

## **Szintetizált anyaggal végzett magyar magánhangzófelismerési kísérletek**

A hang- és általában a beszédszintézis elsődleges célja az analízis eredményeinek igazolása, bizonyítása. A spektrografikus analízis megismerteti a beszédhangok akusztikai szerkezetét alapvetően meghatározó tényezőkkel (a magánhangzók formánsfrekvencia-, formánsamplitúdó-, formánssáv szélesség- és alapfrekvencia-értékeivel, valamint a mássalhangzókat meghatározó magánhangzó-formánsátmenetekkel és locus-értékekkel). E tényezők értéke akkor tekinthető azonban végérvényesnek, ha a szintézis síkján is igaznak bizonyul. Ilyenképpen tehát a szintézis az analízis útján kapott eredmények próbájának is tekinthető. A beszédszintézis azonban ennél jóval távolabbra mutat. A hangok különböző akusztikai tényezőinek egymás utáni, illetőleg egyidejű módosítása bő lehetőségeket nyújt annak kikísérletezésére, hogy melyek a hangokat alapvetően meghatározó akusztikai tényezők, valamint annak megállapítására, hogy hol van az egyes hangok érthetőségének a határa. Röviden: módot ad a beszédhangok releváns, illetőleg redundáns sajátosságainak elhatárolására, valamint széleskörű hang-, illetőleg beszédfelismerési kísérletek folytatására.

A szintetizált beszédhangokon végzett érthetőségi kísérletek nem újkeletűek. A Haskins-csoport 1952-ben már szép eredményekről számol be (vö. H. S. COOPER—P. C. DELATRE—A. M. LIBERMAN—J. M. BORST—L. J. GERSTMAN, *Some Experiments on the Perception of Synthetic Speech Sounds*. J.A.S.A. 24 : 1952. 597—606) szintetizált amerikai-angol zárhangokkal és magánhangzókkal végzett érthetőségi kísérletei alapján. Azóta is időszerű probléma ez és a kísérleti fonetikusok nagy számát foglalkoztatja. A szintézis- és érthetőségekísérletek többsége a magánhangzókra vonatkozik. Egyes szerzők a formánsoknak a frekvenciaskálán elfoglalt helyét tartják az érthetőség kulcsának (vö. F. S. COOPER—P. C. DELATRE—A. M. LIBERMAN—J. M., BORST—L. J. GERSTMAN: i. m. 503; W. E. KOCK, *The Problem of Selective Voice Control*. J.A.S.A. 24 : 1952. 633; A. CARPENTER, *Discrimination of Vowel Sounds Synthesized by Harmonic Addition*. *Language and Speech* 5 : 1962. 151). P. LADEFOGED módosítja a formánsfrekvenciának a magánhangzófelismerésben betöltött vezetőszeropét: nem az abszolút formánsfrekvencia-értékeket tartja döntőnek a felismerésben, hanem a relatív értékeket, amelyek az egymás környezetében álló magánhangzók formánshelyzeteinek kölcsönös függőségi viszonyából születnek egyazon személy beszédében (vö. *The Perception of Vowel Sounds*. Ph. D. Thesis. University of Edinburgh, 1959. 1.). Ez a nézet kristályosodik ki később D. E. BROADBENTtel közösen írt tanulmányában (D. E. BROADBENT—P. LADEFOGED, *Vowel Judgements and*

Adaptation Level. Proc. Royal Soc. 1960. 151, 384). Más szerzőknél is találkozzunk hasonló véleménnyel: D. B. FRY, A. S. ABRAMSON, P. D. EIMAS és A. M. LIBERMAN (The Identification and Discrimination of Synthetic Vowels. Language and Speech 5 : 1962. 178) tapasztalata szerint a magánhangzók kevésbé élesen elhatárolhatók, meghatározhatók és abszolútak, mint a mássalhangzók; a magánhangzófelismerés erősen az akusztikai kontextus függvénye B. BOROVICKOVÁ (Physikalische und perzeptionelle Identifikation der tschechischen Vokale. ZfPhon. 17 : 1964. 142) döntőnek tartja az első két formánsnak a frekvenciaskálán elfoglalt helyét a magánhangzófelismerésben, de szintén nem az abszolút frekvenciaértékeket, hanem azok egymáshoz való viszonyát („ihr gegenseitiges Verhältnis”). CARPENTER, aki fent említett cikkében a magánhangzófelismerés alapját a formánsok helyzetében látja, további kísérletei során azt tapasztalta, hogy a magánhangzók felismerésében mégsem a formánsok helyzete dönt (A. CARPENTER—J. MORTON, The Perception of Vowel Colour in Formantless Complex Sounds. Language and Speech 5 : 1962. 213—14; J. MORTON—A. CARPENTER, Judgement of the Vowel Colour of Natural and Artificial Sounds. Language and Speech 5 : 1962. 200). Sávszűrő-szerkezet segítségével előállított formánsstruktúra nélküli komplex hangokkal végzett sikeres lehallgatási kísérleteket és arra a megállapításra jutott, hogy a hangfelismerés mechanizmusa nem fut azonos síneken a hangképzés és az ezen alapuló hanganalízis mechanizmusával; az utóbbi a magánhangzókat kizárólag a formánsfrekvencia alapján határozza el egymástól, míg az előbbi enélkül is képes őket elhatárolni, megkülönböztetni. R. L. MILLER (Auditory Tests with Synthetic Vowels. J.A.S.A. 25 : 1953. 114) a formánsfrekvencia mellett a formánsamplitúdót is relevánsnak tartja a magánhangzófelismerésben. Ezt a véleményt erősíti meg P. JANOTA (Zur Wahrnehmung synthetischer tschechischer Vokale. ZfPhon 17 : 1964. 235) is, aki 200 szintetizált magánhangzóval folytatott kísérlete alapján arra a megállapításra jutott, hogy a formánsamplitúdó elhagyása igen döntően befolyásolja a hangok felismerését. A. COHEN—J. F. SCHOOTEN—J. 't HART (Contribution of the Time Parameter to the Perception of Speech. Proc. of the Fourth Intern. Congr. of Phon. Sciences. The Hague, 1962. 557) azt tapasztalták, hogy a szintetizált hangok időtartamának változtatása más-más eredményhez vezet a felismerési kísérletek során. COHEN és 't HART további kísérletei (Speech Synthesis of Steady-State Segments. Proc. of Speech Communication Seminar. Stockholm, 1963. Vol. II. Session F. p. 7) azt igazolták, hogy az inherens módon rövid magánhangzók (amelyeknek nincs hosszú párjuk) hosszúra szintetizált variánsai nem felismerhetők. Ugyanakkor azt is tapasztalták, hogy a hanglejtés (amelyet az egyes hangokra vonatkoztatva „micro-intonation”-nek neveznek, szemben a prozódia síkján számon tartott, nagyobb nyelvi egységet jelölő „macro-intonation”-nel) hozzáadása lényegesen javítja a lehallgatási kísérletek eredményét. W. TSCHESCHNER (Ergebnisse bei der Analyse von deutschen Sprachlauten. ZfPhon. 17 : 1964. 6) szerint a hangok alapfrekvenciája szintén számottevően ható tényező a felismerésben: kísérletei alapján a 200 Hz fölötti alapfrekvenciával előállított magánhangzók felismerhetetlenek voltak. KOCK (i. m. 633) ezzel szemben úgy látja, hogy az alapfrekvencia, mint paraméter, igen kevés szerepet játszik a hangfelismerésben.

Ez a dolgozat 11 magyar magánhangzó-fonéma (*i, e, ε, a, u, y, ø, y, : , ø* : ) izolált és kontextusba (egyszavas mondatokba) ágyazott szintézi-

sének, valamint a szintetizált anyaggal végzett felismerési kísérleteknek az eredményeit mutatja be. A szintézis a londoni University College „Spyl”-nek nevezett beszédszintetizáló berendezésével készült (leírását l.: G. F. ARNOLD — P. DENES — A. C. GIMSON — J. D. O'CONNOR — J. L. M. TRIM, *The Synthesis of English Vowels. Language and Speech* 1 : 1958. 114).

A munka első fázisában a *magánhangzókat izoláltan*, egyéb hangoktól elszigetelve szintetizáltam. A szintézis a saját ejtésű hangok spektrogramjainak mintájára készült. Az előállított hangok három formánssal rendelkeztek (a berendezés a következő formánsfrekvencia-tartományokon belül működött:  $F_1$ : 170—900 Hz,  $F_2$ : 470—2500 Hz,  $F_3$ : 1500—2700 Hz), alapfrekvenciájuk 130 Hz volt (a berendezés alapfrekvencia-tartománya 80—180 Hz volt), időtartamuk annak a hanghosszúságnak felelt meg, ahogy az ember egy izolált hangot természetes módon kiejt, a rövid hangok időtartama 0,07—0,09 sec., a hosszú hangoké 0,18—0,20 sec. volt (a berendezés időtartam-tartománya szegmentumonként 0,13—3,70 sec. volt). Az emberi és az annak mintájára előállított hangok formánsfrekvencia- és relatív formánsamplitúdó-értékeit az 1. táblázat szemlélteti. (A dB értékek megállapításánál a természetes hang első formánsamplitúdóját [ $A_1$ ] vettem viszonyítási alapul.) Ha a táblázaton a természetesen ejtett hangokra vonatkozó mérésadatokat összevetjük a nagy anyag alapján kapott középértékekkel (vö. MAGDICS KLÁRA, *A magyar beszédhangok akusztikai szerkezete. NytudÉrt.* 49. szám.), akkor ezeket az adatokat valószerűeknek ítélni lehetjük (vö. 2. táblázat).

A szintetizált hangok formánsfrekvencia- és formánsamplitúdó-értékei igazolják a természetes hangok méréseredményeit — és ugyanakkor a nagy anyag alapján kapott középértékeket is (vö. 1. és 2. táblázat). A szintetizált magánhangzók harmadik formánsának frekvenciaértéke néhány esetben nem éri el a természetes hang esetében kapott értéket, és ennek oka a berendezés korlátolt,  $F_3$ -frekvencia-tartományában rejlik. A harmadik formáns az előképzett magánhangzók hangszínét befolyásolja (vö. R. K. POTTER — G. A. KOPP — C. W. GREEN, *Visible Speech*. New York, 1947), s az  $F_3$ -nak a „normánál” alacsonyabb értéke a szintetizált *i*, és kisebb mértékben az *e* : esetében észrevehető volt. Az *i* szintézisének tökéletlenségét az  $F_2$ -nek — a fenti okból származó — a „normánál” alacsonyabb helyzete is erősíti.

A szintetizált izolált magánhangzókkal a továbbiakban a következő kísérleteket folytattam: 1. Mindegyik szintetizált magánhangzót lehallgattattam 20 — idegen nyelvet vagy nyelveket is ismerő — alannyal (közülük 10 fonetikailag képzett volt, 10 nem részesült alaposabb fonetikai képzésben, de egyetemet, illetőleg főiskolát végzett; mindegyikük legalább tíz éve Budapesten él és a köznyelvet beszéli), akiknek ezúttal arra a kérdésre kellett válaszolniuk, hogy melyik magyar magánhangzót hallották. 2. Ugyanezeket a szintetizált hangokat ismét meghallgatták és most azt kellett eldönteniük, hogy az illető hangok a magyar köznyelvhez szokott fül számára természetesnek vagy természetellenesnek hatnak-e, s ha természetellenesnek hatnak, miben áll ez a természetellenesség. 3. Mindegyik szintetizált magánhangzó első, második és harmadik formánsán (egymás után) frekvenciatorzítást hajtottam végre az  $F_1$ -et 50, 100, 200, 300 Hz-cel, az  $F_2$ -t 200, 400, 600, 700 (—800) Hz-cel, az  $F_3$ -at 300, 500, 700, 900 (—1000) Hz-cel süllyesztettem, illetőleg emeltem. Az egyes formánsoknak a hang megértése tekintetében releváns „módosítás-sávja” eltérő volt:  $F_1$  esetében 50—300,  $F_1$  esetében 200—700 (—800),  $F_3$  esetében 300—900 (—1000) Hz. Minden egyes torzított variánsról szonogramot

1. táblázat

A természetes (t) és a szintetizált (sz) hangok formánsfrekvencia ( $F_n$ ) és relatív formánsamplitúdó-értékei ( $A_n$ )

	$F_n$ (Hz)			$A_n$ (dB)		
	$F_1$	$F_2$	$F_3$	$A_1$	$A_2$	$A_3$
<i>i</i> t.	300	2600	3200	-3	-10	-13
sz.	300	2500	2700	-3	-10	-12
<i>e</i> : t.	400	2550	3100	-1,5	-6	-11
sz.	400	2500	2700	-1	-5	-
$\varepsilon$ t.	750	2000	3000	0	-6	-10,5
sz.	750	2000	2700	-0,5	-6	-10
<i>a</i> : t.	1000	1600	2400	+2	-4	-7
sz.	900	1600	2500	+1,5	-3	-8
<i>a</i> t.	700	1400	2500	0	-2	-9
sz.	650	1350	2500	0	-2	-8
<i>o</i> t.	500	900	2750	0	-4	-9
sz.	500	900	2700	0	-3	-8
<i>u</i> t.	300	800	-	-2	-4,5	-
sz.	300	800	-	-1,5	-4	-
$\emptyset$ t.	350	1650	2600	-0,5	-3	-8
sz.	400	1650	2600	0	-2	-7
$\emptyset$ : t.	400	1850	2500	-1	-5,5	-7
sz.	450	1800	2500	0	-4	-7
<i>y</i> t.	300	1800	2500	-1	-5	-9
sz.	300	1700	2500	0	-5	-8
<i>y</i> : t.	300	1900	2800	-1	-6	-11
sz.	300	2000	2700	-1	-5,5	-10

és magnetofonfelvételt készítettem. Azután mindegyik torzított variánst meghallgattattam az alanyokkal (az egyes hangok torzított variánsai előtt mindig a torzítatlan, „stabil” hang szólalt meg, amelyhez a torzított variánsokat viszonyítani lehetett), akiknek egyrészt arra kellett válaszolniuk, hogy *a*) a hallott hang érthető vagy érthetetlen; *b*) ha érthető, megegyezik-e a viszonyítási alapul szolgáló, elsőként hallott hanggal, vagy más hang; *c*) ha megegyezik vele, természetesnek hat-e a magyar köznyelvhez szokott fül számára, vagy nem; *d*) ha más hang, magyar hang-e, vagy idegen. 4. Az *e* ; *o*,  $\emptyset$ ,  $\emptyset$  ; *y* : hangokon a három formáns amplitúdótorzítását is végrehajtottam (értéküket 2, 4, 6, 8 (-10) dB-lel csökkentettem, majd növeltem). A lehallgatási kísérlet menete ugyanaz volt, mint a frekvenciatorzítás esetében. 5. Az *e* ; *o*,  $\emptyset$  hangok alapfrekvenciáját 10, 20, 30, 40 (-50) Hz-cel módosítottam (süllyesztettem, majd növeltem). A lehallgatási kísérlet menete olyan volt, mint a formánsfrekvencia- és formánsamplitúdó-torzítás esetében.

Az első kísérlet (1) osztatlanul pozitív eredménnyel zárult: mind a húsz alany felisterte az izoláltan hangoztatott szintetikus magánhangzókat.

A második kísérletnél (2) már néha megoszlottak a vélemények. A kapott válaszokat a 3. táblázat szemlélteti.

2. táblázat

Formánsfrekvencia- és formánsamplitúdó-középtértékek (k) és a szintézishez modellként használt magánhangzók egyedi értékei (e)

	F <sub>n</sub> (Hz)			A <sub>n</sub> (dB)		
	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>
<i>i</i> k.	410	2470	3140	-2	-9	-12
e.	300	2600	3200	-3	-10	-13
<i>e:</i> k.	484	2452	3020	-1,5	-7	-11
e.	400	2550	3100	-1,5	-6	-11
<i>ε</i> k.	695	2088	3000	-0,5	-6	-10
e.	750	2000	3000	0	-6	-10,5
<i>a:</i> k.	1122	1610	2600	+2,5	-3	-8
e.	1000	1600	2400	+2	-4	-7
<i>a</i> k.	792	1300	2660	0	-3	-10
e.	700	1400	2500	0	-2	-9
<i>o</i> k.	642	1205	2701	0	-4	-12
e.	500	900	2750	0	-4	-9
<i>u</i> k.	445	1028	2340	-1,5	-5	-
e.	300	800	-	-2	-4,5	-
<i>ø</i> k.	486	1573	2605	-0,5	-3,5	-9
e.	354	1650	2600	-0,5	-3	-8
<i>ø:</i> k.	412	1624	2685	-1	-5	-8
e.	400	1850	2500	-1	-5,5	-7
<i>y</i> k.	328	2020	2730	-1	-6	-8,5
e.	300	1800	2500	-1	-5	-9
<i>y:</i> k.	292	2165	2753	-1,5	-7	-10
e.	300	1900	2800	-1	-6	-11

3. táblázat

Az izolált szintetizált magánhangzókval végzett 2. lehallgatási kísérlet eredménye.  
(A számok itt és a további táblázatokban, a szavazatok számát jelölik.)

	Természetes	Természetellenes	Miért, természetellenes?
<i>i</i>	10	10	Alacsonyabb nyelvállású, ernyedt képzésű, nem eléggé illabiális.
<i>e:</i>	15	5	Leszorított, préselt hang.
<i>ε</i>	—	—	—
<i>a:</i>	—	—	—
<i>a</i>	20	—	—
<i>o</i>	20	—	—
<i>u</i>	20	—	—
<i>y</i>	20	—	—
<i>y:</i>	20	—	—
<i>ø</i>	20	—	—
<i>ø:</i>	20	—	—

Szint. hg	Mód. ir.	A formánsfrekvencia - módosítás mértéke (Hz-ben)									
		50					100				
		É				Éh	É				Éh
		V		-V			V		-V		
		T	Te	M	I		T	Te	M	I	
<i>i</i>	↑			16 : e		4			19 : e	1	
<i>e:</i>	↑		6		12 : ε 1 : ə	1		5	15 : ε		
	↓		4	12 : i		4		1	15 : i	4	
<i>ε</i>	↑		6	2 : a:		12		7	9 : a:	4	
	↓	5	12			3		3		17	
<i>a:</i>	↓	18	2					12		8	
<i>a</i>	↑		11	2 : a:	1 : A	6		4	8 : a:	2 : A	6
	↓	18	2					10	10 : o		
<i>o</i>	↑	19	1					9	5	4 : a	2
	↓	19	1					18	2		
<i>u</i>	↑	8	12					4	16		
<i>y</i>	↑		3			17			2	2 : ε	16
<i>y:</i>	↑		5			15			2 : ε	4 : œ	14
<i>ø</i>	↑	10				10	10	2			8
	↓		6	6 : y		8		4	6 : y		10
<i>ø:</i>	↑			3 : o:	3 : ə:	14			1 : o:	2 : ə:	17
	↓			10 : y:		10			8 : y:		12

Izolált szintetizált magánhangzókkal  $F_1$ -torzítás mellett végzett érthetőségi vizsgálatok eredménye (Szint. hg = Szintetizált hang; É = Érthető, Éh = Érthetetlen; V = Várt hang, -V = Nem várt hang; T = Természetes, Te = Természetellenes; M = Magyar, I = Idegen; Mód. ir. = A módosítás iránya)

Azt látjuk, hogy magas nyelvállású előképzett magánhangzók (*i*, *e*;) esetében, ahol az  $F_2$  és az  $F_3$  (technikai okból) a kívántnál alacsonyabb helyet foglalnak el a frekvenciaskálán, az alanyok 50, illetve 25%-ban a „természetellenes” mellett szavaztak. Az *ε*, *y* : és — kisebb mértékben — az *o* szintézisé-nél is fennáll bizonyos  $F_3$  elégtelenség (vö. 1. táblázat), de itt — az egybehangzó vélemény alapján — ez nem befolyásolja számottevően a hangkvalitást.

A harmadik kísérlet (3) eredményeit a 4–6. táblázatok szemléltetik. Az első formáns ( $F_1$ ) helyének változtatása az *i*, *e*:, valamint az *y*, *y* : , *ø* : , tehát elsősorban az előképzett, magas nyelvállású magánhangzók esetében mutatkozott relevánsnak a hangok felismerése szempontjából. Az  $F_1$  50 Hz-es megemelése az *i*-t már felismerhetetlenné tette, és az alanyok 80%-ában (16) az *e* érzetét keltette; 100 Hz-es megemelése után már az alanyok 95%-a (19) *e*-re szavazott. Az *e* :  $F_1$ -ének 50 Hz-es megemelése az alanyok 60%-ánál (12), 100 Hz-es megemelése az alanyok 75%-ánál (15) *ε*-érzetet keltett; 100 Hz-es  $F_1$ -süllyedésnél az alanyok 75%-a (15) *i* : -re szavazott. Míg tehát az *i*, *e* :  $F_1$ -elmozdítása igen gyorsan más magánhangzó érzetét keltette a hallgatókban,

## láza t

Szi- nt. hg	Mód. ir.	A formánsfrekvencia - módosítás mértéke (Hz - ben)									
		200					300				
		É				Éh	É				Éh
		V		-V			V		-V		
		T	Te	M	I		T	Te	M	I	
i	↑										
e:	↑										
	↓										
ε	↑	6	9 : a:		5		5	9 : a:		6	
	↓			4 : ə	16			14 : ø	2 : ə	3	
								1 : e			
a:	↓		3 : ε	4 : ε	13			9 : ε	4 : ε	7	
α	↑		8 : a:	3 : ɹ	12			10 : a:	3 : ɹ	7	
	↓	1	2 : o								
			19 : o								
o	↑		16 : a		4			18 : a		2	
	↓	7	3 : u		5			5 : u		15	
u	↑		2 : o		17			3 : o		17	
y	↑		5 : ε		15			6 : ε		14	
y:	↑		8 : ε	2 : œ	10						
ø	↑		1 : ε	2 : ə	17			1 : ε	1 : ə	18	
	↓	2	7 : y		11			7 : y		13	
ø:	↑										
	↓										

addig az *y, y:*, *ø:* esetében ugyanez a torzítás inkább érthetetlen, felismerhetetlen, mint új hangot eredményezett. Az *y, y:*, *ø:* esetében az  $F_1$  50–300 Hz-es megemlése 50–85%-ban az „érthetetlen” felé billentette a szavazatok mérlegét. Az *a:*, *a*, *o*, *u* — tehát általában a hátulképzett, alacsony nyelvállású magánhangzók — felismerhetőség szempontjából már sokkal kevésbé mutatkoztak érzékenyek az  $F_1$ -torzításra. 50–100 Hz-es módosítás mellett még túlnyomó százalékban felismerték a hangot a hallgatók. 200–300 Hz-es torzítás kellett ahhoz, hogy a hangok túlnyomó részben érthetetlennek, kisebb százalékban más hangoknak bizonyuljanak. Az *ε* és *ø* mintegy átmenetet képez a két említett kategória között: az *i—e:*, illetőleg *y—y:* —*ø* csoportba már nem illik bele, mivel már nem olyan érzékeny az  $F_1$ -módosulásra, az *a:* —*a*—*o—u* csoportba pedig még nem tartozhat, mert annyira viszont még nem érzékeny az  $F_1$ -torzítás irányában. Az  $F_1$  megemlése és süllyesztése nem egészen azonos módon és mértékben hatott a magánhangzók felismerhetőségére. A formáns megemlése általában jobban és hamarabb zavarta a megértést, mint a formáns süllyesztése (vö. 4. táblázat). A második formáns ( $F_2$ ) helyének módosítása általában egyetlen magánhangzó megértése, felismerhetősége szempontjából sem volt közömbös. A módosítás fülhallható eredménye kétféle volt: *a)* más magyar hang, *b)* meghatározhatatlan, azonosíthatatlan hangjelenség. Az *a)*-eredményhez vezetett az *i*, *e:*, *ε*, *a:*, *y*, *y:*, *ø:*, a *b)*-ered-

Szint. hg	Mod. ir.	A formánsfrekvencia-módosítás mértéke (Uz-ben)									
		200					400				
		É			Éh		Ë			Éh	
		V		-V			V		-V		
		T	Te	M	I		T	Te	M	I	
i	↓		7		4 : i	9			1 : e	1 : e	18
e:	↓	1	7		7 : ə	5		2	9 : ø:		9
ε	↑		11		3 : e	6		9	2 : e	3 : e	6
	↓		11	3 : a:		6		2	1 : a	1 : A	13
a:	↑		9		4 : ε	7		4	5 : e	3 : e	8
	↓	16	4					15	5 : e	5 : e	
α	↑		10	6 : a:	4 : A				8 : ε		12
	↓		6	5 : o		9		3	2 : a:		
o	↑	5	13		2 : ə		3	15	10 : o		7
	↓	20					6	7	1 : ə		1
u	↑		10	1 : y		9		9	5 : u		
y	↑	17	3				10	10	3 : y		8
	↑	16	4					3			17
y:	↑		3	4 : i:		13		2	8 : i:		10
	↓	10	10				4	9			11
ø	↑			2 : e	3 : ə	15			4 : e	3 : ə	13
	↓	19		1 : o				8	3 : o	5 : e	3
									1 : e		
ø:	↑			9 : y:		11			15 : y:		5
	↓			1 : o:	1 : ə:	18			10 : o:		6
									4 : u:		

Izolált szintetizált magánhangzókkal  $F_2$ -torzítás mellett végzett érthetőségi vizsgálatok eredménye

ményhez az *o*, *u*  $F_2$ -módosítása; az *a*, *ø* esetében a szavazatok körülbelül fele-fele arányban oszlottak meg az *a*) és *b*) kategóriák között. Azt mondhatjuk tehát, hogy az előlképzett magánhangzók általában más magyar hangkvalitásba mentek át az  $F_2$ -módosítás hatására, míg a hátulképzett magánhangzók érthetelenné, felismerhetelenné váltak. A más magyar hangkvalitásba való átmenet tekintetében sem volt egységes a kép: az *i*, *e* : 200 Hz-es  $F_2$ -módosításnál részben természetellenesnek, részben érthetetlennek hatott. 400 Hz-es módosításnál a szavazatok döntő többsége az „érthetetlen”-re esett, de ugyanakkor, ha halványan is, már felbukkant a más magyar hangkvalitás érzete. 600 Hz-es módosítás már a „más magyar hang” rovatba sűrítette a szavazatokat. 700 (–800) Hz-es módosításnál az utóbbi kikristályosodása mutatkozott. Az *ε*, *a*: 200 Hz-es  $F_2$ -módosítása az alanyok többsége szerint természetelle-



*l á z a t*

Szint. hg	Mód. ir.	A formánsfrekvencia-módosítás mértéke (Uz-ben)									
		600					700(-800)				
		É				Éh	É				Éh
		V		-V			V		-V		
		T	Te	M	I		T	Te	M	I	
i	↓			15:y	1:ø	4			17:y	1:ø	2
e:	↓			15:ø:		5			16:ø:		4
ε	↑		7	4:e	3:e	6		4	5:e	3:e	5
				6:α		10			3:i		
	↓			4:a:					18:α		2
a:	↑			10:ε	2:ε	8			12:ε		
	↓			8:ε		12			13:ε		7
a	↑			10:ε		10			10:ε		10
	↓		2	10:o		8		1	12:o		7
o	↑		3	6:ø	2:e	9			3:ø	4:ø	13
	↓	1	5	5:u		9	1	4	6:u		9
u	↑		2	3:y		15					
y	↑			6:i	1:i	12			13:i		7
	↓			1:t							
				10:u		10			16:u		4
y:	↑			12:i:		8			12:i:		8
	↓		7	4:ø	3:ø:	6		5	11:ø:	1:ø:	3
ø	↑			10:ε	1:ø	9			10:ε	1:ø	9
	↓			4:o	3:ø	10			8:o	5:ø	7
				2:a:							
				1:α							
ø:	↑			15:y:		5					
	↓			16:o:		4					

nessé tette a hangokat. 400 Hz-es módosításnál a szavazatok a „természetellenes”, „érthetetlen” és a „más magyar hang” rovatok között oszlottak meg. 600 Hz-es módosításnál megerősödött a megváltozott hangkvalitás érzete, de többségben volt az „érthetetlen”-re adott szavazatok száma. A 700 (-800) Hz-es módosítás kikristályosította az új hangkvalitás érzetét. Az *y* 200 Hz-es módosítása lényegében nem változtatott a hangkvalitáson. 400 Hz-es F<sub>2</sub>-emelkedésnél még mindig nem mutatkozott számottevő különbség, ugyanilyen F<sub>2</sub>-süllyedés azonban a hang felismerhetetlenségéhez, érthetlenségéhez vezetett. 600 Hz-es módosításnál körülbelül egyforma arányban kaptam szavazatokat az új hangkvalitás, illetőleg az érthetetlennek ítélt hang mellett. A 700 (-800) Hz-es módosítás erőteljesen megnövelte az új hangkvalitásra adott szavazatok számát. Az *y*-, *ø*-, *ø*: esetében már 200 Hz-es F<sub>2</sub>-módosításnál jelentkezett az új hangkvalitás érzete, a szavazatok zöme azonban az „érthetetlen”-re esett. A 400 Hz-es módosítás csökkentette az „érthetetlen”-re adott szavazatok számát és növelte az új hangkvalitás érzetét. Ez utóbbi

tovább erősödött 600 Hz-es, és szavazati többséget ért el 700 (–800) Hz-es módosítás mellett (vö. 5. táblázat). A formáns emelkedése ezúttal is kissé erőteljesebben befolyásolta a hangfelismerést, mint a formáns süllyedése. A *harmadik formáns* ( $F_3$ ) helyének változtatása elsősorban az előképzett magánhangzók ( $i, e ; \varepsilon, y, y ; \phi, \phi :$ ) esetében volt észrevehető. Az  $i, e ; y$  : a szavazatok túlnyomó többsége szerint 300 Hz-es  $F_3$ -módosításra érthetatlenné vált, s a további — igen kisszámú — szavazat más magyar hangkvalitásra esett; ez az arány fokozottabb  $F_3$ -módosítás mellett sem változott számottevően. Az  $\varepsilon$  300–500 Hz-es módosítás mellett igen kevésbé változott, elsősorban természetességéből vészett a hang és az alanyok 10%-a ítélte csak érthetatlenné. 700 Hz-es módosításnál még mindig a „természetellenes” rovatban gyűlt össze a legtöbb szavazat, az alanyok további százaléka érthetatlenné, illetőleg más magyar hangkvalitásnak ítélte a hallott hangot. 900 (–1000) Hz-es módosításnál az alanyok fele más magyar hangkvalitásra szavazott, fele pedig érthetatlenné tartotta a hallott hangot. Az  $y$  alig változott 300–500 Hz-es módosítás mellett, legfeljebb természetellenesnek ítélték. 700–900 (–1000) Hz-es módosítás mellett a „természetes” és „természetellenes” rovaton kívül az „érthetetlen” rovatba is jutott szavazat. Az  $\phi, \phi :$  300 Hz-es módosításnál alig változott, a további módosítás megosztotta a véleményeket: az alanyok egy része természetellenességet tapasztalt a hangban, mások új hangkvalitást éreztek, ismét mások felismerhetatlenné, meghatározhatatlannak ítélték. Az  $\phi :$  esetében az  $F_3$ -süllyedés nem változtatta meg az  $\phi_:$ -érzetet, legfeljebb természetellenességet éreztek benne a hallgatók. A formánsmódosítás iránya (a formáns emelése, illetőleg süllyesztése) ezúttal irrelevánsnak mutatkozott: sem az emelkedés, sem a süllyedés nem bizonyult a másiknál erősebb hatótényezőnek (vö. 6. táblázat).

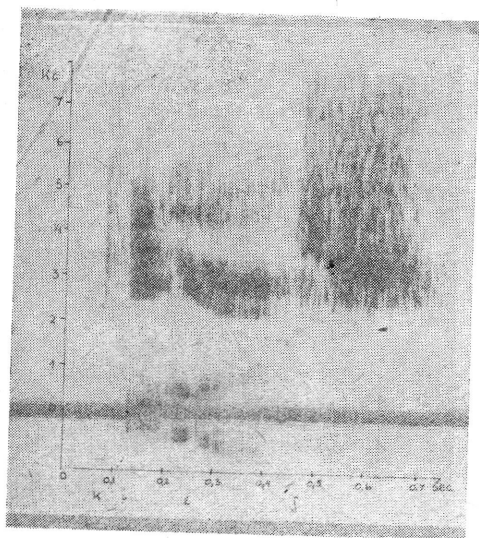
A 4. kísérlet eredményét a 7. táblázat szemlélteti. Az *első formánsamplitúdó* ( $A_1$ ) értékének 2–4 dB-lel történő módosítása (növelése, illetőleg csökkentése) általában nem változtatott nagymértékben a hangok kvalitásán, legfeljebb természetellenes, a szokásostól kissé eltérő szint adott. 6 dB-es módosítás az  $e :$ -t és az  $\phi^1$ -t általában felismerhetatlenné, érthetatlenné tette. 8 (–10) dB-es  $A_1$ -emelés mellett mindegyik hang felismerhetatlenné vált, ugyanilyen  $A_1$ -süllyedés mellett azonban az  $o$  és  $\phi :$  érthető maradt a szavazatok többsége szerint. A *második formánsamplitúdó* ( $A_2$ ) értékének 2 dB-lel való módosítása az alanyok többsége szerint csak az  $e : \uparrow$  és  $\phi : \uparrow$  felismerését akadályozta. A 4 dB-es  $A_2$ -emelés az  $e ; \phi :$ -n kívül az  $y ; \phi$ -t is felismerhetatlenné tette; az ugyanilyen  $A_2$ -süllyedés mellett mindegyik magánhangzó felismerhető maradt. 6 dB-es  $A_2$ -emelkedésnél felismerhetatlenné bizonyultak a hangok, ugyanilyen  $A_2$ -süllyedés mellett az  $e ; o, \phi :$  érthető maradt. 8 (–10) dB-es  $A_2$ -emelkedésnél felismerhetatlenné voltak a hangok, ugyanilyen  $A_2$ -süllyedés mellett csak az  $\phi :$  bizonyult érthetőnek. A *harmadik formánsamplitúdó* ( $A_3$ ) 2–4 dB-es módosítása az alanyok túlnyomó többsége szerint nem akadályozta a hangok felismerését. 6 dB-es  $A_3$ -emelkedés mellett csak az  $o$  maradt felismerhető, ugyanilyen  $A_3$ -süllyedés viszont az  $\phi$ -n kívül nem befolyásolta a magánhangzók felismerését. 8 (–10) dB-es  $A_3$ -emelkedés mellett csak az  $o$ , ugyanilyen  $A_3$ -süllyedésnél az  $o, \phi :$  őrizte meg érthetőségét. A for-

<sup>1</sup> A felfelé mutató nyíl — itt, és a továbbiakban — a hang valamelyik akusztikai tényezőjének — ebben az esetben a formáns-amplitúdónak — az értékemelkedését, a lefelé mutató nyíl azok értékcsökkenését jelöli.

mánsamplitúdó-értékek növelése lényegesen erősebben befolyásolta az érthetőséget, mint az értékek csökkentése.

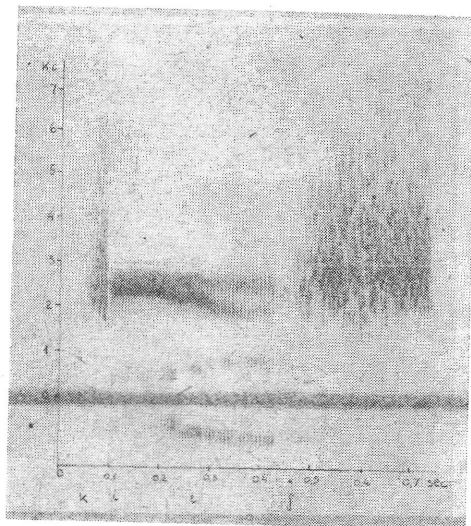
Az 5. kísérlet eredményét a 8. táblázat összegezi. Az *alapfrekvencia* ( $F_0$ ) 10–40 (–50) Hz-cel történt módosítása (emelése, illetőleg süllyesztése) nem befolyásolta számottevően az  $\phi$  és  $o$  érthetőséget. Az  $e$  : esetében azonban már a 10 Hz-es  $F_0$ -emelkedés is jelentős mértékben akadályozta a hang felismerését; az  $F_0$  süllyedése ugyanakkor csupán a hangszín természetességét, megszókottkvalitását vette el. Ez a tapasztalat erősíti TSCHESCHNERNEK (i. m. 6) a német magánhangzók érthetőségére vonatkozó kísérleteredményeinek azt a pontját, mely szerint az alapfrekvencia emelése befolyásolhatja a magánhangzók érthetőségét.

A munka második fázisában a fenti magánhangzókkal *egyszavas mondatok*at alkottam úgy, hogy az illető magánhangzók hangsúlyos helyzetben legyenek, ezeket az egyszavas mondatokat saját kiejtésemmel magnetofon-szalagra vettem, analizáltam és a kapott spektrogramok mintájára elkészítettem szintézisüket (vö. például az 1–5. ábrákat). Az egyszavas mondatok a következők voltak: Kies., Szép., Ezt., Sápadt., Azok., Most., Utas., Sütött., Öt., „Ó” és „ú”. A mondatok hanglejtése — a magyar normatív kijelentő mondat követelményeinek megfelelően — ereszkedő volt. A kísérlet tárgyát képező — természetesen ejtett és szintetizált — magánhangzók formánsfrekvencia ( $F_n$ )-, formánsamplitúdó ( $A_n$ )- és alapfrekvencia ( $F_0$ )-adatait a 9. táblázat szemlélteti. (A dB értékek megállapításánál a természetes hang első formánsamplitúdó-értékét vettem viszonyítási alapul.)



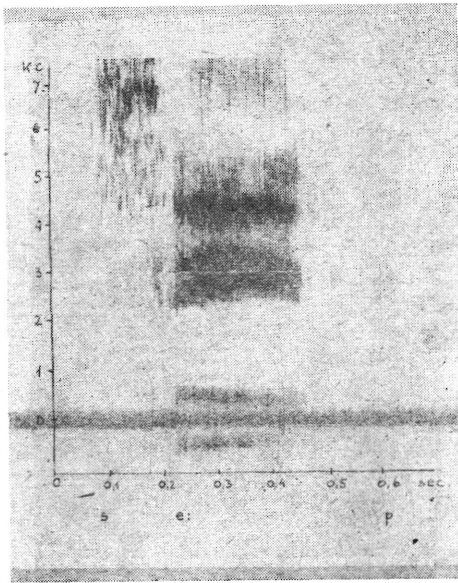
1/a ábra

A *Kies* egyszavas mondat szonogramja a szerző kiejtése alapján



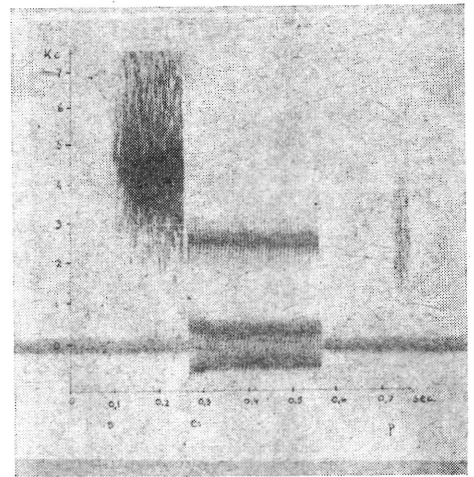
1/b ábra

A *Kies* egyszavas mondat szintetizált változatának szonogramja



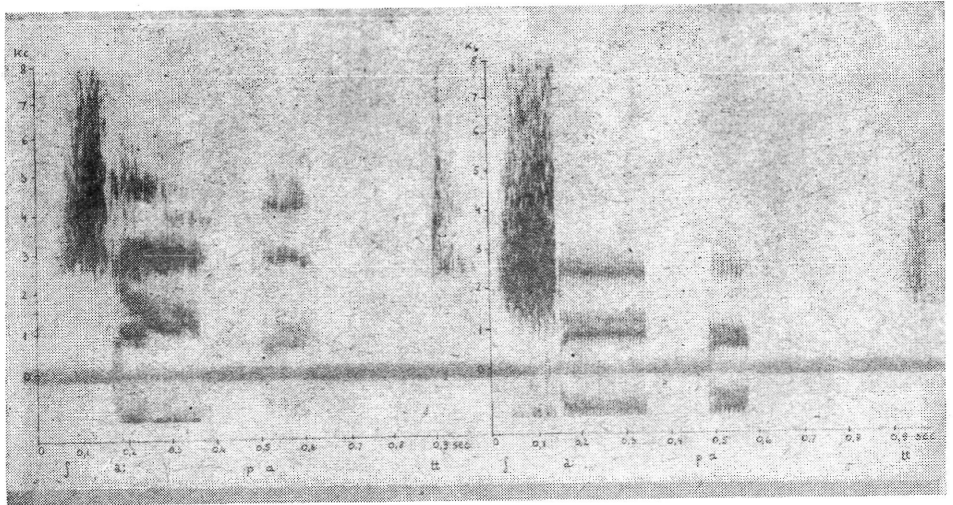
2/a ábra

A *Szép* egyszavas mondat szonogramja  
a szerző ejtése alapján



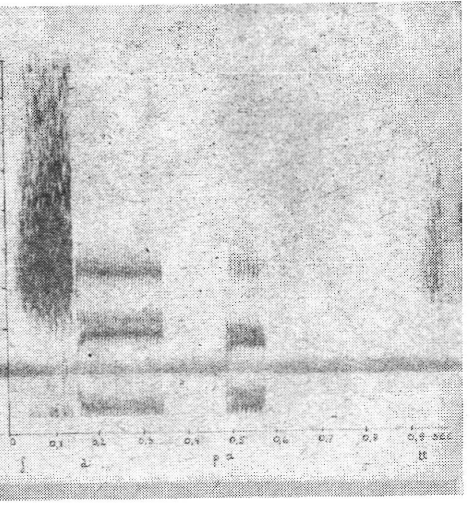
2/b ábra

A *Szép* egyszavas mondat szintetizált  
változatának szonogramja



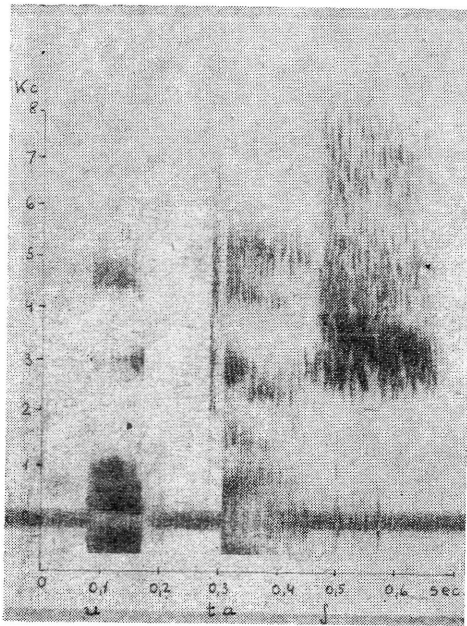
3/a ábra

A *Sápadt* egyszavas mondat szonogramja  
a szerző ejtése alapján



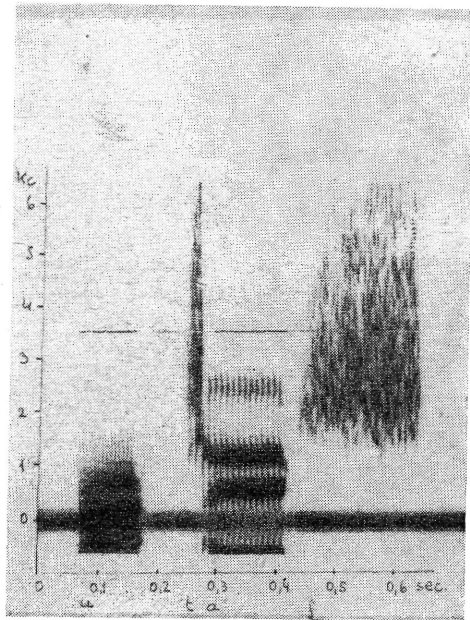
3/b ábra

A *Sápadt* egyszavas mondat szintetizált  
változatának szonogramja



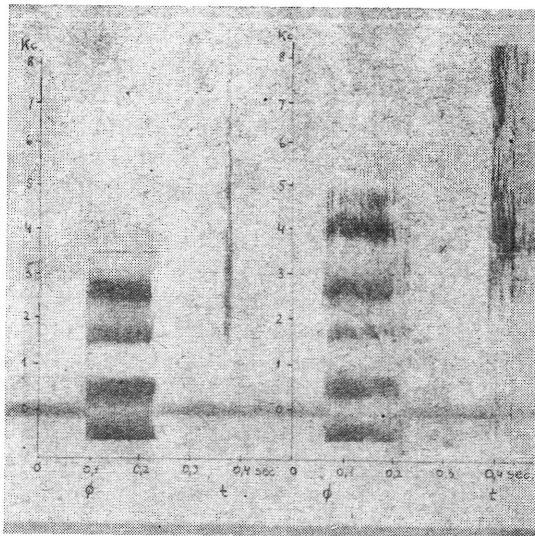
4/a ábra

Az *Utas* egyszavas mondat szonogramja a szerző ejtése alapján



4/b ábra

Az *Utas* egyszavas mondat szintetizált változatának szonogramja



5/a ábra

Az *öt* egyszavas mondat szonogramja a szerző ejtése alapján

5/b ábra

Az *öt* egyszavas mondat szintetizált változatának szonogramja

Szint. hg	Mód. ir.	A formánsamplitudó-módosítás mértéke (dB-ben)										
		300					500					
		É				Éh	É				Éh	
		V		-V				V		-V		
		T	Te	M	I		T	Te	M	I		
i	↓			1 : e	1 : ə	18			2 : e		18	
e:	↓			4 : ε	2 : ə	14			5 : ε		15	
ε	↓	13	7				6	12			2	
a:	↕	20					20					
		18	2				12	8				
a	↕	20					20					
		20					20					
o	↓	20					20					
u	↓	20					20					
y	↕	18	2				8	8				
		19	1				15	5				
y:	↓		3		8 : ə :	9		1		9 : ə :	10	
ø	↓	20						7	4 : y		9	
ø:	↕	10	10				8	10	2 : y :			
		10	10				10	10				

Izolált szintetizált magánhangzókkal  $F_3$ -torzítás mellett végzett érthetőségi vizsgálatok eredménye

A szintetizált egyszavas mondatokban, pontosabban azok hangsúlyos magánhangzóiban, ugyanazokat a módosításokat hajtottam végre és a kifejezésekkel ugyanazt a kísérletsorozatot végeztem el, mint a fenti izolált magánhangzókkal (vö. 357. l.).

Az *első kísérlet* (felismerhetők-e a szintetizált mondatok?) során mind a 20 alany érthetőnek és elfogadhatónak találta a hallott egyszavas mondatokat.

A *második kísérlet* (természetesnek vagy természetellenesnek hatnak-e?) a *Kies.* kivételével szintén osztatlanul pozitív eredménnyel zárult: az alanyok egyöntetűen érthetőnek és természetesnek ítélték az egyszavas mondatokat. A *Kies.*-t hárman természetellenesnek érezték, az *i* ernyedtsége, alacsonyabb nyelvvállása miatt.

A *harmadik kísérlet* (a mondatok meghallgattatása  $F_{1-3}$ -módosítás mellett) eredményeit a 10–12. táblázatok szemléltetik. Az *első formáns* ( $F_1$ ) helyének változtatása ezúttal is — akárcsak az izolált hangok esetében — elsősorban az előlképzett, és ezek közül is a magas nyelvvállású, magánhangzók esetében mutatkozott relevánsnak. A hatás folyamata azonban itt számottevően lassabbnak és a hatás foka enyhébbnek bizonyult, mint az izolált hangok esetében. Az izolált *i* már 50 Hz-es módosításnál *e*-kvalitásba ment át a szavazatok 80%-a szerint, míg a kontextusban (*Kies*) szereplő *i* ugyanilyen  $F_1$ -

## I ázat

Szint, hg	Mód. ir.	A formánsamplitudó-módosítás mértéke (dB-ben)									
		700					900(-1000)				
		É				Éh	É				Éh
		V		-V			V		-V		
T	Te	M	I		T	Te	M	I			
i	↓			2 : e		18					
e:	↓			5 : ε		15					
ε	↓		9	3 : a		8			10 : a:	10	
a:	↑↓	20					20				
		12	8				12	8			
a	↑	20					20				
	↓	20					20				
o	↓	20					20				
u	↓	20					20				
y	↑	10	2			8	9	1		10	
	↓	12	3			5	10	3		7	
y:	↓		5			15		5		15	
ø	↓		6	6 : y		8		5	6 : y	9	
ø:	↑		7	5 : y :		8		7	6 : y:	7	
	↓	10	10				10	10			

módosításnál a szavazatok 60%-a szerint még érthetőnek, csak természetellenesnek tűnt, és az alanyoknak csupán 40%-a ítélte érthetetlennek; 100 Hz-es módosításra vált a kifejezés — a szavazók 75%-a szerint — felismerhetetlenné. Az izolált  $\varepsilon$  : 50 Hz-es  $F_1$ -módosításnál a szavazók többsége (60–65%) szerint elvesztette  $e$  : jellegét, míg a Szép még 200 Hz-es módosítás mellett is érthető maradt a szavazatok 80%-a alapján. Az izolált  $\varepsilon$  50 Hz-es  $F_1$ -emelkedésnél az alanyok 60%-ának megítélése szerint érthetetlen lett, majd 100 Hz-es emelkedésnél az alanyok 40%-a más hangkvalitásnak érezte, ezzel szemben az Ezt  $\varepsilon$ -je 200 Hz-es  $F_1$ -emelkedésnél sem vesztette el érthetőségét. Az izolált  $\varepsilon$  100 Hz-es  $F_1$ -süllyesztése 85%-os szavazattöbbséggel érthetetlen hangot eredményezett, míg az Ezt  $\varepsilon$ -jének 300 Hz-es süllyesztése sem gátolta a kifejezete felismerését a hallgatók 90%-ánál. Az izolált  $y$ ,  $y$  : 50 Hz-es  $F_1$ -módosítás ez a szavazatok túlnyomó többsége (70–85%-a) szerint érthetlenné vált, a Sütött ugyanakkor csak az  $y$  200 Hz-es módosításánál bizonyult — a szavazók fele szerint — felismerhetetlennek, s az „ $\hat{O}$ ” és „ $\hat{u}$ ”-t is csak az alanyok 60%-a találta érthetetlennek az  $y$  : 300 Hz-es módosítása mellett. Az  $\emptyset$  azonos módon viselkedett felismerhetőség szempontjából izoláltan és kontextusban: 200 Hz-es  $F_1$ -módosításnál érthetlenné vált a hang és az egyszavas mondat ( $\hat{O}i$ ). Az izolált  $\emptyset$  : 50 Hz-es  $F_1$ -emelkedésnél és 100 Hz-es  $F_1$ -süllyedésnél az alanyok 70–85%-a szerint felismerhetetlen volt, ugyanakkor az „ $\hat{O}$ ” és „ $\hat{u}$ ” a hallgatók felének megítélése szerint érthető maradt, ha az  $\emptyset$  : első formánsát 50 Hz-cel emeltem, illetőleg 100 Hz-cel süllyesztettem. Míg az izolált  $a$  :  $a$ ,  $o$ ,  $u$  200 Hz-es

Szint. gh	A <sub>n</sub>	Mód. ir.	A formánsfrekvencia-módosítás mértéke (Uz-ben)									
			2				4					
			f		Éh		f		Éh			
			V		-V		V		-V			
T	Te	M	I	T	Te	M	I					
e:	A <sub>1</sub>	↑		16		2 : ə	2		13		6 : ə	1
		↓	1	15		3 : ə	1		11		7 : ə	2
	A <sub>2</sub>	↑	2	9		3 : ə	8		5		1 : ə	14
		↓	17				1	1	14			5
	A <sub>3</sub>	↑	1	17			2		9			11
		↓	1	16			3		15		1 : ə	4
o	A <sub>1</sub>	↑	16	4			10	10				
		↓	2	17	1 : a		1	17	2 : a			
	A <sub>2</sub>	↑	4	6	2 : a		8	10	6 : a	2 : ə		2
		↓	8	12				20	20			
	A <sub>3</sub>	↑	20				20					
		↓	20									
ó	A <sub>1</sub>	↑	2	15			3	1	10			9
		↓	1	16			3	1	17			2
	A <sub>3</sub>	↑	10	4			6	1	8	6 : ə		5
		↓	13	4			3	3	17			3
	A <sub>3</sub>	↑	3	12	3 : ə		2		13			7
		↓	18	2				16	4			
ö	A <sub>1</sub>	↑	2	18				2	18			
		↓	20				19	1				
	A <sub>2</sub>	↑		7		3 : ə:	10		5		2 : ə:	13
		↓		15	1 : e:	2 : ə:	2		13	2 : e:	3 : ə:	2
	A <sub>3</sub>	↑	20						10			10
		↓	8	12				8	12			
y:	A <sub>1</sub>	↑		11	9 : u				10	5 : u		5
		↓	20				20					
	A <sub>2</sub>	↑		10	5 : i	5 : ə:			8	10 : i		2
		↓	20					10	10			
	A <sub>3</sub>	↑	10	10				10	10			
		↓	10	15	3 : u:	2 : ə:			11	9 : u:		

Izolált szintetizált e:, o:, ó hangokkal A<sub>1-3</sub>-torzítás mellett végzett érthetőségi vizsgálatok eredménye

F<sub>1</sub>-módosításnál általában felismerhetetlen lett, illetőleg más magyar hangkvalitásba ment át, addig a *Sápadt*, *Azok*, *Most*, *Utas* szómondatok — a *Sápadt* kivételével — még 300 Hz-es módosítás mellett is felismerhetők maradtak; a *Sápadt* 300 Hz-es módosításánál a szavazók 60%-a szerint elvesztette érthetőségét (vö. 10. táblázat). Az F<sub>1</sub>-emelkedés ezúttal is — akárcsak az izolált hangok esetében — erősebben befolyásolta az érthetőséget, mint az F<sub>1</sub>-süllyedés. A *második formáns* (F<sub>2</sub>) helyének 100 Hz-es változtatása az „ő” és „ü”, valamint az *Öt* hangsúlyos magánhangzóiban az alanyok 60–100%-ának véleménye alapján akadályozta a mondatok megértését; a további mondatok a szava-



l á z a t

Szint. hg	A <sub>n</sub>	Mód. ir.	A formánsfrekvencia-módosításmértéke (Uz-ben)									
			6					8(-10)				
			f				f <sub>h</sub>	f				f <sub>h</sub>
			V		-V			V		-V		
			T	Te	M	I		T	Te	M	I	
e:	A <sub>1</sub>	↑ ↓		6		1 : ə	13		4		1 : ə	15
				9		3 : ə	8		7		2 : ə	11
	A <sub>2</sub>	↑ ↓		4			16		4			16
o	A <sub>2</sub>	↑ ↓	N	10			9		8			12
	A <sub>3</sub>	↑	9	5			11		9			11
				5			5		9			11
ø	A <sub>1</sub>	↑ ↓		4	6		10		8			12
	A <sub>2</sub>	↑ ↓		1	17	2 : a		18	2 : a			
	A <sub>3</sub>	↑ ↓			10	17 : a	2 : o	1	17 : a			2
ø	A <sub>2</sub>	↑ ↓			10		10		9			11
	A <sub>3</sub>	↑ ↓		4	12	2 : ø	6	15	12	1 : ø	2 : ø	5
				18	2			5	4			
ø	A <sub>1</sub>	↑ ↓		1	4		16		4			16
	A <sub>2</sub>	↑ ↓			7	4 : ə	9		6		2 : ə	12
	A <sub>3</sub>	↑ ↓			8	3 : ə	9		6		1 : ə	13
ø	A <sub>3</sub>	↑ ↓			8	3 : y	9		2	6 : y		12
				10	3		7	9	2			9
	A <sub>1</sub>	↑ ↓		5	10		10		12			8
ø	A <sub>2</sub>	↑ ↓			3			17	3			
	A <sub>3</sub>	↑ ↓			5		15	5	10	2 : e :	2 : ə :	15
				10	10	2 : e :	2 : ə :	6	10			6
y:	A <sub>1</sub>	↑ ↓			9		11		9			11
				10	10			6	6			8
	A <sub>2</sub>	↑ ↓			10	6 : u	4		5	5 : u		10
y:	A <sub>2</sub>	↑ ↓		8	4		8	8	1			11
	A <sub>3</sub>	↑ ↓			8	5 : i	7		8	5 : i		7
				5	4		11	5	4			11
y:	A <sub>3</sub>	↑ ↓			8		11	1	8			11
					10	9 : u :	1		5	9 : u :		5

zók többsége szerint felismerhetők maradtak. Az y :, ø, ø : izolált állapotban is felismerhetetlenek voltak 200 Hz-es F<sub>2</sub>-módosításnál, ugyanakkor azonban egy új hangkvalitás érzete is jelentkezett. A hangsúlyos magánhangzó 400 Hz-es F<sub>2</sub>-módosítása a *Kies, Szép, Ezt, „Ő” és „ű”, Őt* kifejezéseket a szavazatok 60–100%-a szerint felismerhetetlenné tette. A felismerhetetlenség az izolált i, e :, ε, y :, ø, ø : esetében is jelentkezett, de ugyanakkor igen erőteljes kristályosodás mutatkozott más magyar magánhangzóknál tekintetében, amely 600–700 (–800) Hz-es F<sub>2</sub>-módosítás mellett általános véleménnyé vált. A további kifejezések (amelyek jobbra hátulképzett magánhangzókat tartalmaztak) a szavazatok nagyobb százaléka alapján felismerhetők maradtak 400 Hz-es F<sub>2</sub>-módosításnál. Az izolált hátulképzett magánhangzók esetében hasonló F<sub>2</sub>-módosítás mellett kevés az „érthető”-re adott szavazatok száma. 600–700 (–800) Hz-es F<sub>2</sub>-módosítás hatására a mondatok – kivéve a *Sál-*

Szint. hg	Mód. ir.	Az alapfrekvencia-torzítás mértéke (Hz-ben)																		
		10					20													
		f				f <sub>h</sub>	f				f <sub>h</sub>									
		V		-V			V		-V											
		T	Te	M	I		T	Te	M	I										
e :	↕		8							12		7								13
	↕	1	19								2	18								
o	↕	20									18	2								
	↕	20									20									
ø	↕	10	10								8	12								
	↕	18	2								8	12								

Izolált szintetizált magánhangzókkal alapfrekvencia-torzítás mellett végzett érthetőségi vizsgálatok eredménye

9. táblázat

	F <sub>n</sub> (Hz)			A <sub>n</sub> (dB)			F <sub>0</sub> (Hz)
	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	
i t	300	2700	3300	-3	-9	-10	250
sz	300	2500	2700	-3	-9	-10	180
e : t	400	2600	3200	-1	-5	-10	200
sz	400	2500	2700	-1	-5	-9	150
ε t	700	2100	3100	-0,5	-5	-10	190
sz	700	2100	2700	0	-5	-10	120
a : t	950	1500	2350	+3	-4,5	-8	210
sz	900	1500	2400	+2,5	-4	-7	130
α t	700	1450	2600	0	-1	-8,5	190
sz	650	1400	2600	0	-1	-8	100
ø t	450	800	2700	+1	-5	-9	180
sz	450	800	2700	0	-4,5	-8,5	12
u t	300	750	—	-2	-5	—	180
sz	300	750	—	-2	-5	—	140
ø t	350	1700	2600	0	-2	-3	190
sz	350	1650	2600	0	-2	-3	160
ø : t	400	1850	2500	-1	-5	-7	220
sz	400	180	2500	-1	-4	-7	170
y t	300	1750	2500	-1	-1	-6	210
sz	300	1700	2500	-1	0	-5	130
y : t	250	1900	2800	-2	-6	-10	210
sz	250	2000	2700	-1,5	-5	-10	170

A természetes (t) és szintetizált (sz) egyszavas mondatok hangsúlyos magánhangzóinak formáns-frekvencia (F<sub>n</sub>)-, formánsamplitúdó (A<sub>n</sub>)- és alapfrekvencia (F<sub>0</sub>)-értékei

## I ázat

Szint. hg	Mód. ir.	Az alapfrekvencia-torzítás mértéke (Hz-ben)										
		30					40(-50)					
		É				Éh	É				Éh	
		V		-V			V		-V			
		T	Te	M	I		T	Te	M	I		
e :	↑ ↓	5	10 15					10		5 17	3 : i	11 3
o	↑ ↓	18 20	2						15 20	5		
ø	↑ ↓	6 6	10 14					4	6 2	9 18		5

10. táblázat

Egyszavas mondatokkal a hangsúlyos magánhangzó F<sub>1</sub>-torzítása mellett végzett érthetőségi kísérletek eredménye

Szint hg	Mód. ir.	A formánsfrekvencia-módosítás mértéke (Hz-ben)											
		50			100			200			300		
		É	Éh		T	Éh		É	Éh		É	Éh	
		T	Te		T	Te		T	Te		T	Te	
Kies	↑		12	8		5	15			20			
Szép	↑ ↓		16 20	4		18 16	2 4		10 16	10 4	18 8	12 12	
Ezt	↑ ↓	15 12	5 8		15 10	5 10		5	12 15	8	18	20 2	
Sápadt	↓		12	8		10	10		10	10	8	12	
Azok	↑ ↓		20 4			20 18			19 15	1 5	19 15	1 5	
Most	↑ ↓	18 17	2 3		8 17	12 3		10	15 10	5	5 10	5 15	
Utas	↑	8	12			20			20		20		
Sütött	↑	18	2		14	6		7	3	10	7	3	10
Ű	↑		10	10		9	11		9	11		8	12
Őt	↑ ↓	20			20				5 6	15 14		3 2	17 18
Ő	↑ ↓		10 12	10 8		8 10	12 10		8 10	12 10		4 9	16 10

11. táblázat

Egyszavas mondatokkal a hangsúlyos magánhangzó  $F_2$ -módosítása mellett végzett érthetőségi kísérletek eredménye

Szent mondat	Mód ir	A formánsfrekvencia-módosítás mértéke ( $Uz$ -ben)											
		200			400			600			700(—800)		
		É		Éh	É		Éh	É		Éh	É		Éh
		T	Te		T	Te		T	Te		T	Te	
Kies	↓		10	10		5	15		3	17			20
Szép	↓	2	12	6		8	12		8	12		2	18
Ezt	↑	2	16	2	1	15	4		11	9		6	14
	↓		10	10		10	10		10	10			20
Sápadt	↑		19	1		15	5		10	10		8	12
	↓	15	5		13	7		10	10			13	7
Azok	↑		13	7		12	8		10	10		7	13
	↓		10	10		10	10		8	12		5	15
Most	↑		10	10		10	10		8	12		7	13
	↓	20			20				20		10	10	
Utas	↑		13	7		11	9		11	9		10	10
Sütött	↑	15	5		13	7		12	10			11	9
	↓	18	2		15	5		10	10			8	12
Ű	↑		8	12		7	13		7	13		6	14
	↓	10	10		10	10		7	13			8	12
Őt	↑			20			20			20			20
	↓	20			20					20			20
Ó	↑	15	5		13	7		10	10			8	12
	↓		6	14		5	15		3	17		2	18

*padt, Mo|st, Utas, Sütött* kifejezéseket, amelyeket általában felismertek az alanyok — a szavazatok 60–100%-a szerint érthetetlennek bizonyultak (vö. 11. táblázat). Az izolált hangok közül egy sem maradt felismerhető 600–700 (—800) Hz-es  $F_2$ -módosítás mellett. A *harmadik formáns* ( $F_3$ ) helyének változtatása csak a *Kies, Szép, Ezt* szómondatok esetében akadályozta az alanyok többsége szerint a megértést, a *Kies* és *Szép* esetében 300 Hz-es, az *Ezt* esetében 500 Hz-es módosítás mellett (vö. 12. táblázat). Az izolált hangok közül csak az *a, o, u, y* őrizték meg érthetőségüket 900 (—1000) Hz-es  $F_3$ -módosítás mellett.

A 4. kísérlet (formánsamplitúdó-torzítás) eredményét a 13. táblázat mutatja. Az  $A_1$  értékének 2–8 (—10) dB-es módosítása a szavazatok többsége szerint nem akadályozta a mondatok megértését. Az  $A_2$  értékének 8 (—10) dB-lal történő növelése a *Szép* és *Most* kifejezéseket a szavazatok 65, illetőleg 80%-a szerint felismerhetetlenné tette; a többi mondatot azonosította az alanyok többsége. Az  $A_3$  értékének módosítása általában nem befolyásolta számottevően a mondatok megértését. Az izolált hangok esetében az  $A_1$  és  $A_2$  6–8 (—10) dB-es értékmódosítása általában felismerhetetlenné tette a hangokat (vö. 7. táblázat).

Az 5. kísérlet (alapfrekvencia-torzítás) eredményét a 14. táblázat szemlélteti. A *Szép* alapfrekvenciájának 30 Hz-es emelése a szavazatok 60%-a

12. táblázat

Egyszavas mondatokkal a hangsúlyos magánhangzó  $F_3$ -módosítása mellett végzett érthetőségi kísérletek eredménye

Szint. mondat	Mód. ir.	A formánsfrekvencia módosítás mértéke (Hz-ben)											
		3			500			700			(900(—1000))		
		É	Éh		É	Éh		É	Éh		É	Éh	
		T	Te		T	Te		T	Te		T	Te	
Kies	↓		5	15		5	15		4	16		1	19
Szép	↓		8	12		8	12		5	15		5	15
Ezt	↓	16	4		5	15			10	5		5	15
Sápadt	↑	20			20			20			20		
	↓	19	1		19	1		17	3		16	4	
Azok	↑	20			20			20			20		
	↓	20			20			20			20		
Most	↓	20			20			20			20		
Útas	↓	20			20			20			20		
Sütött	↑	20			19	1		17	3		16	4	
	↓	20			08	2		15	5		10	10	
Ű	↓		14	6		14	6		15	5		15	5
Őt	↓	20			18	2		10	10			14	6
ő	↑	15	5		15	5		10	10		8	12	
	↓	15	5		15	5		10	10		10	10	

szerint érthetlenné tette a kifejezést. Az izolált *e* : esetében a felismerhetőség már 10 Hz-es  $F_0$ -emeléskor jelentkezett az alanyok 60%-ánál (vö. 8. táblázat).

A fenti kísérletek konklúziójaként arra a kérdésre kell feleletet adnunk, hogy melyek a vizsgált magánhangzók  $F_n$ -,  $A_n$ - és  $F_0$ -érthetőség határértékei. ,más szóval: meddig módosíthatók tehát a kérdéses magánhangzók  $F_n$ -,  $A_n$ - és  $F_0$ -értékei anélkül, hogy a módosítás a hangok, illetőleg a mondatok megértését gátolná. A választ a 15. táblázat foglalja össze szemléletesen. Az érthetőség határértékének nevezem az  $F_n$ - $A_n$ - és  $F_0$ -módosításnak azt az értékét, amelyenél a szavazóknak legalább 50%-a szerint felismerhető még a hang, illetőleg a mondat. A \*-gal megjelölt számok azt jelentik, hogy az illető hangot vagy mondatot az alanyoknak legalább a fele a módosítás legutolsó fokán is érthetőnek tartotta. A táblázaton látható nullák (0) azt jelentik, hogy az érték az érthetőség keretein belül nem módosítható.

Azt látjuk, hogy az  $F_n$ -,  $A_n$ - és  $F_0$ -értékek emelése általában erősebben befolyásolta a megértést, mint az értékek süllyesztése. A mondatok lényegesen nagyobb toleranciát mutattak mindennemű módosítással szemben, mint az izolált hangok. A három akusztikai tényező közül kétségkívül a formánsfrekvencia, mégpedig elsősorban az első két formáns frekvenciája, volt a legszámottevőbb az izolált hangok és a mondatok megértése terén. A módosításokkal szemben az előlépzett magánhangzók sokkal nagyobb érzékenységet mutattak, mint a hátulképzettek. Az izolált *i* és *e*: azonnal elvesztette érthető-

13. táblázat

Egyszavas mondatokkal a hangsúlyos magánhangzó  $A_{1-3}$ -torzítása mellett végzett érthetőségi kísérletek eredménye

Szint. mondat	Mód ír	A <sub>1</sub> formánsamplitúdó-módosítás mértéke (dB-ben)												
		2			4			6			8(-10)			
		É		Éh	É		Éh	É		Éh	É		Éh	
		T	Te		T	Te		T	Te		T	Te		
Szép	A <sub>1</sub>	↑		20				20					16	4
		↓	4	16		4	16						20	
	A <sub>2</sub>	↑	2	18		2	18						7	10
Most	A <sub>2</sub>	↓	2	18		2	18		2	18		5	2	18
	A <sub>3</sub>	↑	4	16		4	16		4	16			4	16
		↓	5	15		3	17		3	17			3	17
Őt	A <sub>1</sub>	↑	8	12		8	12		6	14			10	10
		↓	4	16		4	16		4	16			12	8
	A <sub>2</sub>	↑		10	10		10	10		10	10		10	10
Ő	A <sub>2</sub>	↓	20			13	7		8	8		5	10	10
	A <sub>3</sub>	↑	15	5		10	10					8	12	8
		↓	15	5		15	5		15	5		15	5	
Ő	A <sub>1</sub>	↑		17	3		17	3		15	5		12	8
		↓	5	10			20			20			20	
	A <sub>2</sub>	↑	5	15		5	15			20			20	
Ő	A <sub>2</sub>	↓	10	10		10	10			15	5		15	5
	A <sub>3</sub>	↑		15	5		10	10		10	10		10	10
		↓	9	11		10	10		10	10		10	10	
Ő	A <sub>1</sub>	↑	2	18		2	18		2	18		2	18	
		↓	20			20			20			20		
	A <sub>2</sub>	↑		15	5		15	5		15	5		15	5
Ő	A <sub>2</sub>	↓		20			20			20			20	
	A <sub>3</sub>	↑	5	15		5	15		5	15		5	15	
		↓												
Ő	A <sub>1</sub>	↑		16	4		16	4		16	4		9	11
		↓	20			20			20			20		
	A <sub>2</sub>	↑		15	5		15	5		16	4		16	4
Ő	A <sub>2</sub>	↓		18	2		18	2		18	2		18	2
	A <sub>3</sub>	↑		16	4		16	4		15	5		15	5
		↓	12	8		12	8		10	10		10	10	

ségét, ha az első három formáns bármelyikét elmozdítottuk helyéről. Igen érzékeny volt az izolált  $\gamma$ : csak az  $F_2$ -súlyesztést tűrte el 200 Hz-en belül, bármely más  $F_n$ -módosítás felismerhetetlenné tette. Az izolált  $\phi$ : nem bírt el  $F_{1-2}$ -módosítást, az  $F_3$ -módosításra azonban nem volt különösebben érzékeny. A legnagyobb toleranciát az izolált  $o$  és  $u$  mutatta az  $F_{1-3}$  módosítással szemben. A formánsamplitúdó-módosítás kevésbé hatott a hangok érthetőségére. Az  $A_2$ -módosítás mutatkozott még a leghatékonyabbnak. Az  $A_2$  értékének a legcsekélyebb emelése felismerhetetlenné tette az izolált  $e$ -t és  $\phi$ -t. Az alapfrekvencia-módosítás még kevésbé befolyásolta az érthetőséget, csak az előlétezett magas nyelvvállású  $e$ :  $F_0$ -emelése gátolta a megértést.



Az alábbiakban *összegezhetjük* a kísérletek eredményét:

1. Az értelemmel bíró kifejezés érthetősége jóval tovább megmarad a magánhangzó akusztikai tényezőinek módosítása során, mint az izolált magánhangzóé. Az  $F_1$  50–250, az  $F_2$  0–600, az  $F_3$  0–900 Hz-cel tovább módosítható (emelhető, illetőleg süllyeszthető) mondatba ágyazott magánhangzóknak, mint izolált hangoknak anélkül, hogy érthetőségét elvesztené.
2. A három akusztikai tényező ( $F_n$ ,  $A_n$ ,  $F_0$ ) közül a formánsfrekvencia befolyásolja leginkább az érthetőséget, a formánsamplitúdó kevésbé és az alapfrekvencia a legkevésbé. A formánsfrekvencia tekintetében az első két formáns, a formánsamplitúdó tekintetében az  $A_2$  módosítása a leghatékonyabb.
3. Az  $F_1$ -módosításnak a hang megértése szempontjából releváns alsó határa 50, felső határa 300 Hz, ezek a határok az  $F_2$  esetében 200, illetőleg 700 (–800), az  $F_3$  esetében 300, illetőleg 900 (–1000) Hz-nél vannak.
4. Az  $F_n$ -,  $A_n$ - és  $F_0$ -értékek emelése általában relevánsabb az érthetőség szempontjából, mint az értékek csökkentése.
5. Az akusztikai tényezők módosításában szemben általában az előlképzett magas nyelvvállású hangok a legérzékenyebbek, kísérleteink alapján az *i*, *y*: és  $\emptyset$ .
6. Az izolált hangok az akusztikai tényezők módosítására más felismerhető (magyar vagy idegen) hangkvalitásba mennek át vagy meghatározhatatlan, felismerhetetlen hangjelenséggé válnak. Az új hangkvalitás elsősorban az  $F_1$ - és  $F_2$ -módosítás hatására keletkezik; az  $F_3$ -,  $A_n$ - és  $F_0$ -módosítás következménye inkább a felismerhetetlen, meghatározhatatlan hangjelenség.
7. Az új hangkvalitás gyakrabban lép fel előlképzett (magas) hangok esetében, mint hátulképzetteknél (mélyeknél).

MAGDICS KLÁRA