

A magyar magánhangzók első két formánsának meghatározása szintetizált hangmintákat felhasználó percepció kísérlet segítségével

A magyar magánhangzók akusztikai szerkezetét ez ideig számos kutató vizsgálta behatóan. Figyelemreméltó eredményeiket több cikkben, könyvben ismertették (vö. Magdics 1965, Molnár 1970, Tarnóczy 1974, Szende 1976, Bolla 1978, Vértés O. 1982). Ezirányú kutatásaik nagy része a magyar magánhangzók formánsszerkezetének meghatározására irányult. Bennünket különösen az első (F1), és a második (F2) formánsra vonatkozó méréseik érdekeltek. Az eredmények ellenére úgy érezzük, hogy egy ponton érdemlegesen kiegészíthetjük az eddigi vizsgálatokat. Ugyanis a felsorolt kutatók — Bolla Kálmán kivételével (lásd még később) — vizsgálódásaikat csupán az analízis oldaláról végezték el. Továbbá az MTA Nyelvtudományi Intézetének fonetikai osztálya laboratóriumában az utóbbi években lezajlott korszerűsítés nyomán rendelkezésünkre álló számítógépes beszéd szintetizáló rendszer szinte kínálja magát, hogy a fent említett méréseket a „szintézises elemzés” (analysis by synthesis) módszerrel is kiegészítsük. Sőt az előbbi módszer felhasználására előkészített hanganyagot auditív—percepció teszteléssel próbáljuk még egzaktabbá tenni.

A „szintézises elemzés” módszerére magyar nyelven először Fónagy Iván hívja fel figyelmünket Laziczius Gyula Fonetika című művé 1963. évi kiadásának utószavában, a következő szavakkal: „A szintetizáló készülékek tudományos feladata az analízis. Ezt sokkal pontosabban látják el, mint a hangszínelemző, hangszíníró készülékek. Hiszen a legpontosabb spektrogram sem árulja el, melyek a beszédhang felismeréséhez nélkülözhetetlen jegyek. Az analízis szintézis útján (analysis by synthesis) a beszédakusztika legújabb jelszava.” (I. m. 192)

Ezek ismeretében elmondhatjuk, hogy munkánk elé azt a célt tűztük, hogy a legkorszerűbb módszerek egyikével végrehajtott, lehetőleg a szubjektív mozzanatoktól minél inkább mentes kísérlet során kapjunk választ a magyar magánhangzók formánsszerkezetére (különösen az F1, F2 formánsra tekintettel).

Munkánk emellett még más, témánkhoz kapcsolódó kérdésre is megpróbál választ adni.

Hogyan lehet ezt a kissé „túl jól hangzó” módszert a gyakorlatban alkalmazni? A válasz a következő: az eddigi bőséges irodalom adatait figyelembe véve, számítógépes beszéd szintetizáló rendszerünkkel előállítunk viszonylag nagyszámú (magánhangzó jellegű) hangmintát tartalmazó tesztlelésre alkalmas hanganyagot. Az egyes hangminták első és második formánsának frekvenciáját úgy állítjuk be, hogy a hangminták összessége „lefedje”

az összes magyar magánhangzó F1, F2 koordináta-rendszerben ábrázolt szórás képét. Majd a hangmintákat percepciós kísérlettel végzett „ellenőrzésnek” vetjük alá. Kellően nagy létszámú csoportot kérünk meg, hogy jelöljék meg azokat a hangmintákat, amelyeket valamely magyar magánhangzóval azonosítani tudnak és tüntessék fel, hogy melyik ez a magánhangzó. A válaszokat összesítve — a frekvenciaméréseket végző kutatóktól eltérő módszert alkalmazva — előttünk is kirajzolódik a magánhangzók szórás képe az F1, F2 koordináta-rendszerben. Ezáltal az eddigi kutatások eredményeit összehasonlíthatjuk saját eredményeinkkel.

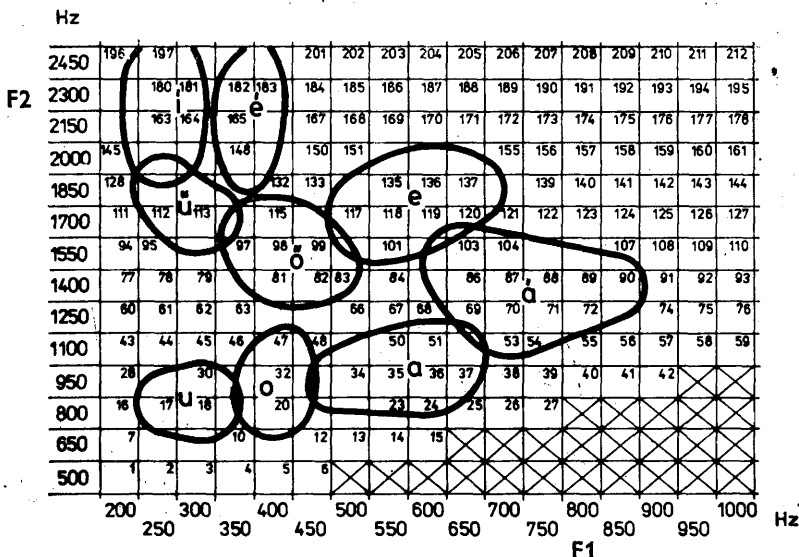
Mint a fentiekben utaltunk rá, a hangmintákat a fonetikai laboratórium INBERE nevű (vö. Kiss—Olaszy 1982/a) ún. interaktív beszédszintetizáló rendszerével készítettük. A szintetizáló rendszerben az OVE III beszédszintetizátort egy PDP 11/34 típusú számítógép vezérelte a saját fejlesztésű FOPRO program (vö. Kiss—Olaszy 1982/b) segítségével.

Kísérletünk megtervezése során nagymértékben merítettünk Franco E. Ferrero és Emanuela Magno Caldogneto, a padovai kísérleti-fonetikai központ (Centri di Studio per le Ricerche di Fonetica del CNR) két munkatársának Risultati sperimentali di una prova di percezione di suoni vocali sintetici (A szintetizált magánhangzók percepciós próbájának kísérleti eredményei) című, 1976-ban publikált cikkéből.

A fenti cikkben leírtakhoz hasonlóan 212 hangmintát állítottunk elő, amelyekben az első formáns (F1) és a második formáns (F2) frekvenciáját változtattuk. Az első formáns értéke 200 Hz-től 1000 Hz-ig terjedt 50 Hz-es ugrásokkal, összesen 17 lépésben, a második formáns értéke 500 Hz-től 2450 Hz-ig, 150 Hz-es ugrásokkal, összesen 14 lépésben. Az így elkészíthető $17 \times 14 = 238$ hangmintából csak azt a 212-t használtuk fel, amelyek esetében az első formáns frekvenciáértéke nem haladta meg a második formáns frekvenciáértékét. Végül is az elkészített 212 hangminta F1, F2 koordináta-rendszerben való elhelyezkedését az 1. ábra mutatja. Az ábrába berajzoltuk Tarnóczy (vö. Tarnóczy 1974) mérési eredményeit.

A hangminták akusztikai leírásával kapcsolatban meg kell jegyeznünk, hogy a harmadik formáns (F3) frekvenciáját mindegyik hang esetében 2700 Hz-nek választottuk, míg a negyedik formáns frekvenciáját szintetizáló rendszerünk automatikusan 3500 Hz-en állította elő. Az elkészített hangminták alaphangértéke (F0) 116 Hz volt. Mivel a szintetizált hangminták természetesen egymáshoz képest igen nagy intenzitásbeli szórást mutattak (vö. Olaszy 1974), ezért dinamikai korrekciót végeztünk rajtuk. Így közel egyenlő hangerősséggel hallották a kísérleti személyek a mintákat. Egy-egy hangminta időtartama 500 ms volt. A szintetizált hangminták időtartamának azért választottuk ezt a spontán beszédben előforduló magánhangzók időtartamát többszörösen is meghaladó időtartamértéket, mert Gósy M. percepciós kísérlete szerint a magánhangzók időtartamának növekedésével az egyes fonémakategóriák formánsainak realizációs frekvenciasávja csökken (vö. Gósy 1984). Így a hosszabb hangmintákra adott percepciós válaszok pontosabb értéket tükröznek. Kísérletsorozatunk kiindulási pontja tehát 212 akusztikailag magánhangzó jellegű, 500 ms-os gépi hang volt. Ezek a hangminták nincsenek „akusztikailag” hozzágörbítve egy organikus nyelvi hangkörnyezethez. Azonban fonémarealizációs minősítését ezeknek a hangeseményeknek alátámasztja az a tény, hogy a kísérleti személyek által magyar magánhangzókkal azonosíthatóaknak bizonyultak.

A percepció tesztelés előkészítése során a mintákat magnetofonszalagra vettük fel — véletlen elrendezésben — oly módon, hogy mindegyik mintát háromszor rögzítettük egymás után, közöttük 500 ms-nyi szünettel. A hangszalagon a három azonos hangmintát 20 másodperces szünet követte. A tesztelés folyamán a kísérleti személyektől azt kértük, hogy a két minta közötti



1. ábra. Az F1, F2 koordináta-rendszerben az előállított 212 hangminta elhelyezkedése. Az ábrán feltüntetettük Tarnóczy Tamás mérési eredményeit

20 másodperces szünet alatt döntsék el, hogy az előzőekben hallott három egyforma hangminta melyik magyar magánhangzóval azonosítható, illetve azonosítható-e egyáltalán valamelyikkel. Az általunk ily módon összeállított kísérleti anyagot 20 kísérleti személlyel hallgattattuk meg. Ezzel próbáltuk saját szubjektív döntésünket a kísérletből kirekeszteni, és azt a felmerülő hibalehetőséget munkánkból kizárni, melyről Bolla Kálmán a következőképpen ír: „Ezután került sor a szórási sáv megállapítására. Kiindulásul mérési adataink szolgáltak, de a küszöb és a plafon helyét minden esetben a szubjektív észlelés döntötte el. Ott, ahol fülünk már nem a vizsgált hangtípushoz kapcsolta az előállított hangzást vagy abban egy, a magyar köznyelvben nem használatos realizációt hallott, próbáltunk megkeresni egy optimális határvonalat. *Nem állítom, hogy ez így teljesen problémamentes megoldás*, de az eddigieknél mindenképpen egzaktabb alapokon nyugszik, s ami talán a leglényegesebb, hogy nyelvtől releváns eredményt kaptunk.” (Vö. Bolla 1978. 57) Az általunk kiválasztott 20 kísérleti személy fele férfi, fele nő volt. Többségük életkora 20 és 30 év között mozgott. Mindegyikük legalább középfokú iskolai végzettséggel rendelkezik, budapesti születésű, és iskoláikat is a fővárosban végezték. Fonetikailag nem voltak képzettek. Mindnyájan a köznyelvet beszélték. A tesztelés során, mivel a tesztanyag majd másfél óra volt, három ízben hosszabb szünetet tartottunk.

Kísérletünk megkezdésekor tudatában voltunk már annak, hogy két fontos kérdésre nem kaphatunk választ. Az első az a kérdés, hogy miben tér el a férfi és a női ejtésű magánhangzók formánsszerkezete. Ugyanis az általunk beállított 116 Hz-es alaphang férfi ejtésnek felel meg, így a kísérleti személyek természetesen „férfi” ejtésben hallották a szintetizált hangmintákat. (A későbbiekben tervezzük ugyanennek a hangmintaanyagának az előállítását női ejtésnek megfelelő alaphanggal — pl.: 220 Hz — értékkel is.) Így az első kérdésre a válasz a két kísérlet eredményeinek összehasonlításával adható meg.

Másodjára felmerül az a kérdés, hogy a rövid-hosszú magánhangzó-párok (*i—í, o—ó, u—ú, ü—ű*) rövid és hosszú tagjai között mi az akusztikai különbség, különösen az F1, F2 formánsok elhelyezkedését tekintve. Erre a kérdésre — úgy gondoljuk —, hogy az eddigi méréseket kiegészítve a választ a miénktől eltérő, de jellegében hasonló észlelési kísérletsorozat adhatja meg.

Egyébiránt felmerülhet az előbbi két kérdésen kívül még egy további kérdés is: a kísérlet miért csak az F1 és az F2 meghatározására terjedt ki, és miért nem tér ki a további formánsok, a harmadik (F3), a negyedik (F4), stb. formánsok vizsgálatára. A szakemberek jó részének az a véleménye, hogy a magánhangzók minőségét elsősorban az első két formáns elhelyezkedése határozza meg. Többek között ez Tarnóczy Tamás véleménye is (vö. Tarnóczy 1974. 184) „...nem helyes túl nagy jelentőséget tulajdonítani a felsőbb formánsoknak. Általában az a korszerű felfogás, hogy az első két formáns határozza meg a magánhangzó kvalitását, a többiek az egyéni hangszín és a járulékos közlési szándék (pl. emóció) hordozói.” Ettől kissé eltérően nyilatkozik Gordos Géza és Takács György. Digitális beszédfeldolgozás című könyvükben ez áll: „Jó minőségű beszédfelvételeket elemezve általában az első 2—6 formáns emelkedik ki az alapzajból. Az első három formáns frekvenciaértéke a magánhangzók jellegét határozza meg, a magasabb formánsfrekvenciák főleg a beszélő személyre jellemzőek.” (I. m. 56) Tehát Gordosék a harmadik formánsnak is jelentőséget tulajdonítanak a magánhangzó minőségének a kialakításában.

Kísérletünk első felének, az auditív percepciós tesztelésnek a befejeztével az egységes formátumú tesztlapokról a válaszokat (az adatokat) számítógéppel dolgoztuk fel. A feldolgozás első lépéseként kifejlesztettük az ITALA nevű számítógépprogramot. A programot FORTRAN nyelven írtuk meg, terjedelme mintegy 550 sor. E program segítségével helyeztük el a feldolgozandó adatokat egy ún. számítógépes adatbázisban, és ennek a programnak a segítségével tettünk fel kérdéseket, dolgoztuk fel, végeztünk számításokat a számítógépben levő adatokkal. Az előbb említett adatbázis tulajdonképpen két-dimenziós, 212-szer 20-as nagyságú mátrix. E mátrix egyes soraiban a húsz kísérleti személy részéről az egyes hangmintákra adott válaszok találhatóak, míg oszloponként az egyes kísérleti személyek által a hangmintákra vonatkozólag adott 212 válasz olvasható le. A korábbiakban említettük, hogy a hangmintákat véletlenszerű elrendezésben hallgattattuk meg a kísérleti személyekkel. Az adatbázist is ezzel a sorrenddel építettük fel, majd a teljes felépítés után az ITALA program egy szubrutinja állította elő az eredeti, tényleges sorrendnek megfelelő állapotot. Az adatokat úgy helyeztük el ebben a mátrixban, hogy az 1—10. oszlopban férfi, míg a 11—20. oszlopban a női kísérleti személyek válasza található meg. Így arra a kísérlet végrehajtása során felvetődött kérdésre, hogy a férfi, illetve a női kísérleti személyek válasza némiképp eltér-e egymástól, szintén választ kaphatunk.

Nevezzük a továbbiakban az üres magánhangzónak azokat a hangmintákat, amelyeket a kísérleti személyek egyetlen magyar köznyelvi magánhangzóval sem azonosítottak. Az értékelés során a fenti mátrixból kiindulva összeszámoltattuk számítógépes programunkkal, hogy a $20 \times 212 = 4240$ észlelési adatból hány esik az egyes magánhangzókra. A számítás eredményét az 1. táblázat tartalmazza. Az üres magánhangzóra 1360 válasz jutott, ez az összes válaszok 32 százaléka. Átlagban tehát egy kísérleti személy a 212 szin-

1. táblázat

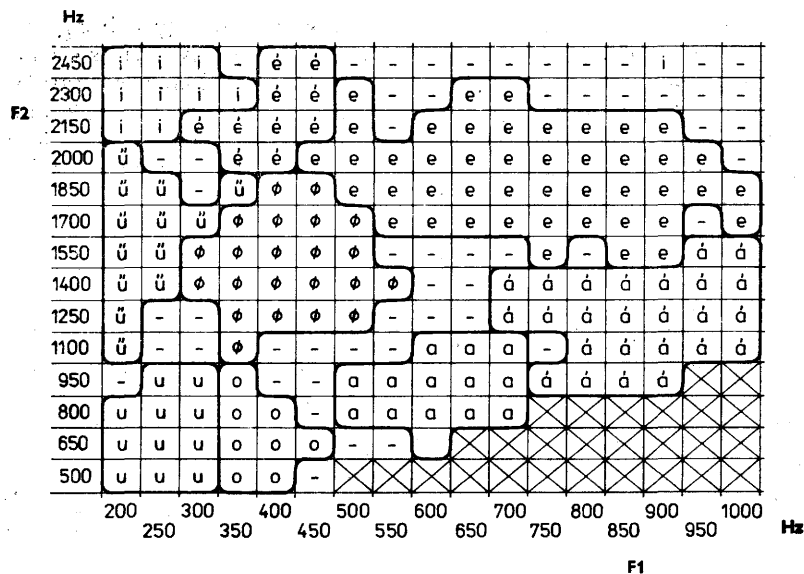
A 4240 válasz közül az egyes magánhangzókkal történő azonosítás száma

Sorszám	A magánhangzó	A 4240-ból azonosított minta	Egy személy által átlagban azonosított minta	Százalékban
1.	üres	1360	68,0	32
2.	e	614	30,7	14,5
3.	ö	520	26	12,3
4.	á	510	25,5	12
5.	a	317	15,85	7,5
6.	é	256	12,8	6
7.	o	188	9,4	4,4
8.	ü	183	9,15	4,3
9.	i	159	7,95	3,8
10.	u	133	6,65	3,2
		4240	212	100

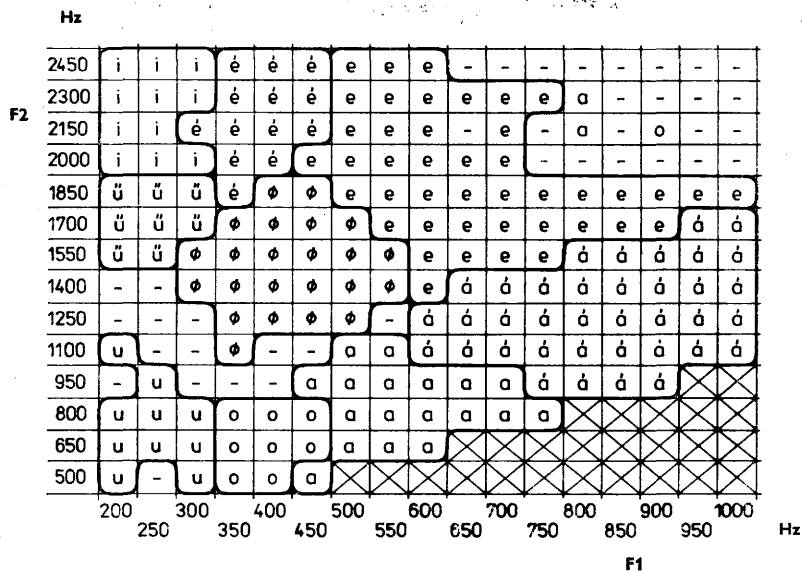
tetizált hangmintából 68-at nem azonosított egyetlen magyar magánhangzóval sem. A táblázatot tovább szemlélve megállapíthatjuk, hogy az „e” hang szórása a legnagyobb, az „e” hang fedi le a legnagyobb területet az F1, F2 tengelyek által kifeszített negyedsíkban. Az „e” hang vezető szerepét azzal magyarázzuk, hogy ennek a hangnak az észlelése során az esetlegesen két fonémát realizáló nyílt és zárt „e”-t egyaránt apperceptiálták a kísérleti személyek, ennek a hangnak ezért ilyen széles az akusztikai spektruma.

Továbbá, ha elfogadjuk, hogy valamely magánhangzó első formánsának frekvenciája szorosan összefügg: mégpedig fordítottan arányos a hangképzés közben a nyelv emelkedésének fokával, valamint a második formánsa a magánhangzónak egyenesen arányos a hang képzése közben a nyelv előretoltságának mértékével, akkor a következő megállapításokat tehetjük: a szájüregben az „e” hang képzése igen széles területen realizálódhat, míg az „u” hang adja a legkisebb mozgásszabadságot a nyelvnek a hangképzés közben. Az előbbi táblázat értékei ezt tükrözik.

Elkészítettük a kísérleti személyek válaszaiból összeállítható 20 szórás-képet, amelyeket szintén az F1, F2 koordináta-rendszerben helyeztünk el a számítógép. Mind a 20 szórás-képet jól reprezentálja a J. B. (férfi) és az F. A. (nő) kísérleti személyek válaszaiból kialakított két szórás-kép, amelyeket a 2. és a 3. ábrán mutatunk be. Az ábrákon az adott sorszámú szintetizált hangminta (lásd 1. ábra) F1, F2 frekvenciáját leolvashatjuk a két tengelyen, ugyanakkor a két érték kereszteződésében láthatjuk meg, hogy a hangmintát mely magánhangzóval azonosította a kísérleti személy. Ahol a „—” jelet találjuk, azt a hangmintát nem azonosította a kísérleti személy egyetlen magyar magánhangzóval sem. Az ábrákról jól leolvasható, hogy a kísérleti személyek egye-



2. ábra. A J. B. (férfi) kísérleti személy válaszaiból kirajzolódó magánhangzó szóráskép

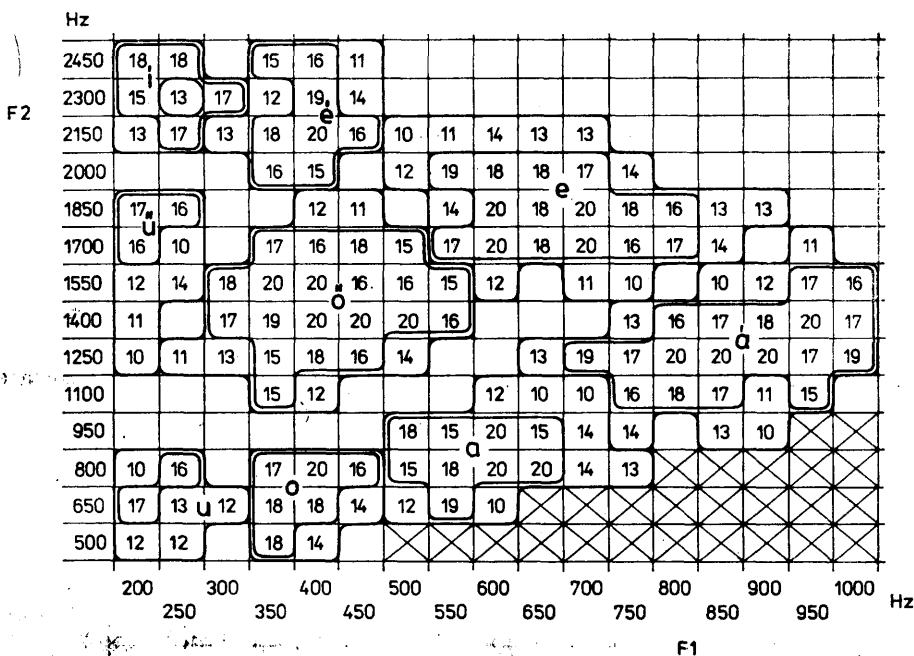


3. ábra. Az F. A. kísérleti személy válaszaiból kirajzolódó magánhangzó szóráskép

magánhangzóinak elhelyezkedése viszonylag zárt, kompakt szórásképet mutat, mindegyik hang esetében. Ez azt bizonyítja szerintünk, hogy minden magyar anyanyelvű egy jól körülhatárolt, konkrét „hangképpel” rendelkezik, amelynek realizációs halmaza akusztikailag pontosan leírható.

Feldolgozásunk következő lépéseként számítógépes programunk elkészítette az előbbieken említett 20 szóráskép összesített eredményét. Ennek

nyomatott képét a 4. ábra tartalmazza. Az ábrán az egyes hangmintáknál csak azt a magánhangzót tüntettük fel, amelyiket a kísérleti személyek legalább 50 százaléka észlelt. A jobb megfigyelhetőség érdekében két szintvonalat is berajzoltunk az ábrára. Az egyik szintvonalat azon hangminták köré húztuk, amelyeket a kísérleti személyek 75–100 százaléka az adott magánhangzónak jelzett, vagyis legalább 15 személy azonosította ugyanazzal a hang-



4. ábra. A magyar magánhangzók szórásképe (a 20 kísérleti személy válaszaiból összesített eredmény)

gal. A második szintvonalat pedig azon hangminták köré húztuk, amelyeket a kísérleti személyek legalább 50 százaléka, azaz legalább 10 személy az adott magánhangzónak jelzett.

Meglepőnek és megjegyzésre érdemesnek találjuk, hogy az egyes hangok szórási területének közepétől a szélek felé haladva nem fokozatosan csökken az azonosított hangminták száma, hanem kb. az 50 százalékos azonosítási értéknél, vagyis a legalább 10 személy által az adott magánhangzónak jelzett hangminták után, hirtelen lecsökken az azonosítások száma. Úgy gondoljuk, hogy ez annak a következménye, hogy a 20 személy által megadott szórásképek igen közel állnak egymáshoz. Így tehát az összesítés „nem mosta el” a jellemző adatokat, nem hozott ki félrevezető átlageredményeket, hanem releváns értéket tükröz. Ezt bizonyítja az a tény is, hogy minden magánhangzó esetében — kivéve az „i”-t, „ü”-t és az „u”-t — volt olyan hangminta, amelyet mind a húsz kísérleti személy ugyanannak a magánhangzónak jelzett. És itt talán választ adhatunk arra a korábbi kérdésre is, hogy a harmadik formáns hozzátartozik-e a hang kvalitását meghatározó tényezőkhöz. Az előbbieket figyelembev-

telével azt kell mondanunk, hogy a magas nyelvállással képzett magánhangzók — különösen az „i” és az „u” — esetében a harmadik formánsnak kvalitásmeghatározó szerepe van. Hiszen, ha a harmadik formáns frekvenciáját nem rögzítjük egységesen 2700 Hz-ben, hanem változtatgatjuk, akkor bizonyára az „i” és az „u” esetében is találtunk volna olyan hangmintát, amelyet a többi magánhangzóhoz hasonlóan mind a húsz kísérleti személy a fent említett két hang valamelyikével azonosított volna. A többi magánhangzó esetében — kísérletünk tanúsága szerint — a harmadik formáns frekvenciájának nincs döntő jelentősége a magánhangzó minőségének meghatározásában.

A 4. ábrán az „üres” magánhangzó szórás képét figyelve szemünkbe ötlük, hogy három nagyobb, jól elkülöníthető csoport alkotja a szórás képét. Az üres magánhangzó szórás képét alkotó egyik csoport az ábra jobb felső háromszög részében helyezkedik el, várakozásunknak megfelelően. Ennél a csoportnál váratlanabb annak a csoportnak a megjelenése és pozíciója, amely az „ö” hang szórás képétől élesen elkülöníti az „o”, „a”, „á” hangok szórás képét. Ez azt bizonyítja, hogy az „ö” hang realizációs halmaza az előbbi három hang realizációs halmazával nem határos. Meglepő módon ilyen éles elkülönítést nem találunk az „ö” hang másik oldalán, az „ö” és az előképzett „e”, „é”, „i” hangok szórás képe között.* Az utóbbi három magánhangzó szórás képének találkozásában a 130. és a 134. hangmintában is realizálódó „üres” magánhangzó jól jelzi a határokat. Az „üres” magánhangzó szórás képét alkotó harmadik nagyobb csoport az „u” és az „ü” hangok szórás képe között helyezkedik el.

A 2. táblázatban azokat a hangmintákat mutatjuk be, amelyeket a kísérleti személyek feltűnő módon több különböző magánhangzónak is azonosítottak. A táblázatba csupán azokat a hangmintákat vettük fel, amelyeket a kísérleti személyek legalább 25 százaléka (legalább 5 fő) egyébnek azonosított, mint a kísérleti személyek egy másik 25 százaléka. Meglepően kevés hangminta került ebbe a táblázatba, összesen 16, vagyis a 212 hangminta 7,5 százaléka. Ez azt bizonyítja, hogy a hangminták igen nagy százaléka egyértelműen tartozott valamely magánhangzóhoz. Érdemes megemlíteni még azt a tényt, hogy a 212 hangminta között egyetlen egy sem volt, amelyet három magánhangzónak azonosítottak volna oly módon, hogy a kísérleti személyek legalább 25, 25, 25 százaléka egységesen azonos magánhangzónak jelezte volna.

Az 5. és a 6. ábrán bemutatjuk az előző ábrák egy bizonyos szempontból való összesítését. Nevezetesen összeszámoltattuk számítógépes programunkkal azt, hogy az azonos F1 frekvenciájú hangmintákra összesen hány azonosítást tettek az egyes magánhangzók vonatkozásában a kísérleti személyek. Ezt az összeszámolást nemcsak az első formánsra, hanem további lépésként az azonos második formáns frekvenciájú hangmintákra is elvégeztettük. Az ábrákat szemlélve elmondhatjuk, hogy a korábbiakban említett 50 százalékos azonosítási érték körüli hirtelen csökkenés itt is jól megfigyelhető. Amennyiben ez nem így lenne, hisztogramjaink jóval laposabb és elnyúltabb haranggörbéhez hasonlítanának. Egyedül talán csak az „ö” hang mindkét, valamint az „e” hang F1 formánsának hisztogramja kivétel, az aránylag széles, elnyúl-

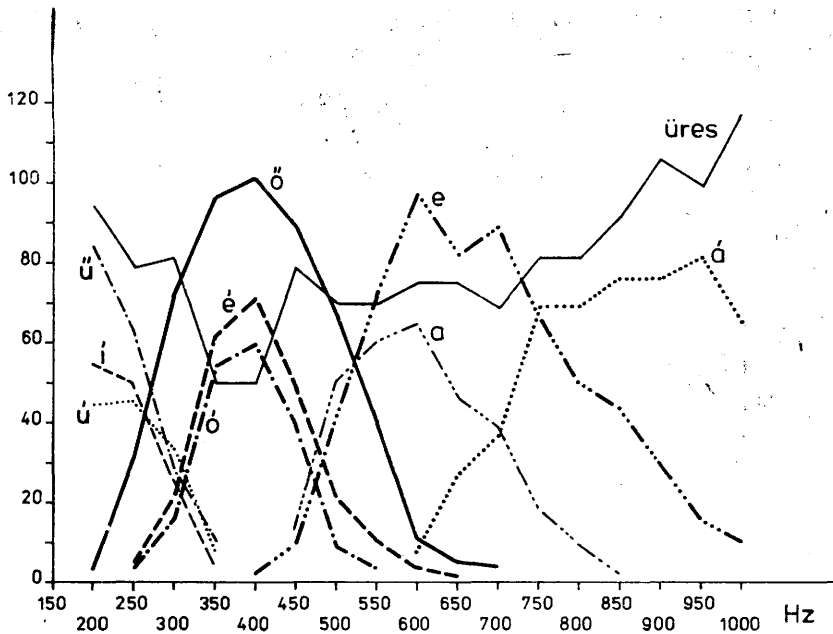
* Ha felállítjuk a magyar magánhangzók sorát a viszonylagos rezonancia szempontjából, akkor — Gombocz Zoltán szerint (ÖM II./1, 27) —, „a sorozatban az á és az ö viszonylagos helye lesz — vidékek és egyének szerint — a legingadozóbb.”

2. táblázat

A kísérleti személyek legalább 25, 25%-a által különbözönek jelzett hangminták

A hang- minta sorszáma	Az adott magánhangzóval történő azonosítások száma										Érték (Hz)	
	a	á	e	é	i	o	ö	u	ü	—	F1	F2
3						7		6			300	500
13	12					5					500	650
40	7	7									800	950
52	10	8									650	1100
53	10	8									700	1100
78							6		6		250	1400
95							5		14		250	1550
105		5	10								750	1550
106		7	6								800	1550
107		10	5								850	1550
113							9		9		300	1700
131							6		8		350	1850
150			5	7							450	2000
168			10	6							500	2150
185			8	5							500	2300
186			7	5							550	2300

tabb képével. De a figyelmes szemlélőnek feltűnhet az is, hogy az 5. ábrán az „e” hang F1 hisztogramjának két jól elkülöníthető maximumértéke van, mégpedig a 600 és a 700 Hz-es frekvenciánál. A következőkkel magyarázzuk ezt az anyagunkban egyedülálló jelenséget: minden magyar anyanyelvűben potenciálisan létezik mindkét „e” hang, a nyílt és a zárt „e” hang realizációs halmazának a képe. Ez a két kép akusztikailag jól körülhatárolható, leírható; ennek eredménye az említett hisztogram speciális, két csúcsot tartalmazó képe. Az üres magánhangzó azonosításainak értéke tulajdonképpen a kilenc magánhangzó hisztogram összegének a fordítottja. Ebből láthatjuk, hogy a magyar magánhangzók viszonylag egyenletesen használják fel az első formáns realizációjára a rendelkezésre álló frekvenciasávot, de leginkább a 350–700 Hz-es tartomány van kihasználva. Továbbá a második formáns a rendelkezésére álló frekvenciasávot nem olyan egyenletesen használja ki, mint az első formáns. Az alacsonyabb frekvenciák felől a magasabb frekvenciák felé haladva a kihasználtság foka egyre inkább csökken. Ezen belül a második formáns két sávban fordul elő a leggyakrabban, mégpedig az 500–800 és az 1250–1850 Hz-es tartományban. Mondhatjuk, hogy a magyar magánhangzók ezeket a kiemelt frekvenciasávokat töltik ki, használják fel többszörösen. Táblázatos formában a hisztogramok legfőbb tulajdonságait a 3. táblázatban mutatjuk be.

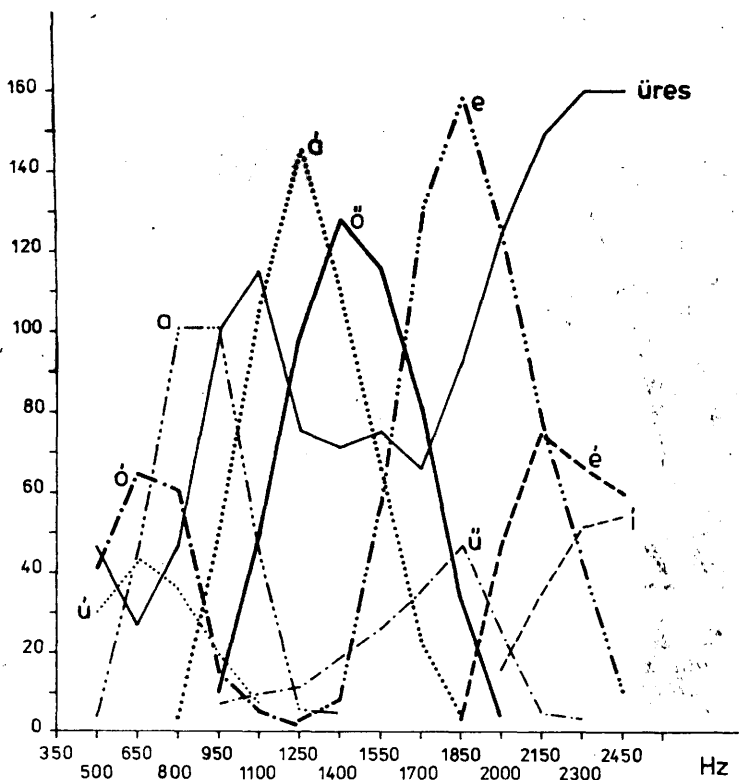


5. ábra. A 212 szintetizált hangmintára adott 4240 válasz közül az egyes magánhangzókkal azonosított válaszok száma az F1 formáns függvényében

3. táblázat

Az adott magánhangzókkal legnagyobb számban azonosított hangminta formánsfrekvenciaértékei

A magánhangzó	Az első formáns (F1) (Hz)			A második formáns (F2) (Hz)		
	a csúcstérték közelében	a csúcstérték 75%-ánál	a csúcstérték 50%-ánál	a csúcstérték közelében	a csúcstérték 75%-ánál	a csúcstérték 50%-ánál
1. a	600	500–600	500–700	800–950	800–950	800–950
2. á	950	750–1000	750–1000	1250	1250–1400	1100–1400
3. e	600–700	600–700	550–600	1850	1700–2000	1700–2000
4. é	400	350–400	350–400	2150	2150–2450	2000–2450
5. i	200	200–250	200–250	2450	2300–2450	2150–2450
6. o	400	350–400	350–400	650	650–800	500–800
7. ő	400	350–450	300–500	1400	1250–1550	1250–1700
8. u	250	200–250	200–300	650	500–800	500–800
9. ü	200	200–250	200–250	1850	1850	1550–1800



6. ábra. A 212 szintetizált hangmintára adott 4240 válasz közül az egyes magánhangzókkal azonosított válaszok száma az F2 formáns függvényében

A korábbiakban ígéretet tettünk arra, hogy megpróbálunk választ adni arra a kérdésre, hogy a férfi és a női kísérleti személyek eltérően értékelték-e a szintetizált hangminta anyagunkat. E kérdés megválaszolása végett elkészítettük a fentiekben már pontosan leírt hisztogram formákat, külön a férfiak, majd külön a nők válaszát összesítve. Elmondhatjuk, hogy különösen szembe-tűnően nem térnek el, nagy vonalakban tendenciájukban megegyeznek a férfiak és a nők válaszainak összesítésének eredményei. Azonban a csúcser-tékeket táblázatba gyűjtve (lásd 3. táblázat) figyelemreméltó tényt vehetünk észre: az „a, á, o, u, i” hang esetében a nők más formánsfrekvenciájú hang-mintákat azonosítottak legnagyobb számban az egyes magánhangzókkal. Sőt a fent említett magánhangzók közül a négy hátul képzett hang esetében a nők a magasabb frekvenciájú mintát azonosították a legtöbbször az adott magánhangzónak, az első és a második formáns tekintetében egyaránt. A jelen-ség még érdekesebbnek tűnik, ha párhuzamba állítjuk azzal az irodalomból jól ismert ténnyel, hogy a női ejtésű magánhangzók formánsainak frekvenciája általában magasabb, mint a férfi ejtésüké. Azt a jelenséget, hogy a férfiak és a nők, ha kis mértékben is, de eltérően észlelték ugyanazt a hangmintaanyagot, nem tekintjük véletlennek vagy kísérletünk megengedhető pontatlanságának,

4. táblázat

Az egyes magánhangzókkal legnagyobb számban azonosított hangminta formánsfrekvenciaértékei a férfi, a női kísérleti személyek esetében, valamint ezen értékek összesítése

A magánhangzó		Az első formáns (F1) (Hz)			A második formáns (F2) (Hz)		
		férfi	női	összesített	férfi	női	összesített
1.	a	550	600	600	800	950	800 – 950
2.	á	850 – 950	950	850 – 950	1250	1250	1250
3.	e	600 – 700	600 – 700	600 – 700	1850	1850	1850
4.	é	400	400	400	2150	2150	2140
5.	i	250	200	200	2450	2300	2450
6.	o	350 – 400	400	400	650	800	650
7.	ö	400	400	400	1400	1400	1400
8.	u	200	250	250	650	650	650
9.	ü	200	200	200	1850	1850	1850

hanem a jelenség egyik lehetséges magyarázata a következő: minden beszélő rendelkezik az anyanyelve magánhangzó-fonémáinak megfelelő realizációs halmaz képével. Ez a kép kissé eltérő a férfiak és a nők esetében, ugyanis ennek a realizációs halmaznak a középpontjában, a magjában az illetőnek a saját maga által ejtett hangképei (magánhangzóképei) állnak. Mivel a férfiak és a nők akusztikailag is kissé eltérő hangot ejtenek ugyanannak a fonémának a realizálása során, ezért törvényszerű, hogy a fonémák realizációs halmazának középpontja is máshol van férfiak és máshol van nők esetében. Mármost, ha az előzőeket elfogadjuk, akkor nem meglepő az említett jelenség, ugyanis a férfiak a hallott hangmintát a saját realizációs halmazukkal hasonlították össze, míg a nők szintén a saját realizációs halmazukkal, függetlenül attól, hogy a minta női vagy férfi ejtésű hang volt-e. Mivel a férfiak és a nők fonéma-realizációs halmaza különböző, a hangminták azonosítása is különböző lett. Ebből adódik tehát, hogy ha csak kissé eltérő, de más formánsfrekvenciájú hangmintát azonosítottak maximális számban az egyes magánhangzóknak a férfiak és a nők. Kísérletünk fő célja nem ennek a bizonyítása volt, mégis, az előbbiekből levonhatjuk azt a tanulságot, hogy igen gondosan kell megválasztani a hangtani percepciók kísérletekben résztvevők csoportját. Erre csak az eredmények tüzetesebb szemrevételezése során bukkantunk rá, de meggyőződésünk, hogy ez irányú észrevételeinket egy kifejezetten erre a célra összeállított kísérletsorozat is alátámasztja majd.

Úgy gondoljuk, hogy a későbbiekben is felhasználható lesz az általunk mesterségesen előállított 212 szintetizált hangminta. Ugyanis e hanganyag alkalmas arra, hogy más anyanyelvűek percepciók tesztjének alapjául szolgáljon, így ennek segítségével lehetőség nyílik különböző nyelvek magánhangzó-rendszerének összehasonlítására is. Ezenkívül egy nyelven belül, a nyelvjáráskutatásban is felhasználhatónak tartjuk.

KISS GÁBOR

Irodalom

- BOLLA K. (1978), A magyar magánhangzók akusztikai analízise és szintézise. Magyar Fonetikai Füzetek 1: 53—68
- FERRERO, F. E.—E. M. CALDOGNETTO (1976), Risultati sperimentali di una prova di percezione di suoni vocali sintetici. In: Gli studi di fonetica e fonologia. Róma
- GOMBOCZ Z. (1940), Magyar fonétika. In Gombocz Zoltán: *ÖM II/1*, 1—61. Budapest
- GORDOS G.—TAKÁCS Gy. (1983), Digitális beszédfeldolgozás. Budapest
- GÓSY M. (1984), A kategoriális percepció kérdése tipológiai vonatkozásban. Magyar Fonetikai Füzetek 13: 121—136
- KISS G.—OLASZY G. (1982/a), Interaktív beszédszintetizáló rendszer számítógéppel és OVE III beszédszintetizátorral. Magyar Fonetikai Füzetek 10: 21—45
- KISS G.—OLASZY G. (1982/b), A magyar beszéd előállítása számítógéppel. *Információ Elektronika* 1982/4: 344—354
- KISS G.—OLASZY G. (1984), A HUNGAROVOX magyarnyelvű, szótár nélküli valósidejű, párbeszédés beszédszintetizáló rendszer. *Információ Elektronika* 1984/2: 98—111
- LAZICZIUS Gy. (1963), Fonetika. Budapest
- MAGDICS K. (1965), A magyar beszédhangok akusztikai szerkezete. Budapest
- MOLNÁR J. (1970), A magyar beszédhangok atlasza. Budapest
- OLASZY G. (1978), Szintetizált magánhangzók formáns-intenzitás és formáns-sáv szélesség értékei. Magyar Fonetikai Füzetek 1: 68—76
- SZENDE T. (1976), A beszéd folyamat alaptényezői. Budapest
- TARNÓCZY T. (1974), A magánhangzók akusztikai vizsgálatának problémái. *ÁNyT* 10: 181—196
- VÉRTES O. A. (1982), A magyar beszédhangok akusztikai elemzésének kérdései. In: Bolla Kálmán (szerk.) *Fejezetek a magyar leíró hangtanból*. Budapest. 71—114

Determining the First Two Formants of Hungarian Vowels by a Perception Test Involving Synthesized Sound Samples

by GÁBOR KISS

The author investigates the variance of the first two formants of Hungarian vowels by a perception test. To that end, varying first and second formants, he produced 212 synthesized sound samples that covered the full range of Hungarian vowels as well (see Figure 1) in the Phonetics Laboratory of the Linguistics Institute of the Hungarian Academy of Sciences. He asked twenty subjects to listen to those samples and see if they could identify them with any Hungarian vowel, and if so, which one. The F1, F2 chart of Hungarian vowels based on the answers of the twenty subjects is shown in Figure 4. Frequency values of first and second formants of Hungarian vowels as determined by the auditive perception test are listed in Table 3. In analysing the results, the author comes to the conclusion that, if only to a small extent, men's and women's responses differ from each other. He explains that fact by assuming that both men and women try to identify the sound sample they hear in terms of the range of their own realisations, irrespective of whether what they hear is a male or female voice. This is shown in Table 4. The author thinks that third formants are significant in the case of close vowels, especially [i] and [u], in determining vowel quality.