

Szegedi Tudományegyetem
ÁOK–TTIK Orvosi Fizikai és Orvosi Informatikai Intézet

TDK-dolgozat

Mobiltelefonos hangrögzítés megbízhatóságának vizsgálata a második formáns meredekségének meghatározásában

Készítette: Ábrahám Ágnes
Fizikus MSc szakos hallgató

Témavezető: Dr. Makra Péter egyetemi adjunktus
SZTE Orvosi Fizikai és Orvosi Informatikai Intézet

Szeged, 2014

Bevezetés

A beszédkésztséget befolyásoló betegségek (pl: agyvérzés) esetén a beszéd spektrális elemzése diagnosztikai lehetőségeket nyithat meg. Erre egy példa a beszéd érthetőségének és a második formánsnak a kapcsolata. Egyes tanulmányok rámutattak arra, hogy (pl. ALS-es) betegek hangfelvételeit összehasonlítva egészséges emberek hangfelvételeivel, a második formáns meredeksége kisebb.

Munkánk célja agyvérzéses betegek beszédrehabilitációjának segítése a beszédminőség objektív minősítésére alkalmas spektrális paraméterek keresésével, lehetőség szerint mobileszközök bevonásával. Feladatom egyrészt a formánskeresés reprodukálhatóságának feltérképezése volt. Mivel a végcélunk egy mobileszközökön is futó alkalmazás kifejlesztése, további feladatom volt összehasonlítani a számítógépes és a mobiltelefonos hangfelvételek spektrumát, továbbá megvizsgálni, hogy a második formáns meredeksége függ-e a hangrögzítésre használt eszköz típusától.

1. fejezet

Elméleti összefoglaló

1.1. A beszéd

1.1.1. A hang

A hang valamely rugalmas közeg mechanikai rezgése, eme rezgés hullámként való szétterjedése. Longitudinális hullám, azaz a közeg részecskéinek rezgési iránya egybeesik a hullám terjedési irányával [1]. Hangforrásnak tekinthető minden olyan rugalmas közeg, amely rezgéseket képes kelteni, illetve külső hatásra rezgésbe tud jönni. Szempontunkból a levegő a legfontosabb közeg, melynek mozgása során nyomásingadozás lép fel. Ennek a nyomásingadozásnak az ismétlődése adja a levegő rezgését, amit a fülünk átvesz, továbbad az agynak, amely hangként észleli a nyomásváltozást. Alapvető hangforrás a gégefőben található hangszalagok. A tüdőből kiáramló levegő szétfeszíti a hangszalagokat. A levegő áthaladása után csökken a nyomás, így a hangszalagok visszatérhetnek eredeti helyzetükbe, és ez a nyitódó- záródó mozgás adja a hangszalagok rezgését, melyet a levegő részecskéinek a rezgései továbbítanak.

Spektrum alapján osztályozva a hangokat beszélhetünk periodikus hangokról, ezen belül is tiszta hangokról, illetve zenei hangokról, továbbá nem periodikus hangokról, ezen belül zörejekről és dőrejekről. A tiszta hang spektruma egyetlen vonal, a zenei hangé vonalas, többnyire egy alaphang és a felharmonikusai. A zörejek és dőrejek spektruma folytonos, mindkettő rövid ideig tart, az utóbbi lökésszerű.

A gégefőben keletkezett hang, amelyet zöngének nevezünk, felhangokban gazdag. A beszédhangok egy csoportjánál a gégefőben a hangszalagok nyugalmi állapotban maradnak, így nem keletkezik zöngéhang. A tüdőből kiáramló levegő egy, a hangképző szervek által kialakított akadály feloldásával jut ki a szájüregből, így jönnek létre a zöngétlen mássalhangzók. A magánhangzók tehát fizikai szempontból zenei hangnak, a zöngétlen mássalhangzók zörejhangoknak, a zöngés mássalhangzók pedig a zenei hangok és a zörejhangok keverékeinek tekinthetők [2].

1.1.2. A beszédhangok képzése

A hang beszédünk legkisebb egysége, ami a tüdőből kiáramló levegő beszédsszervek általi rezgésbe hozásával keletkezik. A hangképzésben szerepet játszó szerveket beszédsszerveknek nevezzük. A beszédsszervek harmonikus együttműködése szükséges a hangképzés megvalósulásához. A beszéd hangjait kilégzéssel képezzük. A tüdőből kiáramló levegő a légcsőn keresztül a gégefőnél találkozik a hangszalagokkal [3]. Ez utóbbi beszédsszervek a kannaporc által állíthatók az éppen szükséges helyzetbe, ugyanis a hangszalagok feszessége a kannaporc elfordulásával változtatható [4].

A toldalékcső azon üreg, amely a gégefőben létrehozott hangot formálja. Magát a toldalékcsövet három üreg alkotja: a garatüreg, a szájüreg és az orrüreg, valamint az ezek körül elhelyezkedő szervek (pl. ajkak, fogak, stb...). A különböző üregek rezonátorként működnek. Az egyes beszédhangok képzése során bizonyos beszédsszervek aktívak, mások passzívak.

Abban az esetben, ha valamely beszédhangunkat akarjuk jellemezni, be kell mutatnunk a hangképző szervpárok részvételének módját is.

A magánhangzó alapeleme a zöngé, nyíláshang, ami azt jelenti hogy a szájüregben nem ütközik akadályba a tüdőből kiáramló levegőáram, minőségét a szájüreg alakja és nagysága, továbbá az ajaknyílás alakja határozza meg.

A mássalhangzó akadályhang, ami azt jelenti, hogy a tüdőből kiáramló levegőáram akadályba ütközik. Ezen akadály leküzdésének módja határozza meg a létrejövő hang egyik minőségi jegyét. Lehet zár, rés és zár- rés. A mássalhangzók ejtések a hangszalagok nem feltétlenül rezegnek [2].

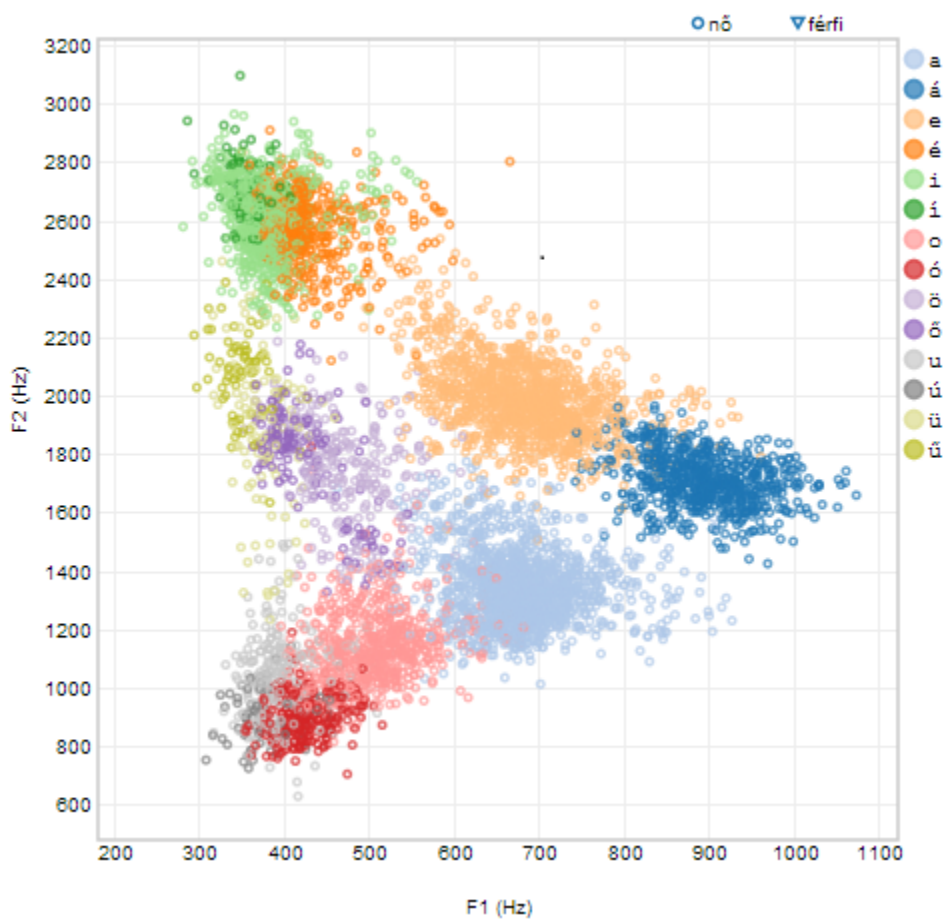
1.2. A formánsok

1.2.1. A formáns fogalma

A rezonátor bizonyos frekvenciákat átenged, másokat nem, tehát szűrőként viselkedik. A beszédhangok kialakulását tekintve a rezonancia, a rezonátor és ennek áteresztőképessége jelentős szereppel bír, mivel a hangforrás alatti és feletti üregek rezonátorként viselkednek. Az alsó üregek (mellüreg és légcső) az egyéni jellegzetességekhez járulnak hozzá, míg a felsők, a garatüreg, a szájüreg és az orrüreg a hangzó jellegét határozzák meg. Ennek menete a következő: a zöngéhang a felső rezonátorüregbe kerül, s ekkor az üregek felerősítik a saját rezgésszámuknak megfelelő vagy annak közelébe eső felhangokat. Ezen felerősített hangok neve formáns, amely tulajdonképpen energiakoncentrációs helyet jelent. A formánsokat az alaphangtól (f_0) való távolságuk alapján számozzuk. Az alaphanghoz legközelebb eső formáns az első formáns (f_1), a távolabb találhatóak f_2 , f_3 , stb [2].

1.2.2. A magánhangzók frekvenciaszerkezete

A magánhangzók formánsszerkezetének megadásához általában az első három formáns frekvenciaértékének ismerete szükséges. Az első (f_1) és a második (f_2) formáns a magyarban a magánhangzó minőségét egyértelműen meghatározza, a harmadik formáns(f_3) általában csak másodlagosan jellemzi a magánhangzót, azonban a beszélő személy egyéni ejtésére vonatkozóan adhat információt. Ha egymáshoz képest ábrázoljuk az első és a második formánst, szemléletes képet kaphatunk a viszonyokról. Az f_1 és f_2 alkotta koordináta-rendszerben a sűrűsödési ellipszisek az egyes magánhangzókra jellemző frekvenciaértékeket tartalmazzák. A sűrűsödési ellipszisekből látható, hogy például az [u, ú] pár esetében a rövid mássalhangzó első formánsának szóródási sávja teljes egészében, a második formánsának a szóródási sávja pedig félig lefedi hosszú párjáét. Ehhez hasonló átfedések az [i, í] és az [ü, ő] esetében is megfigyelhetők [5].



1.1. ábra. f_1/f_2 eloszlás női beszélő esetén [6]

Az alábbi táblázat a magyar magánhangzók formánsainak tájékoztató adatait tartalmazza [5].

Magánhangzó	Első formáns f_1 [Hz]	Második formáns f_2 [Hz]	Harmadik formáns f_3 [Hz]
/á/	800	1400	2500
/a/	600	1100	2400
/o/	450	900	2300
/ó/	400	950	2300
/u/	300	650	2300
/ú/	280	600	2100
/ö/	380	1600	2400
/ő/	400	1550	2500
/ü/	280	1750	2200
/ű/	260	1800	2150
/e/	550	1900	2500
/é/	420	2100	2800
/i/	280	2400	3000
/í/	260	2500	2850

1.2.3. A mássalhangzók frekvenciaszerkezete

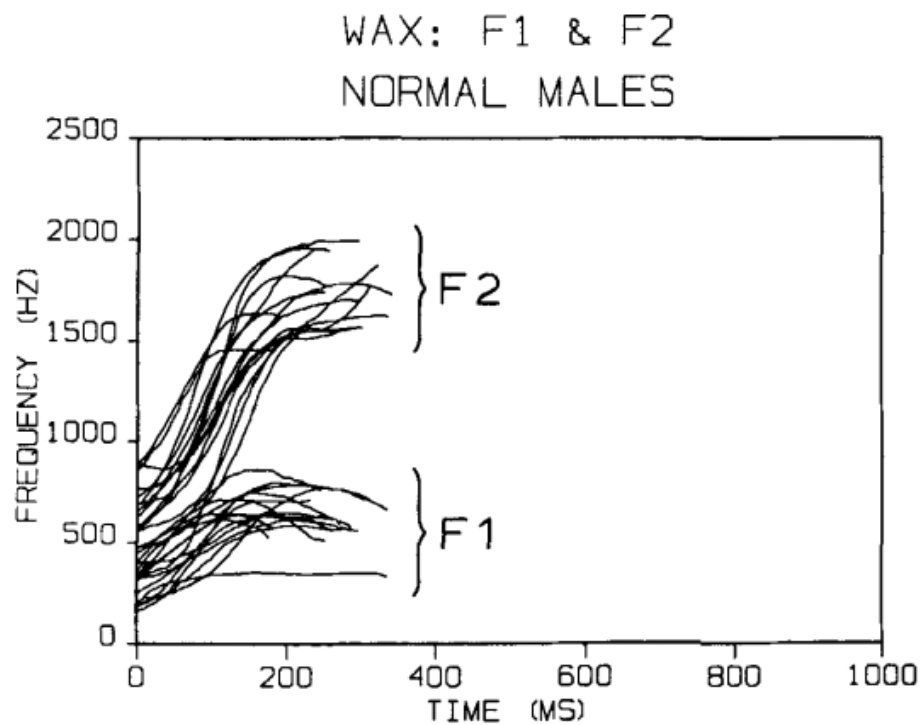
A zöngés mássalhangzók gerjesztésekor – a magánhangzókhoz hasonlóan – a hangszalagok rezgésbe jönnek, tehát ezen hangok esetében is beszélhetünk formánsokról, erre azonban a mássalhangzó képzésétől függően valamilyen zörej is ráakódik. A zörejes mássalhangzók esetében nem formánsokról, csak zörejfrequencia-helyekről beszélünk.

Az alábbi táblázat a magyar mássalhangzók formánsainak tájékoztató adatait tartalmazza [5].

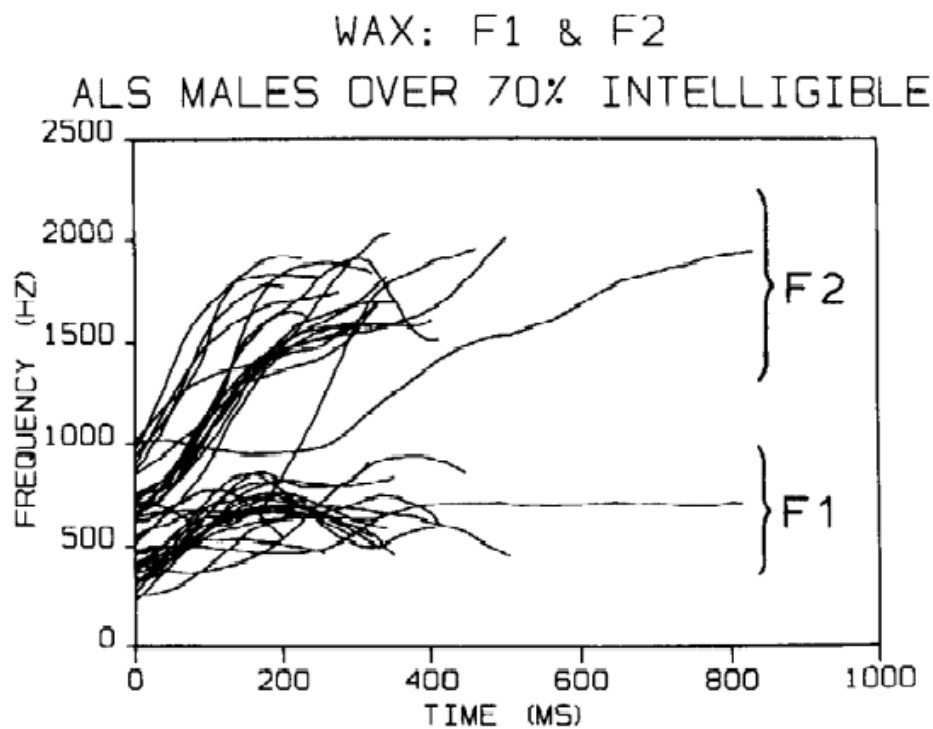
Mássalhangzó	f_1 [Hz]	f_2 [Hz]
/b/	300–400	600–1600
/d/	400	1300–1800
/dz/	300	1400–1800
/dzs/	300	1400–1800
/g/	300	900–2800
/gy/	300	2100
/j/	370	2200
/l/	350	1300–1600
/m/	300–400	800–2200
/n/	300–400	800–2200
/ny/	300–400	800–2200
/r/	500	1400
/v/	300	800–1600
/z/	250–300	1300–1800
/zs/	250	1300–2100

1.3. A második formáns diagnosztikai szerepe

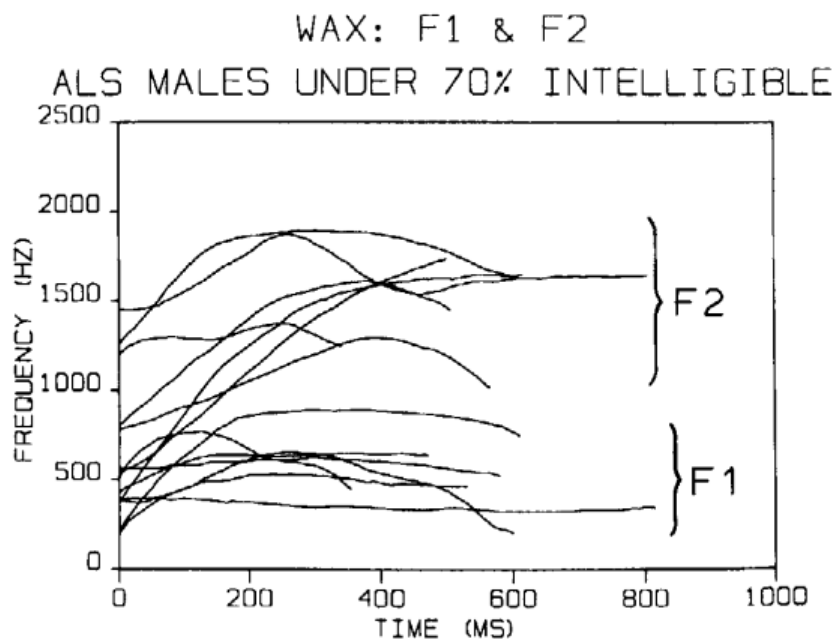
Kent és munkatársai 1989-ben elvégeztek egy mérést, ami a beszéd érthetősége és a második formáns meredeksége közti kapcsolatot vizsgálja. A beszéd érthetőségét az alapján számszerűsítették, hogy a felkért hallgatóság a betegek által felolvasott szavak hány százalékát tudta helyesen beazonosítani. Dizarthriás betegekkel meghatározott szavakat mondtak ki, és ezekben a szavakban vizsgálták az első és második formánsok meredekségének alakulását. Az alábbi ábrák a *wax* szó (f_1) és (f_2) formánsgörbéit mutatják be. 3 esetet különböztetnek meg: az 1.2 ábrán egészséges férfiak eredményei, míg az 1.3 ábrán olyan alanyok formánsmeredekségei szerepelnek, akiknél a beszédérthetőség 70% feletti, az 1.4 ábra pedig a 70%-os beszédérthetőség alatti eredményeket ábrázolja.



1.2. ábra. Az angol *wax* szó formánsmeredeksége egészséges beszélők esetén [7]



1.3. ábra. Az angol *wax* szó formánsmeredeksége 70% feletti beszédérthetőség esetén [7]

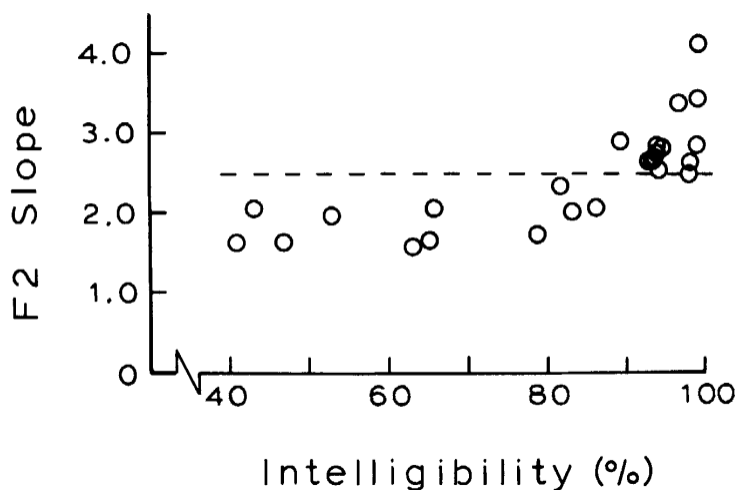


1.4. ábra. Az angol *wax* szó formánsmeredeksége 70% alatti beszédérthetőség esetén [7]

Mindkét ALS-es csoportnak hosszabb lefutású F2 pályáik vannak (tekintve, hogy esetükben a szavak kimondása több időt vesz igénybe) és kevésbé meredek a második formáns növekedése, mint a kontrollcsoport esetén.

Ezekből az eredményekből Kenték azt feltételezték, hogy a második formáns meredekségének csökkenése összefügg a beszéd érthetőségével.

Az alábbi ábrán a második formáns meredeksége és a beszéd érthetőségének kapcsolata látható. A szaggatott vonal alatt találhatók az alacsonyabb beszédérthetőségű csoport, míg felette a magasabb beszédérthetőségű csoport formánsmeredekségei [7].



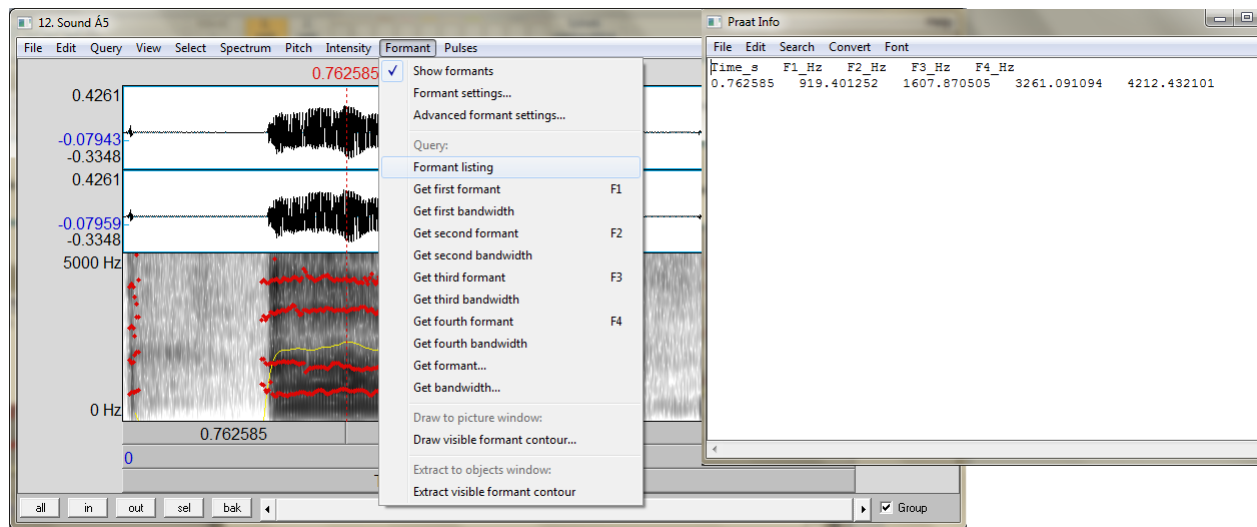
1.5. ábra. Az f_2 meredekségének és a beszéd érthetőségének kapcsolata [7]

2. fejezet

Saját mérések

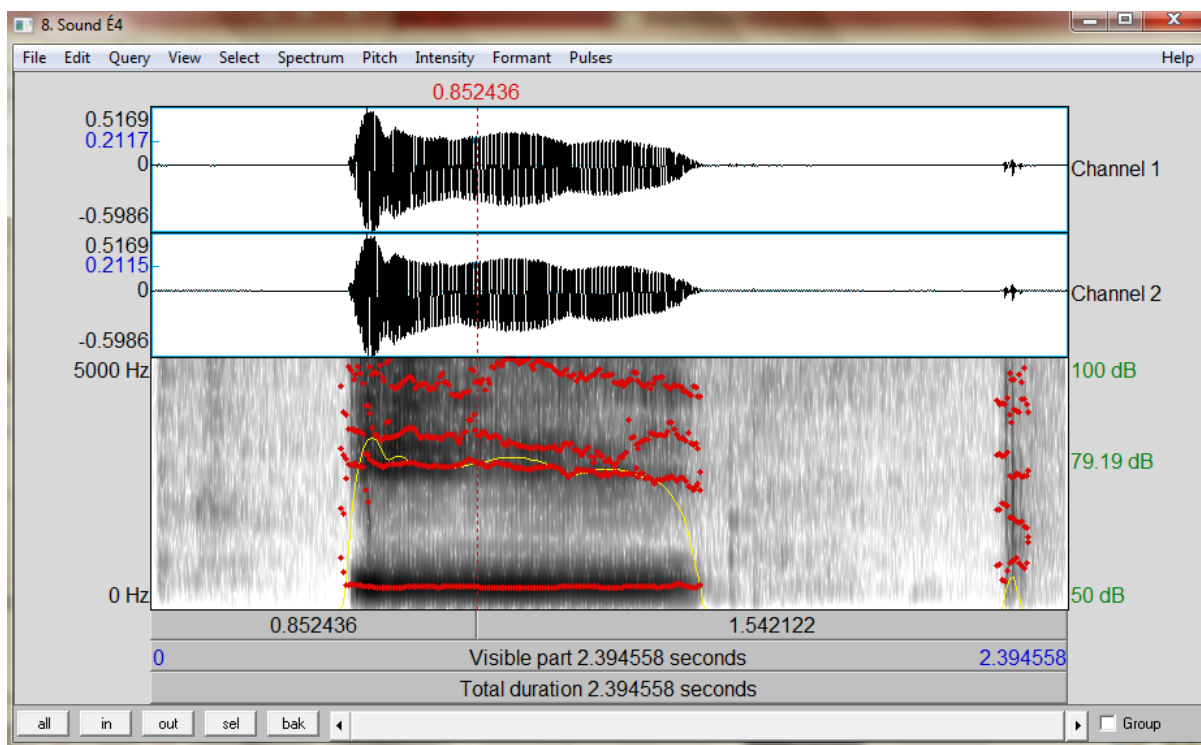
2.1. A formánsok követése

Az Audacity nevű ingyenes program segítségével sikerült a saját hangomon minden magyar hangzóról 15 egymástól független felvételt készítenem. Ezután a Praat nevű program segítségével minden felvételt külön kielemeztem. Vizsgálatom az első négy formánsig terjedt, átlagolással egy referenciaértéket határoztam meg, továbbá az egyes formánsok különböző felvételekből adódó szórását is vizsgáltam. Minden felvételnél csak egy pontban mértem, leginkább ott, ahol az intenzitás a maximális értékű volt.



2.1. ábra. Formánsok kiíratása

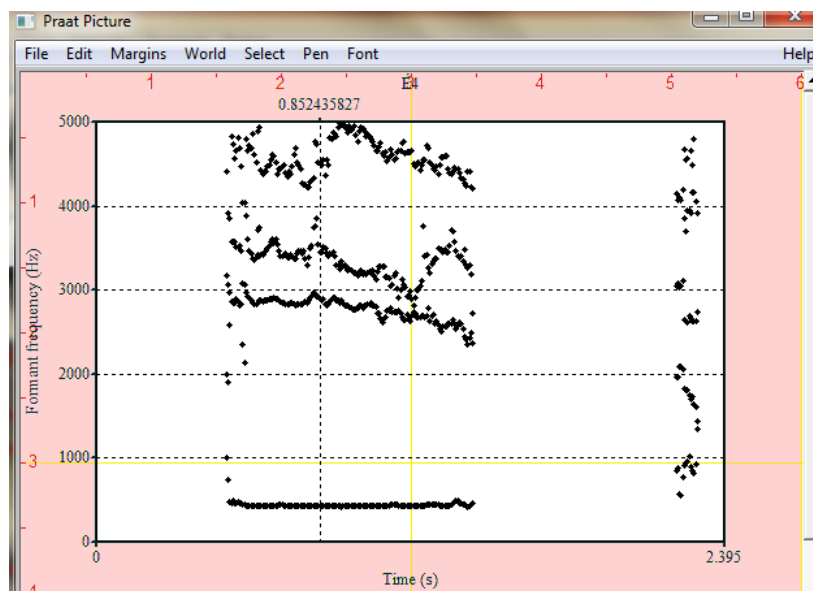
A magánhangzók esetén többnyire jól elkülönülnek a formánsok.



2.2. ábra. Az /é/ formánsai

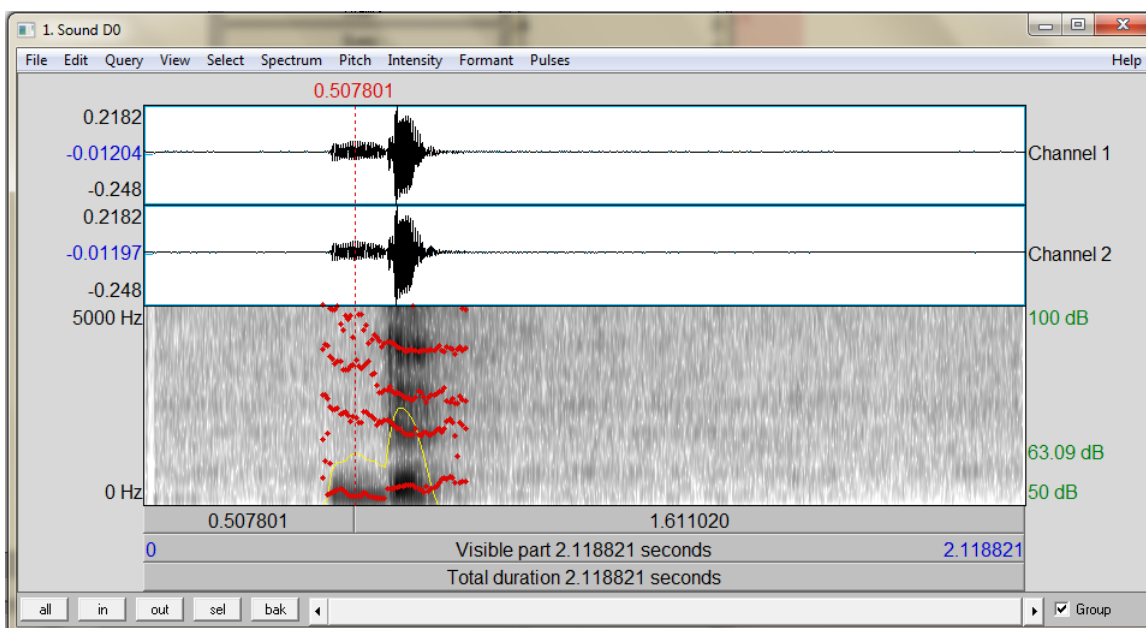
A fenti ábrán az /é/ hang formánsai láthatók. Tisztán kivehetők, azonban az is látható, hogy a harmadik formáns szórása igen nagy. A sárga színű vonal az intenzitást mutatja.

A program lehetővé teszi a formánsok sávszélességének ábrázolását is.



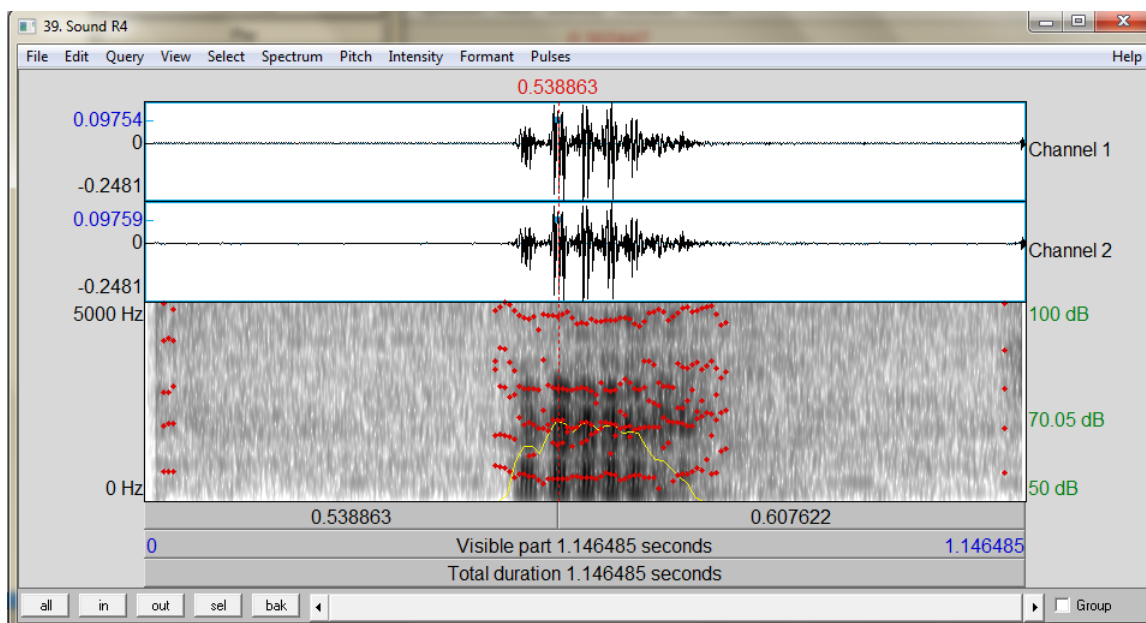
2.3. ábra. Az /é/ formánsainak sávszélessége

A felpattanó zárhangoknál a spektrumon a zöngés zár-szakasz, valamint a zárfelpattanás is jól látszik. Az alábbi ábrán a /d/ hang spektruma, illetve formáns szerkezete látható.



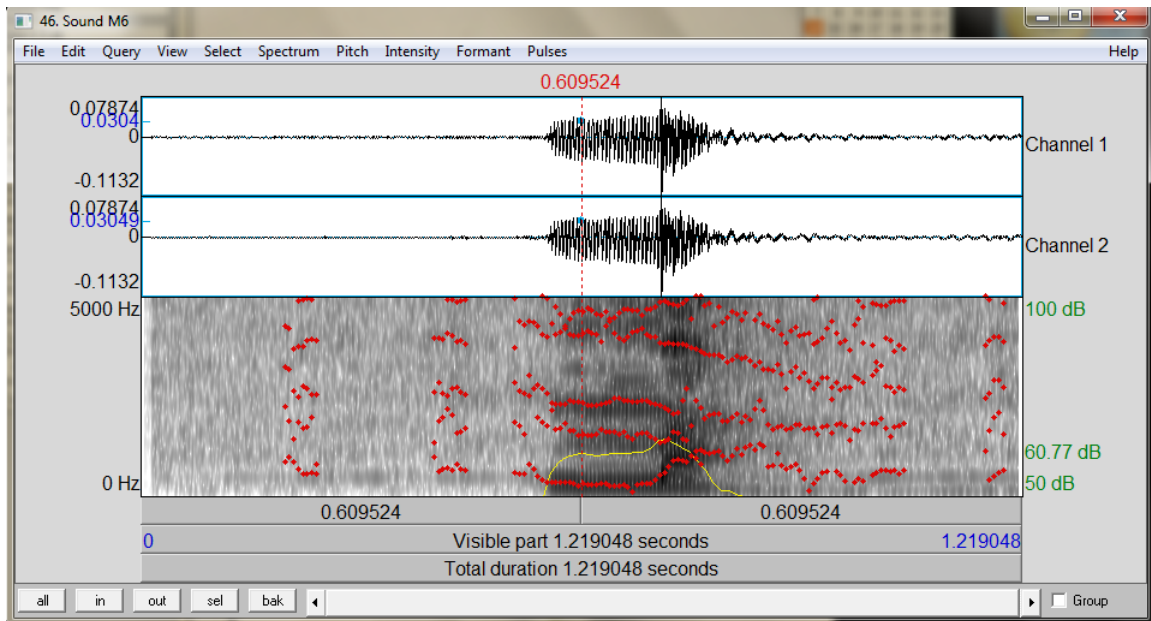
2.4. ábra. A /d/ hang formánsai

Az alábbi ábra az /r/ pergőhang Praat-tal való ábrázolása, a pergőhang periodikusságát és formáns szerkezetét mutatja be.

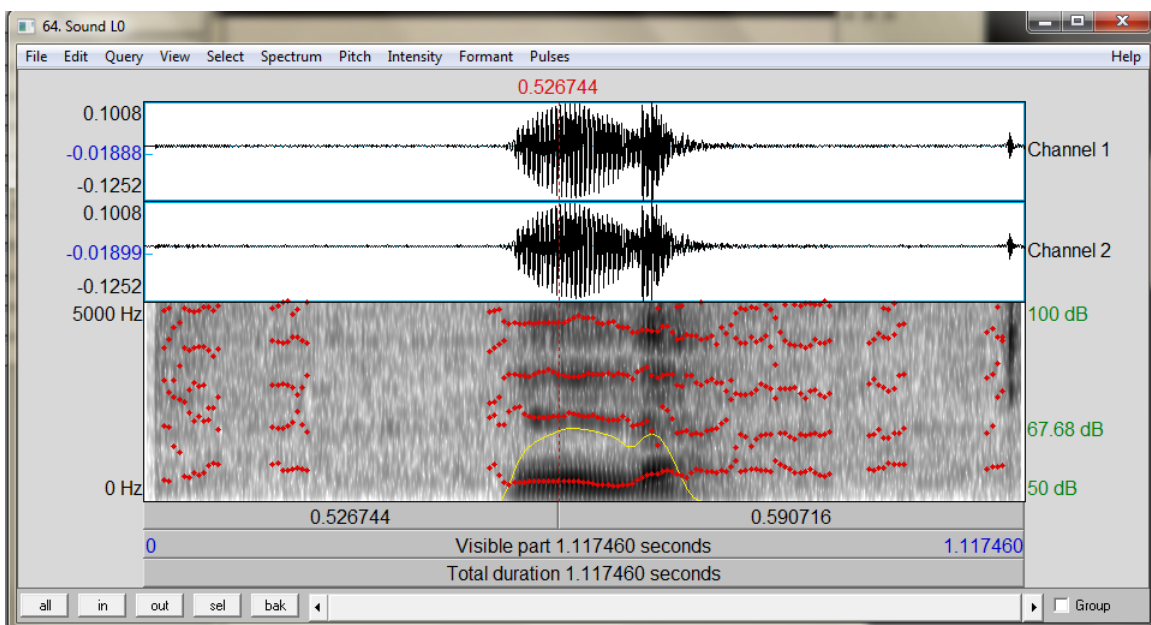


2.5. ábra. Az /r/ hang formánsai

A nazálisoknál és a laterálisoknál is jól elkülönülnek egymástól a különböző formánsok. Az alábbi ábrákon az /m/ nazális hang és az /l/ laterális hangok spektrumai és formánsszerkezetei láthatók.



2.6. ábra. Az /m/ hang formánsai



2.7. ábra. A /l/ hang formánsai

2.1.1. Reprodukálhatósági vizsgálatok saját hangfelvételeken

Előzetes méréseim során sikerült beazonosítanom a magyar beszédhangok formánsait, valamint a formánsazonosítás reprodukálhatóságát is feltérképeztem. A felvételeket számítógéppel készítettem.

	f_1 [Hz]	f_2 [Hz]	f_3 [Hz]	f_4 [Hz]
á	924,21	1575,54	3028,60	4091,72
szórás	66,62	183,28	660,58	434,07
a	620,74	986,65	3184,68	3929,31
szórás	39,51	34,26	257,69	352,42
o	571,19	1040,61	3278,82	4137,71
szórás	26,39	85,40	63,08	161,71
ó	428,46	730,24	2665,67	4077,60
szórás	16,24	73,27	172,37	105,90
u	411,75	739,13	2771,80	4135,46
szórás	16,70	67,13	162,88	82,03
ú	447,94	699,31	2954,74	4114,83
szórás	67,98	39,13	209,32	124,13
ö	577,78	1669,62	2906,89	4284,08
szórás	24,69	133,59	100,61	137,79
ő	395,29	1688,97	2905,38	4219,64
szórás	14,54	349,61	397,04	317,75
ü	388,06	2016,13	2695,78	3802,49
szórás	15,47	92,10	143,53	168,06
ű	381,92	2246,27	2871,20	4094,76
szórás	45,97	183,20	271,98	329,64
e	836,23	1964,85	3155,69	4274,65
szórás	50,11	90,99	96,58	210,34
é	413,48	2805,77	3355,92	4442,13
szórás	13,20	271,60	302,31	232,56
i	349,84	2739,24	3557,11	4228,86
szórás	22,10	235,61	337,83	125,13
í	297,98	2851,25	3653,84	4175,62
szórás	25,45	47,81	215,76	94,25

A.1. táblázat: Magánhangzók formánsai és szórásai saját hangfelvételeimben

	f_1 [Hz]	f_2 [Hz]	f_3 [Hz]
b	231,92	1629,1	2907,88
szórás	43,94	248,00	348,7
d	308,74	2142,5	3333,50
szórás	46,61	311,87	510,24
gy	253,55	2155,97	3203,78
szórás	68,99	268,47	317,19
g	416,06	1895,24	2613,40
szórás	25,32	107,82	77,2
j	351,6	2715,22	3331,36
szórás	27,67	93,17	189,49
v	258,25	1834,02	2888,12
szórás	76,05	451,95	193,83
z	327,90	1986,83	3208,74
szórás	52,30	206,15	280,01
zs	309,25	2474,37	3154,71
szórás	45,35	277,41	201,7
dz	374,24	2016,51	3004,22
szórás	35,19	168,097	142,75
dzs	381,12	2323,49	3136,86
szórás	29,57	159,34	161,82
r	597,82	1650,69	2412,67
szórás	68,66	179,46	338,11
m	323,023	2046,70	2960,69
szórás	37,28	287,46	461,94
n	365,80	2240,53	3134,18
szórás	11,14	172,95	321,39
ny	363,47	804,73	2445,99
szórás	17,66	88,44	137,93
l	397,27	1980,27	3065,60
szórás	46,06	122,01	193,58

A.2. táblázat: Zöngés mássalhangzók formánsértékei és szórásai saját hangfelvételeimben

2.1.2. Az egyes hangfelvevő eszközök összehasonlítására vonatkozó mérések

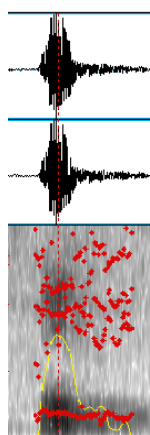
Arra is választ kerestem, hogy vajon másnak adódnak-e a formánsértékek a telefonos hangfelvételeknél, mint a számítógépeseknél. Ez a kérdés azért vetődik fel egyáltalán, mert a telefon hangrögzítésre is ugyanazt a csatornát használja, amit a telefonáláskori hangtovábbításra,

és ebben pedig jelentős jelfeldolgozási lépések vannak, pl spektrális tömörítési algoritmusok, frekvenciaszelektív szűrés. A kérdés megválaszolásához számítógéppel, egy Samsung S6802 Galaxy Ace és egy Sony Xperia E C1505 okostelefonnal 16 különböző korú és nemű alannal hangfelvételeket készítettem. Az alábbi táblázat az alanyok adatait tartalmazza.

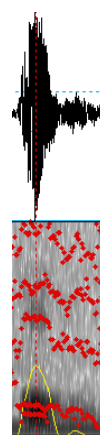
#	Nem	Kor [év]	Mag. [cm]	Tömeg [kg]	Születési hely	Dohányzik?
1	♂	36	169	75	Hódmezővásárhely, HU	nem
2	♀	20	165	62	Budapest, HU	igen
3	♀	19	168	58	Szekszárd, HU	nem
4	♂	22	183	80	Dunaújváros, HU	nem
5	♀	50	174	69	Szeged, HU	nem
6	♀	26	172	58	Budapest, HU	nem
7	♀	20	164	52	Budapest, HU	nem
8	♀	43	170	85	Szeged, HU	nem
9	♀	49	165	58	Szabadka, SRB	nem
10	♀	23	163	53	Szeged, HU	nem
11	♀	49	165	60	Kiskunhalas, HU	nem
12	♂	31	168	70	Cegléd, HU	nem
13	♂	18	180	80	Zenta, SRB	nem
14	♀	44	168	60	Zenta, SRB	igen
15	♂	47	179	90	Törökkanizsa, SRB	igen
16	♂	26	180	75	Győr, HU	nem

A mobiltelefonos hangfelvételek a Smart Voice Recorder telefonos alkalmazással készültek. A feladat a magyar hangok egymás utáni felsorolása, illetve a *rozs*, *baj*, *galád* szavak felmondása volt. Azért esett ezekre a szavakra a választásom, mert kizárólag zöngés mássalhang zókat tartalmaznak. Ez azért fontos, mert így van értelme a formánsátmenetek vizsgálatának. Minden alannal készülékenként 3 felvételt készítettünk. Később ezeket a felvételeket kielemeztem. Az összetartozó méréseknél, a különböző eszközöknél azonos időpillanatban listáztattam ki a formánsokat, így lehetőség nyílt az eltérések vizsgálatára. Minden hangot megvizsgáltam, így sikerült kimutatnom az eszközök közötti eltéréseket és egy kisebb adatbázist is sikerült létrehozni.

A tapasztalatom az, hogy a legtöbb esetben egyeznek a formánsértékek, azonban felfedeztünk egy igen érdekes jelenséget. A Praat program jellemzően az /o/, /ó/, /u/ és /ú/ hangoknál az alapbeállítás esetén (az alapbeállítás 5 formáns kilistázása) a számítógépes hangfelvételen a legalacsonyabb formánsokat (f_1 és f_2) egynek észleli, míg a telefonos hangfelvételeken elkülöníti.

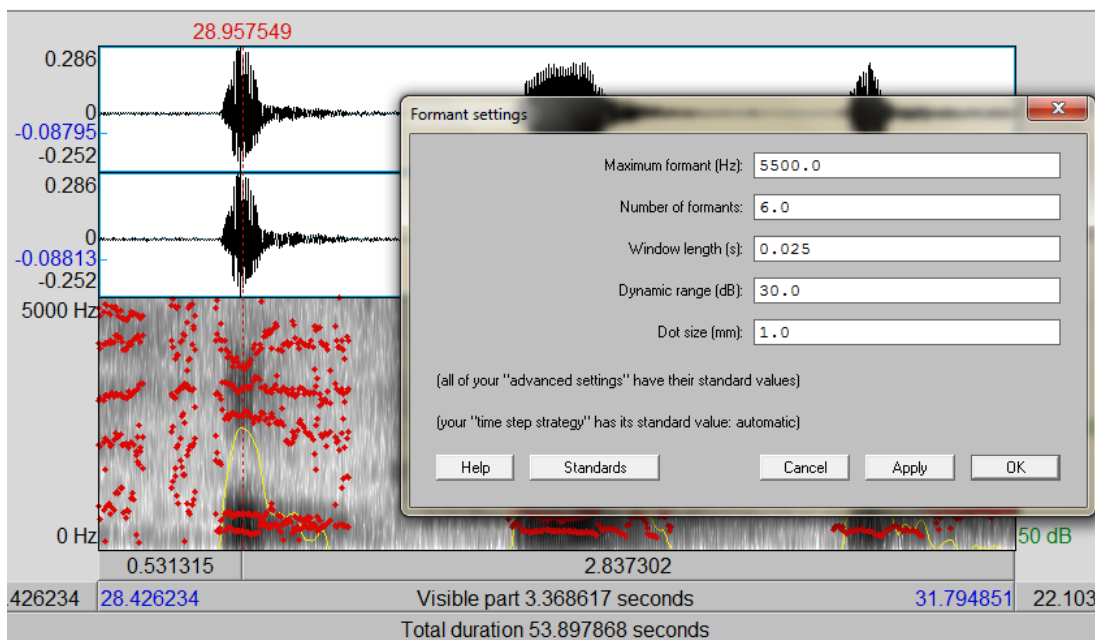


2.8. ábra. Az /o/ hang spektruma számítógépes hangfelvétel esetén



2.9. ábra. Az /o/ hang spektruma Samsung telefonos hangfelvétel esetén

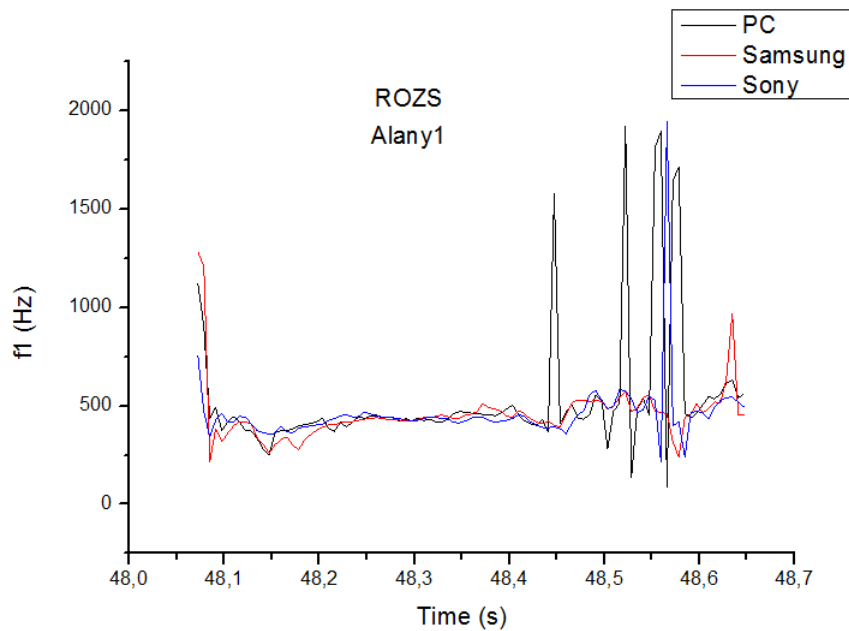
Azonban, ha megváltoztatjuk az alapbeállítást a számítógépes hangfelvétel elemzésekor, és 5 formáns helyett 6 formánst íratunk ki, akkor a számítógépes hangfelvétel spektrumában is elkülönül a két legalsó formáns.



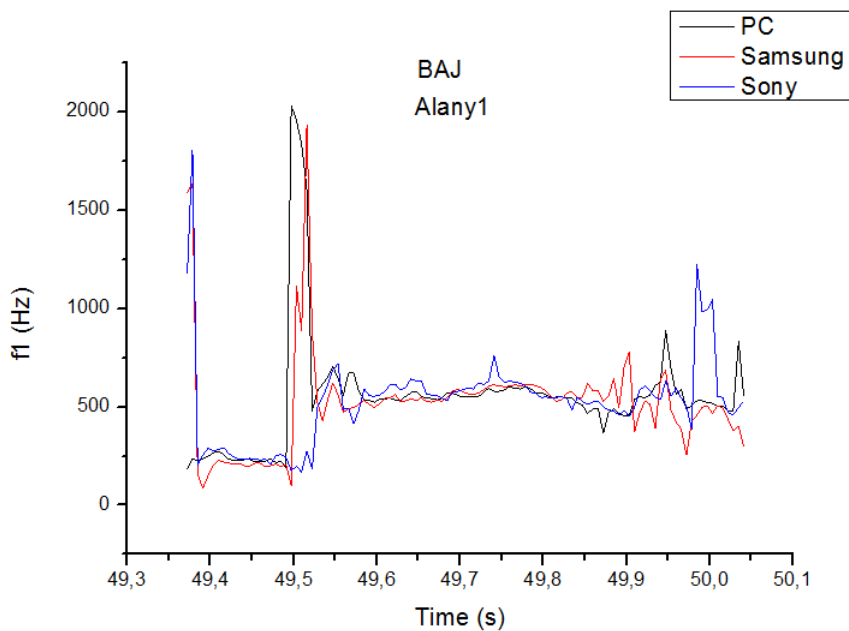
2.10. ábra. Az /o/ hang spektruma számítógépes hangfelvétel esetén, az alapbeállításokat megváltoztatva

Az (A.5) táblázat azokat a mérési eredményeket, ahol „formánselcsúszás” tapasztalható, szürke színnel kijelöltük. Megvizsgáltuk az /a/ hangra (Dr Nyári Tibor szíves segítségével), hogy mutatkozik-e szignifikáns különbség az egyes eszközök között (ismételt mérések varianciaanalízis, szignifikanciahatár, $\alpha = 0,05$). A tapasztalat az, hogy a szűretlen adatokon szignifikáns különbség mutatkozott az eszközök között, de a formánsok fölhasadásából adódó elcsúszott formánsértékek eltávolítása után nincs az egyes eszközök között szignifikáns különbség.

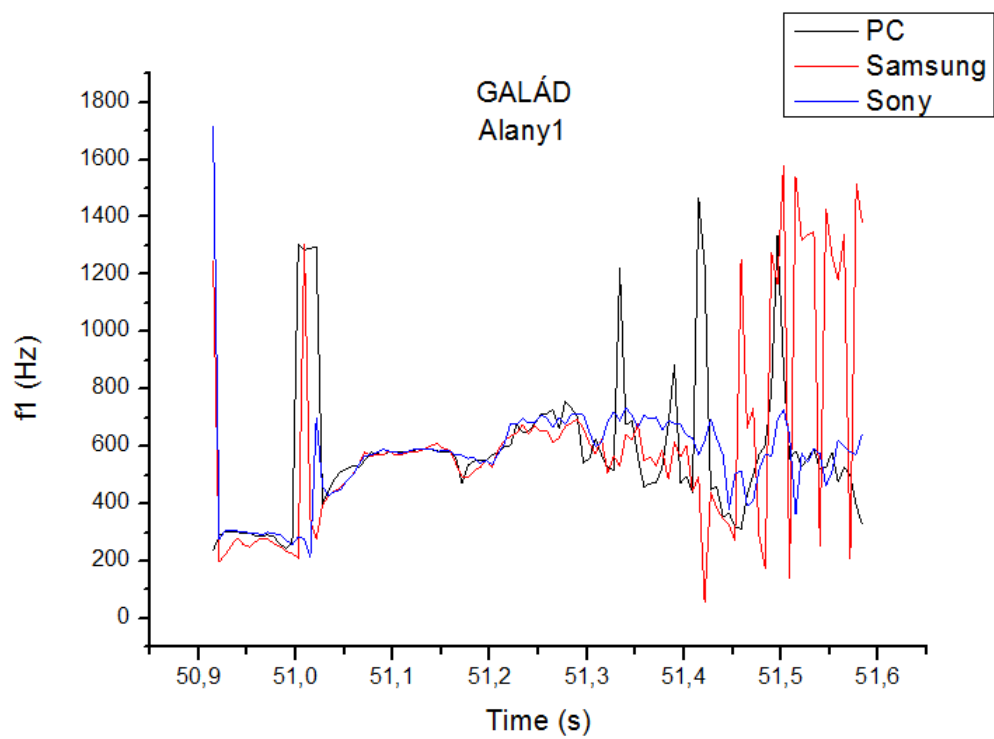
Amint azt már korábban láttuk, az artikulációs nehézségekkel küzdő betegek beszédmínőségének megállapításában nagy szerepet játszik a második formáns meredeksége a szóformálás folyamata során. Ezért a fent említett szavakban megvizsgáltuk, mennyire tér el az egyes eszközök esetén a második formáns időbeli változása. Ugyanazon a grafikonon ábrázoltam az azonos méréshez, különböző eszközhöz tartozó szavak formánsátmeneteit. Az alábbi hat grafikon az első alany hanganyagának elemzése.



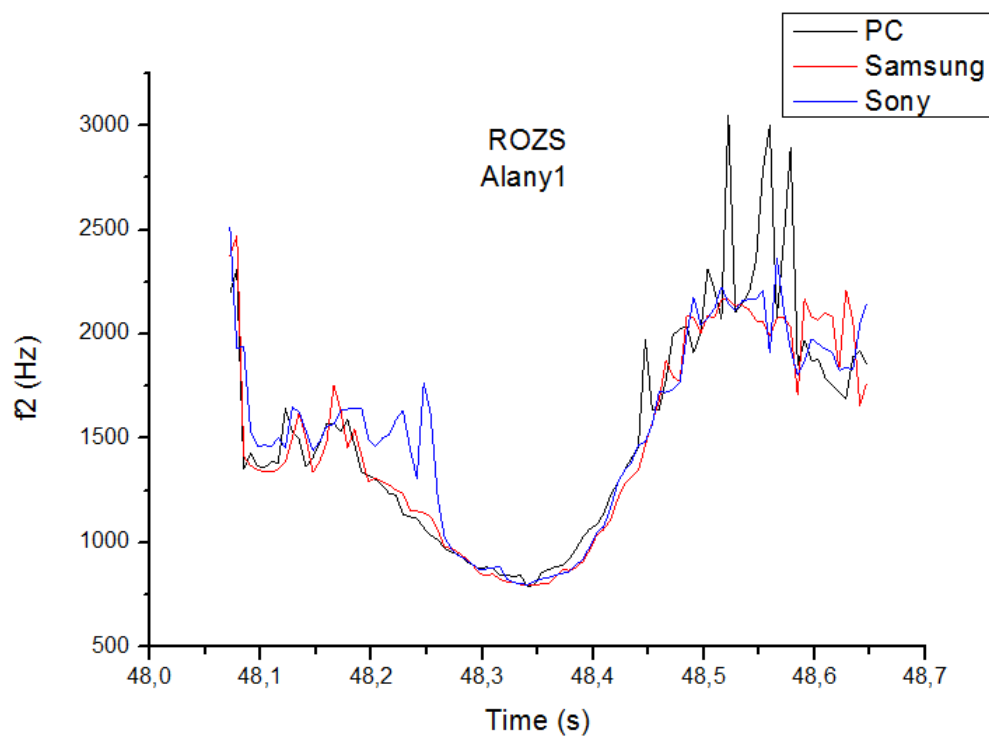
2.11. ábra. Az első formáns változásai a *rozs* szóban



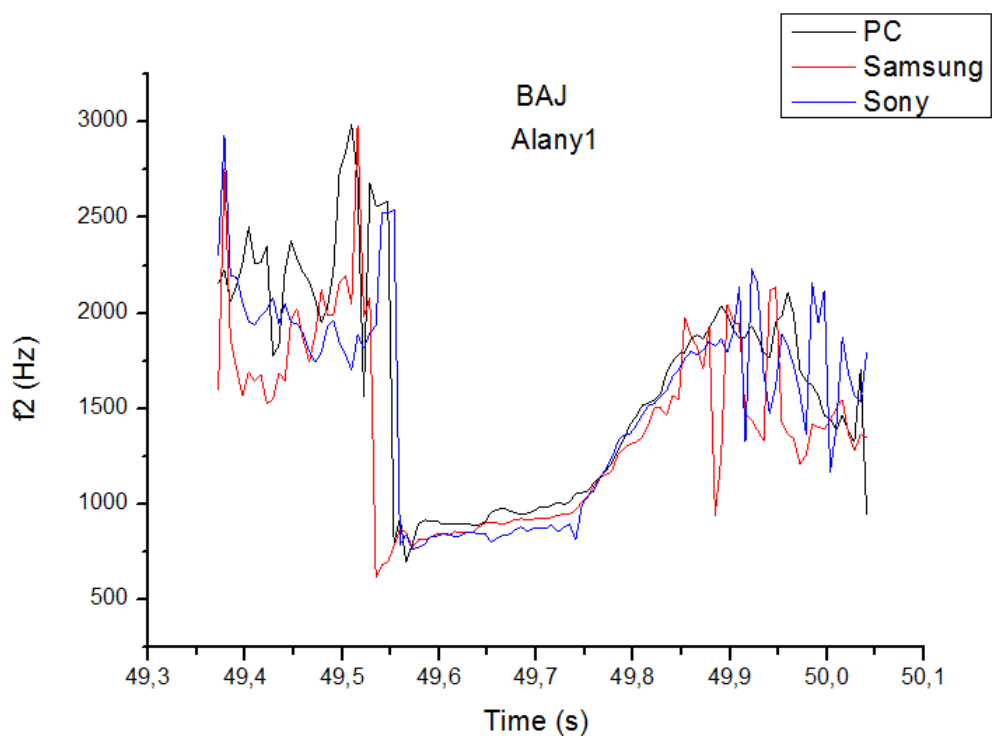
2.12. ábra. Az első formáns változásai a *baj* szóban



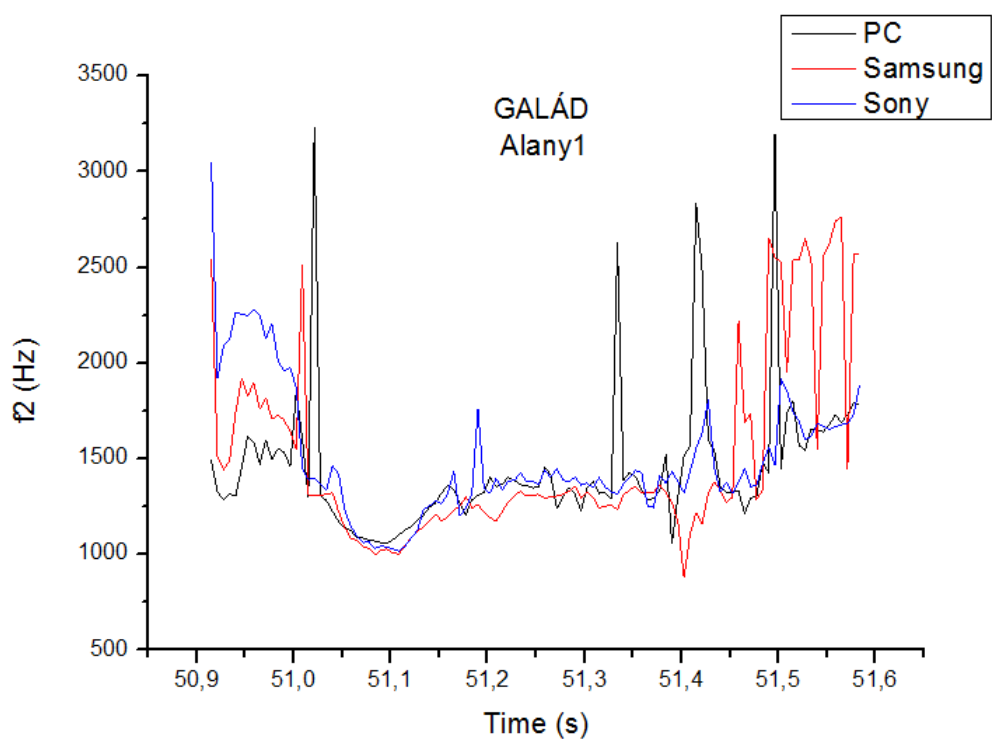
2.13. ábra. Az első formáns változásai a *galád* szóban



2.14. ábra. A második formáns változásai a *rozs* szóban



2.15. ábra. A második formáns változásai a *baj* szóban

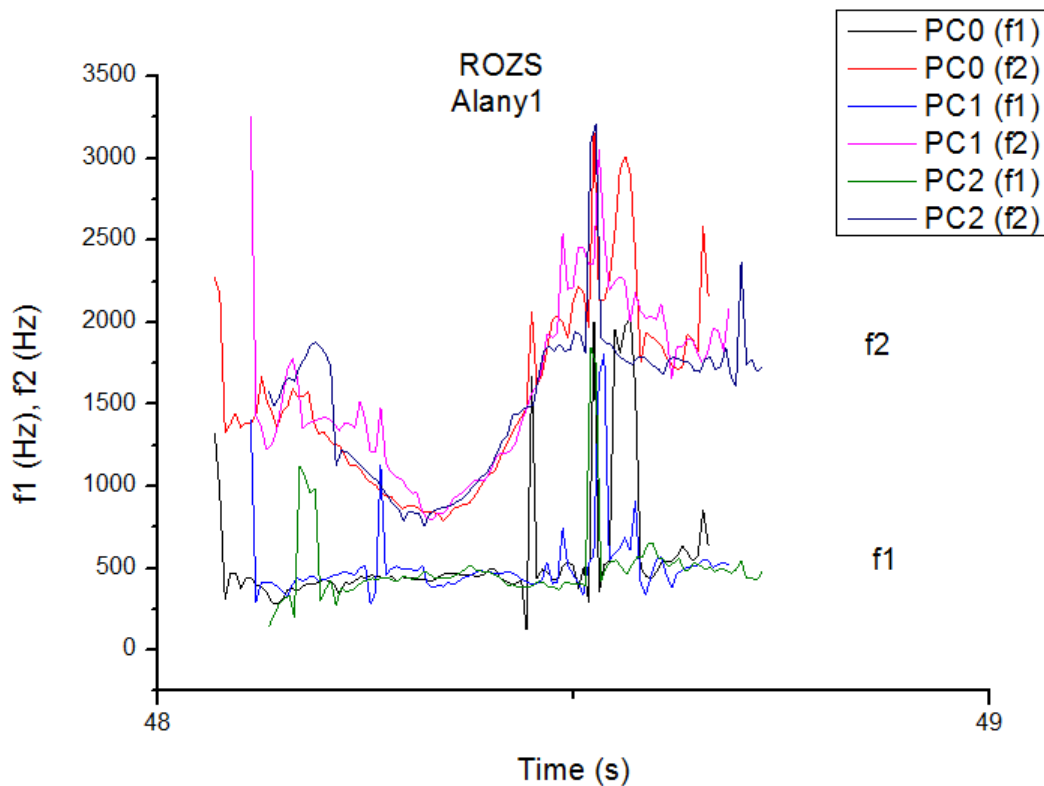


2.16. ábra. A második formáns változásai a *galád* szóban

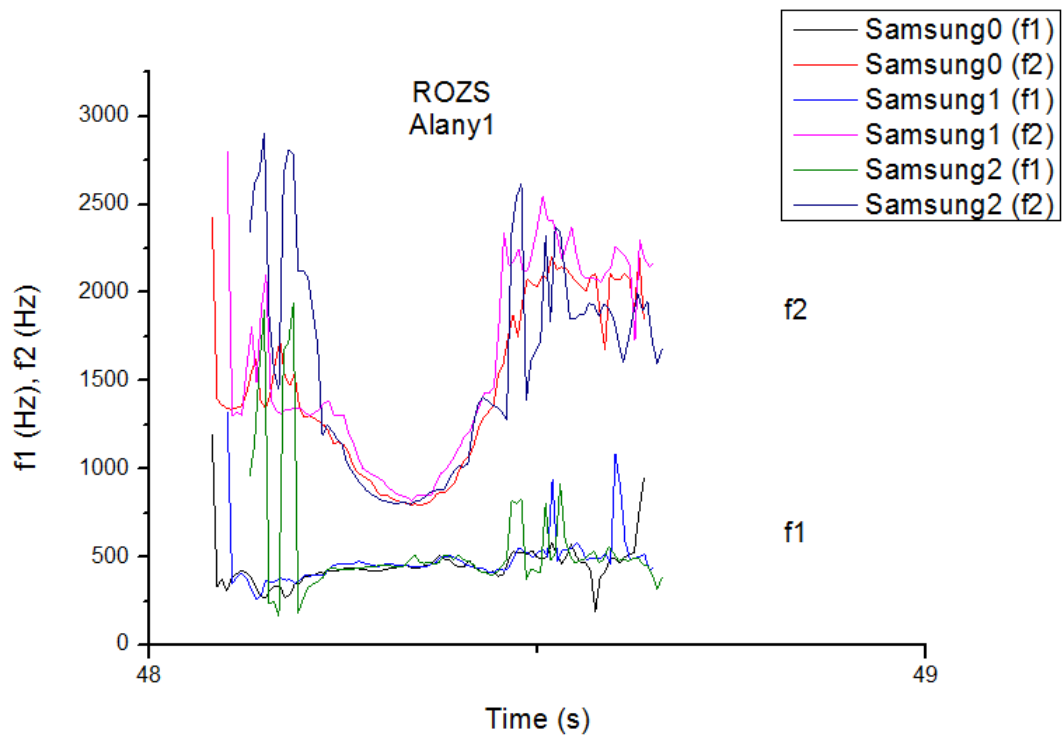
Látható, hogy a különböző eszközökhöz tartozó eredmények elég jól átfedik egymást, a szóra jellemző formánsminták szépen kirajzolódnak. A hangfelvételek zajossága, illetve az eszközök

különböző pozíciója miatt helyenként kiugró frekvenciaértékeket tapasztalhatunk.

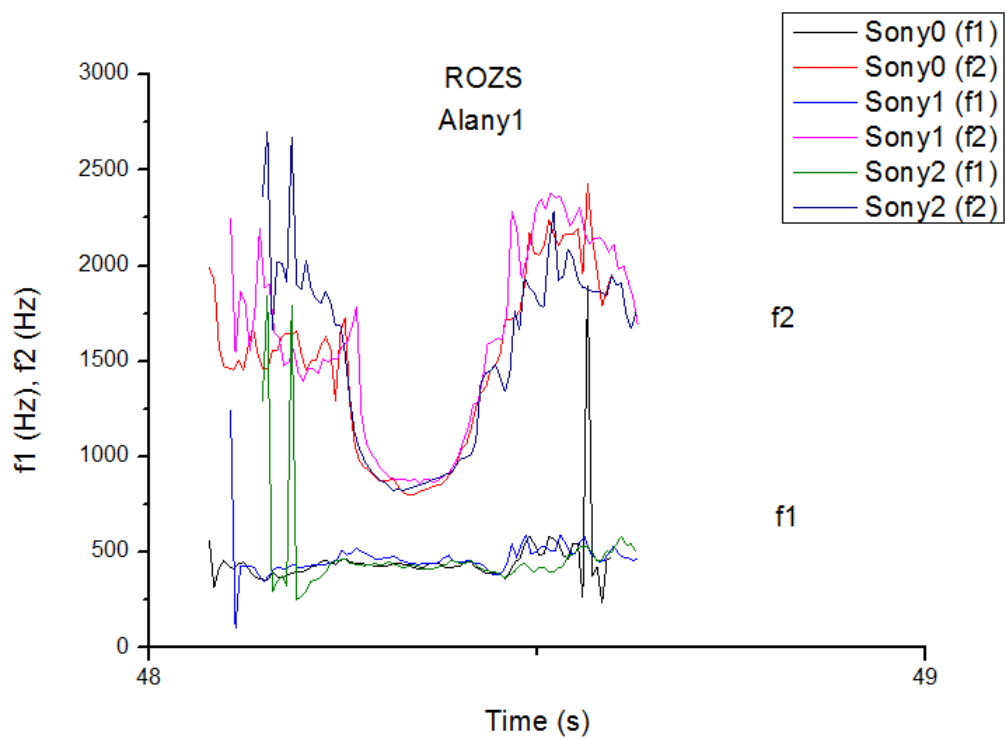
Az ugyanazon eszközön belüli szórás szemléltetéséhez az előzőekhez hasonló grafikono-
kat készítettem. Megvizsgáltam, hogy az alanyonkénti három mérés, azonos eszközt tekintve
mennyire ad különböző eredményt. Ezt az alábbi grafikonok szemléltetik.



2.17. ábra. A *rozs* szó három számítógépes hangfelvételének összehasonlítása



2.18. ábra. A *rozs* szó három samsung-os hangfelvételének összehasonlítása



2.19. ábra. Az azonos eszközzel végzett mérések összehasonlítása

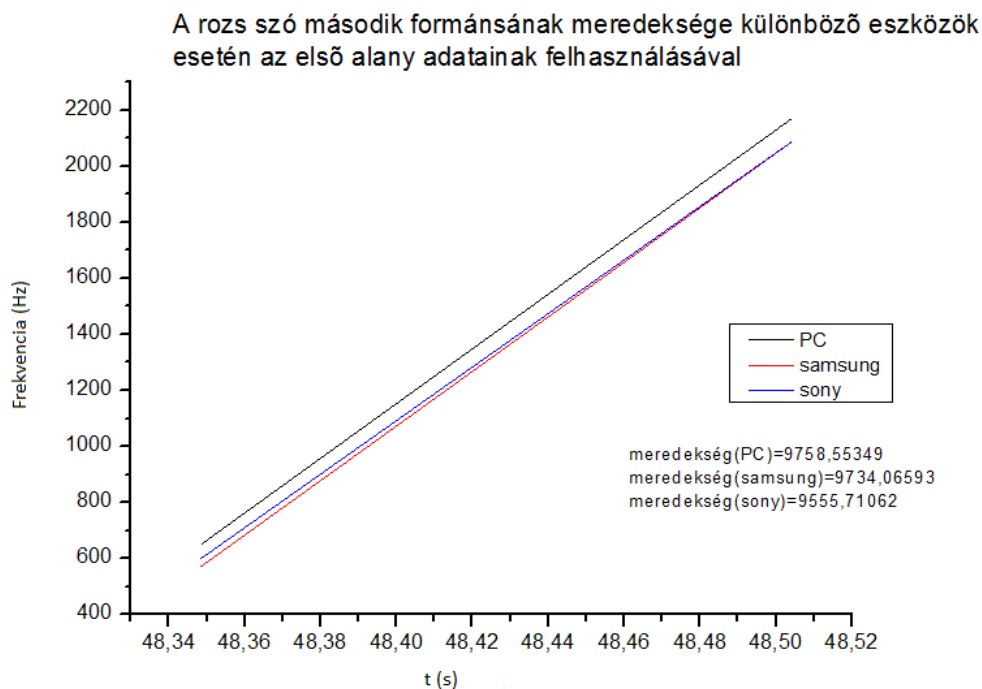
2.1.3. A második formáns meredekségének meghatározása

A meredekség számszerűsítésének és összehasonlíthatóságának érdekében az egyes szavakon belül a közel lineáris szakaszokra illesztettem egyenest. Ezeket a méréseket a rozs és baj szavakra végeztem el mivel a galád szóban nincs meredek váltás a második formánsban.

Az így kapott eredményeket az alábbi táblázatban foglaltam össze.

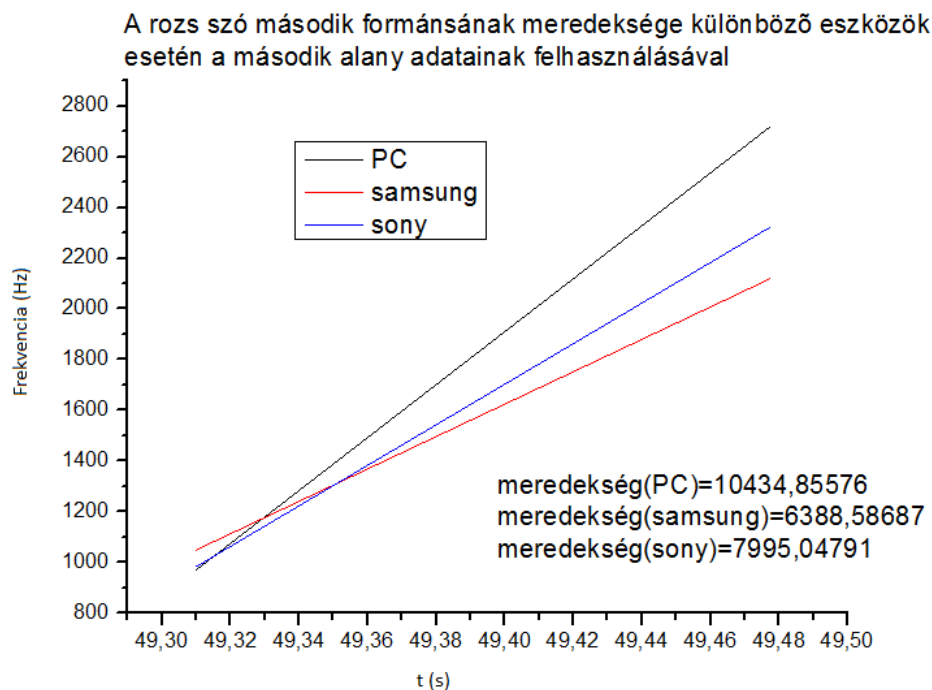
Alany	Szó	PC		Samsung		Sony		t_1 [s]	t_2 [s]
		Meredeks.	Tengelym.	Meredeks.	Tengelym.	Meredeks.	Tengelym.		
Alany1	rozs	9758,55	-471160,75	9734,07	-470055,76	9555,71	-461404,06	48,35	48,50
	baj	7080,02	-351176,03	5125,39	-253915,22	6399,93	-317338,62	49,75	49,89
Alany2	rozs	10434,86	-513572,07	6388,59	-313971,94	7995,05	-393252,20	49,31	49,48
	baj	7752,05	-390983,97	7916,87	-399391,09	7237,00	-364958,46	50,58	50,70
Alany3	rozs	7702,60	-342013,14	10833,66	-481614,82	7882,13	-349974,54	44,53	44,62
	baj	9498,04	-434762,34	13013,97	-596577,59	5429,64	-247842,01	45,90	46,00
Alany4	rozs	9647,33	-330398,73	8324,43	-284907,13	7201,91	-246330,10	34,32	34,40
	baj	7985,16	-281888,67	7177,53	-253257,41	5968,40	-210345,56	35,42	35,56
Alany5	rozs	9386,65	-335716,93	9930,18	-355295,78	9750,77	-348826,37	35,85	35,94
	baj	6378,72	-234645,06	8190,41	-301684,21	5874,66	-216029,60	36,92	37,11
Alany6	rozs	17491,93	-657317,50	15632,83	-587402,91	10551,30	-395942,07	37,62	37,71
	baj	8610,43	-333726,73	7385,40	-286234,68	7535,90	-291862,13	38,86	39,98
Alany7	rozs	7199,79	-279322,83	10116,14	-393016,17	10506,55	-408307,14	38,96	39,03
	baj	5613,74	-223049,34	7270,33	-289253,75	8798,91	-350465,80	39,95	40,05
Alany8	rozs	8330,94	-311029,89	7868,99	-293840,59	13988,04	-523186,56	37,45	37,58
	baj	13056,35	-501575,86	9511,14	-365118,33	8272,09	-317512,07	38,49	38,54
Alany9	rozs	8478,81	-242570,86	8913,41	-255130,94	9176,22	-262640,42	28,71	28,81
	baj	8964,42	-265261,48	11108,64	-329105,98	10172,95	-301192,88	29,71	29,82
Alany10	rozs	6386,64	-211033,43	5332,68	-176221,37	9817,35	-325045,32	33,21	33,26
	baj	6563,27	-223105,41	15851,95	-540901,49	8117,25	-276122,46	34,16	34,25
Alany11	rozs	8212,50	-326440,80	12404,57	-493605,99	16397,22	-652943,18	39,84	39,93
	baj	14257,21	-581407,02	6236,46	-253068,67	13914,23	-566942,08	40,80	40,93
Alany12	rozs	25805,99	-1151110,00	14793,32	-659553,28	12975,40	-578379,77	44,63	44,69
	baj	10967,55	-508418,34	11335,10	-525566,02	11491,53	-532835,41	46,46	46,55
Alany13	rozs	nagyon kusza a grafikon, nincs lineáris szakasz							
	baj	nagyon kusza a grafikon, nincs lineáris szakasz							
Alany14	rozs	8642,66	-369408,89	6001,53	-256241,73	9212,46	-393927,19	42,85	42,92
	baj	6687,73	-297948,55	7746,22	-345450,54	7968,28	-355406,11	44,69	44,88
Alany15	rozs	nagyon kusza a grafikon, nincs lineáris szakasz							
	baj	nagyon kusza a grafikon, nincs lineáris szakasz							
Alany16	rozs	9334,91	-471408,66	7922,54	-399954,47	8294,43	-418759,25	50,59	50,66
	baj	15794,54	-819791,44	10160,60	-527044,02	12410,49	-643947,51	51,94	52,01

Egyes eredmények azt mutatják, hogy a különböző eszközök között nincs nagy eltérés, például az első alany adatai alapján kapott meredekség, amit az alábbi ábrán láthatunk.



2.20. ábra. Meredekségek összehasonlítása

Vannak azonban olyan esetek is, amikor az egyes eszközökkel kapott eredmények lényegesebb eltérést mutatnak, például a második alany esetén a rozs szóban, ahogy azt az alábbi ábrán is láthatjuk.



2.21. ábra. Meredekségek összehasonlítása

Összefoglaló

Első lépésben megismerkedtem a Praat nevű programmal, melynek segítségével először a saját hangomról készített hangfelvételek reprodukálhatóságát vizsgáltam, és összevetettem a kapott eredményeket az irodalmiakkal. Azt kaptam, hogy az eredmények viszonylag kis szórással reprodukálhatók, és összevethetők az irodalmi értékekkel. Eltérések természetesen tapasztalhatók, de ez nem nagy probléma, hiszen az irodalmi adatok csak a populációra jellemző átlagos frekvenciatartományt adják meg, amelytől egyedi eltérések lehetnek.

A feladatom ezután az volt, hogy összehasonlítsam a számítógépes és a mobiltelefonos hangfelvételeket, ezek eltérésének a mértékét megállapítsam. A tapasztalatom az, hogy ha kiszűrjük azokat az eseteket, amikor a program valamiért egyes hangoknál a legalsó formánsokat összevonja, tehát a különböző eszközök hangfelvételeinek spektrumát nem egységesen kezeli, akkor nincs szignifikáns különbség a számítógéppel kapott formánsok és a telefonos formánsok között.

A későbbi alkalmazások szempontjából kiemelkedő jelentőségű a második formáns meredeksége mint beszédminőséget jellemző, a beszédnehézségekkel korreláló paraméter. Ezt megvizsgálva azt tapasztaltuk, hogy a mobiltelefonos felvételek összehasonlíthatók a PC-sekkel mind az egyes hangok, mind a meredekségek esetén, tehát használhatók az alkalmazásokban.

A konklúzió mindezek alapján tehát egyrészt, hogy a mobileszközök is jól használhatók a beszédterápiás alkalmazásokban, másrészt azonban további feladatunk a formánselcsúszások okainak feltérképezése és megelőzése.

Függelék

Hang	#	f_1 [Hz]			f_2 [Hz]			f_3 [Hz]			f_4 [Hz]		
		PC	Samsung	Sony	PC	Samsung	Sony	PC	Samsung	Sony	PC	Samsung	Sony
A	1	625,64	638,14	489,87	2241,44	673,50	647,24	3448,57	2734,51	2792,94	?	3563,16	3632,12
	2	538,61	652,23	681,55	2847,53	2724,11	2739,69	3510,77	3719,13	3700,25	4099,97	?	?
	3	525,91	562,94	539,12	2627,17	2614,20	646,95	3564,05	3711,83	2535,69	3780,62	4800,68	3650,41
	Átlag	563,39	617,77	570,18	2572,05	2003,94	1344,63	3507,80	3388,49	3009,63	3940,29	4181,92	3641,26
Á	1	856,49	741,61	796,42	1286,52	1275,06	1342,44	2726,76	2646,23	2739,73	3822,58	3926,79	3984,24
	2	786,03	728,39	750,23	1261,84	1228,23	1245,97	2787,12	2709,64	2683,12	3736,11	3737,22	3896,13
	3	778,98	731,00	749,15	1267,85	1316,87	1225,77	2852,60	2804,42	2693,80	3994,40	3616,65	3986,33
	Átlag	807,17	733,67	765,27	1272,07	1273,39	1271,39	2788,83	2720,10	2705,55	3851,03	3760,22	3955,57
B	1	136,75	1613,10	1119,92	2065,01	2800,36	2353,42	3090,26	3932,94	3741,68	4177,41	4926,10	4834,87
	2	238,26	201,71	277,27	2211,92	1710,45	1864,39	3086,85	3088,88	2873,07	4330,41	3408,34	3984,45
	3	222,11	187,87	241,83	1306,11	1875,93	1781,80	2781,61	2783,43	2974,32	4051,87	4047,07	4216,44
	Átlag	199,04	667,56	546,34	1861,01	2128,92	1999,87	2986,24	3268,42	3196,36	4186,57	4127,17	4345,25
D	1	215,55	201,72	254,05	1794,13	1653,61	1736,39	2841,66	2685,93	2887,44	4118,86	3995,22	3879,77
	2	230,15	94,44	236,36	1463,36	1602,98	1883,62	2583,23	2748,58	2775,97	4007,63	4076,55	3829,18
	3	222,69	128,34	234,36	1641,58	1502,14	1905,33	2854,95	2656,83	3127,41	4048,50	3861,34	4459,10
	Átlag	222,79	141,50	241,59	1633,02	1586,24	1841,78	2759,95	2697,11	2930,27	4058,33	3977,70	4056,01
DZ	1	219,36	155,41	233,64	1480,14	1689,72	1767,02	2862,42	2953,58	2757,00	3965,15	3679,24	3777,11
	2	229,73	225,72	266,30	2182,32	1871,68	1818,56	3223,39	2774,60	3068,97	4189,04	3662,42	4019,26
	3	386,51	420,06	400,13	1561,52	1519,02	1552,10	2941,72	2633,32	2633,28	3831,52	3836,19	3634,67
	Átlag	278,54	267,06	300,02	1741,33	1693,47	1712,56	3009,18	2787,17	2819,75	3995,24	3725,95	3810,35
DZS	1	233,12	201,96	254,65	1474,24	1440,45	1542,76	3034,59	2581,10	2776,24	4072,15	3903,44	3915,70
	2	256,85	231,78	275,18	2237,13	1818,12	1899,01	3624,39	2970,42	3165,73	4320,21	4264,61	4049,03
	3	270,66	246,30	288,75	2458,04	1939,41	1967,95	2863,38	2688,97	3098,80	4430,09	3352,49	4030,88
	Átlag	253,54	226,68	272,86	2056,47	1732,66	1803,24	3174,12	2746,83	3013,59	4274,15	3840,18	3998,53
E	1	517,55	553,08	541,25	1906,24	1945,60	1923,06	2667,35	2630,26	2605,84	3640,13	3947,18	3557,36
	2	507,36	557,15	531,76	1872,30	1930,38	1888,97	2703,93	2805,32	2556,90	3661,70	4628,60	3294,25
	3	518,09	561,74	524,99	1860,81	1953,82	1916,84	2909,01	2828,12	2727,41	3489,47	4441,36	3197,58
	Átlag	514,34	557,32	532,67	1879,78	1943,27	1909,62	2760,10	2754,57	2630,05	3597,10	4339,05	3349,73
É	1	324,15	332,19	320,00	2267,86	2204,17	2219,77	2657,54	2566,85	2561,24	3519,95	3433,92	3617,70
	2	330,79	319,96	358,24	2111,07	2045,93	2023,58	2884,52	2607,57	2555,80	3538,59	3291,49	3622,78
	3	319,20	331,77	329,69	2161,93	2118,46	2148,19	2805,51	2603,58	2538,83	3588,38	3064,20	3365,74
	Átlag	324,71	327,97	335,97	2180,28	2122,85	2130,51	2782,52	2592,67	2551,96	3548,97	3263,20	3535,41
G	1	231,85	205,31	277,55	1413,17	1382,17	1659,60	2815,09	2519,05	2645,78	3892,39	3818,92	3882,52
	2	228,96	152,08	238,04	2115,04	1829,35	1949,55	2930,47	3085,38	2875,81	4272,50	4369,67	3881,63
	3	238,33	147,44	252,24	2106,86	1693,65	1939,66	2526,91	2735,73	2713,15	3650,59	4003,25	3738,91
	Átlag	233,05	168,27	255,94	1878,35	1635,06	1849,60	2757,49	2780,05	2744,91	3938,49	4063,95	3817,69
GY	1	241,99	190,69	283,64	1668,24	1614,51	1824,07	2625,44	2544,18	2665,34	3771,37	3782,12	3935,14
	2	220,23	194,97	230,60	1977,06	1786,26	2104,25	3194,02	3080,25	2571,95	4188,15	4399,60	3713,86
	3	234,75	122,98	259,84	2335,88	1648,34	2099,56	2877,83	2832,39	2714,07	3896,71	3835,86	3876,29
	Átlag	232,32	169,54	258,03	1993,73	1683,04	2009,29	2899,10	2818,94	2650,45	3952,08	4005,86	3841,76
I	1	292,04	254,17	272,71	2291,42	2200,50	2225,07	2931,17	2843,04	2894,75	3770,86	3725,98	3868,89
	2	269,31	259,22	278,63	2321,78	2231,85	2234,09	3015,54	2867,60	2866,60	3674,17	3543,16	3674,79
	3	293,33	268,42	293,67	2307,58	2305,12	2224,26	2917,08	2816,02	2788,85	3902,17	3686,01	3753,49
	Átlag	284,89	260,60	281,67	2306,93	2245,82	2227,81	2954,60	2842,22	2850,07	3782,40	3651,72	3765,72
Í	1	270,41	288,40	284,21	2560,33	2296,17	2297,28	3005,04	2915,22	2979,99	3719,11	3776,64	3763,46
	2	261,51	248,29	275,21	2457,34	2302,38	2295,22	2963,71	2956,12	3007,41	3632,85	2995,14	3517,28
	3	261,83	248,22	273,36	2369,45	2275,69	2256,43	2961,70	2901,61	2929,91	3767,39	3070,10	3672,33
	Átlag	264,58	261,64	277,59	2462,37	2291,41	2282,97	2976,82	2924,31	2972,43	3706,45	3280,63	3651,02
J	1	266,35	256,10	291,74	2167,78	2074,07	2099,30	3124,53	2987,05	3072,81	3695,80	4094,73	3866,39
	2	264,16	246,68	269,96	2461,77	2158,61	2205,98	3053,53	2997,09	3084,32	3749,62	3902,85	3898,82
	3	273,82	242,45	276,08	2650,69	2165,87	2231,64	2909,50	2907,41	2945,71	3618,44	3550,02	3728,13
	Átlag	268,11	248,41	279,26	2426,75	2132,85	2178,97	3029,19	2963,85	3034,28	3687,95	3849,20	3831,11
L	1	347,34	347,06	372,12	1470,43	1494,44	1508,28	2695,81	2601,33	2676,04	3860,68	3948,09	3899,09
	2	382,85	380,24	394,68	1290,99	1294,57	1310,37	2702,39	2682,77	2531,20	3658,23	3598,19	3767,77
	3	355,76	370,14	364,56	1411,78	1358,96	1456,18	2764,41	2629,32	2575,51	3481,32	3328,71	3511,70
	Átlag	361,98	365,81	377,12	1391,07	1382,66	1424,94	2720,87	2637,81	2594,25	3666,74	3625,00	3726,19
M	1	264,40	280,16	275,87	1296,36	1376,29	1505,66	2372,75	2380,91	2348,74	4207,50	3776,11	3899,74
	2	216,69	269,86	270,70	1293,11	1508,43	1456,99	2307,66	2310,28	2313,28	3613,54	3479,29	3829,59
	3	243,95	331,06	259,91	1407,20	1954,22	1785,14	2326,26	2347,25	2311,11	3775,34	4380,38	4263,24
	Átlag	241,68	293,69	268,82	1332,22	1612,98	1582,60	2335,56	2346,15	2324,38	3865,46	3878,59	3997,52
N	1	235,37	262,46	296,13	1175,95	1270,76	1498,00	2444,48	2474,08	2489,20	3551,20	3824,25	3604,86
	2	248,70	244,82	274,17	1220,85	1437,99	1376,67	2405,56	2371,41	2288,56	3410,65	3756,34	3651,07
	3	253,65	262,98	273,90	1315,10	2157,09	2254,82	2544,79	2528,32	2465,15	3493,45	4005,28	4396,14
	Átlag	245,91	256,75	281,40	1237,30	1621,95	1709,83	2464,95	2457,94	2414,30	3485,10	3861,95	3884,03
NY	1	259,35	283,68	305,65	1215,55	1452,34	1888,57	2355,87	2222,37	2472,40	3491,93	3716,33	3583,33
	2	250,85	287,75	304,94	1248,14	1954,11	1870,48	2262,40	2240,47	2286,52	3303,68	3301,59	3616,34
	3	268,48	313,15	313,42	1191,13	1636,48	1674,40	2241,26	2170,01	2233,39	3226,30	2576,74	3570,68
	Átlag	259,56	294,86	308,00	1218,27	1680,98	1811,15	2286,51	2210,95	2330,77	3340,64	3198,22	3590,12
O	1	603,02	489,67	487,50	2343,50	706,11	615,17	3303,66	2703,36	2769,43	3823,33	3582,11	3580,01

A.5. táblázat: Az első alany telefonos és számítógépes hangfelvételeinek elemzése és a formások összehasonlítása

Hang	#	f_1 [Hz]			f_2 [Hz]			f_3 [Hz]			f_4 [Hz]		
	2	501,89	431,86	370,01	2924,25	587,92	469,04	3333,11	2831,06	2834,95	?	3181,62	3363,50
	3	515,42	420,45	450,62	2856,90	610,30	2056,13	3441,69	2764,48	3352,73	5189,78	3327,89	4486,40
	Átlag	540,11	447,33	436,04	2708,22	634,77	1046,78	3359,49	2766,30	2985,70	4506,55	3363,87	3809,97
Ó	1	444,21	344,22	418,91	2735,09	626,53	622,19	3455,33	2635,59	2740,34	5294,61	3619,29	3518,55
	2	471,03	381,37	489,76	2823,96	681,97	542,67	3410,70	2736,65	2755,92	5211,42	3406,20	3656,19
	3	528,41	308,63	539,55	2994,72	718,58	2822,22	3064,24	2827,14	3400,80	3425,88	3243,33	4930,87
	Átlag	481,21	344,74	482,74	2851,26	675,70	1329,03	3310,09	2733,12	2965,69	4643,97	3422,94	4035,20
Ö	1	384,95	389,86	362,99	2022,88	1975,19	1889,21	2593,04	2318,06	2325,88	3827,99	3505,55	3559,86
	2	384,36	421,73	402,43	1867,56	2141,55	2049,44	3013,57	2910,98	2269,97	3658,86	3180,83	3350,54
	3	331,86	389,76	363,24	1774,83	1865,26	2136,68	2976,80	2859,70	2376,04	3475,33	3270,01	3087,91
	Átlag	367,06	400,45	376,22	1888,42	1994,00	2025,11	2861,14	2696,24	2323,96	3654,06	3318,80	3332,77
Ő	1	320,97	356,92	360,91	1959,64	2109,02	2161,70	2815,47	2851,45	2332,39	3235,09	3486,71	3177,22
	2	310,22	329,32	338,55	1829,71	1872,11	1842,31	2945,95	2325,64	2275,96	3441,67	3186,72	3456,76
	3	342,70	363,71	361,52	1784,74	1834,82	1804,57	3217,73	2648,19	2633,23	3542,06	3167,04	3780,96
	Átlag	324,63	349,98	353,66	1858,03	1938,65	1936,19	2993,05	2608,43	2413,86	3406,28	3280,15	3471,65
R	1	574,71	512,14	551,54	1441,27	1306,62	1480,00	2962,42	2398,83	2333,97	3777,65	3529,26	3920,40
	2	596,58	479,26	486,34	1390,03	1342,61	1468,88	2093,44	2392,67	2131,80	3532,80	3638,94	3722,06
	3	551,33	492,91	512,28	1418,32	1411,27	1624,25	3225,64	2560,90	2809,08	4150,21	3116,35	3324,46
	Átlag	574,21	494,77	516,72	1416,54	1353,50	1524,38	2760,50	2450,80	2424,95	3820,22	3428,18	3655,64
U	1	490,89	623,39	586,38	2359,18	2233,80	2475,05	3288,48	3379,38	3492,83	3431,53	4692,92	4889,07
	2	393,70	521,28	396,15	2475,67	2452,30	2363,46	3285,44	3581,82	3334,67	4572,13	4770,03	4648,78
	3	247,74	341,78	492,26	1476,37	1534,15	2648,62	3103,22	2999,19	3726,44	4043,76	4055,39	4923,82
	Átlag	377,45	495,48	491,60	2103,74	2073,42	2495,71	3225,72	3320,13	3517,98	4015,81	4506,11	4820,56
Ű	1	334,17	614,07	518,15	509,64	2539,64	2831,30	2557,72	3404,88	3712,16	3350,32	4927,53	4942,00
	2	430,17	411,71	416,67	2458,82	1719,14	2436,60	3393,08	3114,21	3800,09	5167,55	4434,23	5142,31
	3	360,08	331,93	532,22	1837,35	1248,37	2356,69	3173,42	2824,77	3465,00	4540,81	4059,90	4707,63
	Átlag	374,81	452,57	489,01	1601,94	1835,72	2541,53	3041,41	3114,62	3659,08	4352,89	4473,89	4930,65
Ü	1	291,85	263,37	293,69	1963,90	2032,31	1981,69	2531,53	2233,90	2309,02	3372,62	3260,35	3420,48
	2	267,43	238,79	256,40	1946,50	1974,86	1951,91	2891,83	2372,22	2391,07	3376,74	3215,56	3374,34
	3	281,86	268,00	278,97	1916,40	2058,37	2067,51	2925,02	2206,64	2592,01	3311,49	3055,18	3229,01
	Átlag	280,38	256,72	276,35	1942,27	2021,85	2000,37	2782,79	2270,92	2430,70	3353,62	3177,03	3341,27
Ű	1	263,97	289,42	271,69	1991,67	2025,95	1978,66	2971,65	2538,78	2756,08	3511,52	3308,83	3322,18
	2	271,39	246,07	275,21	1951,25	2060,61	2058,63	2880,38	2882,74	2627,79	3279,53	3196,25	3381,68
	3	257,76	243,77	269,38	1919,58	2042,39	1995,86	2747,72	2722,94	2728,08	3226,14	3010,91	3464,57
	Átlag	264,37	259,75	272,10	1954,17	2042,98	2011,05	2866,58	2714,82	2703,98	3339,07	3172,00	3389,48
V	1	297,57	552,39	390,52	1442,38	1959,27	1795,65	2901,71	3077,41	3118,24	3751,54	4416,62	4502,42
	2	266,45	761,26	292,98	1311,69	2114,05	1636,62	2887,77	3538,10	2891,95	3803,30	4911,62	4030,61
	3	381,19	287,75	312,53	1757,44	1380,54	1991,79	2525,15	2536,70	2274,80	3705,77	3641,91	3892,10
	Átlag	315,07	533,80	332,01	1503,83	1817,96	1808,02	2771,54	3050,74	2761,67	3753,54	4323,38	4141,71
Z	1	273,87	244,57	284,88	1752,99	1740,99	1941,92	3669,44	2900,82	3300,02	4365,58	4390,33	4191,68
	2	269,79	260,06	276,12	2342,97	2079,99	2349,12	3614,47	2539,28	2465,39	4169,65	4208,76	4029,14
	3	333,42	112,17	282,44	2007,48	1644,40	1894,22	3366,28	2744,06	2859,55	4515,67	4609,20	4408,73
	Átlag	292,36	205,60	281,15	2034,48	1821,79	2061,75	3550,06	2728,05	2874,99	4350,30	4402,76	4209,85
ZS	1	290,58	300,26	323,56	2416,89	2456,68	2232,79	2954,43	2618,35	2529,58	4251,71	3351,52	3960,19
	2	237,97	259,65	318,38	2530,47	2267,64	2227,44	2828,54	2768,72	2703,34	4137,66	4590,91	4169,49
	3	288,47	274,52	329,19	2327,21	2246,27	2217,32	2976,46	2675,24	2592,18	4455,40	3324,84	4046,16
	Átlag	272,34	278,14	323,71	2424,85	2323,53	2225,85	2919,81	2687,44	2608,37	4281,59	3755,76	4058,61

A.5. táblázat: Az első alany telefonos és számítógépes hangfelvételeinek elemzése és a formánsok összehasonlítása

Irodalomjegyzék

- [1] <http://hu.wikipedia.org/wiki/Hang>
- [2] R. dr. Molnár Emma, *Leíró magyar hangtan*, Tankönyvkiadó, Budapest, 1989
- [3] <http://www.ngkszki.hu/seged/csej-lyesj/hangkepzh.htm>
- [4] <http://anatomia.uw.hu/ora-047/ora-047.htm>
- [5] Gósy Mária, *Fonetika, a beszéd tudománya*, Osiris Kiadó, Budapest, 2004
- [6] <http://magyarbeszed.tmit.bme.hu/formans/index.php>
- [7] R. D. Kent, J. F. Kent, G. Weismer, R. E. Martin, 'Relationship between speech intelligibility and the slope of second-formant transitions in dysarthric subjects', *CLINICAL LINGUISTICS & PHONETICS*, **1989**, VOL.3, NO. 4
- [8] Yunjing Kim, Gary Weismer, Raymond D. Kent, Joseph R. Duffy, Statistical Models of F2 Slope in Relation to Severity of Dysarthria', *Folia Phoniatr Logop*, **2009** , 61:329-335);