

## Adalékok a magyar nyelv hangtani szerkezetéhez

(Statisztikai vizsgálódások)

A Nyelvtudományi Közlemények LIII. kötetében (107—152) megjelent TARNÓCZY TAMÁS tanulmánya a magyar hangzókapcsolatokról ADY költői nyelve alapján. Felmerül a kérdés, költő, méghozzá mindig, még elbeszélő költeményeiben is (l. Margita élni akar) első személyben beszélő lírai költő nyelve alkalmas-e arra, hogy ennek alapján a magyar nyelv hangzókapcsolataira következtethessünk. Hiszen közismert az a tény, hogy a költő a rím miatt és különböző stílushatások elérése végett, mint MUKAŘOVSKY mondja: „peut changer les relations numériques des phonèmes, en se servant plus fréquemment ou au contraire moins fréquemment que l'on ne le fait dans l'usage normal de certains d'entre eux. Il peut ainsi déformer p. ex. la relation numérique entre les consonnantes et les voyelles, entre les voyelles longues et brèves etc.“ (TCLP. IV, 280.) És éppen ezért „il est donc important... de faire un tableau statistique de la fréquence des phonèmes et des séries de phonèmes *en comparaison*<sup>1</sup> de leur fréquence normale“ (l. c. 280.). De nemcsak hangzóeloszlás, hanem hangzócsoportok szempontjából is más a költő nyelve, mint a próza vagy a mindennapi beszéd nyelve. „Le groupement des phonèmes, surtout consonnantiques, offre au poète des possibilités très variées. Ce groupement est régi, dans une langue donnée, par des principes généraux... de même que par des „affinités“ particulières (certains couples de consonnes peuvent dans une langue donnée facilement former un groupe, d'autres au contraire rarement ou pas du tout; même l'ordre dans lequel se suivent les consonnes d'un groupe n'est pas indifférent à cet égard;... Le poète peut influencer et les principes généraux du groupement et les „affinités“ en choisissant de préférence certaines possibilités ou en évitant d'autres“ (MUKAŘOVSKY l. c. 281.).

Továbbá CARACOSTEA *Expresivitatea limbii române* (București 1942) c. munkájának 69—71 lapján statisztikusan kimutatva látható, hogy a hangsúlyos szótag magánhangzóinak eloszlása menyire különbözik az egyes költők (ALECSANDRI, EMINESCU) használatában a köznyelvétől.<sup>2</sup>

Igen fontosak egy adott nyelvre vonatkozó statisztikai vizsgálatok. Ezt M. COHEN-nal együtt V. POLÁK is hangsúlyozza. (Časopis pro mod. fil. XXXIV, 34): „Jusqu'à présent, les linguistes ne savent pas exploiter les vues théoriques de la statistique, en particulier en ce qui concerne les expressions mathématiques et graphiques des données en question par rapport aux lois de la probabilité mathématique et aux équations mathématiques.“

Az ADYra vonatkozó eredményeknek egy modern, jó magyarságú prózáíró nyelve vizsgálatából adódó statisztikával való összehasonlítására és annak eldöntésére, hogy ADY nyelvének vizsgálatát csakugyan alkalmatlan-e a magyar nyelvre vonatkozó statisztika egyedüli alapjául venni, a Központi

<sup>1</sup> Én emeltem ki.

<sup>2</sup> Azzal azonban nem érthetünk egyet, hogy CARACOSTEA szótári adatok statisztikázását tekinti a köznyelv statisztikázásának.

Fizikai Kutató Intézet Akusztikai és Ultrahang Csoportjának munkaterve keretében nekem kellett VERES PÉTER nyelvének a Próbátétel c. novelláskötet alapján TARNÓCZYNAK ADY feldolgozásához hasonló szempontok szerint való statisztikázását végezniem.

*A statisztika készítésénél figyelembe vett hangmódosulások*

TARNÓCZY munkájának célja a magyar hangzókapcsolatok olyan statisztikai megvizsgálása, hogy ezek segítségével akusztikai és hallásvizsgálatok céljainak megfelelő, a nyelv hangzóeloszlását tükröző, alakilag a magyar szavakhoz hasonló, de értelem nélküli szavak, ú. n. *logatomok* készítése lehetővé váljék. Tehát a hangzókapcsolatokat az érthetőség szempontjából vizsgálja s így igen helyesen állapítja meg, hogy „érthetőségi vizsgálatoknak mindig csak egy nyelv keretén belül van létjogosultságuk“ (108). Tökéletesen egyetértünk azzal az eljárással is, hogy „a feldolgozás mindig hangzók és nem betűk alapján“ történik (111); ezért „a statisztika elkészítésének az alapja a rendelkezésre álló szöveg fonetikai átértékelése“ (112). Igen helyesen választotta általános szempontnak, „hogy csak a fonémváltozásokat kell figyelembe venni, a variánsokat mindig a hozzájuk legközelebb álló fonémához kell jegyezni“ (113). Azaz csak akkor veszünk a statisztikázásnál más hangot figyelembe, mint az írott betű jelölte hangot, ha ejtéskor az írott betűnek megfelelő fonémától eltérő fonémát ejtünk (pl. *aszt* helyett a statisztikában *aszt* szerepel, de *nyomva* és *száma* -*m*-je csak variáns, tehát azonos fonéma). A továbbiakban azonban vitatható a TARNÓCZY követte eljárás. Kérdés, helyesen állapította-e meg a hangzókapcsolatok statisztikai feldolgozásánál figyelembe veendő gyakoribb asszimilációkat, összeolvadásokat és egyéb változásokat.

Kétségtelen és minden eddigi fonetikai vizsgálattal egybehangzó az a megállapítás, hogy a magyar nyelvben van zöngésségi és nazális hasonulás, van „összeolvadás“, két egymást követő magánhangzó között ejtett *j* és ejtés-könnyítő hangkivetés. Azonban azt már nem mernők minden további megfontolás nélkül állítani, hogy ezek a kiejtésben kivétel nélkül minden egyes esetben mintegy kategórikus imperatívusznak engedelmessé meg is való-sulnak. Többek kiejtését ilyen szempontból megfigyelve és véleményét kérdezve, mit hallanak egyes esetekben, arra a meggyőződésre jutottunk, hogy a nazális hasonulás, a vokálisközi jésítés és a hangkivetés esetében TARNÓCZY helyesen oldotta meg a kérdést. (A vokálisközi jésítés talán minden esetben kétirányúnak tekinthető, de általában — ahogy TARNÓCZY helyesen állapította meg — nem annyira, hogy ezt a statisztikázásnál figyelembe kellene vennünk.)

A zöngésségi hasonulás és az összeolvadás esetében azonban állandóan szem előtt kell tartanunk azt a TARNÓCZY említette tényt, hogy „a vizsgálatokból... sohasem küszöbölhető ki a hallgató értékelése, mert ez jelenti a vizsgálatok szempontjából oly fontos nyelvi ítéletet“ (107.).

Kétségtelen, vitathatatlan, hogy a TARNÓCZY által igen szemléletesen csoportosított zöngétlen és zöngés indukálások és összeolvadások (113, 114.) a természetes ejtésű magyar beszédben előfordulnak, azonban erősen függ a hasonulás mértéke a beszéd tempójától, s ezért a nyelvi értékelés alapján arra a megállapításra jutunk, hogy az indukált hang nem megy minden esetben át a megfelelő zöngés, ill. zöngétlen hangba, hanem variánsként fortist, lenist, ill. zöngétlen mediát vagy tenuist hallunk. Tekintettel arra, hogy még

szövegösszefüggés nélkül is meg tudjuk egymástól különböztetni a *hatban* — *hadban*, *hattól* — *hadtól*, *magból* — *makkból*, *magtól* — *maktól* stb. szópárokat, tehát nem tartjuk helyesnek, hogy pl. *b* előtt a *t*-t, *k*-t mindig *d*-nek, *g*-nek, *t* előtt a *d*-t, *g*-t mindig *t*-, ill. *k*-nak tekintsük. Lehetséges, hogy ejtés közben zöngésedő, illetve zöngétlenedő hangok esetében a nyelvi értékelés az ejtés közben megváltozó hang kezdő szakaszához legközelebb eső fonéma jelenlétét állapítja meg.

Hivatkozhatunk arra is, hogy talán még sohasem hallottunk *széddarabol*, *Margicciget*, *áccálló* stb.-féle ejtést. Általában az asszimilációkkal kapcsolatban igaz Lotz JÁNOS megállapítása (Das ungarische Sprachsystem, 1939, Stockholm, 23.): „Manchmal treten die obigen Assimilationserscheinungen in der Aussprache der Gebildeten nicht auf, da das Schriftbild zu stark einwirkt“.

Mindezek figyelembe vételével arra a megállapításra juthatunk, hogy bár „a statisztika elkészítésének az alapja“ feltétlenül „a rendelkezésre álló szöveg fonetikai átértékelése“ (112), TARNÓCZYtól eltérően zöngességi hasonuláskor és összeolvadáskor esetenként kell megállapítanunk a statisztikázásnál figyelembe veendő hangot. Ez különösen fokozott mértékben fennáll, ha a két egymást indukáló hang közé még szótaghatár is esik. Minthogy a szótaghatár kérdésében eddig nem történt elegendő vizsgálat, sem elméleti, sem gyakorlati téren, dolgozatomban a helyesírási szótaghatárt kell a beszédben szereplő szótaghatárként elfogadnom. Mind TARNÓCZY, mind pedig saját IV. táblázatom alapján arra a megállapításra kell jutnunk, hogy szó belsőjében a hármas mássalhangzókapcsolatok viszonylag annyira ritkák, s ezeknek is a felébe (a 2. és a 3. mássalhangzó közé) szótaghatár esik, úgy, hogy gyakorlatilag a soron\* következő XXXIII.—LV. táblázatban feltüntetett szóközépi mássalhangzótalálkozásokat szótaghatárral elválasztot-taknak tekinthetjük. Egyébként a szóközépi hármas mássalhangzócsoportokra még visszaterünk. (I. a folytatásban). Az említett táblázatokban szó elején vagy szó végén feltüntetett mássalhangzótalálkozások közé, természetesen nem eshet szótaghatár. Az eltérő feldolgozási szempontok miatt a táblázatokban a megfelelő zöngés-zöngétlen hangpárokat egymás után vesszük fel, így könnyebb a VERES PÉTERRE vonatkozó eredményeket az ADYRA kapottakkal egybevetni és egyezéseiket vagy eltéréseiket megállapítani.

Külön kell megemlítenünk az *lj*-kapcsolatot. Az *ólja*, *gólja* stb. szavakban semmiesetre sem egyezik az általános ejtéssel, ha az *lj*-t *j*-nek tekintjük a statisztikázásnál, mint TARNÓCZY tette: *ója* v. *ójja*, *gója* v. *gójja*! Ebben az esetben sem fogunk TARNÓCZYval azonos módon eljárni a statisztika készítésekor.

Továbbá eldöntendő kérdés, hogy a talált hangzókapcsolati arányszámokat minden további nélkül elfogadhatjuk-e az érthetőségi szövegminták hangzókapcsolatainak alapjául. Hasonlítsuk össze pl. az *nyh* és az *fr* kapcsolat gyakoriságát. Kétségtelen, hogy mind a kettő, (TARNÓCZY 4. ábrája alapján különösen az első) ritka. Érthetőségi minták készítésekor nem kell-e vajjon arra is gondolnunk, hogy *fr* jóformán csak néhány ritkán használt *f*-re végződő szó (*gróf*, *rőf* stb.) *-ról*, *-ről*, *-ra*, *-re* ragos alakjában és az eléggé ritka *cifra*, *cafrang* stb. szavakban fordul elő. A szóközépi helyzetben ugyancsak nem túlságosan gyakori *nyh* azonban az *ny* végű szavak ragozott alakjain kívül (*aranyhoz*, *érvényhez* stb.) *lanyha*, *enyhe*, *renyhe*, *konyha*, *dunyha*, *kunyhó*, *Menyhért* stb. szavakban is szerepel. Vajjon az előbb említett kapcsolat nem csak azért érthető-e könnyen a magyarul tudók számára, mert morfémahatá-

\* Ezek a táblázatok a dolgozat folytatásában szerepelnek.

## I. TÁBLÁZAT

*A kételemű hangzókapcsolatok eloszlásához feldolgozott anyag adatai*

	Lapok száma				Hangok száma		
	novellánként				összesen	feldolgozott	10 lapon
	(5-) 9-26	27-72	73-122	123-154			
ú	9-26	28-45	74-95	123-144	79,75	321	40
u	9-26	28-45	74-95	123-144	79,75	965	121
ó	9-13	28-31	74-78	124-128	18,75	225	120
o	9-13	28-31	74-78	124-128	18,75	1007	537
a	9-11	28-29	74-75	124-126	9,75	1329	1300
á	9-15	28-33	74-79	124-130	25,75	1164	453
e	9-11	28-29	74-75	124-126	9,75	1240	1270
é	9-18	28-35	74-83	124-133	37,75	1377	364
ö	9-20	28-38	74-85	124-138	49,75	680	137
ő	9-20	28-38	74-85	124-138	49,75	487	98
i	9-13	28-31	74-79	124-129	20,75	1195	578
í	9-13	28-31	74-79	124-129	20,75	102	49
ü	5-26	27-72	73-122	123-154	146,00	768	52
ű	5-26	27-72	73-122	123-154	146,00	188	13
l	9-13	28-32	74-78	124-128	19,75	1144	580
r	9-14	28-33	74-79	124-128	22,75	1235	544
m	9-14	28-33	74-78	124-129	22,75	1208	532
n	9-13	28-31	74-77	124-128	17,75	1225	690
ny	5-26	27-72	73-122	123-154	146,00	1000	68
v	9-18	28-40	74-87	124-135	48,75	1114	229
f	5-26	27-59	74-102	123-154	115,75	1083	94
z	9-20	28-38	74-84	124-135	45,75	1204	264
sz	9-20	28-39	74-85	124-137	49,75	1179	237
zs	5-26	27-72	73-122	123-154	146,00	152	10
s	9-15	28-34	74-80	124-131	28,75	1176	409
j	9-25	28-41	74-88	124-139	61,75	1111	180
h	9-21	28-39	74-85	124-135	48,75	1171	241
b	9-19	28-40	74-84	124-135	46,75	1144	245
p	5-26	27-61	74-96	124-154	109,00	1082	100
d	9-20	28-38	74-85	124-135	48,75	1151	237
t	9-13	28-31	74-77	124-127	16,75	1396	835
g	9-18	28-38	74-84	124-138	46,75	1271	272
k	9-12	28-31	74-77	124-128	16,75	1108	662
c	5-26	27-72	73-122	123-154	146,00	321	22
cs	5-26	27-72	73-105	123-154	129,75	1317	102
gy	9-26	28-41	74-89	124-139	57,75	1129	195
ty	9-26	27-72	73-122	123-154	143,50	131	9

35100

11949

rokon keletkezik és talán nem is realizálódik a szó utolsó hangja és a rag közötti hangkapcsolat a tudat számára, ezzel szemben az említett *nyh* kapcsolat mégis valamivel gyakrabban szerepel rag nélküli szavakban is. Erre a kérdésre jövevényiszavaink összehasonlító tárgyalásával kapcsolatban még visszatérünk.

A statisztikázás technikája olyan, mint ADY verseinek feldolgozásáé. VERES PÉTER Próbátétel c. novelláskötetének mind a négy elbeszéléséből feldolgoztam egy-egy részt, a nagyon ritka hangoknál még az előszót is. Részletesen az I. táblázatban látható, hogy a könyv melyik részéből melyik hangra mennyi adatot dolgoztam fel; általában igyekeztem egy-egy hangra kb. 1000—1200 előfordulást a statisztikákhoz feldolgozni.

### *A statisztikázás matematikai alapjairól*

Statisztikai feldolgozáskor felmerül a kérdés: az eredmények mennyire általánosíthatók és milyen határok között érvényesek. TARNÓCZY dolgozatából nem derül ki, hogy talált adatai, véleménye szerint, mennyire jellemzők a magyar nyelv általános hangzókapsolódásaira, illetve a 4. ábra alatti címből az látszik, hogy ő az ADYNál talált mássalhangzókapsolatok gyakorisági arányát a magyar nyelv mássalhangzókapsolatai gyakorisági arányának tekinti. Ez a merész általánosítás *egy költő* nyelve alapján annál meglepőbb, mert dolgozatában még arra a kérdésre is hiába keresünk választ, hogyha bármely adatának ellenőrzését nem más írónál, hanem ADY összes költeményeinek más helyeiről kísérelnők meg, milyen mértékben kellene a talált adatoknak egyezniük. Bizonyos önellenőrzést végez ugyan, de ennek matematikai értéke igen kérdéses. II táblázatában a magyar nyelv hangzóinak eloszlását vizsgálja 30 lap (az ADY kötet 17—21, 69—73, 146—155, 219—223, 254—258), összesen 34 063 betű, ill. 33 218 hang alapján, a betűstatisztikát 10—10 laponként is közli és azt találja, hogy „a betűk száma 10% eltéréssel megegyezik, a százalékok között pedig nincs 20%-nál nagyobb eltérés“ (118.). TARNÓCZY ezt az ingadozást *meglepően* kicsinek találja (u. o.). A valószínűségi számítás statisztikázására vonatkozó eredményei azonban másképpen szólnak. Matematikailag megállapítható, milyen mértékig nevezhetjük az eltéréseket *szórásnak*, azaz nem hibának, hanem a statisztikázás természetéből adódó eltérésnek, ha viszont az eltérés ennél nagyobb, akkor már kívül esik a hibahatáron, akkor már az eredmények között nem statisztikus eltérések vannak, hanem az eredmények egymásnak ellentmondóak, hibásak!

Nyilvánvaló, hogy nem lehet a hibahatár nagyságát százalékban mérni. Ha pl. 1000 adatból valamilyen jelenségre csak két példát találunk (ami nyilván kevés), akkor természetesen nem tekinthető hibának, ha másik 1000 adatból ugyanerre a jelenségre nem 2, hanem 1, 3, sőt még az sem, ha 4 példát találunk. Itt az eltérés az adatok között, egymáshoz képest meghaladja a 100%-ot, de a vizsgált anyagnak egyik esetben sem teszi ki még a 0,5%-át sem. Természetesen, még jobbnak nevezhető az egyezés, ha nem 1000, hanem mondjuk 1—1 000 000 adatot vizsgálva, találunk az egyes milliós csoportokban 1, 2, 3, 4 megvalósulást. Lényegesen nagyobbnak látjuk az eltérést, ha ugyanezek a különböző előfordulások 100—100 adat vizsgálatából adódnak, hiszen itt már nem tizedszázalékok az eltérések, hanem az egész anyag 1—4%-a között ingadoznak.

Egészen másképp ítéljük meg a helyzetet, ha 1000 adatból valamire 500 példát találunk. Ha másik 1000 adatból csak 250 vagy éppen 750 példát

találnánk, ebben az esetben, annak ellenére, hogy az eltérés a középértéktől csak 50%, a hibát mégis lényegesen nagyobbak, hibának és nem eltérésnek minősítenők.

Abban az esetben, ha 1000—1000 eset közül a kérdéses jelenségre 1, 2, 3, 4 példát találunk, azt állapítjuk meg, hogy a keresett példa ritkán jelentkezik, viszont, ha 1000—1000 adat közül 250, 500 vagy 750 előfordulásra bukkanunk, vagy arra gondolunk, hogy a jelenség — úgy látszik — rendszerelen, nem oszlik el statisztikusan, mert nem lehet az egész anyagnak egyszere 25, egyszer 50, máskor meg 75%-a, vagy azt kell hinnünk, hogy csak sokkal nagyobb adatszából lehet statisztikus előfordulási arányt találni.

Annak megállapítására tehát, hogy az eltérések meddig nevezhetőek hibahatáron belüli szórásoknak, és hol kezdődik vagy a módszer helyességét, vagy a feldolgozás pontosságát kétségbevonó hiba, tekintetbe kell vennünk:

1. a talált adatszám nagyságát,
2. az összes megvizsgált adatok számát.

Ahhoz, hogy a statisztikázásnál előforduló hibahatárok reális méreteit megállapíthassuk, néhány alapfogalmat kell tisztáznunk.

A matematikai alapok tárgyalásával PAUL MENZERATH és WERNER MEYER-EPPLER Sprachtypologische Untersuchungen (Studia Linguistica 1950 nr. 1—2, 54—93) c. művének megfelelő fejezetét követjük, de messzemenő további egyszerűsítéseket vezetünk be.

Egyszerűség és könnyebb érthetőség kedvéért a matematikai valószínűségi alapfogalmakat nem tisztán elméleti alapokon vizsgáljuk, hanem a hangok előfordulási gyakoriságával kapcsolatban.

Ha pl. *tudnók* azt, hogy a XX. század első felében hány hangot ejtettek ki és jegyeztek le összefüggő magyar szövegben és ezek közül mennyi esett az egyes hangokra, akkor ismernők az egyes hangok „ideális” gyakoriságát a XX. század első felének nyelvében. Ezt az *ideális gyakoriságot* százalékban is kifejezhetnők, s ezek a százalékok lennének az *ideális relatív gyakoriság* százalékarányai. Minthogy nemcsak az összes elhangzott beszéd, de minden leírt szöveg (levelek, jegyzetek stb.) statisztikai feldolgozása is teljesen lehetetlen, ezért a valószínűségszámítás egészen más oldalról igyekszik az ideális relatív gyakoriság százalékarányait megközelíteni (*megközelíteni* és nem pontosan *megállapítani*).<sup>1</sup>

Feltehetjük, hogy a hangok százalékaránya nem változik ugrásszerűen kisebb korszakokon belül, sőt még csak erősebb ingadozásnak sincs kitéve, s így a százalékarányok megállapításánál nem kell feltétlenül szigorúan két évszám közötti előfordulási gyakoriság megállapítására törekednünk, hanem figyelembe vehetjük a múlt század azon nagy íróit is, akiknek a nyelvét ma is élő nyelvnek érezzük s ezek figyelembevételével a magyar nyelvben vizsgálhatjuk az előfordulási gyakoriságokat. A nagy számok törvényéből adódik, hogy elég nagy adatszám vizsgálatánál a százalékeloszlások kielégítő pontossággal közelítik meg az ideális relatív gyakoriság százalékarányait. Vezessük be a következő jelöléseket:

Jelöljük nagy betűvel mindazon adatot, amely abszolút adatszám, és kis betűvel a relatív adatszámokat, azaz az abszolút adatszámok és az összes adat hányadosaként adódó viszonyszámokat:

<sup>1</sup> Minden számítás logarléccel vagy számológéppel történt.

$N$ az összes vizsgált adat száma, százalékban .....	100%
$H$ egy-egy jelenségre jutó adatszám, neve <i>abszolút gyakoriság</i> (absolute Häufigkeit); ugyanez az összes adatok számához viszonyítva, vagyis $H/N$ , illetve 100-zal szorozva, százalékban kifejezve adja a <i>relatív gyakoriságot</i> (relative Häufigkeit), jele .....	$h$
az ismeretlen, megközelítendő <i>ideális relatív gyakoriság</i> , jele .....	$*h$
$*H$ az $N$ adatból egy bizonyítot vizsgált jelenségre jutó <i>ideális abszolút gyakoriság</i> , azaz adatszám.	

Konkrét példát véve: VERES PÉTER Próbatétel című novelláskötetének 21., 22., 28., 38., 39., 67., 68., 76., 77., 92., 148., 149. lapján 15 398 hangot számoltam össze, tehát  $N = 15\,398$ , ebből pl.  $a$  hang 552, azaz  $H = 552$ , vagy százalékban kifejezve  $552/15\,398 = 0,0358$ , azaz  $h = 0,0358$ , illetve 100-zal szorozva  $3,58\%$  az  $a$  relatív gyakorisága.

A valószínűségszámítás tanítása szerint az ideális relatív gyakoriság, a  $*h$  nem adódik pontosan egy-egy kiragadott és megvizsgált szövegrészben, hanem bizonyos szórás  $\Delta h$  járulhat hozzá (s részben vagy egészben mindig hozzá is járul) s ez a  $\Delta h$  a  $*h$  értékét növelheti is vagy csökkentheti is, vagyis  $*h \pm \Delta h$ , vagy részletesen kifejtve  $*h - \Delta h$  és  $*h + \Delta h$  között várható elég nagy adatszám vizsgálata esetén annak a jelenségnek (pl. az  $a$ -nak a gyakorisága), amelynek ideális, megközelítendő gyakorisága  $*h$ .

Kérdés, hogyan állapítható meg  $\Delta h$  nagysága. A valószínűségszámítás tanítása szerint

$$\Delta h = \sqrt{\frac{*h(1-*h)}{N}} \quad (1)$$

képlet adja. Az (1) alatti képletet azonban módunkban van bizonyos esetekben egyszerűbben is kifejezni.

Két esetet kell egymástól megkülönböztetnünk:

I

A szórás,  $\Delta h$  értéke igen kicsiny, ha  $*h$  nagyon kicsiny, vagy ha megközelíti 1-et (tehát igen kicsiny, vagy 100%-ot megközelítő értékek esetén).

a) Ha  $*h$  igen kicsiny ( $h \ll 1$ , akkor az  $1-*h$  megközelítőleg 1, tehát 1, mint szorzó elhanyagolható és így elég jó megközelítéssel

$$\Delta h = \sqrt{\frac{*h}{N}} \quad (2)$$

kifejezést kapjuk a szórás nagyságának. Ez látszólag kényelmes, de valójában igen nehezen kezelhető formula, mert a relatív gyakoriság értékének és az összes adat hányadosának a négyzetgyöke adja a szórás nagyságát. Különösen figyelmet igényel az az eset, melynél  $N$  10 hatványa, mert páros hatvány (100, 10 000, 1 000 000 stb.) esetén  $\Delta h$  könnyen adódik, egyszerűen

$$\Delta h = \frac{\sqrt{*h}}{\sqrt{N}}$$

azaz egyszerűen  $\sqrt{*h}$ -t kell  $\sqrt{N}$ -nel osztani és ez utóbbi 10 hatványa, mert 100, 10 000, 1 000 000 stb. négyzetgyöke 10, 100, 1000, azaz ismét 10 hatványa, ezzel viszont nem fáradság osztani, de ha  $N$  10-nek páratlan hatványa, azaz 10, 1000 és 100 000, akkor a gyöke 3,12, 31,2, 312 stb. s ezzel kell osztanunk!

A kérdés ezért még világosabban tekinthető át, ha a relatív gyakoriság szórása helyett az abszolút gyakoriság szórását vizsgáljuk. Minthogy a relatív gyakoriságot,  $h$ -t az abszolút gyakoriságból,  $H$ -ból úgy kapjuk, hogy az összes vizsgált elemek számával,  $N$ -nel osztunk, visszatérésnél az ellenkező műveletet kell végeznünk, azaz  $N$ -nel szoroznunk.

$$N \cdot \Delta h = N \sqrt{\frac{*h}{N}}$$

vagy  $N\Delta h = \Delta H$  jelöléssel és  $N$ -et jobboldalt a négyzetgyökjel alá vite:

$$\Delta H = \sqrt{\frac{N^2 *h}{N}}$$

Egyszerűsítve:

$$\Delta H = \sqrt{N *h}$$

Minthogy  $N *h = *H$ ,

$$\Delta H = \sqrt{*H} \quad (3)$$

(Poisson törvénye). Ennek alapján, ha bármilyen, kis százalékban előforduló részhalmaz adatszámát ismerjük, akkor minden olyan érték helyesnek tekintendő, amely ugyanannyi adatból véve az ismert adatszámától kevesebbrel különbözik, mint ennek négyzetgyöke.

b) Abban az esetben, ha a százalékos előfordulás megközelíti a 100%-ot, akkor az (1) alatti formulában, minthogy  $*h$  megközelítőleg 1 ( $*h \approx 1$ ),  $1 - *h$  értéke jó közelítéssel zéróhoz közeledik, tehát ha a tört számlálója 0-hoz tart, a tört, s ezzel együtt négyzetgyökének értéke is 0-hoz konvergál, azaz  $\Delta h$  is igen kicsiny. A szórás nagyságát abban az esetben minden bizonytalansággal jól meghatározhatjuk a halmazt 100%-ra kiegészítő halmaz szórásának a nagyságából. (Előző példákra hivatkozva a VERES PÉTER-nél számbavett kb. 15400 hangból 96,42% nem  $\acute{a}$  hang, ahhoz ugyanannyi szórás járulhat, mint az  $\acute{a}$  hangok 3,58%-ához.)

## II

Sem nem nagyon kicsiny és sem nem nagyon nagy relatív gyakoriság esetén a szórás nagyságának megállapítására az (1) alatti formulához kell visszatérnünk és abból kiszámítanunk az abszolút szórás nagyságát:

$$N\Delta h = N \sqrt{\frac{*h(1-*h)}{N}}$$

A már alkalmazott átalakításokkal  $N\Delta h$  helyett  $\Delta H$ -t írunk, jobboldalon  $N$ -et a négyzetgyök alá visszük, majd egyszerűsítünk:

$$\Delta H = \sqrt{\frac{N^2 *h(1-*h)}{N}} = \sqrt{N *h(1-*h)}$$

Innen

$$\Delta H = \sqrt{*H(1-*h)}$$

Bevezetve  $*h = *H/N$  jelölést (ez az abszolút és relatív gyakoriság definíciója alapján igaz), végeredményül

$$\Delta H = \sqrt{*H(1-*H/N)} \quad (4)$$

szórási képlet adódik.

A (4) alatti formulát

$$\Delta H = \sqrt{*H} \sqrt{1-*H/N} \quad (4a)$$

formában írva azt olvashatjuk ki belőle — összehasonlítva a (3) alatti  $\Delta H = \sqrt{*H}$  formulával, — hogy a szórás közepes relatív gyakoriság esetében a kis vagy nagy relatív



gyakoriság esetében talált szórás  $\sqrt{1 - *H/N}$ -szerese. Minthogy  $*H$ , az abszolút gyakoriság szükségképpen mindig kisebb  $N$ -nél, az összes adatok számánál, ezért a  $*H/N$  tört értéke kisebb egynél, ezt a törtet egyből kivonva ismét egynél kisebb számot kapunk, ennek gyöke ismét egynél kisebb. Tehát (4a) alapján  $*H$ -t egynél kisebb számmal kell szoroznunk, azaz az értéke csökken. Ha mi tehát a továbbiakban a szórás nagyságát minden esetben az abszolút gyakorisági szám négyzetgyökének tekintjük, akkor evvel a szórási határt növeljük, azaz szélesebb sávon fogadjuk el az értékeket helyesnek, mint az a valószínűségi számítás alapján következne. Ezzel azonban nem követünk el nagy hibát, mert a hangkapcsolatok vizsgálatánál, minthogy igen sok kapcsolódási lehetőség van, úgyszólván alacsony százalékok szerepelnek, a számítás pedig lényegesen egyszerűbb, könnyebb, áttekinthetőbb a (3) alatti formulával. Különben sincs semmi kritériumunk, hogy meddig kell a relatív gyakoriságot kicsinek, nagynak vagy közepesnek tekintenünk.

Eddigi megfontolásaink használhatóságát kérdésessé teszi az a tény, hogy a gyakorlatban nem ismerjük az ideális  $*H$  gyakoriságot, vizsgálódásaink csak  $H$  megismeréséhez vezethetnek. Ez azonban feltehetően tartalmazza a szórási értéket, tehát az általunk gyakorlatilag megállapított  $H$  értéke különbözik  $*H$ -tól, eltérése elég nagy adatszám vizsgálata esetén maximum  $\Delta H$  lehet, tehát

$$H = *H \pm \Delta H$$

vagy másképpen

$$H = *H \pm \sqrt{*H}.$$

Az ismert  $H$ -ból az ideális  $*H$ -t akarjuk megismerni. Ha egyenletünket  $*H$ -ra oldjuk meg, (másodfokú, irracionális egyenlet), akkor kicsit hosszabb számolással

$$*H = H + \frac{1}{2} \pm \sqrt{H + \frac{1}{4}}$$

eredményt kapjuk. Minthogy  $H$  is és természetesen  $*H$  is statisztikus vizsgálatoknál általában nagy szám, ezért ezek mellett  $\frac{1}{2}$  és  $\frac{1}{4}$  elhanyagolhatók, tehát a továbbiakban minden eredményt a négyzetgyökével növelve vagy csökkentve tekinthetünk jó megközelítéssel a reális valószínűségi sávnak.

$$*H = H \pm \sqrt{H} \quad (5)$$

Ez azt jelenti, hogy az ideális  $*H$  gyakoriságnak a vizsgálatok közben talált  $H$ -hoz közel eső,  $H - \sqrt{H}$  és  $H + \sqrt{H}$  közötti értékűnek kell a valószínűségi számítás alapján lennie.

Eddig még nem vettük tekintetbe az összes vizsgált adat befolyását az eredmény megbízhatóságára. A valószínűségi számítás tanítása szerint a *bizonytalanság* az összes vizsgált adat négyzetgyökével fordítottan arányos. Azaz 100 adat esetén  $1/\sqrt{100} = 1/10$  vagy százalékban kifejezve 10%, 10 000 adatnál  $1/\sqrt{10\,000} = 1/100$ , azaz 1% a bizonytalanság, tehát a vizsgált adatok számának erősen meg kell növekednie ahhoz, hogy a bizonytalanság kicsit is csökkenjen. TARNÓCZY vizsgálatainál átlag egy-egy hang 5000 adatra vizsgálva  $1/\sqrt{5000}$ , tehát 1,4%, az én VERES PÉTERRE vonatkozó vizsgálataimnál, átlag 1200 adatot véve egy hangra, kereken 3% a bizonytalanság minden talált adatra a szóráson kívül.

Ezek alapján, visszatérve az előbb említett hangzóeloszlási példára, azt mondhatjuk: amennyiben kb. 15 400 hang elég nagyszámú adat ahhoz, hogy ennek alapján a hangok gyakoriságának eloszlására következtethessünk, akkor példánkban az  $a$ -ra talált  $H = 552$  abszolút gyakoriságával egyező eredménynek kell tekintenünk, ha bármilyen más vizsgálatban ugyanannyi adatból  $552 \pm \sqrt{552} \approx 552 \pm 24$ , azaz 528 és 576 hang közötti  $a$  hang fordul elő. Ha még a bizonytalanságból adódó 0,8% ingadozást is tekintetbe vesszük (552-nek 0,8%-át), akkor 524-től egészen 580-ig jól egyeznek egymással az

adatok. Vagy visszatérve a relatív gyakoriságra, illetve ezeknek a gyakorlatban szokásos százalékos kifejezésére, akkor szórás és bizonytalanság figyelembevétele után azt állíthatjuk, hogy a magyarban az *á* hang ideális relatív gyakoriságának százalékaránya valószínűleg 3,4% és 3,74% között található. Ellenőrzésképpen a Próbatétel alapján nyert eredményt egybevetethetjük mávizsgálatokból adódó százalékarányokkal (I. a II. táblázatot).

A statisztika készítője	<i>á</i> %-a	Bizonytalansági és szórási határ	
		alsó	felső
TARNÓCZY .....	3,5	3,38	3,62
TOLNAI .....	3,60	3,45	3,75
MIKES .....	3,70	3,65	3,75
NEMES .....	3,4	3,22	3,77
VÉRTES .....	3,58	3,40	3,76

Az öt statisztika középértéke

3,56%.

Az összeállításból nyilvánvaló, hogy mind az öt feldolgozás eléggé megközelíti az ismeretlen ideális értéket. Az *á* ideális relatív gyakoriságának a felső határa, az ADY alapján nyert statisztika kivételével, szinte teljesen egyezik, az alsó határnál már nagyobbak az eltérések. Van ugyan egy látszólag ellentmondó adatszám is: MIKES eredményei alapján az *á* ideális relatív gyakoriságának 3,65%-nál nagyobbak kell lennie, TARNÓCZY statisztikája alapján viszont 3,62%-nál kisebbnek. Ha tekintetbe vesszük, hogy mind a két statisztika speciális stílusra (gyorsírási szövegek, illetve lírai költő) vonatkozik, akkor már kisebb a számszerűen amúgy is kicsiny ellentmondás statisztikai jelentősége. Valószínűségszámítási feladatoknál sok adat esetén egyébként sem szokás tekintettel lenni egy-egy, a többitől erősebben eltérő jelenségre. Nem mond ellent az *á* ideális relatív gyakoriságáról kapott itt látható összeállításnak az a II. táblázatban látható adat sem, hogy a Próbatételben az egyes hangoknak tíz lapra kikövetkeztethető előfordulási gyakoriságát (I. az I. táblázatot) aránybaállítva az *á* 3,88%-os relatív gyakorisággal szerepel, hiszen ehhez is járul szórás és bizonytalanság és ezek figyelembevételével 3,59% és 4,01% között kell az ideális értéket keresnünk. Ez viszont tökéletesen egyezik a nem kikövetkeztetett, hanem valóságos eloszlások számlálásából adódó eredményekkel.

### I. Az ADYra és VERES PÉTERre kapott eredmények kritikai egybevetése

Vizsgáljuk meg ezek után TARNÓCZY II. táblázatában szereplő, tíz-tíz lapra vonatkozó részleteredményeket és betűre történt összesítésüket, alkalmazva a valószínűségszámításból adódó eredményeket a kritikában. Tíz-tíz lapon 11 807, 10 889, ill. 11 367 betűt talált, azaz összesen 34 063 betűt. Így a középértéket véve arra lehet következtetni, hogy ADY összes verseinek az általa vizsgált kiadásában tíz lapon 11 354 betű szerepel. Ehhez hozzájárulhat  $\pm$  ennek négyzetgyöke, azaz 107, mint szórási érték és  $\pm$  1%, azaz 114, mint bizonytalanság, tehát 11 133 és 11 575 közötti betűszámot kell találnunk tíz-tíz laponként. Vagyis 442 betű ingadozással még azonosnak tekinthetők

az eredmények. Ehhez a valószínű értékhez képest az első tíz lapon sok, a második tizen kevés és csak a harmadik tíz lapon van megfelelő számú adat. Az eredmény nem meglepő, hiszen ha a verseskötetben lapozgatva néhány lapot összehasonlítunk, akkor inkább azon csodálkozánk, ha tízlaponként jó közelítéssel átlagot lehetne mondani a rajta található betűk számáról.

Ugyancsak a II. táblázat alapján  $33\ 218/3 = 11\ 073$  hang van átlag tíz lapon. (TARNÓCZY I. táblázata szerint, tévesen, 100 lapra 102 500 hang esik) Más helyen (111.) TARNÓCZY azt mondja, hogy az ADY-kötet 532 lapjából 520-on van szöveg s ezen, a cikkben nem közölt módon történő becslés alapján, 535 000 hang található. Ebből viszont az következik, hogy tíz lapon 10 288 hang van. Ha viszont a mássalhangzók 100 lapra eső adatszámából számítjuk ki a tíz lapra eső hangok számát, 8 680-at kapunk. Továbbá az I. táblázatból az is kiderül, hogy a szavak és szótagok hangzóeloszlásához feldolgozott 19 lapon 18 118 hangzó van, ennek alapján tíz lapra 9 535 hang adódik (TARNÓCZY az I. táblázatában — tévesen — 100 lap hangszámának 103 400-at ír 95 350 helyett). Valószínűségi hibahatárok figyelembevétele nélkül is látható, hogy 11 073, 10 288, 8 680 és 9 535 egymással semmiképpen sem egyeztethető átlagok tíz lap hangszámára vonatkozóan.

A Próbatételből a következő eredményekre jutunk :

Lapszám .....	19	20	30	31	106	107	150	151
Rajta lévő hangok száma	1175	1352	1360	1190	1231	1149	1161	1261

azaz összesen 9879 hang, egy lapra tehát 1235 hang esik átlagban  $\pm 35$  mint szórás,  $\pm 13$  mint bizonytalanság, tehát 1187 és 1283 között minden érték egyformán jónak tekinthető. Tehát a vizsgált nyolc lap közül csak három van a valószínűségi számítás hibahatárán belüleső, jó érték (31., 106., 151. lapon). Viszont a többinél is kisebb az eltérés, mint ADYNál tízlaponként. Az összes hangok tíz lapra eső értékéből egy lapra 1192 hang adódik, s ez teljesen jó érték, sőt ennek hibahatára alapján viszont a 19., 107., 150. lapon található hangzószám is megfelelő, így csak a 20. és 30. lapon van a valószínűségi alapon várhatónál több hang. Ha két-két laponként foglaljuk össze az adatok számát, akkor 2527, 2550, 2380 és 2422 hang jut két-két lapra, középérték  $2470 \pm 50$  szórás  $\pm 25$  bizonytalanság, tehát 2395 és 2545 közé esnek a statisztikusan teljesen helyes értékek. A 19—20. és 150—1. lapon levő hangok száma ennek meg is felel, a 30—1. lapon 5-tel több a hang, mint a valószínűségi számítás alapján várnók, a 106—7. lapon viszont 15 hang hiány van. Ezek alapján megállapíthatjuk, hogy a Próbatételben laponként még nem szigorúan statisztikus a hangok eloszlása, de kétlaponként már nagyjából megfelelőnek tekinthető. A feldolgozás során ennél ügyis mindig lényegesen több lapot vettünk figyelembe. Megerősíti ezt az eredményt a II. táblázatban 12 lapról feldolgozott 15 398 hang vallomása (I. lent), mert ennek alapján egy lapra átlag  $15\ 398/12$ , azaz 1283 hang jut, s ez még az 1236-tól 1324-ig terjedő hibahatárának tekintetbevétele nélkül is beleesik az itt 8 lap statisztikája alapján megállapított 1187 és 1283 közötti gyakoriságba.

#### A hangok eloszlása

Abból, hogy ADYNál nem lehet valószínűségi határokon belül maradó hibával megmondani, hogy tíz lapra mennyi betű vagy hang esik, még semmi sem következik a betűk, illetve hangok viszonylagos eloszlásának a pontos-

ságára. Ezt külön meg kell vizsgálnunk. Ragadjuk ki példának TARNÓCZY II. táblázatából az *m*-et. Az összesített statisztika szerint ez a betűk 4%-a, tehát 10 000 betű közül  $400 \pm 20 m$  található. (Nem követünk el hibát, ha a tíz-tíz lapon levő adatokat 10 000 nagyságrendűnek vesszük, az igaz, hogy 11 000 adatra úgy nem 380–420 között, hanem 419–461 darab *m* lenne várható, de ez mindkét esetben 3,8–4,2%-ot jelent, tehát ilyen kis, nagyságrendileg elhanyagolható eltéréseknél nem hibázunk, ha nem a vizsgált adatszámra, hanem annak nagyságrendjére vonatkoztatjuk az eltérések vizsgálatát. Ezzel szemben 10 000 adat esetén a bizonytalanság 1%, 11 000 adatnál csak 0,95%, tehát az eltérést ez is kiegyenlíti.) Ehhez hozzávéve 1% bizonytalanságot, adott esetben 400-nál 4-et, akkor azt kapjuk, hogy 376 és 424, vagy százalékban 3,76 és 4,24% között kellene az *m* betű százalékarányának tíz-tíz laponként lennie. Ezzel szemben 3,57, 3,72, 4,58% a talált százalékarány. Tehát mind a három érték kívül esik a hibahatárokon!

Az <i>m</i> -en kívül még a <i>z</i> esik kívül 3 esetben, összesen ..	2 betű
<i>l, r, k, s, cs</i>	2 „ „ .. 5 betű
<i>o, é, ü, t, b, v, gy</i>	1 „ „ .. 7 betű
	azaz ..... 14 betű

százalékaránya esik kívül a valószínűségi számítás nyújtotta hibahatáron.

Jó értéket ad az *u, a, á, e, i, ö, n, ny, d, p, g, f, sz, zs, j, h, c, ty, dz, dzs*, vagyis 20 betű eloszlási vizsgálata.

Eléggé gondolkodóba ejt az egyezés helyességét illetően az a tény, hogy 34 betűből csak 20-nál, tehát 59%-nál találjuk azt, hogy a részleteredmények a hibahatáron belül megegyeznek a középérték alapján felállított ideális eredményekkel. Hasonlóan nem megfelelő az egyezés az *o-ó, a, á* és *e* hangra az említett táblázatban közölt új és régi magánhangzóstatisztika között sem.

Örvendetes módon sokkal jobb eredményre jutunk, ha nem a tízlaponként nyert eredmények átlagát hasonlítjuk össze a tíz-tíz lapra vonatkozó részleteredményekkel, hanem ha a 30 lapra nyert eloszlási statisztikát összehasonlítjuk azzal a statisztikával, amelyet úgy készíthetünk, hogy TARNÓCZY I. táblázatában az egyes hangoknak átlag 5000 adatig történt vizsgálatából 100 lapra eső kikövetkeztetett értékeit<sup>1</sup> arányba állítjuk egymással, a mássalhangzók összes előfordulási arányát TARNÓCZYNAK a 118. lapon történt megállapítása alapján a hangzóstatisztikában 58,15%-nak véve az összes hangokhoz képest. Így a II. táblázatban látható eloszlási arány adódik.

A II. táblázatban közölt adatokból azt láthatjuk, hogy, ha a 30 lap alapján nyert eloszlást *s* annak hibahatárait kellene irányadóknak vennünk, akkor *o-ó, ny, l, r, p, t, b, g, sz, z, cs* és *ty*, azaz 12 hang eltérései esnének ki a hibahatáron, azaz még így is jobb az eredmény, mint tíz-tíz lap eredményeinek összehasonlításánál. Azonban, minthogy itt két különböző módon nyert (és nem ugyanannak a statisztikának alkatrészeit képező) statisztikáról van szó, módunkban áll feltenni, hogy mindegyikhez  $\pm$  hozzájárulhat a számításból adódó szórási és bizonytalansági érték, tehát az egymástól különböző értékek-

<sup>1</sup> A számításhoz a TARNÓCZY I. táblázatában közölt 100 lapra eső adatszámot használtam fel, annak ellenére, hogy az ugyanott közölt adatszámok és a feldolgozott lapok hányadosaként *ty*-nél kisebb, *r, ny, o*-nál nagyobb adatszám jön ki 100 lapra eső értéknek, mint ott szerepel. Viszont a cikk alapján nem állapítható meg, hogy a lapszám, az egyes hangok adatszáma vagy a száz lapra eső adatszám téves-e.

II. TÁBLÁZAT  
A magyar nyelv hangzóinak összehasonlító eloszlása (%)

Adat	Tarnóczy adatai	Tarnóczy 100 lapról	Saját adataim	Saját 10 lapról	Tolnai* adatai	Mikes adatai	Nemes adatai	Átlag
Adat	33 218		15 398		25 000	238 779	12 000	324 395
ú	1,5		0,45	0,33	0,44	1,19	1,4	1,33
u			0,85	1,01	0,82			
ó	5,3	1,0	1,12	1,01	0,92	4,61	5,0	5,10
o		4,95	4,19	4,50	4,36			
a	9,9		10,71	11,40	9,61	9,37	9,9	9,90
á	3,5		3,58	3,80	3,60	3,70	3,4	3,56
e	10,8		10,28	10,63	10,67	11,21	11,7	10,93
é	3,5		3,48	3,06	3,47	3,84	3,45	3,55
ő		1,37	1,14	1,15	1,27	2,15	2,1	2,06
ö	2,05	0,69	0,75	0,82	0,83			
i	4,5		4,83	4,85	3,86	4,91	4,5	4,73
í			0,47	0,39	0,60			
ü	0,8		0,42	0,43	0,55	0,86	0,8	0,75
ű			0,07	0,11	0,25			
l	5,9	6,3	4,98	4,87	5,95	6,27	6,1	5,84
r	4,4	4,05	4,50	4,56	4,19	4,03	3,95	4,21
m	4,1	4,3	4,28	4,47	4,17	3,69	4,35	4,12
n	5,75	5,8	5,67	5,78	5,49	5,81	5,6	5,46
ny	0,7	0,55	0,69	0,57	0,87	0,68	0,8	0,74
v	2,3	2,3	1,97	1,92	2,20	1,93	1,9	2,06
f	0,9	0,95	0,85	0,79	0,97	0,87	1,0	0,92
z	2,0	2,65	2,38	2,22	2,17	2,93	2,15	2,32
sz	1,85	2,15	2,08	1,98	1,82	1,64	2,1	1,90
zs	0,1	0,08	0,10	0,08	0,06	0,02	0,05	0,06
s	3,8	3,8	3,66	3,43	3,77	4,41	3,35	3,80
j	1,8	1,8	1,50	1,50	1,91	1,65	1,9	1,75
h	1,7	1,8	2,05	2,02	1,69	1,57	1,8	1,76
b	1,8	2,0	1,89	2,05	2,34	1,95	2,35	2,06
p	0,9	0,7	0,70	0,84	0,79	0,94	0,9	0,84
d	2,0	2,0	2,18	1,99	2,30	1,89	2,3	2,13
t	7,15	6,2	7,62	7,00	8,07	7,70	8,0	7,70
g	2,25	2,45	2,68	2,28	2,50	2,52	2,6	2,51
k	5,55	5,75	5,29	5,55	5,16	5,73	4,8	5,30
c	0,3	0,25	0,18	0,18	0,19	0,23	0,2	0,22
cs	0,7	0,5	0,66	0,85	0,59	0,28	0,6	0,56
gy	1,65	1,67	1,61	1,63	1,42	1,35	1,5	1,50
ty	0,2	0,11	0,14	0,07	0,12	0,00	0,05	0,10
Össz.	99,65	msh. 58,16	100,00	100,12	99,99	99,93	100,6	99,77

\* TOLNAI adatait Nyr. XXXV. 422—39 alapján, de csak két tizedes pontossággal idézem, MIKES és NEMES adatait TARNÓCZY cikkéből.

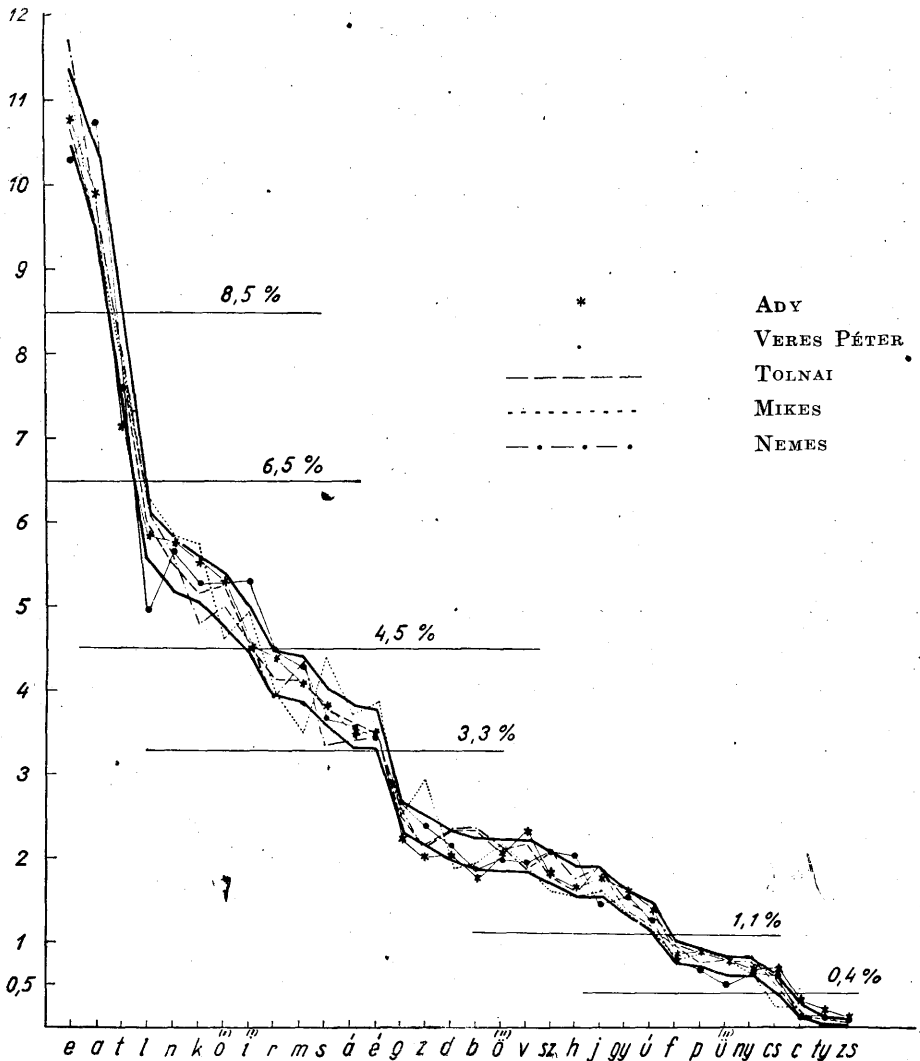
nél a kisebbhez hozzáadva és a nagyobból levonva ezt a javításként vehető értéket, abban az esetben sem tekintendő ellentmondónak két eredmény, ha a bizonytalansági sávuk fedi egymást, vagy éppen még találkozik. A számításokat tízezres nagyságrendű adatokra eső szórásokra elvégezve azt találjuk, hogy így csak a *t*-re kapott értékek nem egyeznek (7,15, illetve 6,2%), azonkívül az *a* és *z* hang százaléktételei vannak még esetleg az egyezés nem-egyezés határán. Ez igen megnyugtató eredmény, különösen ha még azt is tekintetbe vesszük, hogy a *t*-nek mint tárgyragnak és főként mint múltidőjelnek gyakorisága a legnagyobb mértékben a szöveg milyenségétől függ, tehát ez fogható meg legnehezebben statisztikailag. Így arra a megállapításra juthatunk, hogy ADY verseskötete tízlaponként még nem szolgáltathat jó átlagot, de úgy látszik, 30 lap már eléggé, s minthogy az egyes hangok vizsgálata mindig ennél sokkal több lap alapján történt, érdeklődéssel várhatjuk a további kritikai vizsgálat eredményeit. Ezzel egyúttal azt is igazoltuk, hogy helyes a 100 lapra való hanggyakorisági következtetés.

Hasonló módon a Próbátétel 1000—1200 adatig statisztikázott hangjából a tíz lapra eső hangszámot kiszámítva (az 5., 7., 9., 27., 73. és 123. lapot három negyed, a 72., 122. és 154. lapot negyed lapnak tekintve és a 8.-at kihagyva) és ebből eloszlási statisztikát készítve ez az eloszlás a 12 lapról (21—2., 28. 38—9., 67—8, 76—7, 92., 148—9. lap) 15 398 hangból készített statisztikával hibahatárokon belül egyezik, mindössze a *g*-nél van 0,02%-kal nagyobb eltérés, mint a szórás és a bizonytalanság figyelembevétel után megengedett lenne.

Az első ábrán TOLNAI, MIKES, TARNÓCZY, NEMES adatai és saját VERES PÉTERRE vonatkozó adataim középértékéből nyert előfordulási gyakorisághoz tartozó tízezres nagyságrendben adódó szórási értékek és bizonytalanság alapján kiszámított bizonytalansági sávot láthatjuk a hangok csökkenő gyakoriságának sorrendjében. Az egyes szerzők eredményeit különböző jelekkel feltüntetve, világosan látható, mely adatok vannak hibahatáron belül és melyek esnek ezen kívül. Azért nem vettem az egyes szerzők vizsgált adatszámának eltérését figyelembe, azaz azért nem végeztem súlypontos középértékszámítást, mert a feldolgozás szempontjai mindenütt különbözők voltak, s így ez nem adna reális eredményt.

Vastagon kihúzva azt a sávot látjuk, amelyen belül TOLNAI, MIKES, NEMES, TARNÓCZY és saját statisztikám alapján kiszámított középértékek alapján az egyes hangzók előfordulási gyakoriságának esnie kellene (*Unsicherheitszone*). A vízszintes tengelyen a magyar hangzók található gyakorisági sorrendben a rövid és hosszú magánhangzókat együtt számítva. A csillagok az ADYRA, a pontok VERES PÉTERRE vonatkozó eredményeket jelölik. Az *a* hangnál az alsó pont a szavakban, a felső pont a szavakban és az *a* különálló névelőben előforduló *a* hangok együttes százalékszámát jelenti. Ha TOLNAI, NEMES, MIKES eredményeit is különböző jelekkel vezettük volna be a rajzba, akkor az egész áttekinthetetlenné vált volna. Ezért inkább folytonos, egymástól különböző jelölésű vonalakkal jelöltük az egyes szerzők eredményeit. Ebből zavar nem származhat, mert az ábra értelmezése lehetetlenné teszi azt, hogy ezekhez a vonalakhoz másutt jelentést társítsunk, mint éppen az egyes hangoknál, ezzel szemben az az előny származik a vonalak használatából, hogy —ha két vagy több vonal éppen egy hang felett metszi egymást — abból világosan látható, hogy annál a hangnál két vagy esetleg több szerző azonos eredményre jutott.

A VERES PÉTERRE vonatkozó adatokból azt láthatjuk, hogy a magyar nyelvi átlagnál jóval kevesebb *l*-et, sokkal több *i*-t (a sok *is* és gyakori főnévi igenév miatt), több *h*-t (a sok *hogy* miatt) és kevesebb *ü*, *ű*-t használ. A kiugró értékek közül ellenőrzésképpen az *l*-et és az *i*-t átszámoltam másodszer is, de



1. Hangzóeloszlási grafikon

az ellenőrzés, valamint az egyes hangok 1000–1200 előfordulásából számított átlageloszlás is azt igazolta, hogy VERES PÉTER nyelvében csakugyan léteznek ezek az átlagnál ritkább, ill. gyakoribb hanghasználatok. Az *a* és *e* gyakoriságával kapcsolatban TARNÓCZY MNy. XXXVIII, 355–6, XXXIX, 370–1

utalunk, ahol már szó van *a* és *e* gyakoriságának a modernebb íróknál történő cseréjéről.

Az ADYra vonatkozó számadatok 9 helyen lépnek ki a bizonytalansági sávból, de sehol sem annyira, hogy ezt érdemes volna különösebben megvizsgálni.

MIKES irodai és NEMES ugyancsak minden bizonytal nem nyelvészeti-irodalmi szövegeket feldolgozó százalékarányai több helyen kívül esnek a bizonytalansági sávon, ez különösen feltűnő MIKESnél. Megkíséreltem TARNÓCZY ötlete alapján MIKES statisztikáját betűstatisztikának tekinteni és ebből a TARNÓCZY ADYNál kapott eredményei szerint minden betűt hangra átszámítani. Ezek az értékek helyenként a rajzon feltüntetetteknek jobbnak, másutt viszont sokkal rosszabbnak bizonyultak, s ezért ábrázolásukat elhagytam.

Az öt statisztika alapján felállított nyelvi átlagot legjobban TOLNAI nem nagyon nagy terjedelmű, de gondosan válogatott, több írótól származó, vegyes tartalmú szövegekből készült statisztikájának értékei érik el. Mindössze csak a *b* hangnál találunk nála valamivel nagyobb előfordulási arányt. Ezzel igazolva látjuk azt a TARNÓCZYval történt viták során sokszor kifejtett nézetünket, hogy statisztikából nyelvi törvényekre jobban tudunk következtetni, ha pl. tíz író dolgozunk fel kevés hangra, mint ha ADY hangeloszlását hangonként 5000-ig vizsgáljuk.

Az öt különböző statisztika eredményét egybevetve azt láthatjuk, hogy nem annyira a hangok, mint inkább egyes hangcsoportok gyakorisági sorrendjét lehet megadni. Leggyakrabban, mindenütt 9,5% fölött, fordul elő az *a* és *e*. A következő igen széles sávba csak egy hang jut, a *t*. 6% alatt már nagyobb csoportokat találunk. *l*, *n*, *k*, *o*-ó és *i*-í eddig minden vizsgálatnál 4,5%-ot meghaladó értéket mutatott. A következő hangcsoportba 3,3%-nál mindig sűrűbbnek bizonyult, de 4,5%-ot sohasem meghaladó hangok tartoznak: *r*, *m*, *s*, *á*, *é*. A következő meglehetősen nagy csoport 1,2%-nál gyakoribb, de az előző csoport alsó gyakorisági határán eléggé alul marad: *g*, *z*, *d*, *b*, *ö-ő*, *v*, *sz*, *h*, *j*, *gy*, *u-ú*. Némi frekvenciaesés után következnek a ritka, de azért 0,4%-nál még gyakoribb hangok: *f*, *p*, *ü-ű*, *ny*, *cs*. A legritkább hangok: *c*, *ty*, *zs* (*dz*, *dzs* annyira ritka, hogy fel sem tűntethetjük). Feltehetjük az eddigi, százezreket felölelő statisztikázási eredmények alapján, hogy egyes írók, tárgyak szerint a hangok gyakorisága változhat, de minden hang megmarad abban a gyakorisági csoportban, amelybe a grafikon alapján itt soroltuk. Talán *i-í* és *r* esetében nem állíthatjuk ezt teljes biztonsággal (v. ö. a rajzon található eredménnyel), de a többi hangra igen valószínűnek látszik. Ez természetesen nem jelenti azt, hogy szándékosan nem lehet valamely hang gyakoriságát erősen csökkentő vagy növelő, a talált statisztikától szándékosan eltérő szövegmintát készíteni (v. ö. TRUBECKOJ TCLP. 1939. 7. 233.).

#### A kezdő és véghangok eloszlása

A III. táblázatban az ADY és a VERES PÉTER vizsgálatából adódó szókezdő és szóvégi hangzókapcsolati arányt, valamint a különálló hangok arányát hasonlítjuk össze. Az eltérések szókezdeten is és szóvégződésnél is szembeszökően nagyok. Szókezdeten csak *cs*, *e*, *é*, *f*, *h*, *j*, *k*, *n*, *ny*, *ő*, *p*, *t*, *ty* és *u*, azaz 14 hang, szóvégen meg éppen csak *cs*, *é*, *f*, *g*, *l*, *ny*, *ő* és *ty*, azaz csak 8 hang (az elő nem forduló szóvégi *h*-val együtt 9 hang) százalékarányai felelnek meg egymásnak a valószínűségszámítási hibahatáron belül. Nem áll módunkban az eddigi vizsgálatok alapján arra a kérdésre felelni, vajon 3–5000 szó kezdetének, illetve végződésének vizsgálata kevés-e ahhoz, hogy ebből a



## III. TÁBLÁZAT

Szókezdő, szóközépi és szóvégi hangzóeloszlás és egyhangzóú szavak (%)

	Szó elején		Szó- középén	Szó végén		Min. 2 hangú szavak	Egyhangzóú szavak	
	Ady	Veres	Veres	Ady	Veres	Veres	Ady	Veres
Adat	5120	2964	9171	5120	2964	15 091	454	307
ú								
u	2,07	1,07	0,40	0,53	0,03	0,46		
ó	0,31	0,84	1,15	1,33	0,03	0,87	4,8	—
o	0,47	0,10	1,30	—	1,73	1,15		
a	5,90	1,45	6,56	—	—	4,27		
á	5,90	9,05	9,17	7,48	8,60	9,03	63,2	93,7
e	1,90	0,50	5,69	0,18	0,54	3,66		
é	5,72	6,29	13,16	5,05	6,26	10,48	2,6	—
ö	5,31	4,85	4,17	0,27	0,31	3,56		
ő	1,35	0,44	1,17	—	—	1,16		
i	0,25	0,31	0,72	0,82	0,88	0,67	2,6	4,6
í	3,72	5,55	4,54	—	5,52	4,92		
ü		0,74	0,55	4,77	—	0,48		
ű	0,43	0,10	0,67	0,49	—	0,42		
		—	0,09		0,10	0,07		
l	3,70	2,55	6,14	4,15	4,40	5,08		
r	2,16	1,31	5,82	3,13	4,14	4,59		
m	11,65	12,55	1,88	7,33	3,79	4,37		
n	4,93	4,55	5,52	10,20	7,85	5,78		
ny	0,37	0,44	0,88	0,62	0,46	0,70		
v	6,38	5,72	1,45	0,10	0,03	2,02		
f	3,84	3,77	0,20	0,04	0,03	0,87		
z	0,37	0,17	2,46	3,52	4,57	2,42		
sz	5,32	3,30	2,20	1,60	0,91	2,12		
zs	0,14	0,13	0,12	0,02	0,03	0,11		
s	2,85	2,45	2,21	6,99	9,70	3,70	26,9	1,6
j	2,01	2,05	1,64	1,41	0,71	1,53		
h	6,65	7,14	1,13	—	—	2,09		
b	3,73	3,23	1,98	1,17	0,44	1,92		
p	1,58	1,65	0,60	0,60	0,10	0,71		
d	1,48	1,85	2,64	2,77	1,29	2,22		
t	4,63	5,08	5,26	12,13	18,38	7,79		
g	0,57	1,28	2,91	3,89	3,66	2,74		
k	6,98	6,74	3,12	15,75	11,15	5,40		
c	0,14	0,03	0,23	0,04	0,17	0,18		
cs	2,20	2,35	0,28	0,21	0,20	0,68		
gy	0,71	0,40	1,17	3,54	4,37	1,64		
ty	0,01	0,03	0,22	0,02	—	0,14		
Össz.	99,83	100,06	99,40	100,15	100,38	100,00	100,1	99,9

Tekintettel arra, hogy a kételemű hangzókapcsolatok legnagyobb része csak a szóközépen szereplő hangok viszonyaira vonatkozik, továbbá minthogy az egyhangzóú szavak gyakorisága a hangzókapcsolatokat nem befolyásolja, nem látszik fölöslegesnek a csak szóközépi hangzóeloszlási arány és az egy hangzónál hosszabb szavak összes hangzóira vonatkozó eloszlási arány megadása. Az *a* hang gyakorisága sokkal kisebb, ha a névelőt nem vesszük figyelembe, az *e* gyakorisága alá esik, szóközépen még nagyobb a különbség az *e* javára.

hangonkénti eloszlás százalékarányát megállapíthatjuk, vagy pedig a stílus, a tárgy stb. annyira befolyásolja a szókezdeteket és a szóvégződéseket, hogy csak nagyságrendileg sokkal nagyobb számokra vonatkozó vizsgálatok vezethetnek esetleg megnyugtatóbb eredményekre.

Azt az érdekességet azonban megállapíthatjuk, hogy a közepes gyakoriságú *g* és *l* kivételével a többi statisztikusan egyező szóvégek, mind amúgy is ritka hangok, meglehetősen ritka szóvégzódéseire vonatkoznak.

Az eltérések közül négynek viszont magyarázatát is tudjuk adni. A szóvégi *m* és *k* gyakorisága ADYNál az egyes- és többszám első személyének mind az ige-, mind a birtokos személyragozásban való rendkívül gyakori előfordulása miatt adódik, a szóvégi *s* hang ritkasága az *és*-nek *s*-sel, tehát különálló szóval való helyettesítéséből származik. A szóvégi *t*-nél, VERESnél megszámlálva a múlt időt jelölő *-t*, *-tt* szóvégeket a hangstatisztikához feldolgozott részen, ez a szóvégek 5,28%-a, tehát így 18,28% helyett csak 13,10% lenne a szóvégi *t* (persze így az összes többi szóvégen eloszlaná az arányszámokat növelve a levett 5,28%) s ez lényegesen közelebb állna az ADYNál talált eredményhez.

### Hangcsoportok eloszlása

A IV. táblázatban TARNÓCZY IV. táblázatához hasonló módon a könyv nyolc lapjáról (19., 20., 30., 31., 106., 107., 150., 151.) származó adatok alapján összesen 9879 hang eloszlási arányát mutatjuk meg. Minthogy TARNÓCZY

#### IV. TÁBLÁZAT

##### Hangzóeloszlás egyes szavakban (adatszám)

Hangzókapcsolatok	Külön- állók	Szó elején	Szó közben	Szó végén	Összesen		
Magánhangzók	<i>v</i>	182	606	2880	456	4124	4176
	<i>vv</i>	—	1 →	21	← 4	52	
	<i>vvv</i>	—	—	—	—	—	
Mássalhangzók	<i>k</i>	3	1345	1117	1199	3664	5703
	<i>k̄</i>	—	—	264	125	389	
	<i>kk</i>	—	8 →	600	← 175	1566	
	<i>kkk</i>	—	—	27	← ← 1	84	
Összesen	185	1960	5774	1960	9879		

→ azt jelenti, hogy a jelzett számú hangot egy nyíl esetén egyszer, két nyíl esetén kétféle szorozva hozzá kell adni a szó közepén előforduló hangok számához, hiszen két, illetve három hang közül csak 1 számítható szorosabb értelemben véve szókezdő, illetve szóvégző hangnak.

ADYra vonatkozó eredményei is és az itt közölt, VERES PÉTER nyelvéből nyert eredmények is abszolút számokban vannak megadva, az összehasonlítást lépésről-lépésre és százalékarányban végezzük el.

Először nézzük meg a magánhangzók és mássalhangzók arányát:

	<i>Magánhangzó</i>	<i>Mássalhangzó</i>	<i>% összeg</i>
ADY .....	41,25	58,75	100%
VERES PÉTER	42,3	57,7	100%

Az egyezés annyira jó, hogy még a szórási határ figyelembevétele nélkül, a bizonytalansági határon belül is jó. Az egy és két magán-, illetve mássalhangzós szókezdetek arányának a vizsgálatánál ADY nyelvére talált eredményeket oly módon is arányba állíthatjuk, ahogy azok TARNÓCZYNÁL szerepelnek, de figyelembe vehetjük azt is, hogy ADYNál a versben *s* kötőszó szerepel majdnem minden olyan esetben, amikor prózáiról és kötőszót használna. A táblázatban található 150 *s* kötőszót és-sel helyettesítve megnövekedik ADYNál is a szókezdő magánhangzók száma és közelebb esik a VERES PÉTERNél található (rövidítés: mh = magánhangzó, msh = mássalhangzó):

	<i>mh</i>	<i>msh</i>	<i>2 msh</i>	<i>2 mh</i>	<i>% összeg</i>
ADY (s)	26,81	72,55	0,6	—	99,96%
ADY (és)	29,90	69,80	0,6	—	100,3 %
VERES	30,95	68,60	0,4	0,05	100,—%

Az eltérések, akár az *s* mint önálló szó elhagyásával, akár és-sel való helyettesítésével számított sort nézzük, ADY és VERES PÉTER között, feltétlenül a szórási határon belül mozognak.

A szóvégződésnél hasonló módon járhatunk el. ADY *s* kötőszavát mint különálló hangot elhagyhatjuk, de és szónak számítva, szaporíthatjuk vele a mássalhangzós végződéseket.

	<i>mh</i>	<i>2 mh</i>	<i>3 mh</i>	<i>msh</i>	<i>Kettőzött msh</i>	<i>2 msh</i>	<i>3 msh</i>	<i>% összeg</i>
ADY (s)	18,50	0,55	0,05	67,10	4,80	9,10	0,05	100,15
ADY (és)	17,79	0,55	0,03	68,25	4,69	8,64	0,03	99,98
VERES	23,22	0,20	—	61,20	6,37	8,94	0,05	99,98

Érdekes módon szóvégződéseknél ADY *s* szavainak és-sé alakítása nem közelebb, hanem távolabb viszi egymástól ADYra és VERES PÉTERRE talált arányokat, bár még evvel a távolodással is tökéletesen megegyeznek az adatok a valószínűségszámítás által megadott hibahatárokon belül.

Hasonlóképpen jó egyezést kapunk, ha az egy, két és három magánhangzóból álló csoportokat hasonlítjuk össze :

	<i>Egy mh</i>	<i>Két mh</i>	<i>Három mh</i>	% összeg
ADY	97,28	2,68	0,04	100%
VERES	98,75	1,25	—	100%

A két magánhangzóból álló csoport nagyobb arányszáma az ADYNál elég gyakori birtokos személyragozásból adódik (több birtok). (Eredményeink a TARNÓCZYNál közölttől csak látszólag térnek el, ő pl. *-ei-t* I csoportnak, én két hangnak számítottam.)

A mássalhangzók arányánál az egyes és ikerített mássalhangzókat az ejtési ingadozásra való tekintettel együtt is arányba állítjuk :

	$(msh + \overline{msh}) = msh + \overline{msh}$	<i>2 msh</i>	<i>3 msh</i>	% összeg
ADY.....	66,68 + 4,90 = 71,58	27,15	1,27	100%
VERES .....	64,30 + 6,80 = 71,10	27,70	1,45	100,25%

(A számításnál az előforduló hangokat és nem a helyzeteket vettük figyelembe, tehát pl. a *kertben* szóban kétszer egy és egyszer három mássalhangzó szerepel, ebben 2 : 3 az egy és három msh aránya.)

Nem érdektelen megnézni, hogy intervokális helyzeteket tekintve hányszor találunk két magánhangzó között egy, kettőzött, két vagy három mássalhangzót. Ebben az esetben a két, illetve három mássalhangzót egy csoportnak tekintjük.

	<i>1 msh</i>	<i>Ikerített msh</i>	<i>2 msh</i>	<i>3 msh</i>	<i>A számított intervokális helyzetek száma</i>
ADY.....	57,90	9,80	31,0	1,23	3581 (99,93%)
VERES .....	55,40	13,70	29,70	1,33	2008 (100,13%)

VERES PÉTER igen gyakran használt múlt ideje még ebből az összehasonlításból is kiviláglik az intervokális ikerített mássalhangzók aránylag nagyobb előfordulási számából.

Érdekes azon az alapon a nyílt és a zárt szótagok számát is saját IV. táblázatomból és TARNÓCZY IV. táblázatából összehasonlítani, hogy nyílt szótag magánhangzós szóvégén, két magánhangzó között (három magánhangzó között kétszer is) és intervokálisan szereplő egy mássalhangzó előtt van, ezzel szemben mássalhangzóra végződő szóvégeken kívül minden geminátában két, illetve három mássalhangzót tartalmazó intervokális helyzetnél zárt szótag végét találjuk. Ezt figyelembe véve :

	Nyílt szótag	Zárt szótag
ADY.....	41,85	58,15
VERES .....	42,65	57,35

egymással igen jól megegyező arányokat kapunk.

Különös figyelmet érdemel az a tény, hogy

	ADYNál	VERESnél
<i>mh</i> : <i>msh</i> .....	41,25 : 58,75	42,3 : 57,7
<i>nyílt</i> : <i>zárt szótag</i> .....	41,85 : 58,15	42,65 : 57,35

vagyis ugyanabban a szövegben a magánhangzó-mássalhangzóarány csaknem azonos a nyílt és zárt szótagok arányával.

TARNÓCZYNak IV. táblázatából és az itt közölt, VERES PÉTER nyelvére vonatkozó hangzóeloszlási táblázatból azt láthatjuk, hogy a magánhangzók, mássalhangzók s ezek kettes és hármas csoportokat alkotó gyakoriságának aránya nem függ az egyéni nyelvhasználatától, hanem a nyelvre jellemző módon meglehetősen állandó eloszlást mutat.

Annak ellenére, hogy a Próbátételben a hangok laponként még nem szerepelnek elég nagy mennyiségben ahhoz, hogy statisztikus eloszlást mutassanak, a IV. táblázat adatai már csaknem laponként is igazolhatók. Az egy hangból álló szavak már laponként is jól oszlanak meg, a közvetlenül egymást követő mássalhangzók szintén. Kedvező a helyzet a szóvégen is, a magánhangzóra és a két mássalhangzóra végződő szavak eloszlása hét, az egy és a kettőzött mássalhangzóra végződőké a nyolc lapból haton a kívánt hibahatáron belül van. A szókezdeten a mássalhangzós szókezdés hat, a magánhangzós szókezdés négy lapon felel meg a valószínűségi követelményeknek. A legkevésbé a szóközépi helyzetek eloszlása statisztikus laponként, mert a szóközépi kettőzött mássalhangzók hat, magánhangzók csak négy, egy és két mássalhangzóból álló szóközépi csoportok viszont csak három lapon oszlanak meg statisztikusan. Tehát a szóközépen, mely a legrendezetlenebb eloszlást mutatja, kapunk nyolcból három lapon jó eredményt, úgy mint a hangeloszlásnál, a többi helyen mindenütt jobbat. Nem beszéltünk a hármas mássalhangzók részeloszlásáról. Összesen 28 hármas mássalhangzócsoporthoz, azaz 84 mássalhangzó van nyolc lapon hármas kapcsolódásban. Ez olyan ritka előfordulás, hogy ennél nem csodálható, hogy szó belsejében 0, 8, 6, 1, 0, 0, 3, 9 a nyolc lapon az eloszlás, szóvégen egyszer szerepel.

Az ADYNál megvizsgált 18 118 hang közül 270 magánhangzó és 150 mássalhangzó alkotott önálló szót, ezzel szemben a VERES PÉTERNél megvizsgált 9879 hangból 182 magánhangzó és csak 3 önálló szót alkotó mássalhangzót találtunk. Ez a különbség valószínűleg jellemző a költői és a prózai nyelvre. A vizsgált hangok száma ADYNál csaknem kétszer annyi, mint VERES PÉTERNél, az önálló szót alkotó magánhangzók száma mégis csak a felével több a versben, mint a prózában, annak ellenére, hogy a költői nyelvben több, a prózában alig használt egy magánhangzóból álló indulatszó szerepel, viszont versben a határozott névelő, gyakran a szórend választékosabb fordulatában elmaradhat, ezzel szemben, mint már említettük, és helyett a prózában alig használt s gyakrabban szerepel.

A hangok számát a szavak számával osztva (ezt a szókezdetek, illetve szóvégzések számából + önálló szavakat alkotó hangok számából kapjuk

meg) ADYNÁL 4,62, VERESNÉL 4,60 hang egy szó átlagos szóhosszúsága. Ha az egy hangból álló szavakat nem vesszük figyelembe, akkor ADYNÁL 5,06, VERES PÉTERNÉL 4,95 hang az átlagos szóhosszúság. A szavak hosszúságának egybevetésére a következő táblázattal kapcsolatban még vissza kell térnünk.

TARNÓCZYHOZ HASONLÓ MÓDON A PRÓBATÉTEL STATISZTIKÁJA ALAPJÁN IS MEGÁLLAPÍTHATJUK A HASZNÁLT SZAVAK ÁTLAGOS SZÓTAGSZÁMÁT. Ha az összes magánhangzók számát osztjuk az összes szavak számával, akkor 1,95 (ADYNÁL 1,98) szótag, a különálló hangok leválasztásával 2,04 (ADYNÁL 2,05) szótag a használatban az egyes szavak átlagos szótagszáma.

Az átlagos szótaghosszúság (összes hangok száma osztva az összes magánhangzók számával) VERES PÉTERNÉL 2,36, illetve a különálló hangok nélkül 2,42, ADYNÁL VISZONT 2,44, illetve 2,46 hang.

*Szótagtípusok. A szavak hosszúsága*

V. TÁBLÁZAT

*Szótagtípusok eloszlása (adatszám)*

Szótagtípusok	v	lv	2v	v1	lv1	2v1	v2	lv2	Összesen	
Egyszótagú	182	79	—	280	280	1	14	87	923+1*	
Kétszótagú	1.	81	267	2	83	217	—	11	9	670
	2.	—	215	—	7	402	—	1	45	670
Háromszótagú	1.	40	157	2	45	94	2	3	1	344
	2.	6	184	—	9	143	—	—	2	344
	3.	1	88	—	—	242	—	—	13	344
Négyzótagú	1.	22	66	—	15	43	1	—	3	150
	2.	4	86	—	4	55	—	—	1	150
	3.	2	82	—	6	60	—	—	—	150
	4.	3	55	—	—	83	—	2	7	150
Ötszótagú	1.	6	23	—	3	14	—	—	—	46
	2.	—	26	—	2	18	—	—	—	46
	3.	2	26	—	4	14	—	—	—	46
	4.	2	29	—	—	14	—	—	1	46
	5.	—	14	—	1	31	—	—	—	46
Többszótagú	4	29	—	2	14	—	—	1	50	
Összesen	355	1426	4	461	1724	4	31	170	4176	

\* A pénz szó lv3 alakú.

Az V. táblázatban TARNÓCZY V. táblázatához hasonló módon oszlopokban egymás alatt az azonos típusú szótagokat adjuk meg, az ottani beosztással azonos módon a magánhangzót *v*-vel,<sup>1</sup> előtte, illetőleg utána arab számmal jelölve a található mássalhangzók számát. Az egymás alatti sorokban az egy-, két-, három-, négy-, öt- és többszótagú szavak szótagtípusai találhatóak meg. Soronként azt is megállapíthatjuk, hogy valamelyik szótagtípus a többszótagú szó első, utolsó vagy valamelyik középső szótagjában fordul-e gyakrabban elő.

Az ADYra és a VERES PÉTERre vonatkozó eredmények között érdekes egyezéseket és eltéréseket találhatunk.

Összehasonlítva az	1	2	3	4	5	6	7	8 tagú	szavakat
ADYNál . . . .	1330	1492	683	225	29	5	—	3	összesen 3767
VERESnél . .	924	670	344	150	46	7	—	1	« 2142 van

A két eloszlási arány még abban sem egyezik meg, hogy melyik a leggyakoribb szótagszámú szótípus. ADYNál a kéttagúak száma több, mint 10%-kal nagyobb az egytagúakénál, ezzel szemben VERES PÉTERNél nem haladja meg sokkal az egytagúak számának  $\frac{2}{3}$ -át. Míg ADYNál a három szótagúak alig tesznek ki többet a kéttagúak egyharmadánál, VERES PÉTERNél még a felére sem esik le. Annak ellenére, hogy VERES PÉTERNél vizsgált szavak száma az ADYNál vizsgáltaknak csak 56%-a, mégis több öt- és hattagú szót találunk. Abban megegyeznek, hogy nyolctagú szó VERES PÉTERNél is, ADYNál is évszám. Ábrázolva VERES PÉTERNél mintegy logaritmikusan esik a szavak gyakorisága a növekvő szótagszámmal, ADYNál viszont a kéttagúaknál van a frekvenciagörbe maximuma. Arra kell gondolnunk, hogy ennek a nagy eltérésnek a próza és vers ellentéte az oka. Hiszen a kötöttebb ritmusú vers nyilván lehetőleg kerül a hosszabb szavakat. Az egytagúak ritkább volta a versben, mint azt már az önálló hangokkal kapcsolatban is láttuk, a versben lehetséges névelőelhagyásból származhat.

A kérdés tisztázására kb. kétezer szóig megvizsgáltam ARANY, PETŐFI, GÁRDONYI nyelvét és ADY prózáját. A feldolgozott anyag GÁRDONYINál: A láthatatlan ember (Bp. 1915, Singer és Wolfner 84—5., 258—9. lap, Egri csillagok, Bp., 1909 Singer és Wolfner II. kötet, 94—7., 224—5. lap. ADY: Jólások Magyarországról c. kötetéből Julián magyarjai, Goga Octavián rab-sága, Európa Kairója és Istenek alkonya; ARANY: Rege a csodaszarvasról, Toldi I. éneke, Szibinyáni Jank, V. László; PETŐFI: János vitéz I., II. éneke, Bolond Istók 1—37. versszak, Kiskúnság, Nemzeti dal és Az örült. A rege a csodaszarvasról a feldolgozott szövegek közül az egyetlen, ahol ADY költészetéhez hasonlóan több a kéttagú szó, mint az egytagú, ARANY JÁNOS egyéb vizsgált verseiben szinte azonos az egy- és kéttagúak száma, a többi író és költő vizsgált szövegeiben több az egytagú, mint a kéttagú szó. Különösen érdekes Az örült szövegeinek eloszlása, mert 172 egytagúval szemben csak 83 kéttagú szó van, a háromtagúak száma viszont 94. A vizsgált szövegek közül ez az

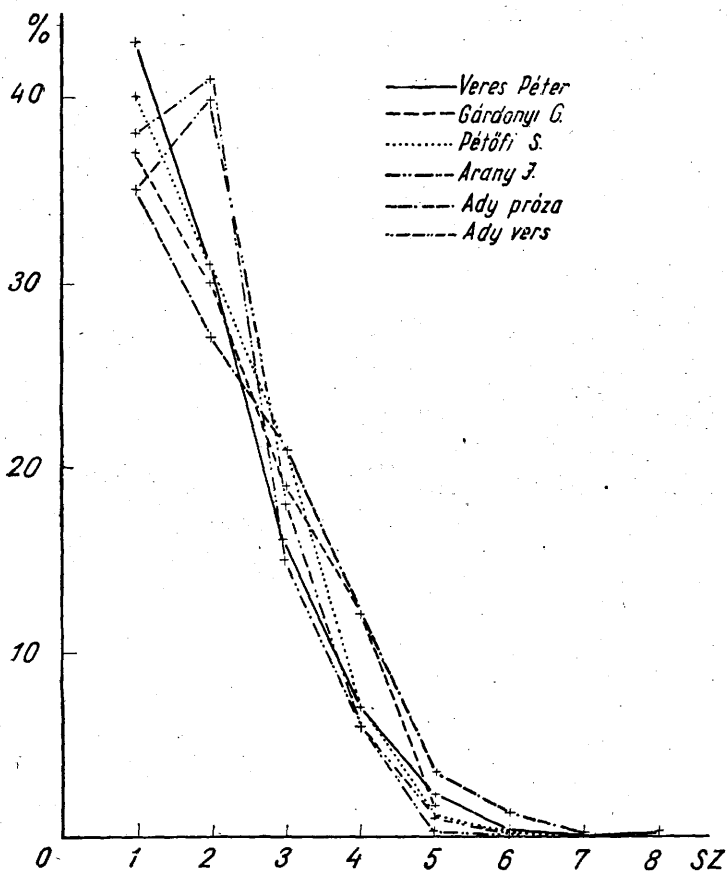
<sup>1</sup> Tekintettel arra, hogy az *a* névelő többször szóba kerül, félreértések elkerülése végett helyesebbnek látszik visszatérni MENZERATH *v* (*vokális*) jelöléséhez TARNÓCZY *a* magánhangzó jele helyett, hasonlóképpen *k* (*konzonáns*) rövidítést alkalmazok a TARNÓCZYNál előforduló *b* mássalhangzó jelölés helyett.

egyetlen, ahol több a háromtagú szó a kéttagúnál. Az eredményeket százalékban közölve a következő eloszlást kapjuk:

VI. TÁBLÁZAT

Író	Adat- szám	1 tagú	2 tagú	3 tagú	4 tagú	5 tagú	6 tagú	7 tagú	8 tagú	9 tagú
VERES P. ....	2143	43%	31%	16%	7%	2,2%	0,3%	—	0,05%	—
GÁRDONYI ....	2298	37%	30%	19%	12%	1,7%	0,4%	0,05%	—	—
ADY próza ....	2017	35%	27%	21%	12%	3,5%	1,3%	0,2%	—	0,05%
ADY vers ....	3767	35%	40%	18%	6%	0,8%	0,13%	—	0,1%	—
ARANY .....	2212	38%	41%	15%	6%	0,1%	0,05%	—	—	—
PETŐFI .....	2002	40%	31%	21%	7%	1%	0,3%	—	—	—

Feltűnő, hogy aránylag ADY prózájában szerepel a legtöbb hosszú szó. A fenti táblázat és az ezt ábrázoló 2. ábra arra enged következtetni, hogy a



2. A szavak szótaghosszúság szerinti eloszlása néhány írónál



prózában csakugyan több az egy-, mint a kéttagú szó, tehát ilyen szempontból nem ad ADY költészetének vizsgálata jó képet a magyar nyelv sajátosságairól. Nem célunk itt eldönteni, hogy a versben a versmérték, a hangulat és még több más számbavehető tényező befolyásolja-e és mennyire az előforduló szavak szótagszámát. Annyi azonban eddigi vizsgálatainkból is kiderült, hogy a sormetszet és az előforduló legnagyobb szótagszám között szoros összefüggés van.

A franciában, figyelmen kívül hagyva a technika és a tudomány nyelvét, A. DAUZAT megfigyelése szerint (Le génie de la langue française, 1944, Paris, 62.) a 2–7 hangú, kétszótagú szó a leggyakoribb. Ha a szóvégi néma *e*-t nem számítjuk, akkor utána az egy-, majd a háromszótagúak következnek, a szóvégi néma *e* figyelembevételével még több a kéttagú szó, majd a háromtagúak következnek és csak a harmadik helyen az egytagúak. A négytagú szavak ritkák, öt- és ennél többtagúak csak elvétve fordulnak elő. Ez a megállapítás nem hasonlítható össze a magyar írók nyelvéből adódó eredményekkel, mert DAUZAT csak névszókat és igéket figyelt meg, tehát az egytagú szavak zömét nem számította. Kijelentése különben is becslésen és nem statisztikán alapul.

VERES PÉTER nyelvében a szavak szótagszám szerinti megoszlása már szinte laponként meglehetősen állandó, a vizsgált nyolc lapról az egytagúak 6, a kéttagúak mind a nyolc, a háromtagúak hét, a négytagúak szintén nyolc és az öttagúak öt lapon már a statisztikus valószínűséget kielégítő módon oszlanak meg.

Az ADY-ra vonatkozó vizsgálatokhoz hasonlóan a Próbátétel nyelvében is a legtöbb az *lv1* típusú szótag, a következő az *lv* típusú. Az egyes szótag-típusoknak az ADYNál és VERESNél talált számát arányba állítva típusok szerint a következő eredményre jutunk :

	<i>lv1</i>	<i>lv</i>	<i>vl</i>	<i>v</i>	<i>lv2</i>	Összes szavak száma	Összes szótagok száma
ADY.....	3311	2424	711	621	352	3767	7362
VERES .....	1724	1426	461	355	170	2143	4177
Hányados ....	1,915	1,705	1,55	1,75	2,06	1,758	1,79

Az arányszámok meglehetősen egyeznek, a ritkább szótagtípusoknál amúgy sem várhatunk egyezést, ezért nem is nézzük ezek arányszámát.

Ugyanazokat a »rendellenességeket« tapasztaljuk VERES PÉTER nyelvében az egyszótagúakkal kapcsolatban, amelyekre TARNÓCZY (124.) ADY vizsgálatánál mint különös érdekességekre felhívta a figyelmet. Feltűnik az egytagúaknál a többszótagúakhoz viszonyítva a gyakoriságával a *vl*, a *v* és az *lv2* szótagtípus. Ezt azonban nem tekintjük rendellenességnek és nem tulajdonítunk ennek nyelvészeti fontosságot, mint TARNÓCZY (u. o.), hanem ez a statisztikai eredmény természetes következménye annak, hogy a határozott és a határozatlan névelő, az és egy, valamint az és kötőszó gyakori szereplése a *vl* típust, az *a* névelő viszont a *v* típust duzzasztja meg, a legtöbb gyakran használt egytagú szó pedig vagy az amúgy is gyakori *lv1* típusú vagy *lv2* típusú. Minthogy a mássalhangzó + magánhangzó kapcsolatokból nem hasz-

nál fel mindent a nyelv szóalkotására, az összes lehetséges esetek száma 23 mássalhangzót és 14 magánhangzót véve számítási alapul, 23.14, azaz 322 lenne, ennyi *lv* típusú szavunk pedig nincs, a meglevők sem annyira gyakoriak, mint a névelő, így érthető, sőt várható is, hogy az egyszótagúakban ritkább az *lv* típusú szó, mint többtagú szavak elején, ahol még végtelen sok további kapcsolódási lehetőség van a 322-féle kezdethez, tehát sok további lehetőségből válogathat a nyelv.

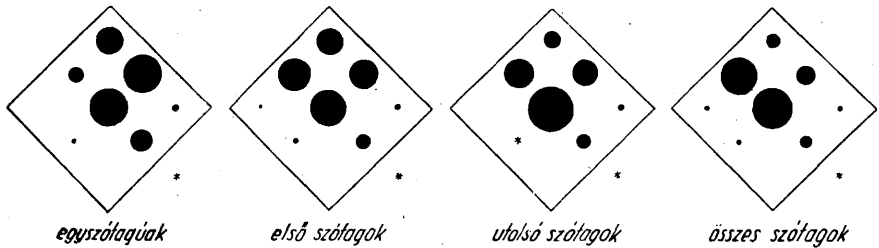
Nem érthetünk egyet TARNÓCZYVAL abban sem, hogy az V. táblázat adataiból feltárt érdekesség a *v* és *vl* szótagtípus fokozatos csökkenése és ezzel párhuzamosan az *lv1* és *lv2* szótagtípus szaporodása a növekvő szótag-számmal. Ez a megállapítás semmiesetre sem engedi sejtetni „hogy a nyelvi statisztika módszereivel a nyelv természetének eddig nem ismert mélységeibe pillanthatunk be” (125.), hanem ez a jelenség egyrészt a magyar szótagolás, másrészt a ritkán előforduló egymást követő magánhangzó-előfordulás következménye. Hiszen nem első szótagban csak akkor találunk *v*, *vl*, azaz magánhangzó kezdetű szótagtípust, ha a szóban két magánhangzó következik egymás után. Ez pedig a több birtokra utaló birtokos személyragozáson kívül magánhangzóra végződő szavak *-ért*, *-ig* ragos és *-é* jeles alakjain kívül (ezek sem nagyon gyakoriak) többnyire csak magánhangzóra végződő szónak vokális kezdetű szóval alkotott összetételénél fordul elő és idegen szavakban, minthogy a *kiált*, *siet*-féle szavak a magyarban elég ritkák. Tekintettel arra, hogy az V. táblázat TARNÓCZYVAL is, itt is ugyanannak a szövegrésznek a feldolgozása, mint a IV. táblázat, pontosan ugyanannyi *v* és *vl*, *v2* szótagot kell nem első szótagban felmutatnia, mint amennyi kettős, illetve hármas magánhangzó-kapcsolat szerepel a IV. táblázatban.

Az *lv1* és *lv2* szótagtípusnál véleményünk szerint, ellentétben TARNÓCZY megállapításával, nem növekvő szótagszámmal nő az előfordulási gyakoriság, hanem kizárólag az utolsó szótag mutat maximumot a többi szótagban való előfordulások számához képest. Ennek magyarázatát a ragok gyakori *lv1* típusú alakja (*-ban*, *-ben*, *-ból*, *-ből*, *-hoz*, *-hez*, *-höz*, *-nak*, *-nek*, *-ról*, *-ről*, *-tól*, *-től* stb.), több képző (*-ság*, *-ség*, *-mány*, *-mény*, *-vány*, *-vény*, *a-lom*, *e-lem*, *-da-lom*, *-de-lem* stb.) és szótagoláskor egy mássalhangzóval bővülő képző (pl. *ha-lás*z), a többszámú alakok stb. okozzák, az *lv2* típus szóvégi gyakori előfordulását a mássalhangzós tőhöz járuló tárgyrag, multidőjel és néhány képző (*-szt*, *-nt*, *-ng* stb.) magyarázza.

A szótagtípusokról beszélve, arra is ki kell térnünk, hogy nem tarthatjuk szerencsésnek és egyes nyelvek közötti különbségre vagy egyezésre döntő módon jellemzőnek a MENZERATH kezdeményezte összehasonlítási módot, az egyszótagú szavak szótagtípusainak egybevetését (Typology of Languages, Acoustical Society of America 1950, 698—701), hiszen a magyarban is, mint a 2. ábra mutatja, az egytagúak az összes szavaknak mintegy 35—45%-át teszik ki, minthogy a nem egytagúak 55—65%-a hosszabb, így a beszéd keltette akusztikai benyomásban még kisebb szerep jut az egytagúaknak, mint 35—45%. A többi nyelvre vonatkozóan meg éppen nem ismerünk semilyen statisztikát sem, amiből az egytagúaknak a többihez viszonyított arányára következtethetnénk, mindössze a németre közöl MENZERATH (Sprachtypologische Untersuchungen 65.) eloszlási arányt.

VERES PÉTERRE vonatkozó statisztikánk alapján a 3. ábrán látható módon összeállíthatunk négy lehetőséget. Az elsőbe csak az egytagú szavakat vesszük fel, a másodikba az összes első, a harmadikba az összes utolsó és végül

a negyedikbe az összes szótagot. Ennek megfelelően különbözőek a szótag-négyszögek. Az egytagú szavaknál azonos mennyiségű *lv1* és *v1* típusú szótag szerepel (ADY alapján *lv1* több lenne), majd *v* s végül kisebb gyakorisággal *lv2*, *lv*, *v2* következik, *2v1* és *lv3* csak szórványosan fordul elő. Az első, utolsó és összes szótagok vokális négyszögében mindenütt *lv1*, *lv*, *v1*, *v*, *lv2* a gyakorisági sorrend, de más-más arányban, a szórványosan előforduló típusoknál már nincs is egyezés a gyakoriság sorrendjében. (A csillag százalékszerűen ki sem fejezhető ritka előfordulást jelöl).<sup>1</sup>



3. Az egytagú szavak, első szótagok, utolsó szótagok és az összes szótag eloszlását feltüntető szótag-négyszögek (VERES PÉTER nyelve alapján)

### Kételemű hangzókapcsolatok

#### Magánhangzókapcsolatok

A 4. ábra TARNÓCZY 3. ábrájához hasonló módon a magánhangzókapcsolatokat mutatja. Célszerűnek látszott a csak összetételekben vagy csak idegen szavakban szereplő kapcsolatokat másképpen jelölni.

A Próbatételben csak idegen szavakban szereplő magánhangzótalálkozásokat a 4. ábrán az előfordulási gyakorisággal arányos területű háromszög ábrázolja (*feudalizmus*, *direktórium*, *szocializmus* stb., többnyire közismert latin szavak).

A szóösszetételekben előforduló magánhangzótalálkozásokat négyzet ábrázolja. Minthogy a magyarban *o*, *ö*-re nem végződik szó, *i*-, *ü*-re is csak elvétve, ezért az *o*, *ö*, *i*, *ü* sorában, azaz az olyan összetételek között, melyek első tagja *o*, *ö*, *i* vagy *ü* lenne, nem látható négyzet. A feldolgozott szövegben *o*, *i*, *ü* csak összetétel második tagjának kezdőhangjaként került két magánhangzó találkozásakor második helyre (*kioszt*, *beír*, *odaül* stb.).

A nem idegen szóban és nem összetételben egymás mellé kerülő magánhangzótalálkozások tarthatnak nyelvészeti szempontból nagyobb érdeklő-

<sup>1</sup> MENZEBATH szimmetrikusnak, illetve egyik vagy másik irányban aszimmetrikusnak az olyan szótagtípust nevezi, amelynél a magánhangzó előtt és mögött álló mássalhangzók száma azonos, illetve különböző. Ez nem azonos a zárt-nyílt szótagellentéttel: A nyílt szótagok minden szótag-négyszögnél csak a vonallal jelzett bal felső részen és

$$\begin{array}{r}
 v \\
 \hline
 lv \quad v1 \\
 2v \quad \hline 2v1 \quad 1v2 \\
 \hline
 2v2
 \end{array}$$

ennek folytatásában található, az összes többi szótagtípus zárt, tehát akár az olasz, akár a magyar szótag-négyszögben három nyílt és hat zárt típus szerepel. (Másképpen TARNÓCZY v. ö. 126.).

désre igényt. Ezeket előfordulási gyakoriságukkal arányos területű kör ábrázolja. A ritkán előforduló s így magánhangzós szóvéghez is csak néha járulók -an, -en, -ul (-ül nem fordul elő) ragon, néhány személyragos főnévi igenéven (innik, enniök) és a leány és lilom szóban lévő magánhangzótalálkozásokon kívül valamennyi körrel jelzett magánhangzótalálkozás gyakoribbnak mutatkozik. Három típust különböztethetünk meg: a) a második tag é, b) a második tag i, c) az első tag i. Ezeket sorban megvizsgálva azt találjuk, hogy

a) é mint a magánhangzókapcsolat második tagja *Dankóék, hazáért, Imréék, kié* stb. féle alkalmazásokban fordul elő;

második tag

	ú	u	ő	o	a	á	e	é	ö	ó	i	í	ü	ű
ú					•	•	•							
u					••									
ő			•		•		•			••	•			
o					•									
a		••			•	•	•	•	•		•	•	•	
á		•			•	•	•	•	•		•	•		
e		•	••		•	•	•	•	•	•	•	•	•	•
é						•					•			
ö														
ó			•		•		••	•			••			
i	•	•	•	•	•	•	•	•	•	•	•	•	•	•
í														
ü														
ű							•				••			

4. Magánhangzókapcsolódások VERES PÉTERNÉI

▲ idegen szó ■ szóösszetétel ● minden egyéb eset

b) az *i* birtokos személyragokban (*falai, gyerekei, ürgéi, dísznöi* stb., stb.) és az *-ig* rag néhány helyzetében (*hazáig* stb.) szerepel magánhangzótalálkozás második tagjaként;

c) a valóban szó belsejében adódó vokális találkozásoknak, a Próbatétel alapján azt mondhatjuk, túlnyomó többségében *i* az első tagja (*fiú, fióka, miatt, kiabál, kiált, hiába, siet, miért* stb.). Tulajdonképpen a magyar nyelvi érthetőségi viszonyok szempontjából ez az egyetlen figyelemreméltó magánhangzótalálkozás, mert a ragok esetében a megértést a nyelvi rendszer segíti, csak a szó belsejében adódó találkozásokról állítható, hogy pusztán az akusztikai viszonyok döntik el az érthetőség milyenségét.

*A mássalhangzók kvantitatív kapcsolatai magán- és mássalhangzókkal*

Az ikerített mássalhangzókapcsolatok gyakoriságával az ejtés ingadozásai miatt nem foglalkozunk részletesebben.

Az itt következő táblázatok az egyes mássalhangzóknak magán-, illetve mássalhangzóhoz való kapcsolódását vizsgálja százalékban kifejezve. Csak százalékos alapon válik lehetővé az ADYRA és VERES PÉTERRE vonatkozó

eredmények összehasonlítása. Az egyes hangok összes előfordulását vesszük 100%-nak, a hatodik adatsorban a szókezdő-, szóközépi- és szóvégi előfordulási arányszám látható. A következő sorban a Próbatétel adatainak durva ellenőrzéseként a hangzóstatisztika egyes hangokra lényegesen kisebb adatszámot feldolgozó adataiból<sup>1</sup> adódó szókezdő-, szóközépi- és szóvégi előfordulási arányszáma van zárójelben megadva. A zárójelet az adatok kis száma teszi indokolttá. A szóközépen az első és második sor összege egyezik a harmadik és negyedik sor összegével, hiszen mindig ugyanannyi hangnál nézzük az előtte, illetve az utána álló hangot. Az ADYra vonatkozó százaléértékeket TARNÓCZY XIII., XVII., XXI., XXV. és XXVI. táblázata alapján számítottam ki.

Az ADY és VERESre vonatkozó statisztikai eredmények egybevetése az egyes mássalhangzók szókezdő-, szóközépi- és szóvégi helyzetére vonatkozóan, valamint az előtte és utána következő magán-, ill. mássalhangzók arányszámát illetően, ha nem is mindig jó, de mindenütt elfogadható egyezést mutatnak. Nagyobb eltérés csak a kevés adatszámú hangoknál van (s talán az *j*-nél, de erről l. XXXIX. táblázathoz tartozó megjegyzést).

## VII. TÁBLÁZAT\*

l

Ady 4895 Veres 1121 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v l			64,4	67,6	16,6	17,—	81,—	84,6
k l			4,—	4,7	—	—	4,—	4,7
l v	15,—	10,7	38,7	39,—			53,7	49,7
l k	—	—	29,7	33,3			29,7	33,3
Összes helyzet	15,—	10,7	136,8	144,6	16,6	17,—	168,4	172,3
Összes hang	15,—	10,7	68,4	72,3	16,6	17,—	100,—	100,—
(767)		(9,5)		(73,5)		(17,—)		

\* A VII—XXIX. táblázat mellett a csillag azt jelenti, hogy a táblázatokhoz tartozó észrevételek a közlemény végén a táblázat sorszámával jelölve találhatóak.

<sup>1</sup> Az első oszlop utolsó sorában zárójelben szereplő szám az ellenőrzésnél feldolgozásra került hangok száma.

## VIII. TÁBLÁZAT\*

r

Ady 5014, Veres 1235 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v r			66,8	69,4	14,3	16,—	81,1	85,4
s			8,2	8,5	—	—	8,2	8,5
r v	10,7	6,1	47,1	44,7			57,8	50,8
r k	—	—	27,9	33,2			27,9	33,2
Összes helyzet	10,7	6,1	150,—	155,8	14,3	16,—	175,—	177,9
Összes hang	10,7	6,1	75,—	77,9	14,3	16,—	100,—	100,—
(694)		(5,6)		(76,8)		(17,6)		

## IX. TÁBLÁZAT\*

m

Ady 4754, Veres 1208 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v m			20,4	23,8	34,—	15,8	54,4	39,6
k m			6,5	7,2	—	—	6,5	7,2
m v	39,1	53,2	19,8	22,—			58,9	75,2
m k	—	—	7,1	9,—			7,1	9,—
Összes helyzet	39,1	53,2	53,8	62,—	34,—	15,8	126,9	131,—
Összes hang	39,1	53,2	26,9	31,—	34,—	15,8	100,—	100,—
(658)		(56,7)		(26,3)		(17,—)		

## X. TÁBLÁZAT\*

*n**Ady 4352, Veres 1221 adat alapján*

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
<b>v n</b>			37,—	43,6	32,6	22,5	69,6	66,1
<b>k n</b>			14,6	13,6	—	—	14,6	13,6
<b>n v</b>	15,8	20,3	24,9	30,7			40,7	51,—
<b>n k</b>	—	—	26,7	26,5			26,7	26,5
Összes helyzet	15,8	20,3	103,2	114,4	32,6	22,5	151,6	157,2
Összes hang	15,8	20,3	51,6	57,2	32,6	22,5	100,—	100,—
(873)		(15,5)		(58,—)		(26,5)		

## XI. TÁBLÁZAT\*

*ny**Ady 1975, Veres 1000 adat alapján*

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
<b>v ny</b>			53,1	71,2	23,6	10,—	76,7	81,2
<b>k ny</b>			7,4	5,8	—	—	7,4	5,8
<b>ny v</b>	15,9	13,—	48,1	61,5			64,—	74,5
<b>ny k</b>	—	—	12,4	15,5			12,4	15,5
Összes helyzet	15,9	13,—	121,—	154,—	23,6	10,—	160,5	177,—
Összes hang	15,9	13,—	60,5	77,—	23,6	10,—	100,—	100,—
(106)		(12,3)		(76,4)		(11,3)		

## XII. TÁBLÁZAT\*

v

Ady 5000, Veres 1114 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v v			23,6	25,4	0,6	0,5	24,2	25,9
k v			16,6	12,—	0,2	—	16,8	12,—
v v	59,—	62,1	39,6	36,2			98,6	98,3
v k	—	—	0,6	1,2			0,6	1,2
Összes helyzet	59,—	62,1	80,4	74,8	0,8	0,5	140,2	137,4
Összes hang	59,—	62,1	40,2	37,4	0,8	0,5	100,—	100,—
(304)		(56,—)		(43,7)		(0,3)		

## XIII. TÁBLÁZAT\*

f

Ady 5009, Veres 1083 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v f			11,8	7,4	1,2	0,3	13,—	7,7
k f			8,8	9,8	—	—	8,8	9,8
f v	76,5	81,8	14,4	16,6			90,9	98,4
f k	1,7	0,7	6,2	0,6			7,9	1,3
Összes helyzet	78,2	82,5	41,2	34,4	1,2	0,3	120,6	117,2
Összes hang	78,2	82,5	20,6	17,2	1,2	0,3	100,—	100,—
(131)		(85,5)		(13,7)		(0,8)		



## XIV. TÁBLÁZAT\*

z

Ady 3719, Veres 1204 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v z			46,8	59,3	41,7	35,8	88,5	95,1
k z			4,4	2,8	0,6	0,2	5,—	3,—
z v	6,5	1,9	41,7	44,3			48,2	46,2
z k	—	—	9,5	17,8			9,5	17,8
Összes helyzet	6,5	1,9	102,4	124,2	42,3	36,—	151,2	162,1
Összes hang	6,5	1,9	51,2	62,1	42,3	36,—	100,—	100,—
(366)		(1,4)		(61,7)		(36,9)		

## XV. TÁBLÁZAT\*

sz

Ady 4993, Veres 1179 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v sz			36,6	51,7	8,3	6,1	44,9	57,8
k sz			7,9	10,—	0,7	0,8	8,6	10,8
sz v	46,5	30,9	29,2	35,—			75,7	65,9
sz k	—	0,5	15,3	26,7			15,3	27,2
Összes helyzet	46,5	31,4	89,—	123,4	9,—	6,9	144,5	161,7
Összes hang	46,5	31,4	44,5	61,7	9,—	6,9	100,—	100,—
(320)		(28,4)		(63,2)		(8,4)		

## XVI. TÁBLÁZAT\*

zs

Ady 483, Veres 152 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v zs			55,9	34,3	8,9	3,3	64,8	37,6
k zs			6,4	7,2	—	—	6,4	7,2
zs v	28,8	55,2	42,2	34,9			71,—	90,1
zs k	—	—	20,1	6,6			20,1	6,6
Összes helyzet	28,8	55,2	124,6	83,—	8,9	3,3	162,3	141,5
Összes hang	28,8	55,2	62,3	41,5	8,9	3,3	100,—	100,—
(16)		(25,—)		(68,8)		(6,2)		

## XVII. TÁBLÁZAT\*

s

Ady 5000, Veres 1167 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v s			31,6	33,6	43,—	45,9	74,6	79,5
k s			5,4	6,4	0,05	0,1	5,45	6,5
s v	19,9	13,7	24,7	25,5			44,6	39,2
s k	0,05	0,3	12,3	14,5			12,35	14,8
Összes helyzet	19,95	14,—	74,—	80,—	43,05	46,—	137,—	140,—
Összes hang	19,95	14,—	37,—	40,—	43,05	46,—	100,—	100,—
(559)		(13,—)		(36,4)		(50,6)		

## XVIII. TÁBLÁZAT\*

*j, ly**Ady 2782, Veres 1111 adat alapján*

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
<b>v j</b>			49,2	50,1	13,4	5,9	62,6	56,—
<b>k j</b>			11,8	15,8	0,9	0,1	12,7	15,9
<b>j v</b>	24,7	28,1	43,6	49,8			68,3	77,9
<b>j k</b>	—	—	17,4	16,1			17,4	16,1
Összes helyzet	24,7	28,1	122,—	131,8	14,3	6,—	161,—	165,9
Összes hang	24,7	28,1	61,—	65,9	14,3	6,—	100,—	100,—
(231)		(26,—)		(65,—)		(9,—)		

## XIX. TÁBLÁZAT\*

*h**Ady 2682, Veres 1171 adat alapján*

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
<b>v h</b>			16,1	18,6			16,1	18,6
<b>k h</b>			12,2	12,9			12,2	12,9
<b>h v</b>	71,7	68,5	28,2	31,3			99,9	99,8
<b>h k</b>	—	—	0,1	0,2			0,1	0,2
Összes helyzet	71,7	68,5	56,6	63,—	—	—	128,3	131,5
Összes hang	71,7	68,5	28,3	31,5	—	—	100,—	100,—
(316)		(67,—)		(33,—)		—		

## XX. TÁBLÁZAT\*

b

Ady 9959, Veres 1144 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v b			23,9	29,7	6,9	7,4	30,8	37,1
k b			25,3	28,8	0,5	—	25,8	28,8
b v	43,2	34,1	46,5	56,6			89,7	90,7
b k	0,2	—	2,7	1,9			2,9	1,9
Összes helyzet	43,4	34,1	98,4	117,—	7,4	7,4	149,2	158,5
Összes hang	43,4	34,1	49,2	58,5	7,4	7,4	100,—	100,—
(291)		(33,—)		(62,5)		(4,5)		

## XXI. TÁBLÁZAT\*

p

Ady 2600, Veres 1082 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v p			37,—	49,6	20,—	5,2	57,—	54,8
k p			5,—	5,4	—	—	5,—	5,4
p v	36,8	38,—	32,5	38,1			69,3	76,1
p k	1,2	1,8	9,5	16,9			10,7	18,7
Összes helyzet	38,—	39,8	84,—	110,—	20,—	5,2	142,—	155,—
Összes hang	38,—	39,8	42,—	55,—	20,—	5,2	100,—	100,—
(107)		(45,8)		(51,4)		(2,8)		

## XXII. TÁBLÁZAT\*

d

Ady 4900, Veres 1146 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v d			35,3	32,9	14,1	5,1	49,4	38,—
k d			27,2	39,2	5,3	5,1	32,5	44,3
d v	17,7	17,5	53,8	60,3			71,5	77,8
d k	0,4	0,2	8,7	11,8			9,1	12,—
Összes helyzet	18,1	17,7	125,—	144,2	19,4	10,2	162,5	172,1
Összes hang	18,1	17,7	62,5	72,1	19,4	10,2	100,—	100,—
(335)		(16,4)		(72,3)		(11,3)		

## XXIII. TÁBLÁZAT\*

t

Ady 9777, Veres 1392 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v t			25,5	24,4	27,8	28,2	53,3	52,6
k t			19,1	16,6	14,1	19,8	33,2	36,4
t v	13,1	10,9	38,2	33,6			51,3	44,5
t k	0,4	0,1	6,4	7,4			6,8	7,5
Összes helyzet	13,5	11,—	89,2	82,—	41,9	48,—	144,6	141,—
Összes hang	13,5	11,—	44,6	41,—	41,9	48,—	100,—	100,—
(1174)		(12,8)		(41,—)		(46,2)		

## XXIV. TÁBLÁZAT\*

g

Ady 4998, Veres 1261 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v g			50,—	47,—	30,4	25,—	80,4	72,—
k g			12,6	16,6	1,5	—	14,1	16,6
g v	5,—	10,9	49,4	49,2			54,4	60,1
g k	0,5	0,5	13,2	14,4			13,7	14,9
Összes helyzet	5,5	11,4	125,2	127,2	31,9	25,—	162,6	163,6
Összes hang	5,5	11,4	62,6	63,6	31,9	25,—	100,—	100,—
(413)		(9,2)		(64,6)		(26,2)		

## XXV. TÁBLÁZAT\*

k

Ady 5000, Veres 1108 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v k			24,4	28,6	41,5	34,1	65,9	62,7
k k			8,6	14,1	5,1	0,5	13,7	14,6
k v	20,2	22,6	26,—	35,5			46,2	58,1
k k	0,2	0,1	7,—	7,2			7,2	7,3
Összes helyzet	20,4	22,7	66,—	85,4	46,6	34,6	133,—	142,7
Összes hang	20,4	22,7	33,—	42,7	46,6	34,6	100,—	100,—
(815)		(24,5)		(35,1)		(40,4)		

## XXVI. TÁBLÁZAT\*

c

Ady 684, Veres 321 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v c			38,—	39,9	4,7	4,4	42,7	44,3
k c			37,5	33,3	6,9	8,4	44,4	41,7
c v	12,9	14,—	66,9	53,3			79,8	67,3
c k	—	—	8,6	19,9			8,6	19,9
Összes helyzet	12,9	14,—	151,—	146,4	11,6	12,8	175,5	173,2
Összes hang	12,9	14,—	75,5	73,2	11,6	12,8	100,—	100,—
(27)		(3,7)		(77,8)		(18,5)		

## XXVII. TÁBLÁZAT\*

cs

Ady 1548, Veres 1317 adat alapján

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v cs			18,7	22,3	1,3	0,6	20,—	22,9
k cs			13,7	12,1	8,6	7,—	22,3	19,1
cs v	57,7	58,—	30,1	30,3			87,8	88,3
cs k	—	—	2,3	4,1			2,3	4,1
Összes helyzet	57,7	58,—	64,8	68,8	9,9	7,6	132,4	134,4
Összes hang	57,7	58,—	32,4	34,4	9,9	7,6	100,—	100,—
(102)		(68,6)		(25,5)		(5,9)		

## XXVIII. TÁBLÁZAT\*

*gy**Ady 3008, Veres 1129 adat alapján*

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
<i>v gy</i>			39,9	37,3	46,4	49,9	86,3	87,2
<i>k gy</i>			2,6	5,3	0,8	0,1	3,4	5,4
<i>gy v</i>	10,3	7,4	38,1	34,7			48,4	42,1
<i>gy k</i>	—	—	4,4	7,9			4,4	7,9
Összes helyzet	10,3	7,4	85,—	85,2	47,2	50,—	142,5	142,6
Összes hang	10,3	7,4	42,5	42,6	47,2	50,—	100,—	100,—
(248)		(4,8)		(43,2)		(52,—)		

## XXIX. TÁBLÁZAT\*

*ty**Ady 278, Veres 131 adat alapján*

	Szókezdő		Közbülső		Végső		Összes	
	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
<i>v ty</i>			75,7	66,5	0,7	—	76,4	66,5
<i>k ty</i>			21,6	29,—	—	—	21,6	29,—
<i>ty v</i>	2,—	4,5	74,8	87,8			76,8	92,3
<i>ty k</i>	—	—	22,5	7,7			22,5	7,7
Összes helyzet	2,—	4,5	194,6	191,—	0,7	—	197,3	195,5
Összes hang	2,—	4,5	97,3	95,5	0,7	—	100,—	100,—
(21)		(4,8)		(95,2)		(—)		



Nincs értelme ugyanezeket a táblázatokat magánhangzókra is elkészíteni, hiszen az már az eddigiékből is látható, hogy előtte és utána is csaknem az esetek 100%-ában mássalhangzó található. A magánhangzók szókezdő, szóközépi és szóvégi eloszlására vonatkozó statisztikáink kevésbé megbízható eredményt látszanak adni, mint a mássalhangzóstatisztikák, ez látszik a XXX. táblázatból, ahol az első számadatok rövid és hosszú magánhangzó-párok együttesen 1000–1200 adatra történt feldolgozásából, a második, zárójelbe tett arányszámok a hangzóstatisztika legtöbb esetben lényegesen kisebb adatszámú magánhangzóinak eloszlásából származnak.

## XXX. TÁBLÁZAT

*A magánhangzók szóelejei, szóközépi és szóvégi eloszlási aránya %*

	Adatszám*		Szókezdő %		Szóközépi %		Szóvégi %	
	Veres Péter	Ady	Veres Péter	Ady	Veres Péter	Ady	Veres Péter	Ady
ú	321 (70)		47,7 (45,7)		43,9 (52,9)		8,4 (1,4)	
u	965 (131)		17,6 (19,1)		81,9 (80,1)		0,5 (0,8)	
ó	225 (173)		2,2 (1,7)	5,8	58,7 (68,8)	74,4	39,1 (29,5)	19,8
o	1007 (645)		7,0 (6,7)	3,8	93,0 (93,3)	96,2	—	—
a	1125 (1362)		18,2 (19,7)		65,5 (61,8)		16,3 (18,5)	
á	1164 (552)		3,9 (2,7)		94,6 (94,4)		1,5 (2,9)	
e	1239 (1581)		12,7 (11,8)		75,5 (76,5)		11,8 (11,7)	
é	1377 (536)		23,9 (26,8)		73,1 (71,5)		3,0 (1,7)	
ö	680 (175)		14,7 (7,4)	14,3	85,3 (92,6)	85,7	—	—
ő	426 (101)		9,6 (8,9)	16,2	72,1 (65,4)	61,6	18,3 (25,7)	22,2
i	1195 (744)		20,6 (22,2)		56,8 (55,9)		22,6 (21,9)	
í	102 (72)		12,7 (30,5)		87,3 (69,5)		—	—
ü	768 (64)		10,4 (4,7)		89,2 (95,3)		0,4	—
ű	188 (11)		—		77,0 (72,7)		23,0 (27,3)	

<sup>1</sup> Az adatszám Adynál: ó 363, o 1 782, ö 2 190, ő 1 109.

Olyan módon is egybevetjük a fenti eredményeket, hogy megvizsgáljuk, milyen gyakran van valamely hang előtt magánhangzó, mássalhangzó vagy semmi, azaz milyen gyakran fordul elő szókezdeten (XXXI. táblázat). Hasonló módon az utána következő magán-, mássalhangzógyakoriság, illetve szóvégi előfordulás is vizsgálható (XXXII. táblázat).

A XXXI. és XXXII. táblázatból kiviláglik az az éppen nem meglepő tény, hogy ragkezdő mássalhangzók előtt aránylag gyakrabban van mássalhangzó, mint olyan mássalhangzók előtt, amelyekkel nem kezdődik rag, *l*, *r*(!), *ny*, *z*, *s*, *p*, *gy* előtt több mint tízszerannyi esetben találunk magánhangzót, mint mássalhangzót, a *c* és a *cs* előtt kb. ugyanolyan gyakran található magán- és mássalhangzó. Az *f* és a *h* (a *-hoz*, *-hez*, *-höz* és *-hat*, *-het* ellenére is) főként szókezdő hangnak nevezhető. A szóvégen nagyon ritka hangok *h*, *f*, *v*, ezek után a szó belsejében is jóformán mindig magánhangzó következik.

## XXXI. TÁBLÁZAT\*

Mássalhangzók kapcsolódásai az előtte álló hanggal

	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
	<i>l</i>		<i>v</i>		<i>b</i>		<i>c</i>	
v	81,0	84,6	24,2	25,9	30,8	37,1	42,7	44,3
k	4,0	4,7	16,8	12,0	25,8	28,8	44,4	41,7
∅	15,0	10,7	59,0	62,1	43,4	34,1	12,9	14,0
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
	<i>r</i>		<i>f</i>		<i>p</i>		<i>cs</i>	
v	81,1	85,4	13,0	7,7	57,0	54,8	20,0	22,9
k	8,2	8,5	8,8	9,8	5,0	5,4	22,3	19,1
∅	10,7	6,1	78,2	82,5	38,0	39,8	57,7	58,0
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
	<i>m</i>		<i>z</i>		<i>d</i>		<i>gy</i>	
v	54,4	39,6	88,5	95,1	49,4	38,0	86,3	87,2
k	6,5	7,2	5,0	3,0	32,5	44,3	3,4	5,4
∅	39,1	53,2	6,5	1,9	18,1	17,7	10,3	7,4
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
	<i>n</i>		<i>sz</i>		<i>t</i>		<i>ty</i>	
v	69,6	66,1	44,9	57,8	53,3	52,6	76,4	66,5
k	14,6	13,6	8,6	10,8	33,2	36,4	21,6	29,0
∅	15,8	20,3	46,5	31,4	13,5	11,0	2,0	4,5
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
	<i>ny</i>		<i>zs</i>		<i>g</i>			
v	76,7	81,2	64,8	37,6	80,4	72,0		
k	7,4	5,8	6,4	7,2	14,1	16,6		
∅	15,9	13,0	28,8	55,2	5,5	11,4		
	100,0	100,0	100,0	100,0	100,0	100,0		
			<i>s</i>		<i>k</i>			
v			74,6	79,5	65,9	62,7		
k			5,45	6,5	13,7	14,6		
∅			19,95	14,0	20,4	22,7		
			100,00	100,0	100,0	100,0		
			<i>j</i>					
v			62,6	56,0				
k			12,7	15,9				
∅			24,7	28,1				
			100,0	100,0				
			<i>h</i>					
v			16,1	18,6				
k			12,2	12,9				
∅			71,7	68,5				
			100,0	100,0				

Jelmagyarázat: v = magánhangzó, k = mássalhangzó, ∅ = szókezdő  
 Adynál 22% szóvégi *m* elhagyásával

*m* 41,5  
 8,3  
 50,2  
 100,0

XXXII. TÁBLÁZAT\*  
Mássalhangzók kapcsolódásai az utána álló hanggal

	Ady	Veres	Ady	Veres	Ady	Veres	Ady	Veres
v	<i>l</i>		<i>v</i>		<i>b</i>		<i>c</i>	
k	53,7	49,7	98,6	98,3	89,7	90,7	79,8	67,3
∅	29,7	33,3	0,6	1,2	2,9	1,9	8,6	19,9
	16,6	17,0	0,8	0,5	7,4	7,4	11,6	12,8
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
v	<i>r</i>		<i>f</i>		<i>p</i>		<i>cs</i>	
k	57,8	50,8	90,9	98,4	69,3	76,1	87,8	88,3
∅	27,9	33,2	7,9	1,3	10,7	18,7	2,3	4,1
	14,3	16,0	1,2	0,3	20,0	5,2	9,9	7,6
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
v	<i>m</i>		<i>z</i>		<i>d</i>		<i>gy</i>	
k	58,9	75,2	48,2	46,2	71,5	77,8	48,4	42,1
∅	7,1	9,0	9,5	17,8	9,1	12,0	4,4	7,9
	34,0	15,8	42,3	36,0	19,4	10,2	47,2	50,0
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
v	<i>n</i>		<i>sz</i>		<i>t</i>		<i>ty</i>	
k	40,7	51,0	75,7	65,9	51,3	44,5	76,8	92,3
∅	26,7	26,5	15,3	27,2	6,8	7,5	22,5	7,7
	32,6	22,5	9,0	6,9	41,9	48,—	0,7	—
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
v	<i>ny</i>		<i>zs</i>		<i>g</i>			
k	64,0	74,5	71,0	90,1	54,4	60,1		
∅	12,4	15,5	20,1	6,6	13,7	14,9		
	23,6	10,0	8,9	3,3	31,9	25,0		
	100,0	100,0	100,0	100,0	100,0	100,0		
v			<i>s</i>		<i>k</i>			
k			44,6	39,2	46,2	58,1		
∅			12,35	14,8	7,2	7,3		
			43,05	46,0	46,6	34,6		
			100,00	100,0	100,0	100,0		
v			<i>j</i>					
k			68,3	77,9				
∅			17,4	16,1				
			14,3	6,0				
			100,0	100,0				
v			<i>h</i>					
k			99,9	99,8				
∅			0,1	0,2				
			—	—				
			100,0	100,0				

Jelmagyarázat: v = magánhangzó, k = mássalhangzó, ∅ = szóvégződés  
Adynál 22% szóvégi *m* elhagyásával

*m* 75,5  
9,1  
15,4  
100,0

## VII. tábl.

Az *l* szókezdő-, szóközépi- és szóvégi eloszlási aránya meglehetősen egyezik ADY-nál és VERES-nél. Szép egyezést találunk szóközépen az *l* előtt, illetve után álló magán- és mássalhangzókapcsolatok megoszlásánál. A VERES-nél valamivel gyakoribb *l*-et követő mássalhangzóarányt a *föld*, *hold* szó gyakori előfordulása indokolja.

## VIII. tábl.

ADY-nál több *r* szerepel szó elején, mint VERES-nél. Kis különbség található szóközépen az utánakövetkező hanggal való kapcsolódásban. ADY-nál aránylag gyakrabban következik *r* után magánhangzó, a Próbatételben — ADY-hoz képest — aránylag több az *r*-et követő mássalhangzó.

## IX. tábl.

Az *m* hangnál mutatkozó igen nagy különbséget ADY az átlaghasználatot feltétlenül jóval felülmúlóan gyakran használt első személyű birtokos személyragja és első személyben szereplő tárgyasan ragozott igéi okozzák. Megkísértem ADY *m*-jeiből 22%-ot, mindet a szóvégen, elhagyni. Ilymódon szinte tökéletesen egyezést kaptam a VERES PÉTER-nél talált arányokkal.

## X. tábl.

Az *n*-nél érdekes módon a 12 lapról készített hangstatisztikában 873 *n* szókezdő-, szóközépi- és szóvégi eloszlása lényegesen közelebb van az ADY-nál talált eloszláshoz, mint a Próbatétel 1225 *n* adatában szereplő eloszláshoz. A statisztikák készítése közben semmi olyan feltűnő jelenséget nem észleltem, amely ezt az eltérést indokolná. ADY-nál *n*-et szóközben gyakrabban követ mássalhangzó. VERES-nél viszont magánhangzó, de mindkét esetben elég kicsiny az eltérés az 1:1 aránytól.

## XI. tábl.

Az *ny* arányszámai is szépen megegyeznek. Mindössze a szóközépi *ny* ritkább és ennek megfelelően a szóvégi gyakoribb ADY-nál, mint VERES PÉTER-nél.

## XII. tábl.

A *v* arányszámai minden vonatkozásban nagyon jó megfelelést mutatnak.

## XIII. tábl.

Az *f* szóközépi előfordulása másképpen alakul ADY-nál, mint VERES-nél. ADY-nál gyakrabban van *f* előtt magánhangzó, mint mássalhangzó. VERES-nél éppen fordítva ADY-nál a szóközépi *f*-eknek csaknem egyharmadát mássalhangzó követi. VERES PÉTER-nél erre csak szórványosan akad példa.

## XIV. tábl.

A *z* szókezdő-, szóközépi- és szóvégi eloszlásának arányszáma más a két vizsgált szerzőnél. Érdekes a különbség a szókezdeten, ahol ADY használatában lényegesen több a *z*, a szóvégi *z* arányszáma is nagyobb ADY-nál. Lehetséges, hogy VERES PÉTER-nél a szóközépi *z* és egyúttal a *z* gyakoribb kapcsolatát az utána álló mássalhangzóval jelen statisztika készítésénél figyelmen kívül hagyott és TARNÓCZYNál tekintetbe vett asszimilációk okozzák.

## XV. tábl.

A *z*-hez meglehetősen hasonló eltéréseket találunk az *sz*-nél is. ADY-nál csaknem azonos a szókezdő és szóközépi *sz*-ek száma, VERES-nél csaknem kétszer annyi *s* van szóközépen, mint szókezdeten. VERES-nél csaknem kétszer annyi *sz* van szóközépen, mint szókezdeten. VERES-nél az *sz*-et követő mássalhangzó arányszáma is nagyobb, mint ADY-nál,

## XVI. tábl.

A *zs* előfordulása annyira ritka, hogy e kevés adat alapján szinte csodálatosnak találhatjuk, hogy az adatok még ennyire is egyeznek, mint látható.

## XVII. tábl.

Az *s*-nél a szóvégi *s* kicsit ritkább volta ADY-nál az *s* kötőszó és helyett való használatából származik. VERES-nél a szókezdő *s* arányszáma kisebb, mint ADY-nál,

## XVIII. tábl.

A *j* hang eloszlásánál feltűnik, hogy VERES PÉTERNÉL szó elején gyakrabban és szó végén lényegesen ritkábban szerepel, mint ADYNÁL. A szóvégi gyakoribb *j*-t ADYNÁL esetleg a gyakrabban használt parancsoló mód második személye okozhatja.

## XIX. tábl.

A *h*-nál egészen kiváló a két feldolgozás egyezése.

## XX. tábl.

A *b* hangnál minden különbséget a szókezdő és szóközépen szereplő *b*-k más arányszáma okoz. Különben a szóközépi magán-, illetve mássalhangzókapcsolódási arány, mind az előtte, mind az utána álló hangoknál eléggé azonos, minthogy mindkét szerzőnél kb. 1 : 1 a szóközépen *b* előtt álló magán- és mássalhangzók aránya, *b* után pedig rendszerint magánhangzó következik.

## XXI. tábl.

A *p*-nél található különbségeket azzal magyarázhatjuk, hogy VERES PÉTERNÉL a *p* végű szavak, úgy látszik, gyakrabban szerepelnek mássalhangzós kezdetű ragokkal (*kalapja, kapnak* stb.). Ez egyszerre magyarázat a szóvégi *p* ritkább és a szóközépen mássalhangzóval követett *p* gyakoribb voltára. VERES PÉTERNÉL a zárójelben közölt ellenőrző statisztika olyan kevés hang alapján készült, hogy nem csodálkozhatunk, ha a szóvégi *p* ott még ritkább.

## XXII. tábl.

A *d* eloszlásában találunk eltéréseket, de ezek eléggé indokolhatók. ADYNÁL gyakrabban szerepel a birtokos második személyű ragja, mint VERES PÉTERNÉL, ezért gyakoribb a *d* a szóvégen, méghozzá előtte magánhangzóval. VERES PÉTERNÉL gyakran szerepel ragozottan és ragozatlanul a *föld* és a *hold* szó és ez növeli a szóközépen és szóvégen is a mássalhangzót követő *d* hangok számát a magánhangzó utáni *d*-k számához viszonyítva.

## XXIII. tábl.

A *t* kapcsolatainak áttekintésénél figyelembe kell vennünk, hogy VERES PÉTER talán sok más írónál gyakrabban beszél múlt időben. Jelen időben játszódó elbeszéléseiben is egyre multidőbeli események elmondására siklik át. A hangstatisztikához feldolgozott részen, mint már a III. táblázatnál is említettük, megszámláltuk, hogy a végződések közül mennyi a múlt idő jele. Az ott talált arány alapján azt mondhatjuk, hogy az itteni 48% *t* végződésből 13,8% feltehetően a múlt idő jele. Minthogy ADYNÁL is szerepel elég gyakran múlt idő, a két statisztika minden különbségét VERES PÉTER gyakrabban használt múlt ideje teljesen kielégítően megmagyarázza.

## XXIV. tábl.

Érdekes módon VERES PÉTERNÉL szó elején, ADYNÁL szó végén nagyobb a *g* előfordulási gyakorisága.

## XXV. tábl.

Ismét ADY szerfelett sokat használt egyes és többes első személyére kell utalnunk annak indoklására, hogy ADYNÁL miért szerepel a szóvégi *k* mind magán-, mind mássalhangzó után gyakrabban, mint VERES PÉTERNÉL.

## XXVI. tábl.

A *c* hang adatszámát annyira kevés ADYNÁL is és VERESNÉL is, hogy inkább az egyezések jósága feltűnő, mint az, hogy a Próbatételben gyakrabban követi *c*-t mássalhangzó, mint ADYNÁL.

## XXVII. tábl.

*Cs*-nél kifogástalanok az arányok megfelelései.

## XXVIII. tábl.

A *gy* megfelelései sem kívánnak magyarázatot, teljesen azonos képet kapunk ADY és VERES PÉTER vizsgálatánál.

## XXIX. tábl.

A *ty* hangnál ismét csak azt állapíthatjuk meg, amit a többi ritka hangnál, hogy a ritka előforduláshoz képest meglepően jók az egyezések.

(Folytatása következik.)

VÉRTES EDIT