűnően rövid. A hangsor utolsó szótagjában szereplő ugyanilyen hang átlagos hossza 74,6 ms, a szórás 16,93 ms. Ugyanakkor a vizsgálati anyagban más helyzetben található /j/ fonéma realizációjára (ötödik csoport) a következő adatokat kaptuk. Hangsor belseji helyzetben 75 ms az átlagos hossz, a szórás 14,8 ms. A hangsor utolsó szótagjában pedig 79,3 ms az átlag, a szórás 15,8 ms. E két utóbbi realizáció időtartama között nincs szignifikáns eltérés (kétmintás *t*-próba: $p = 0,1216$).



1. ábra

Példa a hiátustöltés hanghatárainak bejelölési technikájára a *fiának* (fenn) és a *kiáltás* (lenn) szavakban

A nyilak mutatják a zöngperiódusok kezdetét jelölő markereket. Ezek léptetésével lehet a meghallgatási ablakot tágítani, szűkíteni a hanghatár meghatározásánál.

Egy korábbi mérésben a felolvasott beszédben Olasz (2006) 55,5 ms-os átlagot kapott egy beszédkorpusz összes [j] hangjára vonatkoztatva (a korpuszban 13 hang/s-os volt az artikulációs sebesség). A fenti adatokból azt a következtetést vonhatjuk le, hogy a /j/ fonéma hossza függ attól, hogy milyen magánhangzóhoz kapcsolódik. Amennyiben az [i] szomszédságában fordul elő, akkor szignifikánsan rövidebb, mintegy fele annak, mint amikor egyéb hangkörnyezetben van (kétmintás *t*-próba: $p < 0,000216$).

A második és harmadik szócsoportok vizsgálataiból a hipotézisünk első gondolatához kaptunk adatokat. A második csoportból a mért VV hangátmenetekben a HT átlagos időtartama 26,7 ms, a szórás 6,49 ms, a harmadikban 33,6 ms, a szórás 5,8 ms. Ez azt sugallja, hogy a minimális időtartamú HT, valamint az erős hiátustöltés közötti különbségtételben az időtartam-különbség ad támpontot. Ezt szignifikanciavizsgálattal vizsgáltuk tovább.

Az első három eset hangidőtartamaira egytényezős varianciaelemzést végeztünk. Szignifikáns különbséget találtunk a hangok hosszában [$F(2, 230) = 28,589, p < 0,0001$]. A három eset páronkénti összehasonlítását a Tukey-tesztel végeztük, a különbség mindhárom párra szignifikáns volt ($p < 0,0293$). Ez egyrészt azt sugallja, hogy az [i]-hez kapcsolódó [j] hang – annak ellenére, hogy feltűnően rövidnek tűnik – az időtartama alapján elválik az erős-

nek mondható hiátustöltéstől, azaz hosszabb, mint az. Ezzel a hipotézis második gondolata látszik igazolódni. Másrészről a szignifikáns időtartam-különbség a minimális és az erős HT kategorizálhatóságát is jelenti. Ha igen/nem feltételt vezetünk be, akkor azt mondhatjuk, hogy a második csoportban nem jellemző a HT (vagyis nincs hiátustöltés), míg a harmadikban igen.

A negyedik szócsoporthoz vonatkozó eredmények szerint (hangszorzó helyzetben) az [i] + V hangátmenetben létrejövő HT átlagos időtartama 40,1 ms, a szórás 7,62 ms. A V + [i] esetre ugyanez az adat 52,7 ms, a szórás 11,6 ms. A két alcsoportban szignifikáns különbséget találtunk a hangok hosszában (kétmintás *t*-próba: $p < 0,001$). Ezek szerint a hangátmenetre fordított idő függ attól, hogy az [i]-ből indulunk a következő magánhangzóba, avagy fordítva. Ez a hipotézis első gondolatához kapcsolható igazolásnak fogható fel.

Az ötödik szócsoporthoz tartozó vizsgálatának eredményei egyértelműen kimutatták, hogy a /j/ fonéma realizációja és az őt közvetlenül követő HT között nagymértékű különbség van a fonéma javára a hanghossz tekintetében. Az átlagok és a szórások a következők: 75,1 ms, 14,8 ms, illetve a HT-re 42,2 ms, 10,76 ms. Az eltérés szignifikáns. Az ilyen hanghelyzetben létrejövő HT hosszabb, mint a harmadik szócsoporthoz mért [i] + V kapcsolatból kapott úgynevezett erős hiátustöltés (ami átlagosan 33,6 ms). A két hiátustöltés időtartamának eltérése szignifikáns (*t*-próba: $p < 0,001$). Ezek szerint a HT időtartamát növeli, ha előtte /j/ fonémarealizáció van.

A hatodik szócsoporthoz tartozó vizsgálatok ugyanazt a hangkapcsolati formát vizsgálták, mint az ötödik. Az egyetlen különbség az volt, hogy most a szó záró részében volt a hangkapcsolat és a [j] hang. A kérdés az volt, hogy változik-e a palatális hangnak, illetve a hiátustöltésnek az időtartama. A [j] hang átlagos hossza itt 79,3 ms, a szórás 7,9 ms. A hiátustöltés átlagos hossza 42,3 ms, a szórás pedig 7,9 ms. Kismértékű időtartamkülönbség csak a [j] hangban mutatkozik, azonban ez nem szignifikáns. A hiátustöltés időtartama ebben a kapcsolati formában nem változik a hangsor belseji hanghelyzethez viszonyítva.

A hetedik szócsoporthoz tartozó vizsgálati eredményeit (a hiátustöltés átlaga 39,1 ms, a szórás 8,26 ms) az 5. csoportéval hasonlítottuk össze. A kérdés az volt, hogy a palatális képzési hely, illetve a képzési mód hat-e jobban a létrejövő hiátustöltés hosszára. Az eredmény szerint a képzési mód határozza meg a jelenséget, a kétmintás *t*-próba azt mutatta, hogy nincs szignifikáns eltérés a két átlag között.

A nyolcadik szócsoporthoz tartozó vizsgálatokban a V + [i] + V kapcsolatokban lezajló elvileg két hiátustöltés viszonyát vizsgáltuk. Az elsőre kapott átlag 48,5 ms, a szórás 10,23 ms. A második átlaga 29,6 ms, a szórás 8,35 ms. Az eltérés szignifikáns (kétmintás *t*-próba: $p < 0,001$). A mért hangösszetételű hármas kapcsolatokra tehát az jellemző, hogy a kapcsolat első hiátustöltő eleme dominál, mert az a hosszabb. A magyarázat az lehet, hogy a hármas magánhangzókapcsolatban az artikulációs mozgások még feszesebbek az első két hang kapcsolódásánál, mint a második kettőnél.

A hármas kapcsolatoknál is megvizsgáltuk a palatális hatást a hiátustöltés kialakulására (például: *röplapjaiért*). A mért hangidőtartam-átlagok az ejtés sorrendjében a következők. A /j/ fonéma realizációinak átlaga 60,1 ms, a szórása 21,88 ms, az első hiátustöltés 43 ms, a szórása 8,96 ms, a második hiátustöltés 30 ms, a szórása 7,6 ms. Ezek szerint a palatális zöngés approximáns mint megelőző hang olyan hatással van az artikulációs folyamatra, hogy az első hiátustöltés jellemzően hosszabb, mint a második. Itt is érvényesül tehát az a tendencia, hogy az első hiátustöltés a domináns.

Megvizsgáltuk az [i] + V + [i] + V helyzetet is egyetlen szóban (*misszióiért*). Az átlagok és szóráások az ejtés sorrendjében a következők. Első HT: 34,9 ms, 9,41 ms; a második HT: 46,8 ms, 9,2 ms; a harmadik HT 26,8 ms, 4,56 ms. Az ilyen négyes kapcsolatokban tehát a középső HT lehet a domináns.

Összefoglalás

A vizsgálat kimutatta, hogy a hiátustöltés és a /j/ fonéma viszonyát egyrésztől a hangsorban elfoglalt helyzet szerint két csoportra lehet osztani. A hangsor belsejében más adatok jellemzők mind a /j/ fonéma hosszára, mind a HT-ére. Mindkét esetben a /j/ fonéma szignifikánsan hosszabb, mint a hasonló hanghelyzetű hiátustöltés. A hiátustöltéseknél pedig a hangsor belseji rövidebb, mint a hangsor végén kialakuló. Másrésztől a hiátustöltést megelőző hangok szempontjából is két csoportot kell megkülönböztetni: palatális, illetve nem az. A megelőző palatális hang nyújtja a hiátustöltést.

Irodalom

- Siptár Péter 2002. Optimális hiátustöltés. *Beszédkutatás* 2002. 70–83.
- Menyhárt Krisztina 2006. Koartikulációs folyamatok két magánhangzó kapcsolatában. *Beszédkutatás* 2006. 44–56.
- Olaszy Gábor 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben*. Nyelvtudományi Értekezések 155. Akadémiai Kiadó, Budapest.
- Olaszy, Gábor – Németh, Géza – Kiss, Géza 2001. Hungarian audiovisual prosody composer and TTS development tool. In Puppel, Stanislaw – Demenko, Grazina (eds.): *Prosody 2000*. Adam Mickiewicz University. Poznan, 167–177.