

## FORMÁNSOK A NYELV ÉS A BESZÉLŐ TÜKRÉBEN

Olaszy Gábor – Abari Kálmán

### Bevezetés

A beszéd a nyelv hangzó formája, a beszélő személy pedig az adott nyelvet beszélők egyike, akin keresztül a nyelv megszólaló hangalakja létrejön. Régi kérdés a fonetikában, hogy egy anyanyelvi közösség tagjainak artikulációs mozgásai mennyire hasonlóak a beszédképzésük során, illetve hogy mennyire függenek az egyéni biológiai sajátosságoktól. Az artikulációs mozgássorok és a beszéd formánsszerkezete között szoros a kapcsolat. Ha az artikuláció egy kicsit is változik, annak hatása megmutatkozik a beszéd hangzásában és a formánsok pillanatnyi elhelyezkedésében is (Fant 1960). Az artikuláció finom mozzanatai a beszélő egyén biológiai rendszerének és agyának pillanatnyi állapotától függenek. Adott nyelv esetében felmerül a kérdés, hogy a formánsok értékeiben kimutathatók-e olyan jellemzők, amelyek az egyénre vonatkoznak, illetve olyanok, amelyek az egyének között közel azonosnak mondhatók, tehát azokat a nyelv magasabb szintű rendszere alakítja ki. Ez utóbbit azt jelentheti, hogy az anyanyelvi artikulációs bázis bizonyos szintig szoros pontossággal működik minden beszélő személynél, ami anyanyelvi sajátosság, e felett pedig egyéni az ejtés.

A fonetikai kutatásokban a formánsok vizsgálata az egyik legrégebbi kutatási témakör. Ezen belül az egyén és a nyelv befolyásának tisztázása fontos téma általánosságban is (Stevens 1971) és speciális alkalmazásokhoz is (például büntügyi fonetika; Becker et al. 2008). A szakirodalom szerint a nyelvi információ szerepe elsődleges az alacsonyabb formánsok ( $F_1$  és  $F_2$ ) kialakulásában, míg  $F_3$  és  $F_4$  nagyobb mértékben tartalmaz a beszélőre jellemző információkat. Az egyéni belüli ejtés hosszú távú vizsgálatában kimutatták, hogy ugyanazon beszélő  $F_1$  és  $F_2$  formánsadatai ugyanabban a magánhangzóban, ugyanabban a kontextusban, többszöri kiejtés során kevésbé térnek el, mint az  $F_3$  adatai (Gósy 2004). Egy spanyol kutatás szerint a formánsok rangsorolása a személyfüggőségi skálán a következő: az  $F_2$  függ legkevésbé a beszélő személytől, utána következik az  $F_1$ . Az  $F_3$  és az  $F_4$  inkább a beszélő személy egyéni jellegzetességéhez köthető (Marrero et al. 2008). Jelen kutatásunkban a magyar beszéden végzünk ilyen irányú vizsgálatokat, megállapításainkat a formánsok hang- és mondatszintű mozgási formáira alapozzuk. A hangokban végzett méréseknél figyelembe vesszük a koartikulációs hatásokat, így kiszűrjük a hangkörnyezet befolyását a formánsmozgásokra. A mondatszintű egységeken végzett mérési formát azért választottuk, mert így az

adatsorok jobban tükrözik a folyamatos artikulációt. A vizsgálatainkban foglalkozunk azzal is, hogy van-e eltérés a nemek között a folyamatos beszédben. Kutatásunkat hangokra jellemző formánsvonulatokra alapozzuk, következtetéseinket a hangok 5 mért pontjával leírt formánsmozgásokból vonjuk le. A mondatok formánsvonulatait is a hangokra kapott adatokból számitjuk.

1. hipotézis: a magyar beszéd formánsvonulati alakjainak statisztikai vizsgálatával kimutathatók olyan tulajdonságok, amelyek inkább az artikulációt meghatározó nyelvre jellemzőek, és olyanok is, amelyek inkább az egyénre vonatkoznak.

2. hipotézis: a magyar beszédben a formánsvonulatok alakjai statisztikailag nem térnek el szignifikánsan a férfi és női beszélők esetében, vagyis a két nem artikulációs mozgásai egyformának tekinthetők.

### **Anyag és módszer**

A mérésekhez két adatforrást használtunk, az ARBA elnevezésű magyar formánsadatbázist (Olaszy–Abari 2015a, 2015b), valamint a formánsgeneráló TTS (text to formant) modellt (Abari et al. 2015a, 2015b), amit külön erre a vizsgálatra kibővítettünk az  $F_3$ -vonulatok jóslási képességével. Az ARBA formánsadatbázis mintegy 12 millió ellenőrzött formánsadatot tartalmaz (5 férfi és 5 női beszélő ugyanazt a 2000 mondatot olvasta fel), a TTS modellel pedig tetszőleges szövegre lehet  $F_1$ ,  $F_2$ ,  $F_3$  formánsadatokat kapni, tehát végtelen mennyiségű formáns származtatható a segítségével. Mindkét adatforrásban beszédhangonként 5 ponton találunk formánsadatokat (10, 25, 50, 75 és 90%). Mindezekkel biztosítottuk, hogy megfelelő sokaságú adat álljon rendelkezésre a statisztikai vizsgálatokhoz.

A hangszintű vizsgálatokban figyelemmel voltunk arra, hogy kiküszöböljük a koartikulációs hatások okozta formánsmozgási különbségeket. Ezért vizsgáltunk 5 hangból álló hangkapcsolati egységeket, azokon belül pedig a középső magánhangzót. Ezzel figyelembe vettük a szomszédos hangok koartikulációs hatását és az eggyel távolabbi hangokét is. Az 5 elemű egységek mérési terének kialakítását két lépcsőben végeztük el. Először az ARBA adatbázisból kiválogattuk az összes 5 elemű hangkapcsolatot, ez 21 003-féle különböző hangötöst jelentett, összesen 265 597 előfordulással. A második lépésben a statisztikai vizsgálatához két kritérium alapján szűrtük tovább a hangötösöket, 1. a minimális előfordulási mintaszámot 50-re állítottuk be hangötösönként, továbbá 2. kizártuk a mássalhangzó-torlódásokat, vagyis csak VCVCV hangötösökre korlátoztuk a mérési teret, ahol a középső V-t vizsgáltuk. A két kritériumot 200-féle különböző hangötös teljesítette, összesen 15 142 előfordulással. Az 1. táblázatban a mérésre kiválasztott hangötösök gyakorisági adatai láthatók a középső magánhangzó függvényében. Külön közöljük, hogy ez a magánhangzó hány egymástól különböző hangötösben szerepel (Különböző), valamint hány előfordulása van összesen (Összes). A hangötösök hangjait magyar SAMPA jelöléssel adjuk meg minden esetben.

1. táblázat: A vizsgált VCVCV hangötösök száma az adatbázisban a középső magánhangzó (Hang) szerinti gyakoriságban

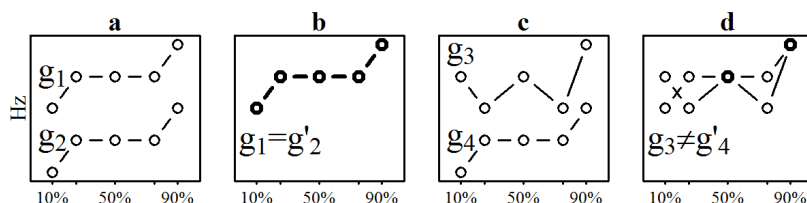
Hang	/O/	/a:/	/E/	/e:/	/i/	/i:/	/o/	/o:/
<b>Különböző</b>	39	16	63	19	19	4	18	5
<b>Összes</b>	3264	1020	4888	1306	1475	240	1356	381

Hang	/2/	/2:/	/u/	/u:/	/y/	/y:/
<b>Különböző</b>	8	3	3	1	1	1
<b>Összes</b>	560	250	182	70	60	90

A fenti táblázat szerint például az /y/ hang összesen 60 hangötösben fordul elő, amely ugyanannak az /e: p y l E/ hangsornak a 6 mondatban való előfordulása a 10 beszélő esetén. A másik véglet az /E/ hang összesen 4888 előfordulással, amely 63 különböző hangötösben fordul elő, hangötösönként különböző előfordulással a 10 beszélő esetén.

Hipotéziseink igazolásához a formánsok mozgásformáinak hasonlóságát vizsgáljuk ugyanazon hangkörnyezetű magánhangzóban. Két formánsmozgás ( $g_1$  és  $g_2$ ) hasonlósága a két görbe azonos  $t_i$  és  $t_{i+1}$  időpontjaiban mért formánsérték-változások hasonlóságán alapul, azaz hogy a  $g_1(t_i)$  és  $g_1(t_{i+1})$  közötti változás mennyire hasonló a  $g_2(t_i)$  és  $g_2(t_{i+1})$  közötti változáshoz. Az 1. ábra a) részében a két görbe egy-egy képzeletbeli formánsmozgást szemléltet. A két görbe által leírt formánsmozgás tökéletesen hasonló, hiszen a pontról-pontra bekövetkező formánsérték-változások azonosak (bár frekvenciában különböznek). Ez könnyen igazolható a b) ábrarészen látható módszerrel, ahol a  $g_2$  görbét addig mozgattuk az y tengely mentén, míg a  $g_1$  és  $g_2$  középső, 50%-os pontjai fedésbe nem kerültek. A  $g_2$  eltolásával kapott  $g'_2$  teljesen lefedi a  $g_1$  görbét, így a  $g_1$  és  $g_2$  formánsmozgások azonosnak tekinthetők, a hasonlóság a legnagyobb mértékű. A c) ábrarész két formánsmozgása ( $g_3$  és  $g_4$ ) azonban már nem hozható fedésbe minden ponton a d) ábrarészen, tehát a formánsmozgások hasonlósága már kisebb mértékű.

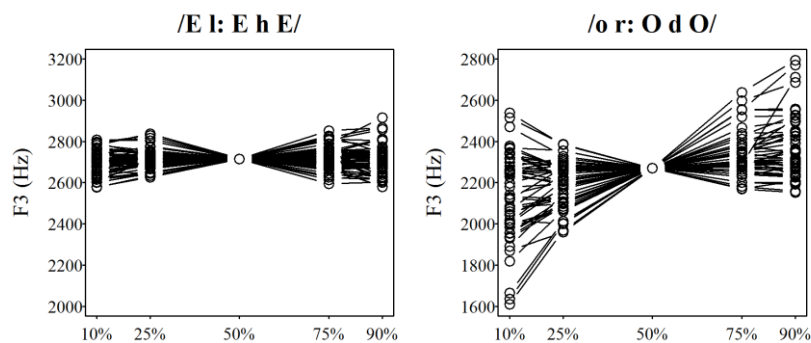


1. ábra

Hangon belüli formánsgörbék összehasonlításának példaábrái 5 mérési pont formánsfrekvencia-értékeivel  
(Az x tengelyen az idő, az y tengelyen a formánsérték szerepel.)

A formánsmozgások fenti értelemben használt hasonlóságának számszerűsítésére két olyan módszert vezetünk be, amelyek a görbék y tengelyen mért távolságát (az esetleges frekvenciakülönbségeket) figyelmen kívül hagyják, csak a görbék menetének azonosságát ítélik meg: 1. a mérési pontokban meghatározott görbékre jellemző szórás ( $s_g$ ), és 2. a Pearson-féle korrelációs együtthatón alapuló TMR értéket (lásd később részletesen).

**Görbékre jellemző szórás.** Az 1. ábrán bemutatott görbeeltolások módszer általánosítható úgy, hogy nemcsak két görbére vonatkozóan, hanem akár egy görbesereg esetén is képes legyen a formánsmozgások hasonlóságának számszerűsítésére. Így például sok ugyanolyan környezetű ugyanolyan hang formánsgörbéi jellemezhetők lesznek egy hasonlósági mértékkel. Itt mindegyik görbe egy adott hangötös középső magánhangzójában realizálódó természetes formánsmozgást ír le. A módszer lényege az 1. ábrán szemléltetett eljárás, mely szerint a görbesereg minden görbáját az y tengely irányában egy közös 50%-os pontba mozgatjuk el, majd a többi 4 pontban (10%, 25%, 75% és 90%-os pontokban), mint a görbesereg függőlegesen vett formánsfrekvencia-értékeinek 4 csoportjában meghatározzuk a szórásokat, majd ezt a 4 szórásértéket átlagoljuk. Ha az így kapott görbékre jellemző szórás ( $s_g$ ) kicsi, akkor a hangon belüli formánsmozgás alakja hasonló a vizsgált mintákban, azaz egyenesnek tekinthető a beszélők között, ha nagy, akkor egyénekenként más-más formánsmenetek vannak jelen. A 2. ábrán az /E l: E h E/ és az /o r: O d O/ hangsor középső hangjában megvalósuló  $F_3$  formánsmozgások közös pontba összetolt görbéit láthatjuk. A bal oldali részábrán a formánsmozgások hasonlóbbak, mint a jobb oldalin. A szórásértékek ezt megfelelően tükrözik: az /E/ hanghoz tartozó  $s_g$  értéke 55,9 Hz, míg az /O/ hanghoz tartozó 145,8 Hz.



2. ábra

Az  $s_{g,F3}$  szórás bal oldalon kicsi (55,9 Hz) az /E/ magánhangzóban, jobb oldalon nagy (145,8 Hz) az /O/ magánhangzóban.

A különböző formánsokhoz tartozó  $s_{g,F1}$ ,  $s_{g,F2}$  és  $s_{g,F3}$  görbékre jellemző szórások összehasonlításához figyelembe vettük a formánstér jellemző nagyságát is, és ezzel korrigáljuk a fent kiszámolt mutatókat. Ezt egyfajta normalizálási céllal tettük, hogy a kis mozgásokat (például  $F_1$ ) és a nagyokat (például  $F_2$ ) egyformán vegyük figyelembe a számításokban. A formánstér nagyságát az ARBA adatbázis 50%-os pozícióban mért formánsfrekvencia-értékei alapján számítottuk ki. A kapott értékek szélső értékeit nem vettük figyelembe, 1-1%-os szűkítést alkalmaztunk. Így a kapott formánstér a férfi és nő beszélők adatai alapján:  $d_{F1, \text{férfi}} = 491$  Hz,  $d_{F2, \text{férfi}} = 1581$  Hz,  $d_{F3, \text{férfi}} = 883$  Hz,  $d_{F1, \text{nő}} = 677$  Hz,  $d_{F2, \text{nő}} = 1941$  Hz,  $d_{F3, \text{nő}} = 1076$  Hz. Példaként megadjuk az első formáns görbéjére kiszámított, a formánstér nagyságával korrigált szórás kiszámítási módját férfi beszélők esetén:

(1)

$$s_{g, F1, \text{férfi}}^* = s_{g, F1} \cdot \frac{1}{d_{F1, \text{férfi}}}$$

Az (1) képletben leírt, férfi beszélőkre jellemző korrigált szórás egy dimenzió nélküli arányszám, amit százalékos formában is értelmezhetünk, és azt fejezi ki, hogy az  $F_1$  formánsterjedelmének hány százalékát teszi ki a formánsmozgások változékonysága. Ha értéke kicsi, akkor hasonlóak a formáns görbék, ha nagy, akkor eltérőek.

#### A TMR érték

A második módszerünk a mondatok formánsmeneteinek hasonlósági vizsgálatát teszi lehetővé, melyet az úgynevezett TMR (Trajectory Matching Rate) értékkel fejezzük ki, azaz a mondat formánsmozgásainak hasonlóságát számszerűsítjük. A számítás a Pearson-féle korrelációs együtthatón alapul. A TMR kiszámításánál felhasználjuk a TTF modell által generált formánsgörbét

is, amely a kitüntetett formánsgörbe szerepét tölti be. Az ARBA adatbázis természetes ejtéshez tartozó formánsgörbéit ehhez a kitüntetett formánsgörbéhez hasonlítjuk a korrelációs együttható segítségével. A TMR értékek meghatározásához a Hermes (1998) által javasolt megoldást vettük alapul, aki görbék hasonlóságát mérte így dallammenetekre. A TMR együtthatót minden formásra ( $F_1$ ,  $F_2$  és  $F_3$ ) külön-külön határozzuk meg, de a mondatból csak a magánhangzókat és zöngés mássalhangzókat vesszük figyelembe. A TMR meghatározásához először a Pearson-féle korrelációs együttható értékét számoljuk ki egy adott mondatra. Tegyük fel, hogy a TTF által generált mondat valamely formánsmenete  $m$  elemű vektor ( $x_{TTF}$ ), és ugyanezen mondat valamely természetes ejtésének ugyanezen formásra vonatkozó formánsvektora ( $y_{term}$ ) szintén  $m$  elemű. A (2) képlet alapján számoljuk ki a két formánsvektor között az  $r(x_{TTF}, y_{term})$  korrelációs együttható értékét:

(2)

$$r(x_{TTF}, y_{term}) = \frac{\sum_{i=1}^m (x_{TTF,i} - \bar{x}_{TTF})(y_{term,i} - \bar{y}_{term})}{\sqrt{\sum_{i=1}^m (x_{TTF,i} - \bar{x}_{TTF})^2} \sqrt{\sum_{i=1}^m (y_{term,i} - \bar{y}_{term})^2}}$$

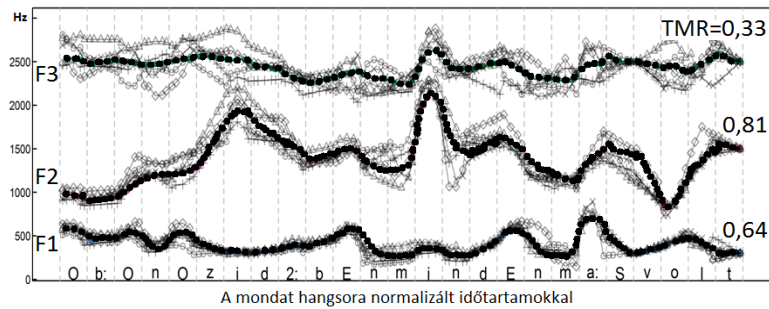
(3)

$$TMR = \frac{\sum_{i=1}^n r(x_{TTF}, y_i)}{n}$$

Amennyiben egy adott mondatra  $n$  db különböző természetes ejtésű mondatrealizáció áll rendelkezésre, akkor az  $n$  db  $r(x_{TTF}, y_{term})$  korrelációs együttható átlagolásával, a (3) képlet alapján kaphatjuk meg a mondatra vonatkozó TMR értéket. A TMR megadja az  $n$  természetes ejtés hasonlóságának mértékét a TTF modellel generált ugyanazon mondatához, mindhárom formáns esetén. Ha a TMR érték +1-hez közeli, akkor nagy a hasonlóság a természetes ejtések és a modell által adott formánsmenet között, vagyis a beszélőkre jellemző variáltság nincs jelen a természetes ejtésű mondatokban sem, tehát a formánsmozgás alakját inkább a nyelvi rendszer határozza meg. Ha valamely mondat formánsmenetére (például  $F_3$ -görbéjére) vonatkozó TMR érték távol van a +1 értéktől, akkor az azt jelenti, hogy a különböző beszélőkhöz tartozó természetes ejtésű formánsmenetek különböznek a TTF modell által generált formánsmenettől, és egyben egymástól is eltérők, vagyis ez a formánsmenet inkább a beszélőtől függ.

A 3. ábrán egy 25 hangból álló mondat szerepel, melyből 23 tartozik magánhangzóhoz vagy zöngés mássalhangzóhoz, így összesen  $m = 5 \times 23 = 115$  formánsadatból álló mondat TTF által generált és természetes ejtésű formánsmenet görbéi láthatók,  $F_1$ -re,  $F_2$ -re és  $F_3$ -ra. A világos színnel jelölt formánsmenetek 5 beszélő természetes ejtéséből származnak, a sötétebb színű görbe a TTF által generált formánsgörbe. Az egyes formánsokhoz tartozó TMR érték az ábráról is leolvashatók: az első formásra 0,64-et kapunk,  $F_2$ -re

0,81-et és  $F_3$ -ra 0,33-at. A 3. ábra formánsmeneteinek vonulataiból vizuálisan is jól kivehető, hogy az  $F_3$  esetében a természetes ejtésű mondatok görbéi sokkal rendezetlenebb hasonlósági képet mutatnak a TTF modell által generált görbéhez viszonyítva, mint az  $F_2$  és az  $F_1$  esetében.



3. ábra

Példa az  $F_1$ ,  $F_2$ ,  $F_3$  formánsmenetre egy vizsgált mondatban  
(A vastag fekete pontok görbéi a TTF modell kimenetei, a világosak az 5 férfi beszélő formánsgörbéi.)

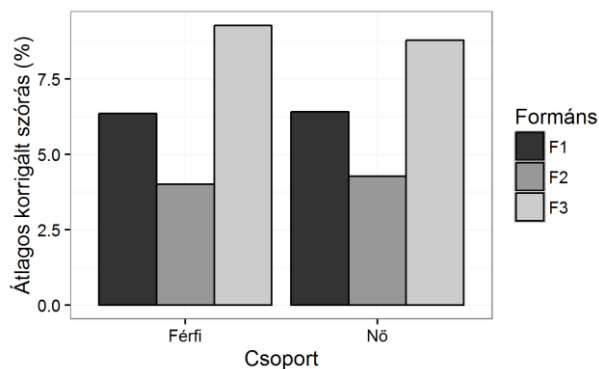
### Eredmények

Az eredmények bemutatását a hangszintű (hangötösre) vonatkozó vizsgálattal kezdjük. A férfi és nő beszélők adatait külön kezeltük, hogy a korrigált szórások kiszámításához használt, nemenként eltérő formánsterjedelmeket megfelelően figyelembe tudjuk venni. Az (1) képletben meghatározott korrigált szórást számoltuk ki a 200 különböző hangötösre a formánsok és nemek által közösen meghatározott csoportokban. Mindegyik csoport 50 vagy nagyobb számú realizációt tartalmazott. Egy csoportba kerültek tehát a különböző beszélők (nemenként 5-5) azonos hangötösre vonatkozó bemondásai. A 2. táblázat az (1) összefüggés alapján kiszámított 200  $s_g^*$  szórás átlagát tartalmazza formánsenként és nemenként.

2. táblázat: Átlagos, formánsgörbékre jellemző korrigált szórás ( $s_g^*$ ) százalékos formában nemenként

	$F_1$	$F_2$	$F_3$
<b>Férfi</b>	6,36	4,01	9,28
<b>Nő</b>	6,41	4,27	8,80

A legkisebb szórás az  $F_2$ -vonulatoknál figyelhető meg, majd az  $F_1$ -ek következnek. Nagy szórást mutatnak az  $F_3$ -vonulatai. Nemenként nincs eltérés a formánsokra vonatkozó szórások között (4. ábra).



4. ábra

A mért magánhangzók formánsmeneteinek átlagos görbékre jellemző korrigált szórása ( $s_g^*$ ) százalékos formában nemenként

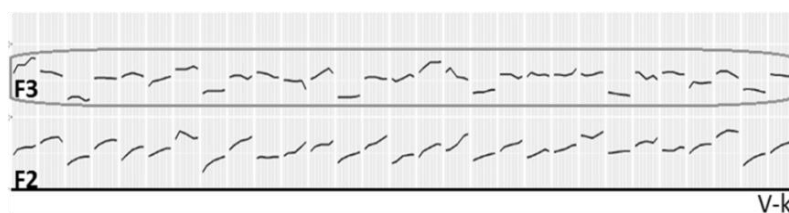
A kapott átlagos szórásoknál részletesebb vizsgálatokat is végeztünk az  $F_3$  vonatkozásában. A kérdés az, hogy az  $F_3$  minden hangötös esetében olyan erősen személyfüggő-e, mint ahogy az a 4. ábrából kiolvasható. Erre a kapott szórásadatok részletezett vizsgálata egyértelmű választ adhat. Amely hangötös esetében kicsi a szórás a vizsgált hangban a kontextust is figyelembe véve, abban kevésbé személyfüggő az  $F_3$  formánsmenet, mint azokban, ahol nagy. A 3. táblázatban a 10 legkisebb szórással rendelkező ötös hangkapcsolatot mutatjuk be.

3. táblázat: A 10 legkisebb átlagos szórással rendelkező hangötös

Hangötös	Gyakoriság	$s_{g,F3}$	$s_{g,F3,férfi}$	$s_{g,F3,nő}$
/E l: E h E/	60	55,9	43,9	65,8
/E l: E n E/	108	59,0	53,4	62,2
/E l E l 2:/	97	61,2	55,1	65,8
/e: t E l E/	69	61,5	49,4	70,2
/E z E l 2:/	50	62,7	54,0	69,8
/E l E v i/	50	63,9	52,7	71,8
/E z E t 2:/	139	64,1	61,8	66,0
/O l E h E/	69	64,7	58,5	63,3
/E j E z e:/	50	65,7	48,0	75,0
/E z E t e:/	68	66,5	69,3	62,8



A legkisebb szórás érték az /E l: E h E/ hangsorban van. Itt a középső /E/ hangban az  $F_3$  közel vízszintes vonulatot mutat a mért 60 ejtésben (5. ábra). Ebben a hangötösben tehát az  $F_3$  vonulati képe is személyfüggetlennek tekinthető.



5. ábra

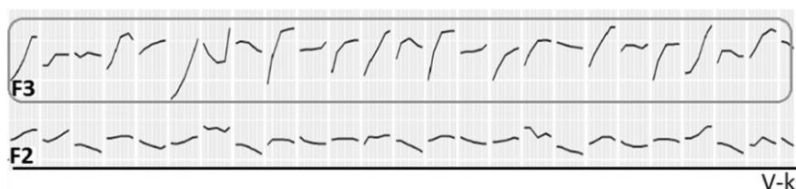
Az  $F_3$  formánsmenet alakjai az /E l: E h E t/ hangsor középső hangjaiban

Feltűnik az is, hogy a kis szórású rangsor elején majdnem mindegyik hangötös közepén /E/ hang szerepel. Tehát az /E/ hang artikulációja  $F_3$  tekintetében egységesebbnek mondható a beszélő személyeknél ilyen hangkörnyezetben, mint a többi hangoké. Ha a legnagyobb szórású értékeket vizsgáljuk meg, akkor a leginkább személyfüggő ötös hangkapcsolati adatokat kapjuk meg az  $F_3$ -ra (4. táblázat).

4. táblázat: A 10 legnagyobb átlagos szórással rendelkező hangötös

Hangötös	Gyakoriság	$S_{g,F3}$	$S_{g,F3,férfi}$	$S_{g,F3,nő}$
/o r: O d O/	50	145,8	143,1	148,5
/O m O d' O/	95	144,6	135,2	154,0
/O S: a: g o/	57	144,0	120,8	150,2
/O m a: S i/	94	141,7	139,7	143,7
/O m u t O/	70	139,3	127,0	143,3
/o h O S E/	119	136,9	116,4	147,7
/O m O t o/	50	135,9	112,4	153,8
/O h o d' O/	79	134,0	115,0	145,4
/O m a: S o/	50	133,1	135,8	129,1
/O m i k o/	80	132,9	111,2	143,8

Az /o r: O D O/ hangötösben a szórás 145,8. Az  $F_3$ -vonulatok változatosak a mért 50 mintában (6. ábra). Vannak szinttartó formák, homorú és domború, emelkedő és csökkenő változatok. Itt tényleg személyfüggő esetekkel állunk szemben.



6. ábra

Az  $F_3$  formánsmenet alakjai az /o r: O D O/ hangsor középső hangjaiban

Az adataink lehetőséget adtak arra is, hogy intraindividuális megállapításokat tegyünk a formánsmenetek jellemzésére használt szórások segítségével. Ha a formánsmeneteket csoportosítjuk hangötösre és azon belül beszélőre is, akkor összesen  $200 \text{ hangötös} \times 10 \text{ beszélő} = 2000$  csoportot képezzünk. Egy adott csoportban tehát a szórás számításához figyelembe vett formánsfrekvenciák ugyanahhoz a beszélőhöz és azonos hangötösköz tartoznak. Minden beszélő és minden hangötös összes előfordulására kiszámoltuk a szórás nagyságát formánsenként. Az egyes beszélőre kapott szórásokat átlagoltuk (5. táblázat). A beszélők közötti eltérések jól láthatók az adatokban, a tendenciák viszont minden beszélőnél ugyanazok. Tehát az átlagos szórás minden beszélőnél az  $F_1$ -ben és  $F_2$ -ben lényegesen kisebb, mint az  $F_3$ -ban. Ez összhangban van Gósy (2004) ide vonatkozó megállapításával.

5. táblázat: Átlagos, formánsgörbékre jellemző korrigált szórás ( $s_g^*$ ) százalékos formában az egyes beszélőkre, azonos hangötösre vonatkozó természetes ejtések alapján

Férfi beszélők	$F_1$	$F_2$	$F_3$	Nő beszélők	$F_1$	$F_2$	$F_3$
f1	5,5	3,7	11,9	n1	5,9	3,3	7,2
f2	6,1	5,2	12,0	n2	7,1	3,8	9,8
f3	4,5	3,0	7,1	n3	7,9	4,8	10,3
f4	4,7	2,9	7,2	n4	6,2	4,1	11,8
f5	5,4	4,0	10,2	n5	9,4	7,3	13,2
Átlag	5,25	3,76	9,67	Átlag	7,30	4,66	10,44

Az első hipotézis teljes körű vizsgálatában a hangszintű elemzések után rátérünk a mondat szintű elemzések eredményeinek ismertetésére. Itt más mértékű adathalmazt kellett feldolgozni, azaz sok hangból összeálló formánsmenetek összehasonlítása adta a végleges eredményt. A formánsmenetek hasonlóságának vizsgálatához a TMR értékeket számoltuk ki (6. táblázat). Az első két formánsnál magas a korreláció, mind a női, mind pedig a férfi ejtésű

mondatokban, az  $F_3$ -nál sokkal alacsonyabb. Az eredmények tehát hasonló tendenciát mutatnak, mint hangszintű vizsgálatok.

6. táblázat: Az átlagos TMR korrelációs értékek az ARBA adatbázis 2000 mondatára a beszélők modelljeinek a függvényében

TTF modell	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>
5.sp. nő modell	0,75	0,87	0,46
5.sp férfi modell	0,76	0,87	0,44

A kétfajta vizsgálat együttes eredményei alátámasztják azt, hogy a formánsmenetek hasonlóságának vizsgálatával is kimutatható, hogy az  $F_3$  inkább egyéni jellemző, az  $F_1$  és  $F_2$  pedig a nyelvre jellemző, továbbá hogy az  $F_2$  áll a legszorosabb kapcsolatban a nyelvi meghatározottsággal. Az első hipotézisben feltett kérdésre a válasz tehát az, hogy az alsó két formáns vonulatai kevésbé függenek a beszélő személytől, míg az  $F_3$  inkább attól függ, de ez a függés bizonyos hangsorok esetében nem mutatható ki.

A második hipotézis vizsgálatához a TTF férfi és női modelljeivel generáltuk le az ARBA adatbázis 2000 mondatát. Így adatokat kaptunk külön a férfi és külön a női ejtésre. A minden mondat minden formánsmenetére kiszámított TMR értékek:  $F_1 = 0,98$ ;  $F_2 = 0,97$ ;  $F_3 = 0,58$ . Ezzel igazolódott a hipotézis,  $F_1$  és  $F_2$  vonatkozásában nincs különbség a nemek közötti ejtésben.

### Összefoglalás

Cikkünkben a fonetikában ismert jelenségeket és tényeket új módszerekkel vizsgáltuk. Célunk az volt, hogy az eddigi eredményeket nagy tömegű adaton végzett statisztikai vizsgálatokkal is alátámasszuk. Kimutattuk, hogy a formánsokra meghatározható inter- és intraindividuális tulajdonságok tekintetében az  $F_2$  formánsmenet kialakítása függ legkevésbé a beszélőktől, ezt követi az  $F_1$ . Ezt a két formánst tehát alapvetően a nyelv határozza meg. Az  $F_3$  tekintetében a beszélőktől függő változatosság mutatható ki. Az  $F_3$  formánsmeneteinek részletesebb vizsgálata ugyanakkor megmutatta, hogy ez a változatosság hangkörnyezet függő, bizonyos hangsorok esetében szinte megszűnik (ez további kutatásokat kíván). Kimutattuk továbbá, hogy vizsgálati eredményeink nem függenek a beszélő nemétől.

### Irodalom

- Abari, Kálmán – Csapó, Tamás Gábor – Tóth, Bálint Pál – Olasz, Gábor 2015a. From text to formants – indirect model for trajectory prediction based on a multi-speaker parallel speech database. In: *Proceedings of Interspeech 2015*. Dresden, Germany. 623–627.

- Abari Kálmán – Csapó Tamás Gábor – Tóth Bálint Pál – Olaszy Gábor 2015b. Szöveg-formáns konverter [Számítógépes program]. <http://magyarbeszed.tmit.bme.hu/ttf/> (A letöltés ideje: 2016. november 30.)
- Becker, Timo – Jessen, Michael – Grigoras, Catalin 2008. Forensic speaker verification using formant features and Gaussian mixture models. In: *Proceedings of Interspeech 2008*. Brisbane, Australia. 2434–2437.
- Fant, Gunnar 1960. *Acoustic theory of speech production*. Mouton de Gruyter, The Hague.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Hermes, Dik 1998. Measuring the perceptual similarity of pitch contours. *Journal of Speech, Language, and Hearing Research* 41. 73–82.
- Olaszy Gábor – Abari Kálmán 2015a. Az artikulációs mozgások akusztikai vetületeinek adatbázisa magyar beszédre. *Beszédkutatás 2015*. 199–210.
- Olaszy Gábor – Abari Kálmán 2015b. Az artikuláció akusztikai vetületeinek adatbázisa magyar beszédre. [Számítógépes program]. <http://magyarbeszed.tmit.bme.hu/artikulacio/> (A letöltés ideje: 2016. november 20.)
- Marrero, Victoria – Battaner, Elena – Gil, Juana – Llisterri, Joaquim – Machuca, Maria – Marquina, Montserrat – De La Mota, Carme – Rios, Antonio 2008. Identifying speaker-dependent acoustic parameters in Spanish vowels. *Acoustics 08*. 5673–5677.
- Stevens, Kenneth 1971. Sources of inter- and intra-speaker variability in the acoustic properties of speech sounds. In: *Proceedings of the 7th International Congress of Phonetic Sciences*. Montreal. 206–227.

### Formant trajectories, aspects of language and speakers

The research intends to show the dependency of certain formant trajectories either on individual patterns of the speaker or on the properties of Hungarian language. For measurements, the ARBA speech database (that contains 2,000 sentences produced by 5 male and 5 female speakers), on the one hand, and  $F_1$ ,  $F_2$ ,  $F_3$  formant trajectories (Text-to-Formant TTF statistical model developed by the authors) were used, on the other. Two types of analysis were carried out. (i) The values of the standard deviations were examined in the mid-point of the middle vowel within VCVCV sequences using the sentences of the ARBA database. (ii) Formant trajectories measured in 2,000 sentences of all speakers were compared among them using Trajectory Matching Rate (TMR). The results show that both  $F_1$  and particularly  $F_2$  trajectories are strongly influenced by the language.  $F_3$  trajectory seems to be speaker-dependent.