

PAPP FERENC

MAGYAR ÉS OROSZ NYELVŰ SZÖVEGEK SZÁMÍTÓGÉPES VIZSGÁLATA*

1. Két évtizeddel—negyedszázaddal ezelőtt, amikor a mi generációnk színre lépett, bizonyos közös vonásokkal lehetett jellemezni az orosz és a magyar nyelv vizsgátságának állapotát. Egyaránt jellemző volt akkor az, hogy igen sok részeredmény felhalmozódott — természetesen területenkénti eltéréssel. Másrészt bizonyos elméleti útkeresést is tapasztalhattunk a legáltalánosabb, távolról sem csupán a két nyelv feldolgozásának konkrét kérdéseit illető problémakörben.

Az akkor induló fiatal kutató előtt az egyik út az lehetett, hogy tovább folytatja az anyaggyűjtést, esetleg kiterjeszti még nem vagy kevésbé érintett területekre — ugyanakkor részt vesz az egyre növekvő anyag szintézisbe hozásában. Egy másik út, melyet e helyzet természetesen diktált: élesen tagadni az eddigieket, a pozitívizmus megnyilvánulásaként elvetni az aprólékos-gyűjtő, elemeket egyenként feldolgozó munkát, megkísérelni valahogy egészen másként megközelíteni a jelenségeket: atomizáltság helyett az egésze, induktív módszerek helyett a deduktív módszerekre támaszkodni, már amennyire ez utóbbi egy konkrét társadalmi jelenségsoporttal foglalkozó tudomány számára lehetséges. Tudjuk, hogy valóban mind a két irányban folytak kutatások, és jelentős eredmények is születtek mind a két nyelv területén.

Volt és van azonban itt kivételesen egy harmadik út is. Az igen nagy mennyiségű apró tényanyag és az elméleti igény közötti feszültséget a számítógépek oldhatják fel. A számítógép használata egyik említett irányzat számára sem kötelező: nagyon is lehet hagyományos szintéziseket létrehozni enélkül (sőt eddig értelmező, etimológiai stb. szótárak például csak így készültek); elméletet alkotni gép nélkül. Az automata mindössze egy lehetséges áthidalást jelent; mindössze hozzásegít ahhoz, hogy míg a nyelvjárásgyűjtő, etimologizáló stb. munkában mindig csak egy-két, legfeljebb néhány tucat vagy néhány száz példát lát maga előtt a kutató, addig a számítógépek segítségével mint egy szputnyikról, az egész vizsgált tájat könnyen át tudjuk tekinteni; bizonyos, olykor a földközelen is érzett vonalakat—összefüggéseket kimutatni; egyéb mélységesen emberi „fogycatóságokat” kiküszöbölni. (Vö. pl. alább az asszociatív szótárról mondandókkal. Hatalmas háttérmemóriánk nem formálisan van szervezve, a benne levőket nem tudjuk formális kritériumok alapján átrendezni stb.)

Számítógépes vizsgálataink, melyek első ötlete itt, egy főleg Kalmár László inspirálta akadémiai tanácskozáson született tizenhat évvel ezelőtt, eleinte csak a magyar

*PAPP Ferenc akadémiai levelező tag székfoglaló előadása az MTA I. Osztályának 1978. április 10-én tartott felolvasó ülésén hangzott el rövidített változatban.

nyelvre, pontosabban: a szótározott magyar szókincsre irányultak. Ebből az anyagból könnyen lehet példát hozni a most mondottakra.

A nagyobb összefüggések közül, amelyekre rá kellett bukkannunk, kettő nyelvtörténeti vonatkozású. Az első: szókincsünk egyes etimológiai rétegeinek magánhangzó-gazdagsága.¹ Kiderült az, amit egyébként tapasztalt etimológusaink korábban is tudtak (éreztek?), hogy ha az egyes lexémákat mai köznyelvi alakjukban tekintjük, a latin–görög eredetűek, de még a szlávok is viszonylag több magánhangzót tartalmaznak a mássalhangzókhöz képest, mint a finnugor eredetűek; hogy tehát szláv eredetű kölcsönszavaink igencsak „megmagyarosodtak”, ha ezen éppen a magánhangzó-gazdagságot értjük. Ennek a ténynek a továbbiakban mondandók számára is jelentősége van éppen úgy, mint annak a körülménynek, hogy mivel hosszú magánhangzók tekintetében a szláv réteg igen szegény, ezért a finnugor eredetű lexémák hangteste időben mérve több magánhangzós elemet tartalmaz, mint a szláv eredetűeké. – A második: szókincsünk az idő mentén ma felé haladva egyre „mélyül”; míg finnugor és török eredetű kölcsönszavaink mai alakjukban enyhe veláris dominancia mellett majdnem fele-fele arányban tartalmaznak magas és mély magánhangzókat, addig a szláv, latin–görög, neolatin eredetűek hangállományának több mint háromnegyede mély.² A környező átható nyelvek szempontjából ez a véletlen műve, hiszen ott a magas–mély oppozíciónak nincs alaktani jelentősége, ám esetleg bonyodalmakat okozhatott a mi nyelvünk szerkezetében. Annál is inkább, mert később kiderült, hogy illeszkedési rendszerünk fokozatos átépülése következtében³ e megnövekedett mennyiségű mély magánhangzó még saját súlyánál is nagyobb mértékben von maga után mély hangrendű illeszkedést. Így tőfőneveink között a szláv, neolatin, latin–görög eredetűek 82–92%-a mutatkozott mély illeszkedésűnek.⁴ Ez pedig felvethette azt a gondolatot, milyen erőben van egész illeszkedési rendszerünk. Leegyszerűsítve a kérdést: vajon nincs-e gyengülőben az egész illeszkedés úgy, hogy a mély illeszkedés kezd uralkodóvá válni? E kérdés megválaszolásához természetesen élő magyar szövegek tömegét kellett feldolgoznunk: a szótár csak bizonyos lehetőségeket ad meg, az életet a szövegek mutatják.

Egy szinkrón alaktani példa kutatásaink e korábbi stádiumából. Vajon mi az oka egyes birtokos személyragos alakjainkban a [j] elem hol rapszodikus, hol egyenesen váratlan megjelenésének (vö. *lába–combja, ablaka–barackja* – a kiemelt mássalhangzócsoportok után váratlan a [j] megjelenése)? E kérdésre a válasz (Tompá József, Károly Sándor munkáiban) gépek nélkül is megjelent; nekünk mindössze sikerült kimutatnunk, hogy valóban: a ritkább, szokatlanabb szóvég követeli meg a [j]-t a bázisvég kiemelése céljából. E bázismegkülönböztető tendencia megállapítása számunkra kivált két szempontból volt érdekes. Ezzel egyrészt felállítottunk egy hipotézist, melynek annyi volt a jelentősége, hogy elfogadásával számos, eddig szétszórta kezelt jelenség számára sikerült egy egységes magyarázatot találnunk. Másrészt a mennyiségi viszonyokból indultunk ugyan ki, méghozzá azok olyan tömegéből, amelyet valóban csak géppel lehetett kezelni, de egy

¹ Vö. Tőszókincsünk etimológiai rétegenkénti hangstatisztikája. *Nyelvtudományi Közlemények* 75. (1973) 3–40. Hivatkozott rész 18 és kk.

² Uő. 21.

³ SZÉPE Gy.: Vegyes magánhangzójú szavaink illeszkedésének kérdéséhez. *Nyelvtudományi Értekezések* 17. (1958) 105–129.

⁴ Vö. A magyar főnév paradigmatiszma rendszere. Bp., 1975. 258.

minőségi-strukturális jellegű megállapításra jutottunk: úgy látszik, az ember számára is fáradtságos minduntalan a háttérmemóriához nyúlni, igyekszik úgy megalkotni a kódot, hogy a szövegek minimális szótár segítségével automatikusan elemezhetőek legyenek. A magyarul beszélő tudatában egész nyelvtörténetünk tanúsága szerint joggal feltételezhetünk egy törekvést a bonyolultabb mássalhangzó-torlódások kikerülésére, a kialakultak kiküszöbölésére. Nem tudjuk, mivel kell mérni e tényező erejét, erre talán a neurolingvisztika ad majd választ. Ám egy relatív erőviszonyt már most megállapíthatunk: a bázismegkülönböztető tendencia az idézett *mbj*-, *-kj*- és ehhez hasonló mássalhangzó-kapcsolatok tanúsága szerint erősebb, mint a torlódást kikerülő-kiküszöbölő tendencia.

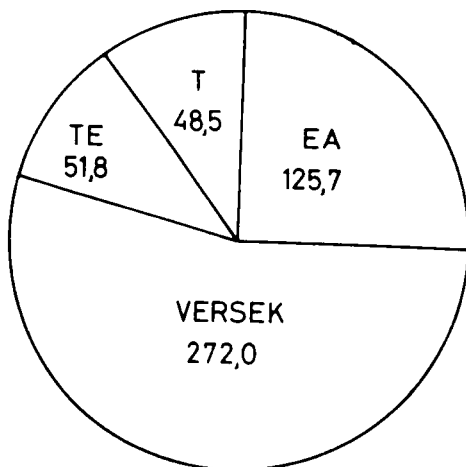
Hasonlóan mennyiségi adatokból indultunk ki és egy strukturális törvényszerűség közelébe jutottunk a következő esetben.⁵ Azokból az eredményekből, melyeket H. H. Josselson több orosz szótár gépi feldolgozásával nyert, rá kellett jönnünk arra, hogy van bizonyos statisztikus törvényszerűség az orosz lexémák szótári alakjának hangsúly szerinti eloszlásában. Nevezetesen: a hangsúly leggyakrabban a szóalak közepe felé foglal helyet, legfeljebb egy szótaggal a szó vége irányában eltérve. Tehát tipikusnak tekinthető az olyan szóalakok, mint *рука'*, *бума'га*, *разобра'ться* *издева'тельство*, *функциони'ровать*. Kiderült, hogy évtizedekkel előttünk költői *s z ö v e g e k r e* vonatkozóan hasonló megállapításra jutott Tomasevszkij és Sengeli (1923). Márpedig a hangsúlynak ilyen szóközépi elhelyezkedése jó összhangban van az orosz hangsúly „szóalakat összetartó”, ebben a magánhangzó-illeszkedésre emlékeztető, funkciójával.

Egyrészt tehát a dolgok belső logikája vitt rá bennünket arra, hogy szövegeket is kezdjünk vizsgálni. A szókincs globális áttekintése olyan kérdéseket vetett fel, amelyekre csak szövegek vizsgálata adhatott választ. Másrészt bizonyos módszertani-kutatási stratégia is ezt indokolta. Nevezetesen: természetesnek tűnt az a gondolat, hogy miután egy elég nagy szótár anyaga, az egyes szavakhoz tartozó szükséges grammatikai információkkal, a gépi memóriában van, úgy majd e szótár s a megfelelő nyelvtani szabályok birtokában automatikusan megoldható a szövegek elemzése is. Részint a nemzetközi tapasztalat, részint talán bizonyos intuíció azt súgta azonban, hogy a szövegekre áttérve érdemes egy lépést hátra tenni: első közelítésben úgy vizsgálni őket, mintha szótárunk nem is lenne; azokat az információkat kivonni belőlük, amelyek onnan szótár nélkül is, minimális emberi előkészítéssel, kinyerhetőek. Ezzel egyben lehetővé vált, hogy orosz (és kisebb mértékben egyéb) szövegeket is feldolgozzunk, ahol az *Értelmező Szótár*éhoz hasonló hatalmas szótári anyag nem állt rendelkezésünkre. 1972 tavaszán lépett át munkánk ebbe a fázisába.

A kutatás egyes egymásra épülő szakaszai részben csak így, félútról visszatekintve, alkotnak ilyen logikusan egymásra épülő sort. A hagyományosabbnak mondható kutatásokkal szemben mi nem láttuk mindig pontosan előre, mi fog a kutatás adott állapotából mint késztermék előállni; milyen stádium fog a most megtettből következni. A vizsgált *a n y a g n a k* kell köszönetet mondanunk: előzmény, nemzetközi tapasztalat nem vagy alig lévén, jól rá kellett hallgatni az anyagra (részben meg az adott gépi lehetőségekre) – az maga lökte az embert egy-egy késztermék megalkotására, egyik kutatási fokról a másikra.

⁵ Bővebben l.: Ф. ПАИП: Два вопроса автоматического синтеза русских парадигматических форм. In: *The Slavic Word*. Ed. by Dean S. WORTH. Mouton, The Hague – Paris, 1972. 283–301 l.

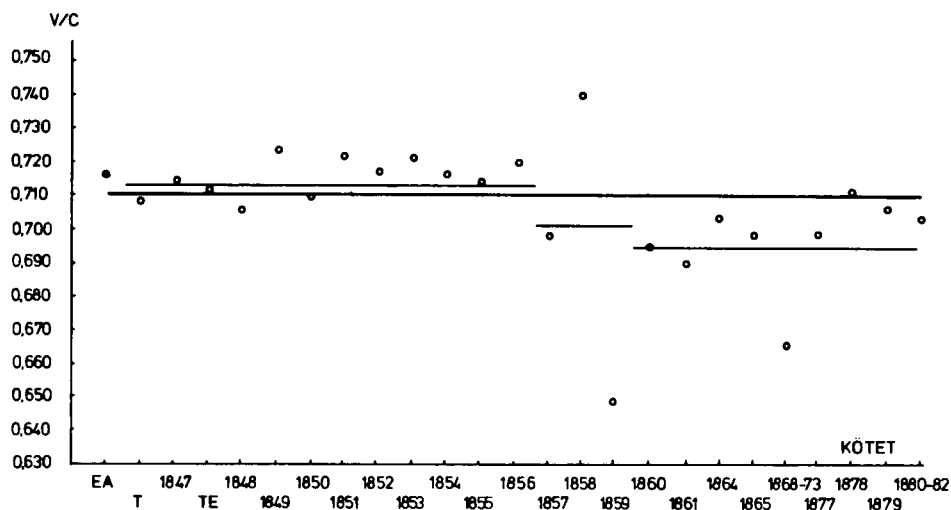
2. Balassi összes műveitől Juhász Ferenc egyes műveiig eddig több millió fonémányi magyar költői szöveget rögzítettünk; ehhez kontrollként egyes magyar szakszövegeket vettünk; rögzítettünk továbbá orosz és részben más nyelvű szövegeket is.⁶ Ennek az anyagnak tekintélyes részét (10–12%-át) teszik ki azok az Arany János-művek, melyek fonéma szintű elemzésének néhány eredményét az alábbiakban tárgyaljuk. Az 1. ábra mutatja, hogy összesen mintegy félmillió fonémányi szövegről van szó. Ennek durván szólva egynegyedét *Az elveszett alkotmány* teszi ki, egy másik negyedét a *Toldi* és a *Toldi estéje*; a fennmaradó részt az ún. „kisebb költemények” 1847-től 1882-ig, az utolsó töredékekig.



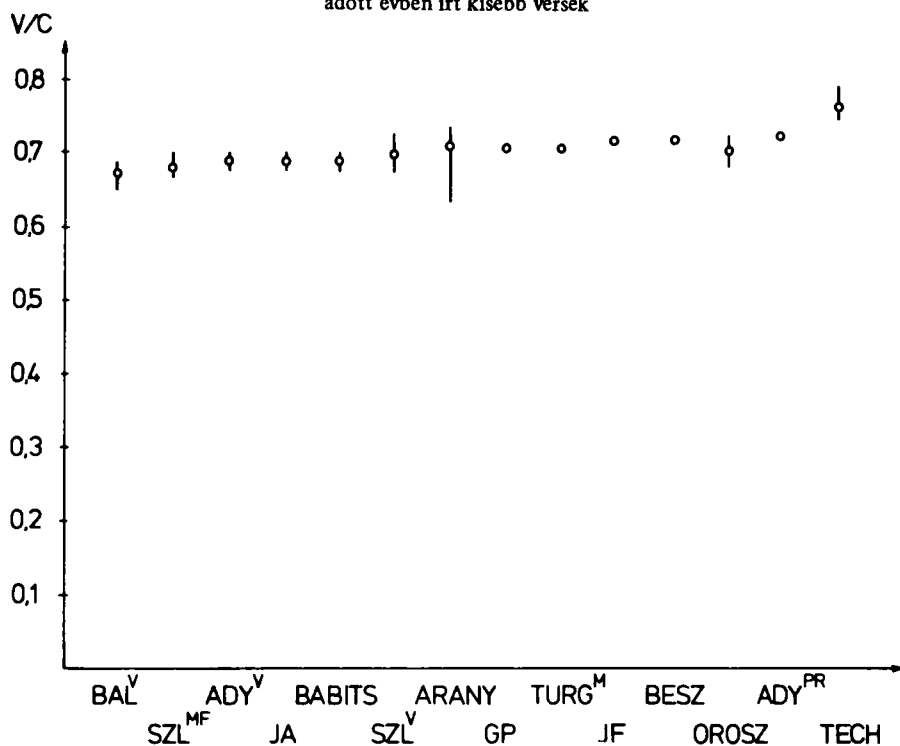
1. ábra. EA = Az elveszett alkotmány, T = Toldi, TE = Toldi estéje, Versek = Kisebb költemények 1847–1882. A számok az egyes korpuszok terjedelmét jelöljük ezer fonémában

A 2. ábra kiemeli még e legsőbb színtről is a legegyszerűbb kérdést: vajon mit mondhatunk magánhangzó–mássalhangzó eloszlás tekintetében Arany Jánosról? Az ábra erre hozzátételgesen időrendben ad választ: előbb következik az EA értéke, majd a *Toldi*, az 1847-es verseké és így tovább. Az egész vizsgált anyagban a magánhangzó–mássalhangzó arány 0,711 volt: 1000 mássalhangzóra 711 magánhangzó esik. (Ugyanez kifejezhető lett volna így is: az összes fonéma 58,44%-a mássalhangzó, 41,56%-a magán-

⁶ Az egyes magyar szerzők szövegeinek rögzítése a kritikai kiadás, annak híján – a legelfogadottabb jelenlegi szövegkiadás alapján történt. (A cikk gerincét képező Arany-szövegeket a VOINOVICH Géza által sajtó alá rendezett akadémiai kiadás megfelelő kötetiből vettük.) A kontrollként vett nem költői szövegek: UNVÁRSZKY Miklós: Fedettívű hegesztés és villamos salakhegesztés. OMK DOK, Házi sokszorosítás, Budapest 1968; Híradástechnika–elektronika–automatika. OMK DOK, Házi sokszorosítás, Budapest 1968; Magyar Hírlap 6. évf. 231. sz., 1973. aug. 23, csütörtök (minta: 53 cikk, ill. hír az újságból, az 1–9. és a 11. oldalról); ZALA Ferenc: Gazdasági mutatók a belkereskedelemben. Budapest 1963, 7–37. A magyar beszélt nyelvi adatok a következő munkából: SZENDE Tamás: Spontán beszédanyag gyakorisági mutatói. Budapest 1973. Hasonlóan vettünk mintát a megfelelő Puskin és Turgenyev kiadásokból; a nem költői szöveggént vett orosz minta a következő forrásból: Материалы XXV съезда КПСС. Издательство политической литературы. Москва 1976.



2. ábra. V = magánhangzó, C = mássalhangzó; EA, T, TE, mint az 1. ábrán; 1847, 1848 stb. – az adott évben írt kisebb versek



3. ábra. V, C, mint az 1. ábrán. A rövidítések feloldása (zárójelben a nagyobb korpuszok terjedelme ezer fonémában): Balassi *Összes versei* (102), Szabó Lőrinc műfordításai (239), Ady *Összes versei* (553), József Attila *Összes versei* (301), Babits Mihály *Összes versei* (406), Szabó Lőrinc *Összes versei* (639), Arany János vizsgált művei (498), Gulyás Pál: *Misztikus ünnepi asztal* (14), Turgenyev magyar műfordítás minta, Juhász Ferenc minta *A szent tűzön regéiből* (24), spontán beszédanyag 1. 1. lábjegyzet (80), Turgenyev eredeti minta, Minta Ady prózájából (407), technikai-politikai jellegű szövegek – 1. 1. lábjegyzet (184)

hangzó.) Az egyes kötetek (egy-éves verstermése) mesterségesnek tűnhet. Valóban, a következőt állapíthatjuk meg még erről az ábráról. Az EA eléggé magánhangzó-gazdag ($V/C = 0,716$), majd következik egy évtized, 1856-ig, amikor a V/C -érték ennél valamivel alacsonyabb, ám az átlagnál valamivel magasabb (0,713); néhány évi (1857–1859) erős ingadozás után ismét kialakul egy egység 1860-tól a költő élete végéig (0,696 körül).

Mindez beszédesebbé válik, ha megnézzük, mivel hasonlíthatjuk össze ezeket az értékeket. A 3. ábrán Balassitól mai technikai szövegekig látjuk a hasonló eredményeket; sőt, a magyar eredmények közé egy helyen beékelünk orosz nyelvi mérésekből kapott adatokat is. Ezen az ábrán az egyes szövegcsoportok, költői életművek növekvő V/C arányuk szerint vannak sorba állítva: középen megjelölve az egész átlaga, alsó és felső határ – a legmássalhangzósból, illetőleg a legmagánhangzósból kötet/évi termés (Balassinál az Eckhardt-féle csoportosítást véve alapul) (az orosz szövegek nincsenek átlagolva, csupán szóródásuk van mutatva). Talán a leglényegesebb, ami ebből az ábrából kiolvasható: a vizsgált szempontból a két nyelv szövegei között nem vonható határvonal – az orosz nyelvű minták a magyar spontán beszédanyag és Ady prózája között foglalnak helyet (ezen belül a legmagánhangzósbabaknak Turgyev vizsgált szövegei mutatkoztak, tompábbak a politikai nyelvből nyert eredmények, legkevesebb magánhangzót viszont Puskin vizsgált részlete tartalmazott). Az ábrán nem tüntettük fel, de ide vonhatjuk a más forrásokból rendelkezésünkre álló lengyel és ukrán adatokat: azok is ezen orosz értékek körül, tehát a magyar értékek közé ékelve szerepelnek.⁷ A második könnyen adódó következtetés: magán a magyar nyelvi szövegeken belül legalábbis Arany Jánostól kezdve e szempontból az idő mentén mérve nincs változás. Arany János vizsgált művei átlagukat tekintve Szabó Lőrinc összes versei és Gulyás Pál egy kötete között foglalnak helyet. A Balassi-eredményekhez hozzá kell fűznünk, hogy ezek azért állnak az élen s mutatják, hogy az elmúlt századok alatt mintha valamivel magánhangzó-gazdagabbá váltak volna mégis szövegeink, mert csupán a költő összes verseit vettük figyelembe. Ha leveleit, drámáját stb. is tekintjük, Balassi egészen belesimul ebbe a XIX–XX. századi csoportba.

Ha időben (és a környező szláv nyelvek tekintetében) nem is mutatkozik különbség, egy váratlan eltérés műfaji szempontból jelentkezik: Ady prózája magánhangzó-gazdagabb volt, mint versei (és egyáltalán: mint bármely eddig vizsgált magyar költői szöveg), aránylag a legtöbb magánhangzót pedig a kontrollként bevont technikai-politikai jellegű korpuszok tartalmaztak. A technikai-politikai szövegek természetesen jóval hosszabb szövegszókából is álltak, mint a széppróza és kivált a költészet, tehát a vizsgált magyar szövegekben bizonyos numeróztatás mutatkozik: a hosszú szavakat tartalmazó minták egyszerű (rövidebb, kevesebb mássalhangzót tartalmazó) szótagokból állnak és megfordítva. Ezek az eredmények összhangban vannak a korábban a szótárból nyertekkel (ahol, mint láttuk, kiderült, hogy ősi tőszavaink aránylag kevesebb magánhangzót tartalmaznak mai alakjukban, mint egyes fiatalabb etimológiai rétegek képviselői). Ugyanakkor ennek a megállapításnak, hogy ti. költői szövegeink mássalhangzó-gazdagabbak, mint a spontán beszéd, mint a nem esztétikai rendeltetésű szövegek, talán további kihatásai lehetnek a nyelvi esztétikum megközelítése szempontjából.

Még egy megjegyzést kell tennünk annak kapcsán, hogy, amint az a 3. ábrából látható, Turgyev eredeti orosz szövege némileg magánhangzó-gazdagabb, mint annak

⁷ Vö.: Д. М. СЕГАЛ: Основы фонологической статистики. «Наука», Москва 1972.

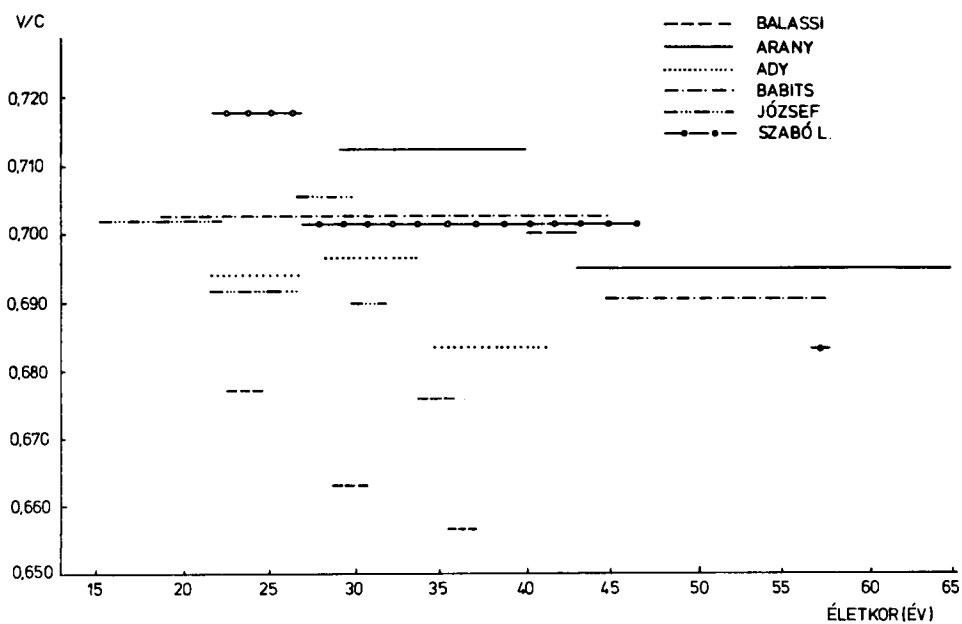
magyar fordítása (V/C: 0,730, ill. 0,718). A különbség nem nagy, a minták ebben az esetben elég kicsinyek voltak, így lehet, az arány nagyobb anyag, más minta esetében meg is fordulna. A lényeg azonban nem itt van. Hanem ott, ha megnézzük, milyen az orosz magánhangzók átlagos hossza és milyen a magyaroké⁸ (mind a két esetben úgy járva el, hogy egyrészt a hangsúlyos–hangsúlytalan, másrészt a rövid–hosszú magánhangzók értékét differenciáltan vesszük figyelembe, de ugyanakkor eltekintünk az egyes magánhangzók fajlagos és környezet-diktálta hosszúságától), akkor közelítőleg azt kell mondanunk, hogy a magyar szövegek összhangzásából mintegy kétszer annyi idő jut a magánhangzók ejtésére, hogy tehát összhangzásukat tekintve a magyar szövegek „kétszer olyan magánhangzósak”, mint a nekik megfelelő oroszok. Ez az eredmény jól megfelel intuitív elképzelésünknek erről a helyzetről; ugyanakkor felhívja a figyelmet arra, hogy a korábban oly szokásos „melyik nyelv a magánhangzó-gazdagabb” kérdést valószínűleg rosszul tettük föl: arra kaptunk választ, hogy darabszámra az olaszban, a finnben stb. van-e több magánhangzó – és nem arra, amit bele akartunk érezni: hogy ti., mely szövegek a ténylegesen „magánhangzósabbak”.⁹

A 4. ábra még egy másik oldalról közelíti meg ugyanezt a kérdést. Nevezetesen: ott több vizsgált költő életművében végeztünk el ahhoz hasonló átlagolást-csoportképzést, amelyet a 2. ábrán Arany számára elvégeztünk. Összefoglaltuk tehát azokat az éveket, amelyek a magánhangzó–mássalhangzó arány szempontjából kisebb ingadozással egységet mutatnak. Erről az ábráról leolvasható, hogy valóban, a többi költő életműve sem éppen évek, hanem nagyobb év-kötetcsoporthoz oszlik meg e szempontból. Ugyanakkor mintha egy olyan tendencia is kiolvasható lenne (vö. kivált Arany, Babits, Szabó Lőrinc eredményeivel – Szabó Lőrincnél nem vettük figyelembe az utolsó éveket összefoglaló *Valami szép* kötetet), hogy megfelelő ingadozás után és mellett a költő nyelve mintegy besűrűsödik – az egész alkotó periódus mentén előre haladva egyre kevesebb magánhangzót találunk. Különösen szembeötlő ez kivált egy alább tárgyalandó mutatóval összevetve; ott valami efféle trend nélküli ingadozást találunk az életmű folyamán. Ez eddig a véletlen műve is lehet; még jóval több költő életművét kell átvizsgálunk ahhoz, hogy akárcsak mint megfigyelt tény is ezt a tendenciát megállapíthassuk.

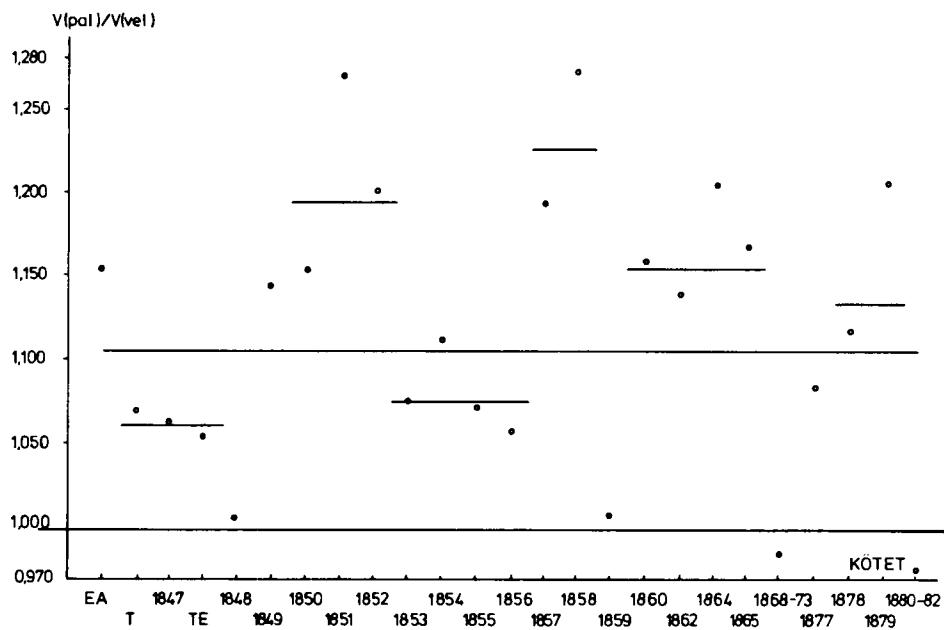
Nyelvünk egész szerkezete szempontjából nagy jelentősége van a hangrendnek, a palatális-veláris oppozíciónak. Az 5. ábra azt mutatja, hogy Arany vizsgált műveiben mi ezzel kapcsolatban a helyzet. Látjuk, hogy itt is kialakíthatók egyes évcsoportok (Arany esetében kisebbek, mint a magánhangzósság szempontjából kialakíthatók voltak): így a *Toldi*–1847-es „kis versek”–*TE* egy aránylag mély hangrendű csoportot alkot

⁸ Vö.: MAGDICS Klára: A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények* LXVIII. (1966) 125–139. KASSAI Ilona: Adalékok a magyar magánhangzók időtartamához. *Nyelvtudományi Közlemények* LXXIII. (1971) 171–190; К. БОЛЛИА: К вопросу соотношения длительности гласных и фонетической структуры слога. *Studia Slavica*. XIV. (1968) 75–87.

⁹ Az a gondolat, hogy a) a melódikusság a magánhangzókkal (esetleg még: a szonánsokkal) kapcsolatos és b) ezek egyszerű előfordulási arányával, újból s újból felbukkan az irodalomban. Egy nemrég megjelent s egyébként igen tanulságos cikk következtetéseiből: „4. A manysi nyelv dallamos [мелодичен], a hangzó manysi beszéd 70%-a szonánsokból és magánhangzókból áll”. (Ю. А. ТАМ-БОВЦЕВ: Некоторые характеристики распределения фонем мансийского языка. *Советское финно-угроведение*. XIII. (1977) 195–198.) A 70%-ot természetesen itt is „darabszámra” kell érteni, a szerző mintegy félmillió fonémát tartalmazó beszélt manysi szöveget dolgozott fel számítógéppel.



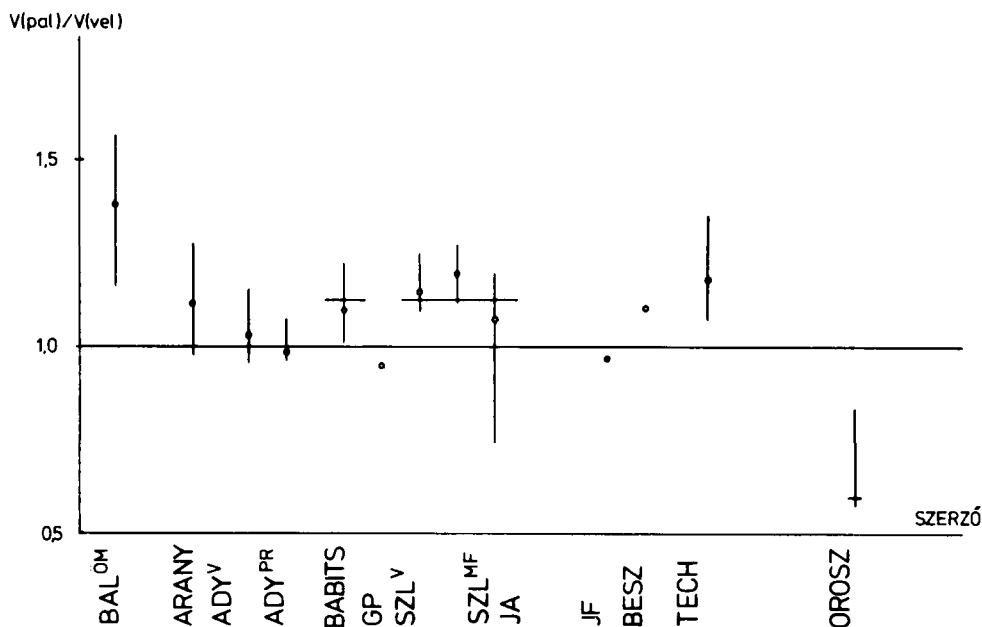
4. ábra. V, C, mint a 2. ábrán



5. ábra. V(pal) – palatális magánhangzók, V(vel) – veláris magánhangzók. Egyéb rövidítések, mint a 2. ábrán

[V(pal)/V(vel) = 1,061], az 1850–52-es „kis versek” egy következő, lényegesen magasabb hangrendűt (1,194), és így tovább. Látjuk, hogy az egész életműben enyhe palatális dominancia tapasztalható, az 1868–73 közötti, valamint a kis terjedelmű 1880–82-es versek, töredékek csoportjától eltekintve. A palatális-veláris Arany-átlag körül fellépő ingadozást tudjuk mintegy normálisnak, különösebb trend nélkülinek elképzelni; hasonló volt e szempontból a helyzet az eddig vizsgált többi költő esetében is.

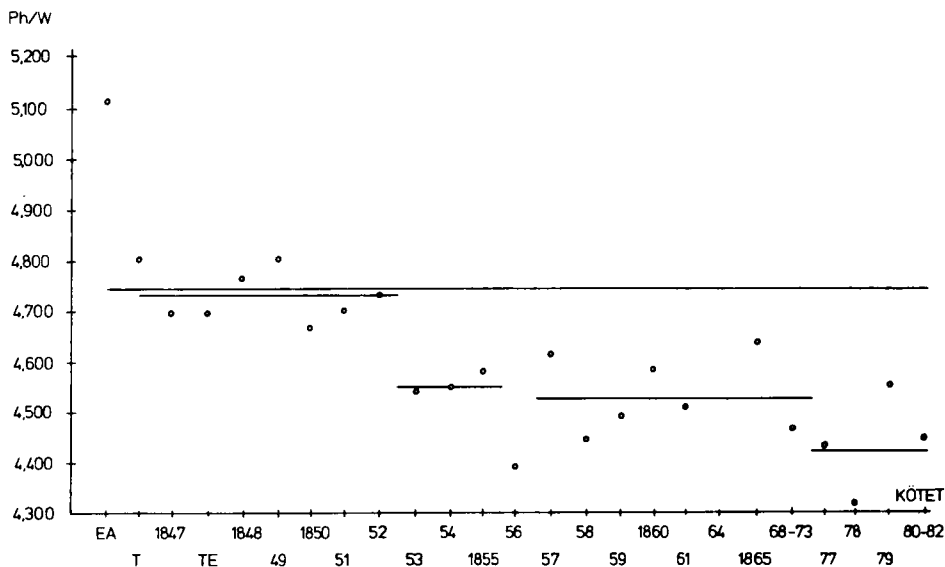
A 6. ábrán az egyes szövegcsoportok keletkezésük hozzávetőleges időrendje szerint vannak sorba rakva. Szembeszökik az idő mentén észlelhető különbség: Balassi összes művei sokkal palatálisabbak, mint a XIX–XX. századi szövegek (BAL: 1,386–XX. sz.: 1,120). Ha visszpillantunk a szókincsben nyert eredményekre: ez a mélyülés részben úgy



6. ábra. Balassi *Összes művei* (272 ezer fonéma), egyéb rövidítések, mint az 5. ábrán

ment végbe, hogy a fiatalabb etimológiai rétegek eleve több veláris magánhangzót tartalmaznak; a veláris illeszkedés még ennél is gyakoribbá válhatott. Részben már Balassi korát megelőzően jellemző volt az egyre mélyülő magánhangzó-összetételű szavak tömeges kölcsönzése. Balassinál korábról csak egészen kis szövegtöredékeket dolgoztunk fel eddig; ám el lehet képzelni, hogy korábban még palatálisabb színűek voltak szövegeink – annak ellenére, hogy, mint láttuk, az ősi szókincs úgyszólván fele-fele arányban tartalmazott mély és magas magánhangzókat. Ha feltételezzük, hogy egy strukturális szabályszerűség akkor van ereje teljében, ha nagy anyagra érvényes, akkor azt mondhatjuk, hogy a századok alatt tapasztalható fokozatos mélyülés előnyére vált az illeszkedésnek: az oppozíció ma mennyiségileg a legnagyobb, a szövegszóknak úgyszólván a fele a magas, a fele a mély hangrend szerint kell hogy illeszkedjék. És ez a „ma” jól látható módon legalább Arany

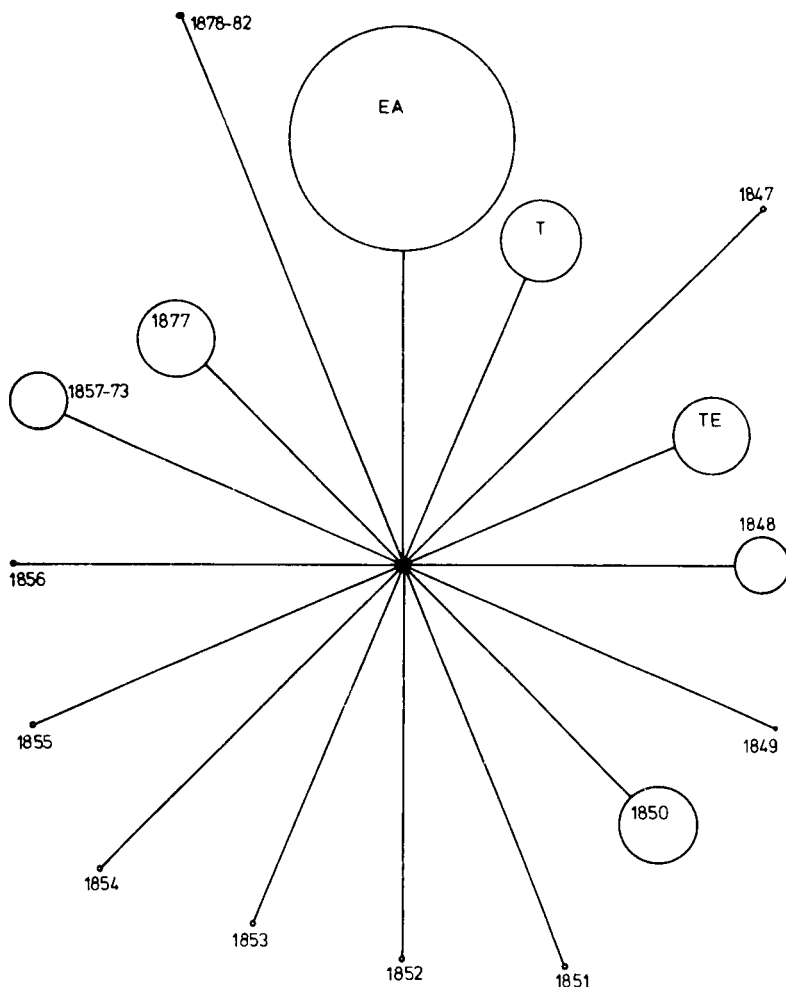
Jánostól (öt közvetlenül megelőző szövegek megint nem állnak még rendelkezésünkre) napjainkig észrevehető megszakítás, törés nélkül tart. Míg Aranynál 1000 veláris fonémára 1116 palatális esett, addig a XX. században kialakult költői átlag: 1000 velárisra 1120 palatális – egy hajszálnyival magasabb. Tekintsük meg egyébként ezt a XX. századi magyar költői átlagot: úgy nyertük, hogy Babits, József Attila összes verseit, Szabó Lőrinc összes verseit és műfordításait egyetlen korpusznak tekintve, ebben a nagy korpuszban állapítottuk meg a mély–magas arányt. (Látható, hogy Ady ennél egy kissé, Gulyás Pál feldolgozott kötete jelentősen mélyebb – ez utóbbi azonban igen kis minta



7. ábra. Ph – fonéma, W – szövegszó, egyéb rövidítések, mint a 2. ábrán

volt, eltérése tehát pusztán a véletlen műve is lehet.) Arany tehát e szempontból nyugodtan besorolhatna Babits és József Attila közé – „nyugodtabban”, vagyis kisebb eltéréssel, mint Ady. Ezt valószínűleg úgy kell értelmeznünk, hogy ebből a szempontból legalább Arany óta nem észlelhető mozgás nyelvünkben: a költői szövegek egyazon, 1,050–1,120-as érték körül mozognak. Jól látható, hogy míg a magánhangzó-gazdagság szempontjából Ady prózája és költészete erősen különbözött egymástól, addig a magas–mély magánhangzóság szempontjából alig: 1,050 a költészet, 0,992 a vizsgált évek prózája. Ehhez a XIX–XX. századi magyar költői maghoz képest mélyebbnek mutatkozik Juhász Ferenc – magasabbnak a spontán beszélt nyelvi meg a technikai szöveg; ám itt a minták kicsisége miatt nem érdemes még következtetéseket levonnunk. (Önmagukban ugyan elég nagyok lehetnének ezek a minták. Amikor azonban egy ilyen alig változó, igen finoman ingadozó jelenségről van szó, kénytelenek vagyunk igen nagy mintákat igen jól megválasztva feldolgozni.) Ismét csak szembeszökő viszont, hogy a magyar mintákhoz képest milyen kiugróan mélyek az oroszok (0,823, 0,587, 0,565). Első pillantásra azt

mondhatnánk, hogy ahol nincs palatális-veláris illeszkedés, ott megengedhető vagy könnyebben elképzelhető az ilyen arányeltolódás valamelyik irányban. A mi mintánk kicsiségét ellensúlyozzák az oroszon végzett korábbi mérések eredményei, melyek a mieinkkel megegyeznek.¹⁰ Ám rendszeresen zavarhatja az eredményeket a mi fonéma-felfogásunk, amellyel a mérés történt; a lágy mássalhangzó-környezet figyelmen kívül



8. ábra. Rövidítések, mint a 2. ábrán. Az egyes körök központjának az ábra központjától mért távolsága arányos a kör által szimbolizált kötet (év) többi kötethez (évhez) mért átlagos távolságával; a körök sugara arányos az egyes kötetek (évek) fonématerjedelmével (Pontos számszerű adatokat l. az 1. táblázaton.)

¹⁰ L. Л. А. БУЛАХОВСКИЙ: Курс русского литературного языка. Том I. «Радянська школа». Киев 1952. 28–30. Ugyanott hivatkozások korábbi hasonló statisztikai vizsgálatokra (БОГОРОГИЧКИЈ, ПЕСКОВСКИЈ).

hagyása stb., tehát célszerűbb az orosz mérési eredményből egyelőre semmiféle következtetést sem levonnunk.

Mintha Arany művein belül az átlagos szóhossz is csökkenő tendenciát mutatna. A 7. ábrán az egyes Arany-kötetekben (években) rejlő szavak átlagos hosszúságát tüntették fel (fonémákban mérve). Az *EA* kitűnik rendkívül hosszú szavaival (átlagos szóhossz: 5,111 fonéma – ezért is lett az Arany-átlag, mint látható, magasabb a további képezhető csoportok átlagánál). Az ezt követő csoportok: a *Tolditól* az 1852-es versekig (4,747), 1853–1855 (4,556) – majd elég nagy szétszóródottság tapasztalható e szempontból, de továbbra is talán csökkenő tendenciával. Meggondolandó, milyen viszony áll fenn az *EA* hexaméteres versformája és nagy átlagos (szöveg)szóhossza között; továbbá e tény s a korábban említett nagyobb magánhangzóarány között.

Végül a 8. ábra az egyes vizsgált Arany-kötetek távolságát és terjedelmét tükrözi, a terjedelmet fonémákban mérve, a távolságon (\bar{D}) pedig az egyes fonémák hasonló vagy eltérő terheltségét ábrázolva. (Tehát: közel van két kötet, ha különböző, páronként megegyező, fonémák arányai a két kötetben közel azonosak; távol, ha nem ez áll fenn.) Szembeötlük mindjárt az *EA*: elég hosszú, 126 ezer fonémányi, mégis viszonylag messze került a többi kötettől.¹¹ „Legaranyibb” hangzású (legközelebb van a többi eddig vizsgált Arany-kötethez) a *TE*, továbbá a *Toldi*, közel vannak a többihez az 1850-es versek (*Emléklapra, Furkó Tamás, Évek, ti még jövendő évek, Koldus-ének, Letésem a lantot* stb.), továbbá az 1877-es versek (*A lepke, Epilógus, Tamburás öreg úr, Honnan és hová, Tengerihántás* stb.). Az említett *EA*-on kívül messze vannak a többtől kivált az 1880–82-es utolsó versek, töredékek, valamint az 1847-es kisebb költemények (*Válasz Petőfinék, A varró leányok. A méh románca* stb.). További ehhez hasonló megfigyelések leolvashatók az 1. táblázatból. Így ott kiderül, hogy ha a kötetek (évi termések) közti hangzásbeli távolságot tekintjük, legközelebb egyes évek a *TE*-hez vannak (amelyhez, mint láttuk, általában minden közel van): így kivált a *Toldi*, továbbá az 1848-as és 1849-es versek; eléggé közel vannak továbbá egymáshoz az 1850-es és 1877-es költemények. A legnagyobb távolságok viszont az utolsó versektől-töredékektől adódnak: ezektől legtávolabb vannak az 1847-es, 1852-es, 1849-es alkotások; viszonylag a legközelebb ezekhez is a *T* meg a *TE* áll. (A szintén igen elszigetelt *EA*-hoz viszont időbeli társai állnak a legközelebb: az 1848-as versek, valamint a *T* meg a *TE*.) Ezek a nyers eredmények nyilván még további nyelvi vizsgálatokra és meggondolásokra kell hogy ösztönözzenek; esetleg azonban már így is mondhatnak valamit az irodalom kutatóinak is, akik viszont hozzásegíthetnek bennünket a jobb nyelvi elemzéshez.

Ahogy a szótári anyagokon, úgy természetesen a szövegekben végzett számlálgatások sem lehetnek öncélúak, már ezen a legalacsonyabb szinten sem. Velük egyrészt igyekeztünk közelebb férkőzni nyelvünk egyes mai szinkrón, valamint diakrón összefüggéseihöz. Másrészt törekszünk bekapcsolódnival egyes költők életművének értékeléséhez; nem csupán a múlthoz, hanem Akadémiánk jelenéhez kívánunk kapcsolódnival, amikor egyelőre igen durva és alacsony szintű (a fonémák szintjén végzett) kutatásaink tárgyául az elsők között olyan szerzőket választottunk, akikről irodalomtörténeti-esztétikai, tehát

¹¹ Vö. ezzel azt a korábbi eredményünket, amely szerint Ady 1899-es „Versek” c. kötete igen távol van a többi Ady-kötettől: JÉKEL Pál–PAPP Ferenc: Ady Endre összes költői műveinek fonéma-statisztikája. Bp., 1974. Akadémiai Kiadó. – 43. l.

1. táblázat

	D/A	EA	T	TE	1847	1848	1849	1850	1851	1852	1853	1854	1855	1856	1857- 73	1877	1878
A	0																
EA	45	0															
T	39	62	0														
TE	38	69	42	0													
1847	92	106	102	84	0												
1848	40	56	52	46	90	0											
1849	60	88	69	46	97	76	0										
1850	37	60	56	67	106	62	83	0									
1851	64	82	83	77	105	57	97	71	0								
1852	53	81	73	71	110	68	81	57	57	0							
1853	59	95	72	59	99	77	65	75	85	73	0						
1854	75	109	91	79	118	90	81	92	91	74	66	0					
1855	65	100	65	84	70	104	84	80	91	74	53	66	0				
1856	69	106	88	75	112	93	72	77	100	77	54	63	55	0			
1857-73	57	77	79	81	127	77	92	55	86	72	84	92	77	80	0		
1877	45	75	64	68	100	75	72	49	88	67	67	78	60	55	61	0	
1818-80	111	123	109	112	150	117	135	119	144	145	116	133	130	122	124	111	0
ΣD		1289	1107	1060	1586	1140	1238	1109	1314	1180	1140	1323	1193	1229	1264	1100	1890
\bar{D}		86	74	71	106	76	83	74	88	79	76	88	80	82	84	73	126
SD		20,1	18,5	17,9	18,5	20,1	19,6	19,8	20,8	22,0	17,5	19,7	20,6	21,3	19,4	17,9	13,0
terjedelem		125,7	48,5	51,8	5,8	30,0	15,0	40,3	18,9	19,5	24,2	9,0	16,4	13,6	32,2	42,8	4,2

Rövidítések: A = az összes vizsgált Arany, D/A = az egyes kötetek távolsága az összes vizsgált Aranytól;

ΣD = az egyes kötetek többitől való távolságainak összege;

\bar{D} = e távolságok átlaga;

SD = a ΣD adatok standard hibája
terjedelem ezer fonémában.

egy igen magas szinten, napjainkban egész monográfia-sorozatok látnak napvilágot (Ady, József Attila). Közvetlenül mérési eredményeink valószínűleg felhasználhatóbbak lesznek az irodalmárok, a költői nyelv kutatói számára, ha a fonémákon nyert eredményeket áttesszük hangokban értékelhetőkké; ha a hangokat több irányú vizsgálatnak vetjük alá. E szempontból sokat ígérőnek tartjuk a hangszimbolika terén folytatott, fonostilisztikai kutatásokat. Az egyes költőkről, korokról felhalmozott fonéma-anyag talán beszédesebbé válik, ha jobban látjuk: egyes hangok, hangcsoportok mit is „jelentenek” esztétikai-asszociációs síkon a mai magyar beszélőnek-hallgatónak.

3. Ha egy szinttel magasabbra, a szavakéra lépünk, elsődleges kiinduló gépi termékünk a konkordancia lehet. (A 9. ábra erre mutat példát: a Toldi „Előhangja”-nak és „Első éneké”-nek az összevont konkordanciájából mutatunk ott egy részletet.) Egy ilyen konkordanciából viszonylag könnyen továbbmehetünk a költői szótárak irányában: hiszen az ilyen listákon a fásasztó, gépies tevékenység el van végezve, az egyes lexémák különféle alakú előfordulásai többé-kevésbé egy-egy csoportba vannak gyűjtve. Ezek után az igazi emberhez méltó munka marad: ezekből az egybegyűjtött előfordulásokból megírni az egyes szócikkeket. Az emberi erővel elkészített, illetőleg készülőlében levő költői szótárak tüzetes tanulmányozása nélkülözhetetlen ahhoz, hogy ezekből a közvetlen gépi termékekből használható költői szótárakat hozzunk létre. (Esetleg nem is éppen a példaként felhozott Arany Jánosból s éppen a *Toldi*ből, hiszen annak szótára, úgy tűnik, legalábbis kéziratban megvan.¹² Bár lehet, hogy legalább részben ezt a munkát újra el kell végezni, vö. alább, 4.)

A tömegesen és könnyen előállítható költői s gyakorisági szótárak révén megint lehetőségünk nyílik bizonyos nagyobb összefüggések együttes szemlélésére. Erre két példát idézünk.

Az elsőben még mind az orosz, mind a magyar anyagot fáradságos kézi munkával állították elő. A rendelkezésünkre álló adatok itt csak azt teszik lehetővé, hogy két egészen más kort, más stílust és jellemet vessünk egy pillanatra össze – Puskin és Juhász Gyulát.¹³ Költészetük legkedveltebb szavai között egyaránt előfordulnak az olyan jelentésűek mint 'lélek', 'éj', 'megy', 'élet', 'ég fn' stb. Szembeszökik ugyanakkor egy különbség is: Puskin, a politikus költő, a romantikus-nacionalista XIX. század költője, háromszor-négyszer olyan kevésszer írta le költeményeiben a *русский* 'orosz' szót, mint Juhász Gyula a *magyar*-t (melléknévi és főnévi használat együtt mindkét esetben). Juhász Gyula költészetében ez a 28. leggyakoribb elem; Puskin költői szókinésében a 124. Másképp szervezett emberi memóriánk könnyen feldob egyes példákat, mint a költőtől gyakran idéztet, rá jellemzőt: *Там русский дух . . . там Русью пахнет* (a *Ruszlán és Ljudmila* előhangjából: Hazai ott a lég s az illat¹⁴) – és viszont: *Magyar táj, magyar ecsettel*. Ám pontos összképet vagy csak fáradságos emberi tevékenység (és akkor csak

¹² Vö. VISKI Károly: Mutatvány az Arany-szótárból. *Magyar Nyelvőr* 72. (1948) 28–30., 70–74.

¹³ Az adatok a következő két forrásból: АН СССР. Институт русского языка. Институт языкознания. Предварительные публикации. Материалы к частотному словрю Пушкина. (Проспект) Москва 1963; BENKŐ László (szerk.): Juhász Gyula költői nyelvének gyakorisági szójegyzéke. *A Szegedi Tanárképző Főiskola Tudományos Közleményei* 1973. Első rész. Szeged 1973. 95–122., kül.: 101.

¹⁴ FODOR András fordítása

- 10 -

I +	MERT VIT<Z VOLT APJA	6-003
ORSZ>G>THA MESSZE,	MESSZE B>MUL, MINTHA	3-003
VASKOS L>BANYUM>TOL	MESSZE RENÜ A PARLAG	13-002
SZ<LES ORSZ>G>THA	MESSZE, MESSZE B>MUL	3-003
G ASZI <JTSZAK>KON,	MESSZIKAL LOHOGVA TE	1-001
&LD HELY'NAGY HAT>R	MEZABEN. + NINCS EGY	1-002
E B>MUL, MINTHA M>S	MEZAKRE V>GYNA E HAT	3-003
KARJA, MAGA V>LVA,	M<G CSAK MEG SE RÉZZ	9-005
D>S VOLT. + = AMIT	M<G E SZÓHUZ GONDOLT	9-003
ASZGS V>LL>N, PEDIG	M<G LEG<NYITOLL SEM P	3-002
EGY PARASZT FI>VAL	M<G SEM >LL KI SENKI	10-005
, MORIHORGAS G<MMEL	M<LYEN N<Z A K>TBA S	2-001
HELY SOV>RGO LELK<T	M<LYEN SZ>NTOGATJA,	6-001
LATTON CIFRA NYERG#	M<NEKI N<ZI MIKLÓS,	7-003
AS >TAT V<GES-V<GIG	M<RTI T>L A TORNYON,	4-001
A? +	M<RT NEM JASZ CSAT>R	12-004
BEN? + SZ<P &CS<M,	M<RT >LLSZ OTT A NA	3-005
+ ILYENFORMA TOLDI	MIKLÓS GUNDOLOTJA, M	6-001
PUSZTAS>GONI TOLDI	MIKLÓS K<PE UGY LOBO	1-001
NYERG# M<NEKI N<ZI	MIKLÓS, N<ZI, S DEHO	7-003
AL, <S EL>MUL RAJTA	MIND EG<SZ HAD>VAL,	10-001
A S ELJANE K<Z<TEK,	MINDEN DULG>T SZEM<	1-004
>JASAT VAGY SZ<PET,	MINDENIK DEROT R> EG	12-002
MOSOLYGO K<PETI +	MINDENIK MUND N<KI N	12-002
YAL>ZATI Z>G, MOROG	MINDENKI, EGY PARASZ	10-005
ZELET, MINT AZ, AKI	MINDJ>RT MEGBIKKOZIK	3-007
><, OLYAN A J>R>SA,	MINT A BARNA <JF<L,	13-003
, SZEME PILLANT>SA,	MINT A S<RIETT VADKA	13-003
OLYAN FORGO SZELET,	MINT AZ, AKI MINU>R	3-007
I MELLETT NATT FEL,	MINT BAR>TJAJ + UYAR	6-004
KANNYEDEN FORGATJA,	MINT CSEK<LY ROTOCSEK	9-004
ALTA R<GIGIS<HEN. +	MINT HA P>SZTORTY>Z <	1-001
R>D VAS KEZ<HEN. +	MINT KOMÜR BIK><, OL	13-003
, + ILY LEG<NYNEK,	MINT TE, OTT VAN >M	12-005
JA SEDESEN HALADVA,	MINTHA F>STOK>DO NAG	3-007
AGJ>NAK. + R<MLIK,	MINTHA L>TN>M TERMET	1-002
K<T SZOLGA HORTYOG,	MINTHA LEGJOBB RENDI	1-003
SSZE, MESSZE B>MUL,	MINTHA M>S MEZAKRE V	3-003
N<Z ELARE HAJOLVA,	MINTHA SZIVE=LÉLKE A	4-004
OL BUO>RA T<R EL, S	MINTHA VASS>> VOLNA K	9-005
IR>LYUUVARBAN. + S	MOG A B<KESEKKEK GY#	6-005
R EZ MAG><RT" LACFI	MOND " AK>KKI; + " E	10-002
K<PETI + MINDENIK	MOND N<KI NY>JASAT V	12-002
, NEKIK JO<JTSZAK>T	MONDANI &R<KHE? + ME	5-005
+ " M>SIK SZ>NVA	MONDJAI: " SZ<P &CS<M	12-006
A T>BOR, CSILLAPUL	MORAJJA: EZT A SZ<L	13-001
N <S GYAL>ZATI Z>G,	MOROG MINDENKI, EGY	10-005
NDENIK DEROT R> EGY	MOSOLYGO K<PETI + MI	12-002
E, A VIL>G<RT SINCS	MOST EGER<SZNI KEDVE	3-007
NEK HINN<TEK. + HA	MOST FÉLT>MADNA S EL	1-004

9. ábra. Részlet a *Toldi* Előhangjának és Első énekének közösleges konkordanciájából (az *m*-mel kezdődő szövegszók *mert*-től *most*-ig). A jobb szélső oszlop első száma a versszakra, a második a mondat sorszáma utal.

egy-két kiemelt költőre), vagy tömeges gépi munka eredményeképpen láthatunk; csak ezt követően merhetjük felvetni a kérdést: vajon véletlen-e ez az eltérés.

Másik példánk nyelvi oldalról rávilágít arra, hogyan van a mi memóriánk szervezve. Az egyes szavak kiváltotta asszociációkról van szó. Egy kísérletről, amelyben a kísérleti személyek előtt sorban felolvassnak egy-egy szót, és ők azt a feladatot kapják: írják le az első erről eszükbe jutó lexémát, kifejezést stb. Amikor egy kisebb kísérleti személycsoportnak hangoztattuk a magyar *bemegy* szót, azok jelentős része így reagált rá: *kijön* (és csak harmadannyian így: *kimegy*, pedig ez utóbbi alak formailag közelebb látszik lenni az asszociáló szóhoz); a *keres*-re a többség *talál*-t mondott, és így tovább. A tömegesen előállítható írói–költői konkordanciák felkinálják a lehetőséget, hogy megvizsgáljuk: milyen egyéb szavak környezetében használja a költő ezt vagy azt a szót. E vizsgálat eredményeit összehasonlíthatjuk a mai magyar beszélőkről készült felméréssel, ahol az ő jellegzetes asszociációikat rögzítettük. E tipikus magyar reakciók várhatóan el fognak térni például a tipikus oroszoktól. A magyar *föld* szóra ilyen válaszokat kaptunk csökkenő gyakoriságuk sorrendjében: *ég, víz, zöld, anya-, eső, fa, kert, -tulajdon, világ* stb. A *земля* 'föld' szóra az orosz anyanyelvűek („orosz anyakultúrájuk”) ugyanakkor így reagáltak: *родная* (ennek megfelelő melléknévünk nem is igen van, ez az „anyaföld”), *круглая* 'kerek', *черная* 'fekete', *планета* 'bolygó', *небо* 'ég', *люди* emberek' stb. A magyar 'föld' asszociációs körében, mint láttuk, elég hamar megjelent az 'eső' – az oroszából ez (*дождь*) majdnem teljesen hiányzott, viszont megvolt a *снег* 'hó': az oroszából hiányzott a 'tulajdon' jelentésű szó, és így tovább.¹⁵ A költő nyilván egyrészt saját népe-nyelve-kultúrája asszociációs körére támaszkodik; másrészt új és új kontextusok létrehozásával igyekszik új jelentéseket adni a szónak.¹⁶ Sőt, nyilván ki kell lépünk pusztán a szavak világából: a művészetek jelei is egyrészt ugyan stílushoz-korhoz kötött jelentésűek egy nagyobb kultúrkörön belül; másrészt azonban mindig nemzetiek is, az adott nép történetéhez, művészeti előzményeihez, hagyományaihoz kötődnek. A művész nem menekülhet: egy-egy nemzetközi áramlatba föl-földobott kőként visszahull; mindig magyar tájat is fest magyar ecsettel, magyar vágyakat is mond el magyar szavakkal. A nyelvvel s a művészetekkel foglalkozó „lágy” társadalomtudományok, valamint a „kemény” társadalomtudományok és a természettudományok közti különbség az itt vázolt fogalomrendszerrel is megragadható. A Földön mindenütt ugyanazok az aerodinamikai törvényszerűségek érvényesek – ám mi éppen azt vizsgáljuk, mit jelent, esetleg társadalmi rétegenként is különböző módon, az *air*, a *Luft*, a *levegő*, a *воздух* stb. az egyes népek képviselői számára: nem magát a jelenséget vizsgáljuk, hanem annak nyelvi tükörképét; az egyes művészetekben megalkotott tükörképét. Egységesen és keményen meg van határozva a 'proletárdiktatúra' fogalma: egyesek nem kívánják magát ezt az objektíve létező jelenséget alkalmazni – mi azt vizsgáljuk, milyen asszociációkat, történetileg-társadalmilag meghatározott válaszokat váltanak ki a *dictature du prolétariat* kifejezés és megfelelői. A jelzett kutatási irány módot nyújt arra, hogy kemény, kísérletes eszközökkel mérjük ezt a jelenségcsoportot, mely természetesen nem lágy, az egyéni tudattól függetlenül objektíve

¹⁵ Az orosz adatok a következő forrásból: Словарь ассоциативных норм русского языка. Под редакцией А. А. ЛЕОНТЬЕВА. Издательство Московского университета. Москва 1977.

¹⁶ Vö. S. MARCUS: A szótári és a kontextusbeli jelentés. – A nyelvi szépség matematikája. Budapest 1977. Gondolat, 129–144.

- 7 -

NY ANYJ>NAK KENNY TOLUL SZEM<RE	, KASZIVH FI>NAK SO	9-002
, OTT FELI<T A T<RDEN TENYER<RE	HAJT>, S ZOKOGOTT	10-002
IK ANYJ>NAK, MINTHA A SZIV<RE	NAGY K<ST M>RTAN>NA	7-007
K + IDE TOLDTNAK A SZO SZIV<RE	NYILALITK, <S OLYA	9-004
FSERVES ANNAK, AKI JUT KF7<RE	, MEGHAIY ANYJ>NAK	12-004
ANGYAT E CSAP>SSAL HRS VEREMRE	TENN<K, ISTEN KENY	15-001
A SZEM<TS F>J A7 FRAS N<Z<SRE	+ <S IIT>NA NYALKA	8-003
, V<ZT, S DEHOGY VFSZI <SZBE	, HOGY A SZEME IS F	8-003
A R<RFSFKKEI, HOVATOM = " DE	GYRGGY >R EZT RIKOI.	7-005
APTG SEM FORPADOZNA RSSZF ; DE	MIDAN A7 BCCSE <PEN	15-001
K HAD>VAL H>NOR>RA KFLNEK ; DE	FELHIT IACKO A B<RE	2-002
DE DE GYRGGY KOPHOLJA DIUV>N ; DE	AZ ELTASZOTJA TESTV	9-002
AMUD>RJA KETEN A MIUNK>NAK, DE	FOGA NEM FELT AHHO	10-004
GYR>V>N, OTT BEMENYKEDIK, DE	GYRGGY KORHOIJA DUR	9-002
AJT>, S ZOKOGOTT MAG>RAN, DE	SENKT SEM HALL> +	16-002
S, NYERS, HARAGOS >RAN, DE	VAN DR>GA DOLOG OTT	1-001
MST<NT, AMI ENGFH ILLFT + .DE	AMI AZ FNYEM, AZT	13-003
>, HOGY KIVUL IS HALLT + IDE	TOLDTNAK A SZO SZIV	9-004
MANKOS, VTHAROS BDR>VAL + ?DE	KI VONA BAJT A7 <GT	12-001
T>N MEGUNTA CY>SZOR SZVEGYS<RE	>GY>T, S M>SNAK AD	4-004
.>TOK <S HAZUD>G MINDEN IGE	SZ>ORAN +	12-001
OST MEGYVENED MAG>VAL, RENYHE	S>SKA N<PNEK PIISZT	6-003
SZU>LL>S LFLKE KPLTATIK REI<JE	+ TOLDT MIKL<SNAK	14-004
MIKL<SNAK SINCS >M GALAMPEK<JE	, BOSSZU>LL>S LELKE	14-004
T TOLDT GYRGGYDE SZ>JA ; LFLKE	GYSLAL>G<N FRAT VF	8-005
ALAMPEK<JE, BOSSZU>LL>S LELKE	K<LTRZTK BEL<JE +	14-004
E HAJTOLVA, MINTHA SZIVE-LFLKE	A SZFM<RFN VOLNA +	5-004
L + <S ANYNT REBER K<L SZHKKE	POR K<D<RAL, UGY K	5-003
KELL + I"SZ<N>T HORN STE<NYKE	KUNN A R<RFSFKKEL,	7-005
MSTKE FEGYVER CSTLLOR, BMSZKE	HANSFRG K<L + JB	5-002
VOLT TEM<PDEK, S ARA BMSZKE	M>JA, SOK NEMES VT	6-002
MSTKE HANSFRG K<L + JBMSZKE	FEGYVER CSTLLOR, R	5-002
<GY H>NYJA, HTNI K<S7HI <KLE	CSONTOS RUZOG>NYA	14-005
ONDTH NCHA I NEM V>IK BELALE	SEMHT, CSAK NAGY L	11-001
DEPKY FAJANKO V>IK <GY BELALE	+ JM>RTSAD TELBE<V	10-003
THE<VAIHA, MIT SE KOD<LT TALE	, MAJD DFR<K FAJANK	10-003
A, MINT A BARKA <IF<L, SZEME	PILLANT>SA, MINT A	14-003
OGY VESZT <S7RF, HOGY A SZEME	IS F>J A7 FRAS N<Z<	8-003
KLF CSONTOS RUZOG>NYA + ISZEME	, MINT AZ AC<L, A	14-005
T, HOGY SENKINEK I>R ALATT NE	LEGYEK, NEM B>NDM	13-001
ADTAM + .NE	MOND, HOGY KI NEM	14-002
N, DE SENKT SEM HALL> + M<NE	FLRUSULVA, N<M>N H	16-002
FS SZOIG>JA, SOK NYFROTA M<NE	, NAGY SFREG KIITY>J	6-002
MEI M<LYFN N<Z A KTRA S BENNE	VIZET K<MEL ; ORI>S	2-001
, MINTHA IEGJOH RENDTN MENNE	DOIGA + .ROGLY>K H	1-003
+ IHA MOST FELT>MADNA S FLJANE	K<7<TEK, MINDEN DO	1-004
HSZTAS>GON ; TOLDI MIKL<S K<PE	UGY LOBNG FEL N<KEM	1-001
+ NYFLVEI A KUVASZ IS F<LDRE	HENGEREDVE, A VIL>	4-003
TT KVV<LY HAD>VAL IACZFT ENDRE	MAGA + JITT VAN IM	8-001
PEN NA NAPONKINT ANYJA RP<M<RE	+ .MOST TENFTO ABL	11-002

10. ábra. Részlet a *Toldi* Előhangjának és Első énekének a tergo konkordanciájából (az *e-re* végződő szövegszók szemébe-től öröme-re-ig). Jobb szélső oszlop mint a 9. ábrán.

éppen úgy létezik és éppen úgy megismerhető, mint a levegő, a proletárdiktatúra maga; nem is kevésbé érdekes és fontos a megismerése. Az meg közművelődésünk feladata, hogy egyrészt megértessük mindenkivel e törvények objektív voltát, másrészt társadalomtörténettől való függését; tudatosítsuk saját kultúránk jeleit s fogékonyra tegyünk a más kultúrákéira.

4. Vessünk még egy pillantást a 9. ábrára. Lehet, hogy egy, ott még aprónak tűnő, vonásban eltér az egyes eddigi, ember által összeállított költői szótáraktól, így alkalmasint a kéziratban elkészült Toldi-szótártól: az önálló jelentéssel nem rendelkező szavakat (az ábrán: *még, mint, mintha* stb.) is közli minden egyes előfordulásukban. Ez nem éppen a szókincs, hanem a nyelvtan kutatójának igen lényeges, sőt elengedhetetlen követelmény.

A 10. ábrán látható a tergo konkordanciák (fordított betűrendes szóalaktárak) egyenesen nyelvtani céllal készülnek. Egy-egy ilyen fordított konkordancián a hasonló jeles-ragos szóalakok kerülnek együvé, így e listák kiváló lehetőségeket nyújtanak valamely szöveg (költői mű stb.) alakitási szerkezetének a vizsgálatára olyan nyelvekben, mint a magyar, az orosz, a latin stb. (Ugyanakkor alig-alig érdekesek egy olyan igen kevésbé ragozó nyelvben, mint az angol – talán ezért maradtak eddig szinte homályban? Az is világos, hogy a nyelvtani viszonyító elemek zömét a szóalakok elején hordozó nyelvekben, amilyenek például a bantu nyelvek, a „közönséges” konkordancia szolgálhat a nyelvtan vizsgálatára, a „fordított”-ból pedig a szókincset lehet vizsgálni.)

A fordított szóalaktárakból kiindítható morfológiai vizsgálatokra két példát hozunk a *Toldi*-ből. (A mű egész nyelvtani feldolgozása még nem készült el; így lehet, hogy az alábbiakon végül némi módosítást fogunk majd végrehajtani, ezek azonban várhatóan nem fogják mondandónk érvényességét befolyásolni.)

(i) Az alanyesetű egyes harmadik birtokos személyragos alakok kérdése (tehát az olyanoké mint *világa, marhája, lelke* stb.; az összesítő mennyiségi adatokat l. a 2. táblázaton). Ennek kapcsán többek között a következő megállapításokat kellett tennünk: a) Szóalakvégen jelentősen visszaszorul az /e/ fonéma és jelentősen előtérbe került az /a/ fonéma – még akkor is, ha az *a* névelőt kihagyjuk a számításból (tehát sem a szövegszók, sem az /a/ végűek közé nem számítjuk be). b) A *Toldi*-ban előforduló mintegy 270 ilyen alak 59%-a mély, 41%-a magas. c) Bár e harmadik személyű birtokos alakok természetesen elég gyakoriak, az összes /a/ és /e/ végű szóalakoknak csupán 8–10%-át teszik ki (ha, amint ez reálisabbnak látszik, az *a* névelőt teljesen kihagyjuk a számításból). d) Míg a mélyek között a három említett típus (*világa, haragja, marhája*), bár más-más arányban, de egyaránt képviselve van, addig a magasak között úgyszólván csak a *lelke*-féle alak található meg; egyáltalán nem volt példa (így, alanyesetben) az *ingje*-változatra (ahol tehát mássalhangzó után *-je* morfban realizálódik e végződés). Ez csupán a véletlen műve: a ragos alakok között nyilván fogunk ilyenre találni, vagy a további ellenőrzések során még az alanyesetűek között is felbukkan egy-két ilyen alak (ez kevésbé valószínű); mindössze arra hívja fel a figyelmünket, amit korábbi kutatásainkban vártunk is: hogy a szövegekben a jelzett aránytalanságot tapasztaljuk, ami az *ingje*-típus teljes hiányához vezethet egyes, nem is túlságosan rövid, szövegekben. Ez elméletileg is érdekes, de esetleg egyes alkalmazásokban (így a magyarnak mint idegen nyelvnek az oktatásában) is megfontolandó következményekre vezethet.

(ii) A tárgyeseti alakok kérdése (tehát olyanoké, mint *utat, virágát, vizet, szavakat, szebbeket, gyermekünket, hangjait, Toldi Miklósét* stb.; az összesítő mennyiségi adatokat

2a. táblázat

A	B	B'	C'	-a	/. . . Cja	/. . . Vja
összes szövegszó	összes /a/ végű a ne-vel	összes /a/ végű a ne nélkül	összes mély 3 PERS			
10 084	1653	984	159	76	60	23
A) 100 = 10 084	16,39		1,58	0,75	0,60	0,23
A') 100 = 9 415		10,45	1,69	0,81	0,64	0,24
B) 100 = 1 653	100,00	59,53	9,62	4,60	3,63	1,39
B') 100 = 984		100,00	16,16	7,72	6,10	2,34
C) 100 = 269			59,11	28,25	22,30	8,55
C') 100 = 159			100,00	47,80	37,74	14,47

Az abszolút értékek A) az összes szövegszó (A' = az a ne nélkül), B) az összes /a/ végű (B' = az a ne nélkül), C) az összes 3 PERS-ok százalékában (C' = mély).

2b. táblázat

A	B	C'	-e	/. . . Cje	/. . . Vje
összes szövegszó	összes /e/ végű	összes magas 3 PERS			
10 084	828	110	106	0	4
A) 100 = 10 084	8,21	1,09	1,05	0,00	0,04
B) 100 = 828	100,00	13,29	12,80	0,00	0,48
C) 100 = 269		40,89	39,41	0,00	1,49
C') 100 = 110		100,00	96,36	0,00	3,64

Az abszolút értékek A) az összes szövegszó, B) az összes (e) végű, C) az összes 3 PERS-ok százalékában (C' = magas).

3. táblázat

A	B	C	SG	PL	SG/SG	SG/PL	PL/SG	{é}
összes szövegszó 10 084	összes /t/ végű 1 380	összes Acc. 701						
A) 100 = 10 084	13,69	6,95	4,51	0,25	2,03	0,03	0,12	0,01
B) 100 = 1 380	100,00	50,80	32,97	1,81	14,86	0,22	0,87	0,07
C) 100 = 701		100,00	64,91	3,57	29,24	0,43	1,71	0,14

Az abszolút értékek A) az összes szövegszó, B) az összes /t/ végű, C) az összes Acc. százalékaiban. SG/SG = tó egyes szám, birtokos egyes szám, SG/PL = tó egyes, birtokos többes; PL/SG = tó többes, birtokos egyes; {é} = a tövet {é} birtokos képző követi.

4. táblázat

PL:	szavakat (VII/9), házfalakat (IX/1), sátrakat (XI/9), krónikákat (XII/19), kutyákat (IX/6), pányvakat (IX/10), lándzsákat (VII/5), dolgokat (VIII/13), nagyokat (IX/5, 16, X/22), farkasokat (VI/5), mézárásokat (IX/10), kortyokat (X/18);	14
szebbeket (XI/8), kincseket (XII/20), köteleket (IX/10), melyeket (XI/4), szemeket (XII/12), öket (V/16, VI/20, IX/2, IX/5 bis), kereszteteket (IV/17).		11
SG/PL:	házakat (IV/6), gyermekünket (IV/18), szívüket (XII/17).	3
PL/SG:	hangjait (0/1), csontjait (X/4), tollait (II/2), rágalmait (VIII/11), farkasait (VI/2), ruháit (VI/5), kutyáit (VII/2);	7
	szemeit (II/16), körmeit (V/7), parittyaköveit (0/2);	3
	legényit (III/2);	1
	csapásit (V/5).	1
{ é }	Toldi Miklósét (VII/1)	1

1. a 3., a ritkább tárgyeseti alakok listáját a 4. táblázaton). Ennek kapcsán többek között a következő megállapításokat kell tennünk: a) /t/ végű szóalak jóval több van, mint akár /a/, akár – kivált – /e/ végű. b) Ezek a szóalakok viszont jól tömörülnek az Acc. körül: a /t/ végűek valamivel több, mint a fele Acc. (Megjegyzendő: a /t/ végű alakokat a hosszú /t:/ végűektől teljesen függetlenül szemléljük, mint ahogy minden¹⁷ más összefüggésben is külön kezeljük a rövid és a hosszú mássalhangzó-fonémákat.)

Az (i) és a (ii) példát összevetve meg kell állapítanunk a következőket. Amint az talán várható is volt, korpuszunkban jóval több (kerekén két és félszer annyi) tárgyeseti alak van, mint egyes harmadik birtokos személyragos; még egyes tárgyeseti alak is másfélszer annyi, mint egyes birtokos harmadik személyű. Ugyanakkor figyelmet érdemel az a körülmény, hogy magán a tárgyeseten belül jóval népesebb az *anyját*, *szolgáját* stb. féle alakok csoportja (ahol tehát a tárgyeseti ragot birtokos személyrag, legtöbbször éppen az egyes harmadik birtokosé, előzi meg), mint a *szavakat*, *kincseket*-féle alakok, többes jellel a tárgyeseti végződés előtt. Várható, hogy az összkép is ehhez hasonlóan fog alakulni: a birtokos személyragos alakok gyakoribbak lesznek a többes számúaknál.¹⁷ A bonyolultabb agglutinatív alakokat, mint *parittyaköveit*, *hangjait* stb., nyilván ritkábban alkalmazzuk, a mindössze kételemes jel-rag tömbök esetében viszont a birtokos alakok lépnek előtérbe – szép összhangban van ez nyelvünk többes számot kerülő szerkezetével.

Amikor orosz szövegeket vetettünk hasonló vizsgálat alá, kivált az egyes funkcionális stílusrétegek okozta tömeges nyelvtani eltérések ötlöttek szembe. Így láthatóvá vált, hogy az anyegini szövegmintában előtérbe kerül az /a/ szóalakvég – mint kiderült, főleg a nőnemű alakok gyakorisága miatt. Ugyanakkor a kontrollként feldolgozott politikai jellegű szövegben kiugrott az /i/ szóalakvég: főleg a számos többes számú igei, névmási stb. végződés miatt. (Természetesen mind az /a/, mind az /i/ alakvégi fonéma számos más nyelvtani jelenséget is rejtett magában.)

Úgy gondoljuk, érdemes olyan zárt korpuszoknak, mint a *Toldi* vagy az *Anyegin* a „véges grammatikáját” elkészíteni. Valójában igen hagyományos feladat ez: egy-egy kódex nyelvtanának az összeállítása hasonlítható hozzá. A tömegesen és könnyen elkészíthető a tergo konkordanciák mindössze lehetőséget nyújtanak arra, hogy a korábbiaknál lényegesen több ilyen véges grammatika készüljön, főleg csak szellemi ráfordítással; hogy a kutató számára egy a korábbiaknál bővebb és rugalmasabb alaktani tár álljon rendelkezésre akár egy-egy szinkrón metszet tekintetében, akár diakrón vonatkozásában.

5. Itt tartunk most. Így jutottunk el hangok-fonémák rendezése, számlálása után az egyes lexémák, majd az egyes nyelvtani alakok csoportos vizsgálatához. Mind e három szinten még igen sok kutatási lehetőség van külön-külön is. Ugyanakkor világos a továbbhaladás egy másik útja: az egyes szinteken, kivált a legutolsón, a grammatikáén, még jobban elszakadni a számoktól és még inkább a szövegeinkben megbúvó, azokat létrehozó törvényszerűségeket kutatni.

¹⁷Vö. NEMES Zoltán hasonló eredményeivel: *Magyar Nyelv* XXVIII. (1932) 30–36, 106–111.

Papp Ferenc tudományos munkássága
(1956–1975)

Önállóan megjelent munkák

- Matematikai nyelvészet és gépi fordítás a Szovjetunióban. (OMK–DOK: Időszerű műszaki dokumentációs kérdések 6.) Budapest, 1964. – 122 pp. – (Ugyanaz angolul: *Mathematical Linguistics in the Soviet Union*. Mouton and Co. The Hague, 1966. – 165 pp.) (Ism.: Petőfi: *Magyar Nyelvőr* 89.2. 257–259., J.-M. W.: *Revue de Phonétique Appliquée*. 1967. 6. 73–75., Bierwisch: *Germanistik* 8. 4. 717.)
- Русь современного русского языка. Budapest, 1968. Tankönyvkiadó. – 669 pp. (Szerkesztő és társszerző BOLLA Kálmánnal és PÁLL Ernával.) (Ism.: Leontyjev: *Русский язык за рубежом*. 1969. 3. 117–119. Krysín: *Studia Slavica* XVI. 141–148., Vascenko: *Romanoslavica* XVIII. 549–553., Chvany: *Soviet and East-European Journal*. XV. 358–363.)
- A magyar nyelv szövegmutató szótára. (Szerk.) Budapest, 1969. Akadémiai Kiadó. – 594 pp. [Ism.: Décsy: *Ungarn-Jahrbuch*. 2. 173., Seilenthal: *Советское финно-угроведение*. VII. 4. 315–317, Rácz: *Magyar Tudomány* XVI. 1. 56–58., Tuomi: *Virittäjä* 69: 401–405., Csapó: *Hungarian Studies in English* (1971) 153–154., Kiss: *Acta Linguistica* XXII. 225–228., Mel'čuk: *Научно-техническая информация* 1972. 8. 24–27.]
- Ady Endre összes költői műveinek fonémastatisztikája. Budapest, 1974. Akadémiai Kiadó. (JÉKEL Pállal közösen.) – 63 + 179 pp.
- A magyar főnév paradigmatisztikus rendszere. (Leírás és automatikus szintézis.) Budapest, 1975. Akadémiai Kiadó. – 329 pp.

Tanulmányok, ismertetések

1956.

- A szófajok meghatározásának kérdéséhez. (Az orosz nyelv anyaga alapján.) – *Acta Universitatis Debreceniensis* III/1. 75–82.

1957.

- Mit tudott Lomonoszov a magyar nyelvről és a finnugor nyelvrokonságról? – *Magyar Nyelvőr* 81/4. 463–468.

1958.

- Az orosz helyesírás fejlődésének fél évszázada és a most megjelent orosz helyesírási szabályzat. – *Magyar Nyelvőr* 82/2. 149–161.
- Kalmár György oroszországi kapcsolatairól. – *Filológiai Közöny* IV/2. 346–349.
- Az új orosz helyesírási szabályzatról. – *Idegen Nyelvek Tanítása* I/3. 92–94.

1959.

- M. В. Ломоносов и финно-угроведение. – *Научные доклады высшей школы. Филологические науки*. I. 56–67.
- Új irányzatok a szovjet nyelvtudományban. – *Nyelvtudományi Közlemények* LXI/2. 392–410.
- Új filológiai folyóirat a Szovjetunióban. – *Filológiai Közöny* V/1–2. 225–229. (Ism.)

1960.

- Вопросы культуры речи. 1–2. – *Magyar Nyelvőr* 84/3. 394–399. (Ism.)

1961.

- A matematikai módszerek alkalmazásáról nyelvtudományunkban. – *MTA I. Osztályának Közleményei* XVII/1–4. 289–306.
- A stíluselemzés egy mennyiségi mutatókra épített módszere. – *Filológiai Közölny* VII/1–2. 69–85.
- Функции и значения местоимения *свой*. (DRAHOS Józseffel közösen.) – *Acta Universitatis Debreceniensis* VII/1. 143–148.
- Количественный анализ словарной структуры некоторых русских текстов. – *Вопросы языкознания* X/6. 93–100.
- Matematikai módszerek a nyelvtudományban. – *Valóság* IV/1. 76–82.
- Трансформационный анализ русских присубстантивных конструкций с зависимой частью – существительным. – *Slavica* I. 53–83.
- Трансформационный анализ русских присубстантивных конструкций с зависимой частью – существительным в родительном падеже. – *Studia Slavica* VII/1–3. 195–206.
- Vita a nyelvtudomány elvi kérdéseiről. (Felszólalás.) – *MTA I. Osztályának Közleményei* XVIII/1–4. 98–102.
- V. K. Чичагов: Из истории русских имен, отчеств и фамилий. – *Magyar Nyelvjárások* VII. 160–162. (Ism.)
- Вопросы культуры речи. Вып. 2. – *Русский язык в школе*. XXII/1. 104–106. (Ism.)
- Újabb orosz nyelvű művelő kiadványok a Szovjetunióban. – *Magyar Nyelvőr* 85/4. 415–418. (Ism.)

1962.

- Морфологическая система глагольных основ в современном русском литературном языке. – *Slavica* II. 127–150.
- Материалы и исследования по русской диалектологии. Новая серия т. I. *Magyar Nyelvjárások* VIII. 158–160. (Ism.)
- A IX. Nemzetközi Nyelvész-Kongresszus előtt (Trends in European and American Linguistics). – *Nyelvtudományi Közlemények* LXIV/1. 126–130. (Ism.)

1963.

- Nyelvészet és nyelvtanítás. – *Modern Nyelvoktatás* 1/3–4. 76–85.
- К построению одной структурной модели словосочетаний современного русского языка. – *Studia Slavica* IX/1–4. 229–237.
- Классификация русских глаголов. – *Русский язык в школе*. 4. 98–101.
- Некоторые вопросы обучения пониманию русской речи на слух. – *Русский язык в национальной школе*. 2. 66–70.
- Классификация русских глаголов. Русский язык в школе. 24/4. 98–101.
- Некоторые вопросы изучения устной и письменной разновидностей языка. – *Slavica* III. 21–30.
- Каковы основные задачи статистического анализа славянских языков? In: *Славянска филология* I. София, 1963. 198–199.
- Каково соотношение статистических методов с другими методами изучения славянских языков? In: *Славянска филология* I. София, 1963. 210–211.
- Как обнаруживается динамика языка при синхронном анализе? In: *Славянска филология* I. София, 1963. 222–223.
- Каковы различия между отдельными лингвистическими традициями в понимании грамматических категорий и каковы возможности унификации этих различных точек зрения? In: *Славянска филология* I. София, 1963. 245–246.
- A. V. Isačenko: Die russische Sprache der Gegenwart. Teil I. Formenlehre. – *Slavica* III. 168–170. (Ism.)

- A. Руд–Е. Чапник: Обучение структурным моделям в английской речи. – *Modern Nyelvoktatás* I/3–4. 185–186. (Ism.)
- Проблемы структурной лингвистики. – *Acta Linguistica* XIII/3–4. 393–399. (Ism.)
- Matematikai és alkalmazott nyelvészet a Szovjetunióban. – *Nyelvtudományi Közlemények* LXV/2. 456–463. (Szemle, KLAUSZER Judittal közösen.)

1964.

- A magyar szókinsz lyukkártya-rendszerű feldolgozása. – *Magyar Nyelvőr* 88/4. 457–464.
- A magyar szószervezet-rendszer néhány sajátosságáról. – *Nyelvtudományi Közlemények* LXVI/1. 129–139.
- Этикет и язык. – *Русский язык в национальной школе* 1. 74–77.
- Mathematische und strukturelle Methoden in der sowjetischen Sprachwissenschaft. – *Acta Linguistica* XIV/1–2. 119–137.
- Nyelvi rendszer, közlési folyamat és ezek néhány matematikai modellje. – *Általános Nyelvészeti Tanulmányok* III. 75–88.
- Felszólalás a matematikai nyelvészeti és gépi fordítói konferencián. – *Általános Nyelvészeti Tanulmányok* II. 292–294.
- A rendezés művészetétől a művészetek tudományáig. – *Valóság* VII/3. 88–90. (Ism.)
- Mathematical and Applied Linguistics at the University of Debrecen. – *Computational Linguistics* II. 213–218.
- О готовящемся обратном словаре венгерского языка. – *Computational Linguistics* III. 205–211.
- Механизация лексикографических работ и обратные словари. – *Computational Linguistics* III. 212–215. (Bibliográfia, DEZSŐ Lászlóval közösen.)
- P. M. Фрумкина: Материалы к частотному словарю языка Пушкина. – *Nyelvtudományi Közlemények* LXVI/2. 462–464. (Ism.)
- И. И. Ревзин: Модели языка. – *Nyelvtudományi Közlemények* LXVI/1. 234–238. (Ism.)

1965.

- A kibernetikai módszerek jelentősége a nyelvtudomány számára. – *Magyar Tudomány* X/11. 704–714.
- A katedra túlsó oldaláról. – *Modern Nyelvoktatás* III/1. 175–179.
- A Magyar Szóvéghmutató Szótár és a vele kapcsolatos munkálatok. – *Magyar Nyelv* LXI/2. 187–200.
- Algoritmus. – *Magyar Nyelvőr* 89/1. 87–93.
- Esztétika és kibernetika. – *Valóság* VIII/7. 29–40.
- Gubera, guberál. – *Magyar Nyelvőr* 89/2. 242–243.
- Kód. – *Magyar Nyelvőr* 89/2. 201–208.
- Modell. – *Magyar Nyelvőr* 89/4. 462–468.
- Szemiótikai jegyzetek. – *Általános Nyelvészeti Tanulmányok* III. 157–176.
- Biró Yvette: A film formanyelve. – *Kritika* III/5. 57–59. (Ism.)
- Первый программированный учебник венгерского языка. – *Computational Linguistics* IV. 177–181. (Ism.)
- Roman Jakobson: Selected Writings. I. – *Computational Linguistics* IV. 167–176. (Ism.)
- И. И. Ревзин: Модели языка. – *Acta Linguistica* XV/1–2. 181–187. (Ism.)
- Változások lesznek az orosz helyesírásban? – *Magyar Nyelvőr* 89/2. 275–277. (Szemle.)
- Pesti levél a pécsi balettről. – *AIföld* XVI/11. 64–66.

И. И. Ревзин–В. Ю. Розенцвейг: Основы общего и машинного перевода. — *Nyelvtudományi Közlemények* LXVII/2. 425–428. (Ism.)

1966.

A magyar főnévragozás három modellje. — *Magyar Nyelv* LXII/2. 194–206.

A magyar szókincs törzsanyagának lyukkártya-rendszerű feldolgozása. — *Információ-elektronika*. I/1. 52–56.

A társadalomtudományok matematizálódásáról. — *Alföld* XVII/2. 41–45.

Kettős és hármas szófajságú szavaink. — *Magyar Nyelvőr* 90/4. 404–408.

Megszóltak a gépek. — *Magyar Nyelvőr* 90/2. 181–188.

On the Depth of Hungarian Sentences. — *Linguistics* 25. 58–77.

Strukturális-matematikai-algebrikus (nyelvészet) — *Magyar Nyelvőr* 90/1. 55–62.

Le vocabulaire du hongrois contemporain sur cartes perforées. — *Cahiers de lexicologie* VII/2. 103–117.

Általános és alkalmazott nyelvészeti szeminárium Besançonban. — *Nyelvtudományi Közlemények* LXVIII/1. 166–168.

Машинный перевод и прикладная лингвистика, вып. 8. — *Általános Nyelvészeti Tanulmányok* IV, 259–265. (Ism.)

Математическая и прикладная лингвистика в Венгрии. — *Машинный перевод и прикладная лингвистика* 9. 45–52. (Szemle, SZÉPE Györggyel közösen.)

Обзор таблиц венгерского словаря, полученных на перфокартных машинах. — *Computational Linguistics* V. 158–167.

1967.

A magyar szókincs gépi feldolgozásának egyes eredményei és további problémái. — *Nyelvtudományi Értekezések* 58. 518–522.

A Magyar Szóvégmutato Szótár és az alaktani ingadozások vizsgálata. (JÁNOSKA Sándorral közösen.) — *Magyar Nyelv* LXIII/2. 138–148.

A strukturális leírás és a nyelvtörténet kapcsolata. — *Általános Nyelvészeti Tanulmányok* V. 237–246.

Die Bearbeitung des ungarischen Wortschatzes auf Lochkartenmaschinen. — *Acta Linguistica* XVII/1–2. 141–172.

Néhány fontosabb adat szókincsünk szófaj és ezen belül hosszúság és jelentés szám szerinti megoszlásáról. — In: *Nyelvfeldolgozás és Dokumentáció*. (Szerk. SZÉPE György.) 104–116.

Tőigéink. — *Magyar Nyelvőr* 91/1. 45–52.

A *buli* és társai az Értelmező Szótárban. — *Magyar Nyelvőr* 91/3. 346–353.

О некоторых количественных характеристиках словарного состава языка. — *Slavica* VII. 51–58.

И. А. Мельчук–Р. Д. Равич: Автоматический перевод 1949–1963. — *Nyelvtudományi Közlemények* LXIX/2. 451–454. (Ism.)

Kiefer Ferenc: Matematikai nyelvészeti tanulmányok — *Magyar Nyelvőr* 91/2. 250–252. (Ism.)

Машинная обработка словарей. — Конференция по машинному переводу. *Тезисы докладов*. Ереван, 1967. — 115–117.

1968.

Сzófaj és végződés. — *Magyar Nyelvőr* 92/1. 29–37.

Amíg a gépi listából olvasható tanulmány lesz. — *Magyar Nyelvőr* 92/3. 321–333.

Hangutánzó és hangfestő szavaink az Értelmező Szótárban. — *Magyar Nyelv* LXIV/1. 32–38.

A szinkronizálás néhány nyelvészeti kérdése. In: *Szinkron és művészet*. Szerk. VERESS József. Debrecen, 1968. 41–58.

Adatok tőszókincsünk etimológiai rétegeinek végződés szerinti eloszlásához. — *Nyelvtudományi Közlemények* LXX/1. 200–209.

- О плане машинной обработки этимологического словаря венгерского языка. – *Советское финно-угроведение* IV/3. 207–214.
 Русская орфография в венгерской школе. – *Slavica* VIII. 195–204.
 Прикладное языкознание и преподавание русского языка. In: *За современное обучение русскому языку*. Budapest, 1968. 100–109.

1969.

- Tő, rag és ami közöttük van. – *Magyar Nyelv* LXV/1. 33–39.
 Filmszinkronizálás és szemiotika. – *Rádió és televízió szemle* 1/4. 83–97.
 Gépi eredményeink felhasználása szótárírásunkban. – *Magyar Nyelvőr* 93/3. 376–381.
 Verzeichniss der ungarischen Suffixe und Suffixkombinationen. Zusammengestellt ... von Wolfgang Veenker. – *Nyelvtudományi Közlemények* LXXI/1. 190–194. (Ism.)
 Az Értelmező Szótár legtöbb jelentésű szavairól. – *Magyar Nyelvőr* 93/1. 94–97.
 „A magyar nyelv történeti-etimológiai Szótára”-nak gépi feldolgozásáról. – *Magyar Nyelv* LXV/2. 145–154.
 Hogyan tanítunk meg az orosz igék parancsoló módjának képzésére egy embert és hogyan – egy elektronikus számológépet? – *Magyar nyelvőr* 93/4. 456–466.
 О машинной обработке одноязычных словарей. – *Научно-техническая информация* 3. 20–29.
 Совместные исследования в области автоматической обработки русских и венгерских текстов. – *Slavica* IX. 65–69.

1970.

- Программа образования повелительного наклонения стандартных русских глаголов на ЭЦВМ Odra 1013. – *Slavica* X. 37–56.
 Деякі кількісні характеристики одного угорського словника. – *Питання структурної лексикології* Київ, 1970. 167–178.
 Inversiu li leksikoni. – *Nyelvtudományi Közlemények* LXXVII/1. 266–269. (Ism., ISTVÁNOVITS Mártonnal közösen.)
 Idegen nyelvű szövegek kézirat-előkészítése, szedése és korrektúrája. – *Magyar Nyelvőr* 94/2. 255–256. (Ism.)
 A számítógépek mint az oktatás modellálásának eszközei. – *Felsőoktatási Szemle* XIX/7–8. 457–461.
 Sur quelques lois statistiques du vocabulaire. – *Actes du X^e Congrès International des Linguistes*, Vol. IV. – Bucarest, 1970. 471–475.

1971.

- A magyar nyelvre irányuló számítógépes vizsgálatok a Kossuth Lajos Tudományegyetemen. – *Magyar Nyelvjárások* XVII. 147–158.
 A szókincs bővítés az orosz nyelvtanárok oktatásában. (HADAS Ferencsel közösen.) – *Felsőoktatási Szemle* XX/4. 217–21.
 Дубляж sub specie *σπειρωτικῆς*. In: *Фонетика–фонология–грамматика*. Москва, 1971. 368–377.
 Szókincsünk *-at/-et* végű névszóirol. (JAKAB Lászlóval közösen.) – *Magyar Nyelv* LXVII/1. 46–57.
 A generatív nyelvészet újabb eredményei és a nyelvoktatás. – *Modern Nyelvoktatás* IX/1–2. 12–22.
 Конференция о славянском слове. – *Slavica* XI. 171–173.
 Russian Derivational Dictionary. – *Slavica* XI. 181–182. (Ism.)

1972.

- Два вопроса автоматического синтеза русских парадигматических форм. In: *The Slavic Word* (ed. by D. S. WORTH). The Hague–Paris, 1972. 283–301.
- Окончательная редакция текстовых единиц, длинее предложений или квази-правильность речи лиц, освоивших иностранный язык. – *Slavica* XII. 27–41.
- Два аспекта анализа художественной структуры повести Ч. Айтматова «После сказки (Белый пароход)». – *Slavica* XII. 141–147.
- A lexémák alaktani szerkezete, szófaja és jelentésgazdagsága. – *Magyar Nyelv* LXVIII/1. 23–33.

1973.

- Hang, fonéma, betű és a teljes magyar ábécé. – *Magyar Nyelvőr* 97/3. 324–331.
- Hagyomány és újítás a nyelvtudományban. – *Magyar Tudomány* XVIII/12. 829–832.
- Néhány gondolat a felsőfokú nyelvoktatás elméletéről. (Hozzászólás.) – *Modern Nyelvoktatás* X/1. 3–40.
- Gépi fordítás és számítógépes nyelvészet. *Tudományos és műszaki tájékoztatás* 20/11–12. 840–844.
- О некоторых общих чертах славянских заимствований в венгерском языке. – *Studia Slavica* XIX/1–3. 225–234.
- Вычислительная лингвистика и образование преподавателей-словесников. – *Computational Linguistics and Computer Languages* IX. 185–192.
- Tőszókincsünk etimológiai rétegenkénti hangstatisztikája. – *Nyelvtudományi Közlemények* LXXV/1. 3–40.

1974.

- A lexémák szófaja, alaktani szerkezete és eredete. – *Magyar Nyelv* LXX/1. 55–68.
- Szövegszó, szóalak, lexéma. – *Magyar Nyelvőr* 98/1. 76–82.
- Stílus és költői nyelv számítógépes vizsgálata. – *Nyelvtudományi Értekezések* 83. 428–434.
- Számológépes vizsgálatok fonetikai-fonológiai síkon. – *Általános Nyelvészeti Tudományok* X. 131–139.

1975.

- Some Questions on the Process of Teaching and Time Factors. In: *Modern Linguistics and Language Teaching*. Ed. by Zs. TELEGDI et. al. Budapest, 1975. 145–150.
- Konkordancia: írói szótár előkészítése számítógépen. – *Magyar Nyelvőr* 99/3. 351–355.
- Automatikus szövegelemzés a legalsó nyelvi szinteken. (GÖRGEY Eszterrel és JÉKEL Pállal közösen.) – *Általános Nyelvészeti Tanulmányok* XI. 115–125.
- Hozzászólás Kiefer Ferenc előadásához. – *Nyelvtudományi Közlemények* LXXVII/2. 386–391.
- Ami, amcsi. – *Magyar Nyelv* LXXI/1. 71.
- Zsilka Tibor: Stilisztika és statisztika. (JÉKEL Pállal közösen.) – *Magyar Nyelv* LXXI/1. 104–107. (Ism.)
- Paul L. Garvin: On Machine Translation. – *Nyelvtudományi Közlemények* LXXVII/1. 297–300. (Ism.)
- Gyakorisági eloszlás. In: *Világirodalmi Lexikon* 4., 95. Budapest, 1975.