

## Adatok a magyar szókincs szerkezetének változásához

1. Nemrégiben úgy adódott, hogy két különböző jellegű munkához is szükségem lett volna olyan adatokra, amelyek a magyar szókincs szerkezetének alakulását időegymásutánban mutatják, egyes kategóriáinak helyzetét, illetőleg változását a diakronia szempontjából konkrétan megfogható módon szemléltetik. Sajnos e szükséglet kielégítésére nyelvészeti irodalmunkból még csak valamelyest tájékoztató adatanyagra sem támaszkodhattam: bizonyos vonatkozásokban semmit sem találhatni, de még a jobbára szókincsünk etimológiai összetételére szorítókozó számítások is rendkívül szórványosak és szegényesek, azonkívül néhány szókincs-csoportra nézve hibás kategorizálási elvekre épülnek föl. Így nem maradt más lehetőségem, mint hogy magam végezzek bizonyos számításokat. — Ismeretes azonban, hogy az ilyen jellegű vizsgálatok mennyire munkaigényesek és hosszadalmasak, hiszen akárcsak egy-egy szónak valamilyen kategóriába való besorolása is gyakran csak tüzetesebb utánanézés alapján végezhető el megnyugtató módon. Így természetesen nem gondolhattam arra, hogy az e téren meglévő tetemes hiányokat nagy nyelvi anyag bevonásával, valamilyen szélesebb körű vizsgálattal pótoljam; csupán arra törekedtem, hogy saját pillanatnyi igényeimet némi hozzávetőleges támpontok nyerésével kielégíthessem.

Egyrészt a számomra adott időbeli lehetőségeknek, másrészt a korlátozott célnak megfelelően négy kisebb terjedelmű szöveget dolgoztam föl a szókincs szerkezeti vizsgálatának szempontjából. A szövegek kiválasztásában két dologra ügyeltem. Az egyik az volt, hogy a kiszemelt szövegek körülbelül azonos stílusrétegből valók legyenek; e tekintetben a prózai szövegek, azokon belül pedig a beszélt nyelvhez viszonylag közel álló stílus kategóriák látszottak legmegfelelőbbeknek. A másik az volt, hogy a kiválasztott szövegek az írásos magyar nyelvtörténetet a nyelvfejlődés gyorsuló ütemének megfelelő szakaszokban szemléltessék, és az egyes szakaszokat a korstílus szempontjából tipikusnak mondható írások képviseljék. E két szemponton kívül a kiválasztásban semmi más előzetes mérlegelés nem vezetett. — A négy megvizsgált szöveg a következő volt: 1. A Halotti Beszéd (a továbbiakban HB.); 2. Bornemisza Péter: Órdögy kisírtetekről...: Az ki nem szentelte az ünnepet. 1578.: ECKHARDT kiad. 172. l. (a továbbiakban Born.); 3. Szemere Pál: Levél Petőfi Sándorhoz. 1846.: Magyar Irodalmi Szöveggyűjtemény II/1, 790—1 (a továbbiakban Szem.); 4. A tizedik pedagógusnapra (köszöntő): Felsőoktatási Szemle 1961: 325 (a továbbiakban Ped.).

E szövegeket öt fő szempont szerint dolgoztam föl: 1. a ténylegesen előforduló és az etimológiailag kihámozható szóelemek egymáshoz való aránya; 2. a szóelemek hangalaki hosszúsága a szótagszám szerint; 3. a szóelemek szófaji kategóriák szerinti megoszlása; 4. a szóelemek eredet szerinti csoportosulása; 5. a szóelemek alaktani szerkezete. Az 5. szempont

alkalmazása azonban — érthető okokból — csak a 4. vizsgálati mód szerint nyert elemek egy csoportjával, a magyar nyelven belül keletkezett szavakkal kapcsolatos, tehát — az előbbiekkal ellentétben — már nem vonatkozik a szövegek szókincsének egészére.

Mint már jeleztem, e szókincsstatisztikai feldolgozás sokkalta kisebb terjedelmű szövegekre épült, hogysem akárcsak a nagyobb, átfogóbb szókatégoriákra vonatkozóan is teljesen biztos, általánosítható adatokkal szolgálhasson, nem is beszélve arról, hogy még az egyazon korból való nagy szöveganyag is eléggé különböző rétegződést mutathat a szöveg tárgyától, stíluskatégoriájától, az író személyétől stb. függően. De ha ez a korlátozott igényekkel, „házi használatra” készült statisztika nem is adhat pontos eligazítást a magyar szókincs szerkezetében végbement történeti változásokról, úgy vélem, több szempontból is helyes lesz, ha a vizsgálat eredményét néhány megjegyzés, kommentár kíséretében e cikk keretében külön is közzéteszem.

Először is az ilyen jellegű számítások, statisztikai adatok a nyelvnek mind diakrón, mind szinkrón vetületét tekintve rendkívül fontos problémákra mutatnak rá, a nyelv szerkezete fejlődésének és állapotának átfogó, fő kérdéseit állítva előtérbe. Ha ez a nagyon szerény igényű és eredményű cikk némi ösztönzést nyújthat a jóval nagyobb anyagon és még sok másfajta szempont szerint is elvégezhető ilyesfajta vizsgálatokra, már elérte célját. — Ezenfelül e kísérlet közzététele alkalmas lehet arra is, hogy bizonyos kiindulási alapot adjon az itt alkalmazható elvek és módszerek továbbfejlesztéséhez, illetőleg vitaanyagul szolgáljon az efféle stúdiumok útkereséséhez. — Végezetül ha e vizsgálat nem is felel meg a statisztika elsődrendű követelményének, a nagy számokból való adat- és arányelvonás kívánalmának, a nyelvtörténetünk négy különböző időszakából feldolgozott szövegek nyelvi anyaga már föltétlenül alkalmas arra, hogy egyes szókatégoriákra nézve a realitás határain belül nyújtsjon tájékoztatót. A sok tekintetben igen érdekes eredményeket óvatos értékelés esetén tehát valószínűleg mások is hasznosíthatják majd a magyar szókincs történeti fejlődése különböző vonatkozásainak megítélésében.

**2.** Mielőtt az említett négyféle vizsgálat eredményeit és legfontosabb problematikáját bemutatnám, eljárás módomból néhány általánosabb szempontjára kell a figyelmet felhívnom.

Az elvégzett vizsgálatokban — a felmerült szókincsstatisztikai szükségletnek megfelelően — csak a „hagyományos” értelemben fölfogott szóra voltam tekintettel. Nem érintettem tehát a szintaktikai-mondattani kérdéseket (mondatrészek, szerkezetek, mondatfajták stb. statisztikája), és nem vizsgáltam az önálló szótest nélküli viszonyítóelemeket: jeleket, ragokat (noha ezek statisztikai feldolgozásának természetesen szintén igen nagy jelentősége lenne).

Ennek a szóra korlátozódó vizsgálatnak a keretein belül számításba vettem minden önálló, külön szónak tekinthető nyelvi elemet. Ennek megfelelően egy-egy alapszónak a szóalkotás különböző módjaival (képzés, összetétel, elvonás stb.) keletkezett továbbalakulásai külön-külön egységekként szerepeltek statisztikámban; tehát például azonos szóelemeknek tekintettem a *néz* ige különböző ragos és jeles alakjait (*néznek*, *néztünk*, *nézne*, *nézzen* stb.), ellenben külön-külön szóelemként vettem tekintetbe a *néz* ige nék efféle továbbalakulásait: *nézés*, *nézdegel*, *nézni*, *néző*, *meg-*

*néz, végignéz* stb. Magától értődik, hogy a nyelvtörténetileg ragos alakok közül is új, önálló szóelemnek, azaz külön egységnek vettem mindenfajta, az adott korban már megszilárdult, megkövült ragos alakulatot (pl. a *hely* mellett a *helyett*-et, a *hát* mellett a *hátul*-t stb.), valamint a határozóragokkal kapcsolatos határozószói típust (*nekem, téled* stb.) is. Ugyancsak önálló szóelemként szerepelt statisztikámban mindenfajta „segédszó” (kötőszó, névelő, tagadószó stb.).

Szótani vizsgálatom egyik vezérelve volt, hogy ha a feldolgozott szövegben egyazon szó többször is előfordult, csak egyszer számítottam be statisztikámba; vagyis például a HB. többször ismétlődő *gyümölcs, lát, hogy, és* stb. szóelemei csupán egyetlen adatként dolgoztattak föl (de persze az előbb mondottaknak megfelelően a HB. *hal* és *holt* vagy *nap* és *bírságnap* szava már két-két adatként szerepelt). Ily módon reálisabb képet nyerhettem magának a szóállománynak, illetőleg egyedeinek arányairól, mivel elkerülhetővé vált, hogy a tárgyi, stilisztikai stb. okokból való szóismétlések vagy egyes minduntalan előforduló segédszók (névelők, kötőszók stb.) alkalmi jellegűvé tegyék a szóelemek belső arányait, bizonyos szempontból eltorzítsák a lexikológiai képet. Ezzel az eljárásommal természetesen korántsem akartam tagadni azt, hogy más — főként stilisztikai — szempontú vizsgálatokban a szóismétlések figyelembevétele is jogosult vagy éppen kívánatos lehet.

A viszonylag kis terjedelmű szövegekre való tekintettel csak a g y o b b , á t f o g ó b b s z ó k a t e g ó r i á k megállapítására és statisztikázására törekedhettem. Ezért főként a 3., 4. és 5. szempont alapján fölvehető, de egymással rokonságban levő összevonható, apróbb részkategóriákat együvé csoportosítottam. Nagy szöveganyag feldolgozása esetén persze az aprólékosabb fölbontás is jobbra reális arányokat tükröztethet, ezért a kisebb kategóriák megkülönböztetése ott nyilvánvalóan kívánatos.

Az egyes szövegekben itt-ott, kis számban előforduló tulajdonnevek nem adhattak volna komolyabb alapot a szóincs-kategóriával kapcsolatos számításokra, ezért munkámban csak a k ö z s z a v a k r a voltam tekintettel.

Az összesítő táblázatokban egyrészt a megvizsgált adatok s z á m s z e r ű m e n n y i s é g é t, másrészt ezek s z á z a l é k a r á n y á t tüntettem föl. Az előbbiekre a számítások esetleges revidálásának, ellenőrzésének könnyebb lehetővé tétele érdekében volt szükség, az utóbbiakra pedig azért, mivel az átfogóan értékelhető, reálisan tájékoztató arányokat elsősorban a százalékszámokból lehet kiolvasni.

3. Az egyes szövegek szóállományára nagyon jellemző az, hogy 1.: mennyi bennük a ténylegesen előforduló, etimológiai fölbontás nélküli, teljes alakjában megjelenő szóelem, és 2.: mekkora bennük azoknak a szóelemeknek a száma, amelyeket az előbbi egységeken belül az etimológiai vizsgálat kihámozhat. Hogy e két kategória különbségét világosabban lássuk, vegyünk néhány példát. A HB. *por, paradicsom, munkás, bírságnap* szavai valóságosan előforduló, etimológiailag nem tagolt állapotukban négy egységnek, négy szóelemnek számíthatók; ez az elsőként emlegetett kategória. Ugyanezen szavakat azonban másként is meg lehet, sőt meg is kell ítélni, mégpedig úgy, hogy a *por* és a *paradicsom* ugyan etimológiailag már tovább nem bontható egyedek, ellenben a két utóbbival már más a helyzet: a szavak történetét, továbbalakulását, elemekre tagolhatóságát tekintve a *munkás* két szóegységet

tartalmaz (*munkás* és *munka*), a *bírságnap* meg éppen négyet (*bírságnap*, *bírság*, *bír* és *nap*); ez a másodikul emlegetett kategória. Mivel az itt szereplő fogalmak megkülönböztetésére később is szükségünk lesz, nevezzük el a *por*, *paradicsom*, *munkás*, *bírságnap*-féle számbavétel útján nyert elemeket tényleges szóelemeknek, a *por*, *paradicsom*, *munkás*, *munka*, *bírságnap*, *bírság*, *bír*, *nap*-féle számbavétel útján nyert elemeket pedig felbontott szóelemeknek.

A mondottak alapján a felbontott szóelemek adatai magukban foglalják a tényleges szóelemeknek plusz az alaktani felbontás útján még nyert elemeknek (különböző rendű alapszók, elő- és utótagok stb.) adatait. — Egyébként az egyes szövegeken belül ismétlődő szavak előbb már ismertetett egyszeri beszámításának elve alapján az alaktani felbontás útján nyert oly szóelemek, amelyek tényleges formájukban is előfordulnak a szövegben, természetesen csak utóbbi értékükkel szerepelnek a statisztikában; így a HB. *halál*, *holt* szavainak *hal* alapszava a felbontott szóelemek számát a ténylegesekhez képest nem gyarapítja, mivel az ige önállóan, „tényleges” sikon is megvan a szövegben (*holz*), és magától értődően ilyen minőségében jön számításba.

Elsősorban a diakrón jellegű és érdekű vizsgálat szempontjából jó a tényleges és a felbontott szóelemek kategóriáját világosan megkülönböztetni, és hasznos egymáshoz viszonyított arányukat az egyes szövegekben megvizsgálni. E két kategória aránya a szövegek időrendjébe állítva nagyon jellemző mozzanatokat árulhat el szókészletünk szerkezeti fejlődésére vonatkozólag.

A négy megvizsgált szövegben egyébként összesen 674 ténylegesen és 1194 felbontottan számbavett szóadat van, ami azt jelenti, hogy az alaktani felbontás során a négy szöveg együttes anyagából még 520 új, „többlet”-szóelem került elő. Ez a néhány szám jól jelzi az egész vizsgálat arányait, diakrón szempontból azonban természetesen nincs lényegesebb jelentősége.

Nézzük ezek után a négy megvizsgált emlékek a tényleges és a felbontott szóelemek viszonyait egymás mellé állító adatait. Az alábbi táblázat „különbség” jelzete az alaktani bontással pluszként nyert elemeket, „a növekedés százaléka” jelzete pedig ugyanezeknek a tényleges elemekhez viszonyított arányait jelenti.

	Tényleges szóelemek	Felbontott szóelemek	Különbség	A növekedés százaléka
HB. ....	123	171	48	39,0
Born. ....	146	221	75	51,4
Szem. ....	192	349	157	81,7
Ped. ....	213	453	240	112,7

A táblázatból kitűnik, hogy időben előre haladva a felbontott szóelemeknek a ténylegesekhez viszonyított aránynövekedése teljesen folyama t o s,

azonkívül meglehetősen nagy fokú és eléggé gyorsuló ütemű. Mindez annyit jelent, hogy a magyar szavak diakrón szempontból gyorsan és egyenes vonalú fejlődéssel haladtak az egyszerű szerkezeti formákból a komplikáltak, többrétűek, több-, „ízületűek” felé. A különböző szóalkotási módoknak (képzés, összetétel, ragszilárdulás stb.) nyelvtörténetünkben való fölerősödése — amely a bemutatott arányeltolódást okozza — úgy általánosságban ismert vagy legalábbis sejtett volt eddig is; az eltolódás jellegét és mértékét azonban a közölt adatok — úgy gondolom — jellemző módon és igen figyelemre méltó eredménnyel szemléltetik.

Úgy látszik, hogy a magyar szóelemek szerkezeti, etimológiai szempontból való komplikálódási, tagolódási foka a HB.-től máig majdnem a duplájára emelkedett, azaz — más szavakkal mondva — a mai magyar szó-kincsben közel megegyezően annyi szó szerkezeti tagolható, külön egységként kezelhető elemet találunk, mint a nyolc évszázaddal ezelőttiben. Az aránynövekedésnek a megvizsgált emlékekben való egyenesvonalúsága arra mutat, hogy ha az itt közölt eredményt több és nagyobb szóanyagú szöveg statisztikai vizsgálata esetleg erre vagy arra némileg módosíthatja is, lényegesen más megállapításra aligha nyújthat alapot.

4. A szóelemek szerkezeti komplikálódásával bizonyos mértékben rokon jelenség alakí kiterjedé s e s e d e s ü k , m e g n ö v e k e d e s ü k i s . Az összefüggés, a rokonság abban áll, hogy a különböző szóalkotási módok fölerősödése — amely a szerkezeti komplikálódást okozza — néhány kivételt (pl. az elvonás) nem tekintve a szó hangalaki meghosszabbodását idézi elő. Éppen e rokonság miatt a fentebb közölt adatok többé-kevésbé rámutatnak a magyar szavak hangalaki növekedésének a mértékére is. Emellett azonban egy másik szempontból és még konkrétan tájékoztat bennünket az efféle fejlődésről a szóelemek s z ó t a g s z á m - s t a t i s z t i k á j a . Itt — a dolog természeténél fogva — csak a tényleges szóelemek szótagszámát kell tekintetbe vennünk, a hangalaki hosszabbodás mértékét ezek egymáshoz viszonyított aránya mutatja.

Íme a megvizsgált négy szöveg szótagszámadatai. A táblázatok legfelső sorában levő számok 1-től 6-ig a tényleges szóelemek szótagszámát jelölik.

	1	2	3	4	5	6
HB. ....	51	56	15	1	—	—
Born. ....	51	62	22	9	2	—
Szem. ....	46	78	40	18	6	4
Ped. ....	32	59	72	31	15	4

Ugyanez százalékszámokban kifejezve:

	1	2	3	4	5	6
HB. ....	41,5	45,5	12,2	0,8	—	—
Born. ....	34,9	42,4	15,1	6,2	1,4	—
Szem. ....	24,0	40,6	20,8	9,4	3,1	2,1
Ped. ....	15,0	27,7	33,8	14,6	7,0	1,9

A közölt adatokból kitűnik, hogy a magyar szavak szótagszáma az idő előrehaladtával fokozatosan és meglehetősen nagy intenzitással emelkedett, szavaink jelentős hányada tehát egyre hosszabb hangalakú, egyre testesebb lett. Míg a HB.-ben az egy- és kétszótagú szóelemek teljesen uralják a szöveget, ez a szerepük a későbbi emlékekben fokról fokra megszűnik, úgyannyira, hogy a mai típusú szövegben a mérleg nyelve már körülbelül a háromtagú szavakon nyugszik. Rendkívül jellemző a háromnál több szótagszámú elemek egyre erősödő jelentkezése is.

A fejlődésnek az a szinte mértani pontosságú egyenesvonalúsága, amelyet a négy szöveg adatai ebből a szempontból is tükröznek, minden bizonynyal nem véletlen; s ha más vizsgálatok egyik-másik kategória arányaira vonatkozóan más adatokhoz fognak is jutni (mint ahogy igen valószínűen úgy is lesz), a változás alapvetően efféle jellegét aligha módosíthatják túlságosan. Hiszen a szókincsünket ért nyelvtörténeti változások ismerete alapján ez az eredmény semmiképpen nem irreális, legföljebb a maga viszonylagos réalitását, konkrétitását kissé meglepő vagy legalábbis szokatlan. Egyébként — ami természetes is — eléggé összhangban áll a szóelemek szerkezeti komplikálódásának vizsgálati eredményeivel.

Talán mondanom sem kell, hogy a magyar nyelv szavainak szerkezeti tagolódásában és ezzel együttjáróan szótagszámában, szótestében bekövetkezett erőteljes gyarapodás mint nyelvünk szerkezetének egyik igen lényeges módosulása kihatással volt a nyelvtörténet számos jelenségének a fejlődésére, egyes hangtörténeti változásoktól kezdve szóalkotásbeli mozzanatokon át egészen a mondat szerkesztés sajátosságaiig. Ezeknek a bőven adódó problémáknak a boncolgatása azonban már nem ennek a cikknek a feladata.

5. A szófaji kategóriák vizsgálatát két szinten: egyrészt a tényleges, másrészt a felbontott szóelemek szempontjából végeztem el. Ez annyit jelent, hogy például a HB. *kínzat* szava az első szempont szerint egy főnévi elemként jött számításba, a második szempont szerint ellenben három szófaji elemként: *kínzat* főnév, *kínoz* ige, *kín* főnév. Tehát a felbontott szóelemek adatait — mint már volt szó róla — itt is a tényleges szóelemek és az alakítási felbontás alapján még nyert elemek (alapszók, elő- és utótagok, igekötők stb.) együttesen teszik ki. Az alakítási felbontás útján újonnan nyert elemek számát az egyes részkatóriáknál már nem közlöm, mert ez

csak bonyolítaná az amúgysem túlságosan egyszerű statisztikát; a kiértékelésben azonban — ha szükséges — számszerűen is utalok majd rá. Ezeknek az újonnan, a szavak felbontása után pluszként nyert elemeknek a száma egyébként a közölt adatok alapján könnyen kiszámítható: az azonos kategóriák felbontott viszonylatú adataiból egyszerűen ki kell vonni a tényleges viszonylatú adatok számát. Tehát például — amint alább látni fogjuk — a HB.-ben a felbontott szóelemek szerinti vizsgálat 53 igét, a tényleges szóelemek szerinti vizsgálat 29 igét tartalmaz; következésképpen 53 mínusz 29 = 24, ami a felbontott vizsgálat során pluszként nyert (*kínó*z típusú) igealakok számával egyenlő. Ennek a problémának a tényleges és a felbontott helyzet összehasonlító vizsgálata terén van jelentősége.

A felbontott szóelemek szerinti statisztikázásra nemcsak azért volt szükség, hogy az egyes szófaji kategóriáknak a különböző szóalkotási módokban való részvételét, illetőleg alapként való szereplését szemléltesse, hanem azért is, mert a tényleges elemek statisztikai adataihoz viszonyított arányai-ból diakrón szempontból érdekes következtetések vonhatók le. Éppen a könnyebb összehasonlítás, az adatok egymás mellett szemléltetése érdekében a kétféle szempontú statisztikát mind az abszolút szám adatok, mind a százalék-számok esetében közvetlenül egymás mellé helyezett táblázatokban közlöm.

Az egyes szófaji kategóriákon belül a kisebb jelentőségűeket, illetőleg ritkábban előfordulókat bizonyos rokonsági csoportok szerint összevontam; erre egyrészt a statisztika elaprózási veszélyének az elkerülése miatt volt szükség, másrészt azért, mert — mint ahogy már jeleztem — a viszonylag nem nagy terjedelmű szövegek a kisebb szófaji csoportokra nézve csak véletlen esetlegességektől függő, tehát az irrealitás veszélyével járó képet nyújthattak volna. Így egy névmási csoportba vontam az összes névmásokat, ugyancsak egy igenévi csoportba a különböző fajta igeneveket; azonkívül a „határozószók” elnevezésű kategóriába foglaltam a valódi határozószókon kívül a módosító szokat, az ige-kötőket és a névutókat, az „egyéb segédszók” elnevezésű kategóriába pedig az összes többi segédzót (névelők, kötőszók, tagadó- és tiltó szók, indulatszók). Különösen ez utóbbi nyelvtörténeti szempontból meglehetősen vegyes kategória, de így összevonva még mindig többet nyújt, mint amit egyenkénti elkülönítésében adhatna.

A szófaji megoszlást tükröztető adatok tehát a következők:

A tényleges szóelemek viszonylatában:

	Ige	Igenév	Főnév	Melléknév	Számnév	Névmás	Határozószó	Egyéb segédszó
HB. ....	29	9	47	9	2	12	11	4
Born. ....	44	7	38	9	2	12	23	11
Szem. ....	32	16	46	23	3	22	36	14
Ped. ....	22	26	86	37	4	10	18	10

## A felbontott szóelemek viszonylatában:

	Ige	Igenév	Főnév	Melléknév	Számnév	Névmás	Határozó- szó	Egyéb segédszó
HB. ....	53	9	66	10	2	13	13	5
Born. ....	78	7	65	14	3	15	27	12
Szem. ....	89	22	107	37	4	26	50	14
Ped. ....	137	32	150	50	6	22	40	10

Ugyanezek százalékszámokban kifejezve:  
A tényleges szóelemek viszonylatában:

	Ige	Igenév	Főnév	Melléknév	Számnév	Névmás	Határozó- szó	Egyéb segédszó
HB. ....	23,6	7,3	38,4	7,3	1,6	9,7	8,9	3,2
Born. ....	30,1	4,8	26,0	6,2	1,4	8,2	15,8	7,5
Szem. ....	16,7	8,3	23,9	12,0	1,6	11,5	18,7	7,3
Ped. ....	10,3	12,2	40,4	17,4	1,9	4,7	8,4	4,7

## A felbontott szóelemek viszonylatában:

	Ige	Igenév	Főnév	Melléknév	Számnév	Névmás	Határozó- szó	Egyéb segédszó
HB. ....	31,0	5,3	38,6	5,8	1,2	7,6	7,6	2,9
Born. ....	35,2	3,2	29,5	6,3	1,4	6,8	12,2	5,4
Szem. ....	25,5	6,3	30,7	10,6	1,1	7,5	14,3	4,0
Ped. ....	30,2	7,4	34,2	11,0	1,3	4,8	8,8	2,0

A fenti statisztikák értékelésében egyrészt a tényleges szóelemek és a felbontott szóelemek külön-külön táblázatainak adatai önmagukon belüli arányaikban, változásaikban érdemelnek figyelmet; másrészt a két szempont szerint vizsgált szóelemek táblázatainak egymáshoz viszonyított — az alak-  
tani bontás további, a tényleges szóelemeken túli eredményeit tükröztető —  
arányaira, változásaira kell figyelniük.



Az első és legáltalánosabb kép, amely mindegyik táblázatból, az adatok egészéből élénk tárul az, hogy a szófaji kategóriák arányai nem mutatnak olyan jelentős mértékű időbeli változást, olyan nagy eltolódást, mint amit az előbbiekből tapasztalhattunk. Mindez a konkrét statisztikai adatok fényében is megerősíti azt, hogy a magyar szófajok rendszere a HB.-től napjainkig nem módosult alapvetően, és belső arányai sem változtak meg túlságosan. A jelenség oka nyilvánvalóan kettős. Egyfelől az alapvető szófajok — fogalmi-logikai és ennek vetületeként nyelvi-grammatikai „kötöttségük” folytán — bizonyos mértékben az egyes konkrét nyelvek fölötti, az emberi nyelv egészét átfogó jelleggel adva vannak, másfelől — de az előbbivel összefüggve — a magyarban a szófaji rendszer alapvető vonásai a HB. idejére már kialakultak; az utána következő módosulások, például a névelők, igekötők kifejlődése az egész rendszert tekintve nem nagy mértékűek, azonkívül nem annyira egyedeik számával, mint inkább gyakoriságukkal hatottak (ez utóbbit azonban a jelen statisztikai vizsgálatok nem tükrözik).

Egy másik feltűnő jelenség, amely a közölt szófaji statisztikában általánosabb jelleggel észlelhető, hogy több kategória a megvizsgált nyelvemlékekben bizonyos — bár viszonylag kisebb mértékű — számbeli, aránybeli ingadozást mutat, időbelileg nézve némi emelkedést, majd visszacsökkenést felváltva. Ezt a kis intenzitású vibrációt szintén két körülmény magyarázhatja. — Az egyik az, hogy annak az egy irányban, egyenes vonalban való eltolódásnak a hiánya, amely eltolódás az előbbi vizsgálati eredményekben tapasztalható volt, vagyis a többé-kevésbé egy szinten való stagnálás ténye elsősorban a megvizsgált, eléggé kis szóanyagú szövegek következtében magától értődően nem eredményezhet a valóságot teljesen híven tükrözhető, kiegyenlített, különösebb ingadozást nélkülöző képet. Éppen egyes szófaji kategóriák diakrón szempontból való vibrációja mutat rá arra, hogy a fejlődés ez esetekben sem az aránybeli növekedést, sem az aránybeli csökkenés irányában nem indult el erősebb mértékben. — A másik magyarázó körülmény tartalmi-stilisztikai vonatkozású. A megvizsgált szövegek a kiválasztás ismertett szempontjai mellett sem egyforma jellegűek: tartalmuk és ennek jórészt függvényeként jelentkező stilisztikai helyzetük még mindig eléggé más ahhoz, hogy bizonyos szófaji kategóriák egyszer gyakoribb, másszor ritkább előfordulásának a forrása legyen. Erről a tényezőről alább még lesz szó.

A közölt táblázatok emez általános kép mellett persze sok érdekes részlet elárulnak az egyes szófaji kategóriák helyzetére, sőt egyikük-másikuk diakrón aránymódosulására vonatkozóan is.

Az igei kategóriában a felbontott típusú vizsgálat adatainak némileg vibrációs stagnálása mellett a tényleges típusú vizsgálat adatainak eléggé jelentős mértékű visszacsökkenése tapasztalható; ez utóbbi vonás egyébként a legszembeűnőbb módosulás az imént jelzett, általában kiegyenlített jellegű összképben. Az adatok azt tükrözik, hogy az ige nagy szerepe a szókincs egészében nyelvtörténetileg viszonylag változatlan maradt, ellenben jelenléte a tényleges igei alakban való megjelenés szintjéről egyre inkább az alakítani felbontás útján nyert szóelemek: alapszók, összetétel-tagok stb. szintjére süllyedt. Más szavakkal mindez annyit jelent, hogy az ige a magyarban igen nagy mértékben és egyre fokozódó intenzitással vált a különböző

szóalkotási módok alapjává. Ez a tényező azonban — elsősorban a szóképzés révén — ugyanakkor magának a „tiszta”, tehát a különböző szóalkotási módok érvényesülése után tekintett igei kategóriának a meglehetősen nagymértékű visszaesését is magában rejtette, hiszen a deverbális képzés nagyon gyakran nem verbumokat, hanem nomeneket eredményezett. — Ez a nyelvtörténeti változás látszik különben a tényleges jellegű *igenevek* diakrón szempontból növekvő számában. Itt viszont a felbontott viszonylatnak a ténylegeshez arányított, jól kivehetően kisebb mennyisége azt jelzi, hogy az *igenévi* kategória — továbbképzett jellegéből természetszerűleg folyóan — igen csekély mértékben szolgált már alapul további szóalkotások számára.

A *főnevek* viselkedése elsősorban az igei kategória imént tárgyalt jellegének bizonyos fokú ellentétével hívja föl magára külön is a figyelmet. Bármennyire nagy, némiképpen még az igét is meghaladó mértékben szerepel a főnév a szóalkotások kiinduló, alapul szolgáló részlegében, tényleges megjelenésének arányai amannál kisebb iramban esökkennek vissza az idők folyamán (a szerényebb visszahúzóást a Ped. tényleges főnévi adatainak a viszonylag nagy száma keresztezi ugyan, ennek azonban stiláris okai vannak, l. alább).

Az *ige* és a *főnév* tényleges és felbontott viszonylatú adatainak egymáshoz való arányait egyébként érdemes külön is számbavenni, mert így a két kategória közti ellentét még világosabb. Az *igénél* a felbontás útján nyert adatok növekedése a ténylegesekhez képest a következő: HB.: 86,7% (azaz 29-hez képest 24 növekedés, mely utóbbit az 53-ból levont 29 adja); Born.: 77,3% (44-hez képest 34); Szem.: 178,1% (32-höz képest 57); Ped.: 522,7% (22-höz képest 115). Ugyanakkor a *főnévi* adatok ilyenmő növekedése: HB.: 40,4% (47-hez képest 19); Born.: 71,1% (38-hoz képest 27); Szem.: 132,6% (46-hoz képest 61); Ped.: 75,6% (86-hoz képest 64). Az ilyen jellegő adatokhoz persze itt és a továbbiakban is hozzá kell számítani, hogy a 3. pontban kifejtettek szerint a felbontott szóelemeknek a ténylegesekhez viszonyított plusz adatai általában is erős növekedést mutatnak: HB.: 39,0%; Born.: 51,4%; Szem.: 81,7%; Ped.: 112,7%. Ennek a komponensnek a szerepe, valamint a százalékszámoknak kisebb, százon belül elosztott volta okozza, hogy a szófaji vizsgálat százalékos táblázata mind az *igénél*, mind a *főnévnél* lényegesen csekélyebb aránynövekedéseket — de természetesen ugyanolyan jellegzetességeket — mutat a tényleges és a felbontott jellegő adatok egymás közti viszonylatában. Itt a felbontott szóelemek növekedése a ténylegesekhez képest az *igéknél*: HB.: +7,4%; Born.: +5,1%; Szem.: +8,8%; Ped.: +19,9%; a *főneveknél*: HB.: +0,4%; Born.: +3,5%; Szem.: +6,8%; Ped.: -6,2%. — Mindez — más szempözből nézve — összegezöen egyúttal annyit is jelent, hogy egyrészt a *főnevek* a nyelvtörténet folyamán végig nagyobb mennyiségben maradtak érintetlenül a képzéstől, tehát maradtak meg egyszersmind biztosan szófajiságuk mellett, mint az *igék*, másrészt *főnevek* diakronikusan növekvöbb számban keletkeztek képzéssel, összetétellel stb., mint *igék*.

A *mellékneveknek* főként a tényleges, de a felbontott viszonylata is lassú, ugyanakkor azonban folytonos növekedést mutat az adatokban. E szófaji kategória szerepének időben előre haladó növekedésében stiláris okok is közrejátszóttak; erről szintén lesz még szó alább. — A felbontott szóelemeknek a ténylegesekhez viszonyított növekedése azonban a *mellékneveknél* nemcsak az *igékhez*, hanem a *főnevekhez* képest is jelentösen kisebb

arányú, és időben is jóval csekélyebb előrehaladású: még a Ped.-ban is csak 35,1% (37-hez képest 13), a százalékos táblázat szerint pedig csak —6,4%. A mellékneveknek a szóalkotásban való „alap”-szerepe tehát minden szempontból elmarad az előbbi két fő szófaji kategória mellett.

Érdekes, bár a szövegek viszonylag kisebb terjedelme miatt ugyanakkor többé-kevésbé véletlen is, hogy a számnevek előfordulásai mennyire egyszintűek. A mind külön tényleges, mind külön felbontott viszonylatban 1—1,5%-os átlag azonban körülbelül jól tükrözi a számneveknek a szófajok közti csekélyebb szerepét.

A névmások és határozószók külön-külön vett tényleges és felbontott viszonylatai elsősorban szerepük, arányaik meglehetősen nagyságánál fogva érdemelnek figyelmet. Ez a két szófaji kategória együttesen az ilyen szempontból megvizsgált adatoknak mindkét viszonylatban körülbelül 15—20%-át teszi ki; egymás közötti arányaik meg a határozószó fölényét mutatják. — A tényleges és felbontott viszonylatú adatoknak az egyes nyelvemlékek közötti összevetése pedig arra enged következtetni, hogy a szóalkotásban való „alap”-szerepük (a felbontott szóelemek plusz adatai közti szereplésük) idővel erősen megnövekedett, elsősorban a határozószóké. Míg a HB.-ben a felbontás útján nyert plusz határozószó abszolút számokban csupán 18,2% (11-hez képest 2), a százalékos táblázat szerint pedig csak —1,3%, addig a Ped.-ban már 122,2% (18-hoz képest 22), a százalékos táblázat szerint pedig +0,4%.

Az egyéb segédszók összevont csoportjának adatairól legfőbbjebb azt érdemes külön is megjegyezni, hogy a ténylegesekhez képest pluszként szereplő felbontott szóelemek alig találhatók meg köztük. Ez a számszerűség az efféle szófaji kategóriáknak a szóalkotás kiindulásaként, alapjaként a nyelvtörténetből egyébként is ismert csekély szerepét csak aláhúzza.

Még néhány olyan stilisztikai szempontú megjegyzés kívánkozik az elmondottakhoz, amely bizonyos ingadozásokra, arányeltolódásokra fényt vethet. — A Born. igen eleven, mozgalmas stílusú szövege kedvező talaj az igék és a határozószók számára; nem véletlen tehát, hogy mind tényleges, mind felbontott viszonylatban itt a legmagasabb például az igék százalékaránya. Elsősorban a Szem.-ben, de még a HB.-ben is erősebb a személyről, illetőleg személyhez szóló jelleg, mint a másik két szövegben; ez okozza főként a névmások, azonkívül — a névmási határozószók (*nekem, tőled, róluk* stb.) révén — a határozószók kategóriájának itteni aránynövekedését. — A Ped. modern tartalmú, illetőleg stílusú, a mai életre és írásmódra jellemző szövegében a szakmai szók szerepe lép előtérbe, ez viszont a névszók számbeli kiterjedésének kedvez. — A melléknevek kategóriájának a nyelvtörténet folyamán való számbeli, aránybeli erősödése alkalmasint összefügg az írói stílus árnyalásra, díszítésre való fokozott törekvésével, melynek a jelzőhasználat gyarapítása lényeges eszköze. De a Szem.-nek a „fennebb stíl” korából való kissé dagályos stílusa, a Ped.-nak pedig ünnepi tartalma, értékelő, magasztaló jellege a jelzői, illetőleg melléknévi használat fokozódását még ezen túlmenően is magyarázza.

6. A szóelemek származását, etimológiai kategóriáit természetesen szintén két szinten: a tényleges és a felbontott szóelemek viszonylatában vizsgáltam. Ez itt annyit jelent, hogy például a HB. *kínzat* szava — mint a szófaji típusban is — az első szempont szerint egy,

mégpedig nyelvünk külön életében keletkezett szóegységként került számításba, a második szempont szerint ellenben három egységként: a *kínzat* és a *kínoz* mint a magyarban keletkezett szó, a *kin* pedig mint jövevényszó. A felbontott szóelemek adatait tehát itt is a tényleges szóelemek és az alakítási felbontás útján még nyert elemek (alapszók, elő- és utótagok, igekötők stb.) együttesen teszik ki. Az alakítási felbontás révén újonnan nyert elemek számát az egyes részkatóriáknál e vizsgálati ágban sem közlöm; de az ugyanúgy, mint a szófaji táblázatoknál is, a megfelelő adatok egybevetése alapján könnyen kiszámítható.

A tényleges szóelemek vizsgálata mellett a felbontott szóelemek szerinti statisztikára természetesen itt is szükség van. Amennyire helytelen volna például a HB. *hatalom* vagy *munkás* szavában csak a *hat* vagy a *munka* alapszót keresni, és ennek megfelelően finnugor, illetőleg szláv eredetűnek minősíteni őket, nem véve tudomást arról, hogy ezek már a magyar nyelv külön életében, képzéssel keletkezett szavak, olyannyira kár lenne az etimológiai statisztikából kirekeszteni e szavak *hat* és *munka* „alap”-jait. Ha kirekesztenék őket, semmiképpen sem kapnánk a valóságnak teljesen megfelelő képet olyan kategóriáknak nyelvünk szóállományában való nagy szerepéről, mint az ősi (uráli, finnugor, ugor) eredetű szókinés vagy a jövevényszavak, hiszen ezek igen jelentékeny hányadát erősen hatáskörébe vonta a belső szóalakítás, következésképpen származékaiban, összetételeiben elmosta eredetibb etimológiai jellegét. — És magától értődik, hogy a tényleges és a felbontott szóelemek szerinti vizsgálat adatainak egymással való összehasonlítása maga is egész sor érdekes problémára világít rá szókinésünk etimológiai tagolódásának körében.

Az egyes kategóriákat itt sem apróztam el, hanem négy fő csoportba (1. ismeretlen és bizonytalan, 2. ősi, 3. jövevény, 4. belső keletkezésű szavak) soroltam a szóelemeket. Az ősi szókinés uráli, finnugor és ugor csoportra való továbbtagolása a besorolás biztonságát illetően is nagy nehézségekkel járt volna, amellett a megvizsgált szóállomány mennyisége alapján sem igen vezethetett volna reális eredményre. A számbelileg legnagyobb csoportot alkotó belső keletkezésű szavakat továbbtagoltam ugyan, e tagolás eredményét azonban nem itt, hanem az 5. szempontként említett alakítási vizsgálat során közlöm. A jövevényszavaknak az egyes konkrét nyelvek, illetőleg nyelvcsaládok szerinti csoportokba sorolását elvégeztem ugyan, eredményét azonban majd a fő kategóriák szerinti vizsgálat bemutatása után közlöm, és csak összesítve, a négy szöveg együttes viszonylatában. A megvizsgált szövegek terjedelmére való tekintettel ugyanis egy-egy ilyen alkategória (török, szláv, latin stb. eredetű szavak) számadatai szinte véletlen esetlegességektől függenek, reális tájékoztató értékük vajmi csekély. Az összesített anyag rövid közreadása azonban — amely mégis már négy szövegnek tényleges viszonylatban közel hétszáz, felbontott viszonylatban pedig majdnem ezerkétszáz szóeleméből kerül ki — legalábbis a magyar szókinés nem diakrón szempontból tekintett állagára vonatkozóan nyújthat némi támpontot, illetőleg adalékot a korábbi ilyen jellegű vizsgálatok eredményei mellé. Ez az összesítés azonban igazán csak egy kis ráadás, semmi több.

A magyar szókinés eredetére vonatkozó eddigi statisztikai számítások (1. pl.: TOLNAI VILMOS: MNy. XX, 55; ZSIRAI MIKLÓS: FgrRok. 47; VERMES STEFÁNIA: NyK. LI, 435) legfőbb hibája az volt, hogy a magyar nyelv külön életében keletkezett szavakat nem választották külön az ősi elemektől, hanem

„eredeti”, „magyar” megjelöléssel általában egybeolvasztották e két csoportot, tetézve ezt még azzal is, hogy a jövevényelemek magyarban való továbbalakításaival nagyon ingadozóan, hol így, hol úgy bántak el. Mindenki tudja, hogy az ősi eredetű és a belső keletkezésű szóelemek elhatárolása — mint általában minden efféle elhatárolás — nem könnyű dolog: gyakran csak nagyon tüzetes utánjárás, megfontolás alapján lehetséges, néha még úgy sem, hiszen vannak egészen bizonytalan határesetek is. E nehézség azonban nem lehet ok arra, hogy ezt a két alapvetően különböző csoportot egybekeverjük. Nem is szólva arról, hogy más etimológiai szókatégoriák elhatárolásában szintén gyakran adódnak hasonló nehézségek, az egybeolvasztás éppen annak a történeti—etimológiai szempontnak mond ellent, amit e kategorizálás szolgáltni akar, azonkívül teljesen lehetetlen helyzetet teremt a jövevénytiszavak belső továbbalakításainak a megítélése terén is. Ha a HB. *munkás* szavával kapcsolatban meg kell különböztetnünk magát a magyarban keletkezett származék-szót a jövevény-elem alapszótól — mint ahogy nyilvánvalóan így kell eljár-nunk —, akkor ugyanezt kell tennünk a HB. *hatalom* szavával is: itt is meg kell állapítanunk külön a tényleges szóelem minőségét (magyarban keletkezett származék) és külön az alapszó minőségét (ősi szóelem). Világos, hogy e példákban csak a *munkás* és a *hatalom* lehet azonos etimológiai kategória, a *munka* és a *hat* sem egymással, sem az előbbiekkal nem kerülhet azonos csoportba. De úgy hiszem, etimológiával foglalkozók előtt nem is kell tovább magyarázni, miért hordoz egészen más problematikát magában, és mennyiben egészen más típus az ősi eredetű (a nyelvrokonságból öröklött) szó és a belső keletkezésű (magának a kérdéses nyelvnek az önálló életében létrejött, ille-tőleg továbbalakult) szó.

Persze az ősi és a belső keletkezésű elemek szétválasztásának szükséges-ségét nemcsak elvi szempontok támogatják, hanem a megkülönböztetésnek gyakorlati hasznossági okai is vannak. Az egybeolvasztással egyfelől teljesen összegabalyítunk olyan szócsoporthokat, amelyeknek külön-külön számbavétele nagyon tanulságos szókinsrendszerünk etimológiai fel-építésének megítélésében, másfelől lehetetlenné tesszük annak megállapít-hatóságát, milyen mértékben vett részt a belső szóalakításban (képzés, össze-tétel stb.) az ősi, öröklött, valamint a jövevény, idegen nyelvekből átvett réteg, azaz mi e két kategóriának szókinsünk életében való további sorsa. Hogy ezek is mennyire lényeges dolgok, aligha kell részletesebben magyarázni.

A bemutatandó adatok elé még csak annyit, hogy a nem teljesen biztos, de még eléggé valószínűsíthető etimológiákat valószínű csoportjukba sorolt-am, a kétes, többféleképpen magyarázható vagy nem kellően megtámogatott származásúakat ellenben már az ismeretlen-bizonytalan kategóriába rak-tam át. A besorolásban tehát nem voltam liberális, nem akartam a konkrétan minősíthető kategóriák (ősi, belső és jövevény elemek) hitelét rontaniésszel egyúttal az egész statisztika belső arányait is bizonyos fokig eltorzítani.

Lássuk mármost a számszerű adatokat, előbb a tényleges szóelemek viszonylatában:

	Ismeretlen- bizonytalan	Ősi	Jövevény	Belső
HB. ....	17	36	11	59
Born. ....	18	27	19	82
Szem. ....	15	28	14	135
Ped. ....	9	17	15	172

A felbontott szóelemek viszonylatának adatai a következők:

	Ismeretlen- bizonytalan	Ősi	Jövevény	Belső
HB. ....	30	63	14	64
Born. ....	31	55	27	108
Szem. ....	36	65	37	211
Ped. ....	47	71	39	296

Most tekintsük át ugyanezeket az adatokat százalékarányokban. A tényleges viszonylat arányai:

	Ismeretlen- bizonytalan	Ősi	Jövevény	Belső
HB. ....	13,8	29,2	9,0	48,0
Born. ....	12,3	18,5	13,0	56,2
Szem. ....	7,8	14,6	7,3	70,3
Ped. ....	4,2	8,0	7,0	80,8

A felbontott viszonylat százalékarányai:

	Ismeretlen- bizonytalan	Ősi	Jövevény	Belső
HB. ....	17,6	36,8	8,2	37,4
Born. ....	14,0	24,9	12,2	48,9
Szem. ....	10,3	18,6	10,6	60,5
Ped. ....	10,5	15,6	8,6	65,3

Ezekben az etimológiai helyzetet mutató táblázatokban ismét határozottnak, egyirányúnak mondható fejlődési tendenciákat, azonkívül meglehetősen erőteljes arányeltolódásokat tapasztalhatunk. A fejlődésnek, az arányok eltolódásának a legáltalánosabb jellemzője mind a tényleges, mind a felbontott szóelemek viszonylatában a belső keletkezésű szavaknak időben előre haladva egyre erősebb előretörése, egyre nagyobb túlsúlyba kerülése, illetőleg — az előbbivel együttjáróan — a másfajta eredetű szavak diakrón tekintetben való visszacsökkenése.

A magyar nyelv külön életében keletkezett szavak a tényleges szóelemek statisztikájában vannak nagyobb arányokban képviselve. Ez teljesen érthető is, hiszen ebben a viszonylatban csak a szóalkotás végső eredményei jönnek számításba. Rendkívül jellemző, hogy e vizsgálati szempont szerint a magyarban keletkezett szavak aránya már a HB.-ben majdnem 50%, tehát legkorábbi szövegműveinkben már csaknem minden második szó valamilyen belső szóalkotással — mint látni fogjuk: ekkor még legtöbbször képzéssel, kisebb részben ragszilárdulással — keletkezett. És az sem kevésbé érdekes, hogy mai szövegben az arányok olyan méreteket öltönek, hogy az összes számításba jöhető tényleges elemek négyötöde belső keletkezésű szó. A szóalkotás belső módjai tehát igen nagy hatókörrel és egyre fokozódó iramban érintették a többi kategóriát, és másították meg azok jellegét. — Az is igen sokat mond a belső keletkezésű szavaknak szókincsünk egészében uralkodó és egyre inkább eluralkodó helyzetéről, hogy még felbontott viszonylatban is a HB. összes elemeinek több mint egyharmada, a Ped. összes elemeinek majdnem kétharmada ebbe a kategóriába tartozik. Hogy az arányok még itt is, a szóalkotási alapok beszámításával együtt is nyolc évszázad alatt közel a duplájára emelkedtek, azt az tette lehetővé, illetőleg az magyarázza, hogy a nyelvtörténetben előre haladva belső keletkezésű szavak is mind jobban és jobban újabb szóalkotási módok hatáskörébe kerültek, vagyis az „alap”-rétegbe is egyre inkább behatoltak. Ez a tény nagyon jól kiviláglik akkor, ha a felbontás útján újonnan nyert belső keletkezésű szóelemeknek a ténylegesekhez mért növekedését nézzük. Számokban kifejezve ez a következő: HB.: 8,5% (59 tényleges elemhez képest 5 növekedés); Born.: 31,7% (82-höz képest 26); Szem.: 56,3% (135-höz képest 76); Ped.: 72,1% (172-höz képest 124). Ugyanez a százalékos táblázat szerint, az általános növekedési komponensnek és a százalékszámok sajátos jellegének a hozzászámításával: HB.: —10,6%; Born.: —7,3%; Szem.: —9,8%; Ped.: —15,5%. Ez utóbbi számokban a mínusz-értékek természetesen csupán relatív jelentőségűek; annyit jelentenek, hogy a többi kategóriában (l. pl. alább az ősi elemeknél) még nagyobb volt a felbontás útján nyert adatok aránynövekedése, mint emezekben.

Ősi eredetű szavainknak a magyar szókincsben való viszonylag nagy szerepe — ha nem is mérhető egészen az előbbi kategória mértékével — világosan kitűnik az adatokból. A HB.-ben még a tényleges viszonylat szerint is a szavak közel egyharmada ősi eredetű, a felbontott viszonylat szerint pedig több mint egyharmada ebből a kategóriából kerül ki. Az a nagy zuhanás, ami az ősi elemeknek az arányaiban diakrón szempontból elsősorban tényleges viszonylatban megmutatkozik, annak tulajdonítható, hogy ezt a szókategóriát a belső szóalkotás legelső sorban pusztította, változtatta át másféle típusra. Ez a tény bizonyos oldalról nézve szintén az ősi elemeknek nyelvünk szókincsében való fontos szerepére utal, hiszen annak a jele, hogy ezek az

elemek igen nagy mértékben kerültek a belső szóalkotás hatása alá, azaz igen nagy mértékben váltak alapjaivá további szókeletkezéseknek. Hogy ez az „alap”-szintre kerülés időben előre haladva milyen fokozódó mértékű volt, azt jól mutatják azok az adatok, amelyek a fölbontás során újonnan nyert ősi szóelemeknek a ténylegesekhez képest jelentkező növekedését jelzik: HB.: 75,0% (36 tényleges elemhez képest 27 növekedés); Born.: 103,5% (27-hez képest 28); Szem.: 132,1% (28-hoz képest 37); Ped.: 317,6% (17-hez képest 54). Ugyanez a sor a százalékos táblázat szerint: HB.: +7,6%; Born.: +6,4%; Szem.: +4,0%; Ped.: +7,6%. A Ped.-ban tehát — amely a mai helyzetet képviseli — a különböző belső szóalkotásoktól érintetlenül maradt, ténylegesnek számítható ősi szavaknál több mint háromszor nagyobb azoknak az ősi szóelemeknek a száma, amelyek csak alaktani felbontás útján elemezhetők ki, vagyis amelyeket már a belső szóalkotási módok hatásuk alá vontak. Ez még a százalékos táblázat szerint is számottevő plusz-értéket jelent, tehát a többi kategória növekedéséhez viszonyítva is kiemelkedő fokot tükröz.

A táblázatok j ö v e v é n y s z ó - a d a t a i időben előre haladva csak abszolút számokban emelkednek, relatíve, százalékos viszonylatban azonban — a HB. aránylag kisebb jövevényszó-állományának kivételével — némi visszaesést mutatnak mind a tényleges, mind a felbontott elemek vizsgálata terén. A visszaesést itt is, mint az ősi elemeknél, a belső keletkezésű szavak előrenyomulása okozta, még közelebbről nézve az, hogy egyrészt nyelvünk belső szóalakítása erősebb ütemű, nagyobb intenzitású volt még a jövevényszavak önmagában szintén jelentős mértékű beáramlásánál is, másrészt a jövevényszavak jelentős hányadát egyre fokozódó mértékben hatása alá vonta, következőképpen eredetibb etimológiai jellegüket elmosta, megváltoztatta a belső magyar továbbalakulás (a *kín*: > *kínoz* típusú fejlődés). Az eredetileg jövevény-jellegű szavaknak magyarban keletkezett szavakra való emez átminősülését jól jelzik a négy megvizsgált szövegnek azok az adatai, amelyek a felbontás útján nyert plusz-jövevényelemeknek a ténylegesekhez mért növekedésére vonatkoznak: HB.: 27,3% (11 tényleges elemhez képest 3 növekedés); Born.: 42,1% (19-hez képest 8); Szem.: 164,4% (14-hez képest 23); Ped.: 160,0% (15-höz képest 24). Ugyanez a sor a százalékos táblázat arányaiban: HB.: -0,8%; Born.: -0,8%; Szem.: +3,3%; Ped.: -1,6%. — Bármennyire számottevő is ez a növekedés, mint ahogy a százalékos táblázat arányai is mutatják, körülbelül csak az alaktani felbontás útján nyert plusz-szóelemek összességének az átlaga körül mozog, de jól érzékelhetően kisebb az ősi elemek efféle növekedésénél, illetőleg — más szempontból nézve — „alap”-szintre szorulásának arányainál. Az ősi és a jövevény-szóelemeknek ez az eléggé különböző arányú fejlődése kettős okra megy vissza. Egyfelől míg nyelvünk ősi szólemeinek a nyelvrokonainktól való elszakadás óta megszűnt az „utánpótlási bázisuk”, azaz ezek tényleges szinten már nem, csak felbontott szinten (belső továbbalakításuk révén) gyarapodhattak, addig a jövevényszavak beáramlása nyelvünk külön élete folyamán végig tartott, sőt nyilván fokozódott, vagyis ezeknél a tényleges szint nemcsak a már meglévő és ilyen jellegüket megtartott (tovább nem alakított) jövevényelemekből adódik, hanem a folytonos gyarapodás eredményét is minden időben mutatja. Másfelől a jövevényszavak, ha nagyon sok új, belső keletkezésű szónak váltak is az alapjává, mégis számos körülménynél (pl. viszonylag újabb voltaknál, a szókinés egészében elfoglalt gyakori periferikus jellegüknél, tárgyi-jelentésbeli háttérüknél stb.) fogva nem vettek olyan arányokban részt a belső magyar



szóalakításban, nem szerepeltek oly gyakran ennek „alap”-anyagában, mint az ilyen szerepre különböző tulajdonságaiknál fogva általában alkalmasabb ősi eredetű szavak.

Most még néhány szót a jövevényszavak beső megoszlásáról, — mint említettem — a megvizsgált szövegek adatait összesítve. A négy szöveg együttes jövevényszavainak aránya a tényleges szóelemek viszonylatában a szóállomány egészének 8,7%-a (674 szóelemből 59); a felbontott szóelemek viszonylatában pedig a szóállomány egészének 9,9%-a (1194 szóelemből 117). A felbontott viszonylat növekedése a tényleges viszonylathoz képest együttesen 98,3% (59 tényleges elemhez képest 58 növekedés). — A továbbiakban már a jövevényszavakon belül maradva a négy szöveg együttese tényleges elemeinek megoszlása ez: iráni 5,1% (3 szóelem); török 25,4% (15); szláv 30,5% (18); latin 30,5% (18); német 6,8% (4); újlatin — (—); görög 1,7% (1); — felbontott elemeinek megoszlása pedig ez: iráni 4,3% (5 szóelem); török 35,0% (41); szláv 30,8% (36); latin 20,5% (24); német 6,8% (8); újlatin 1,7% (2); görög 0,9% (1). E számokban a viszonylag kevés adathból való statisztikázáson kívül természetesen a közel nyolcszáz évet átfogó távlat is torzít; például a német jövevényszavak aránylag csekélyebb számát többek között az is magyarázza, hogy a HB.-ben még nincs német eredetű szóelem (ez időben magában nyelvünk egészében is alig van ilyen). — Legalábbis a három nagyobb létszámú kategóriára nézve adhat némi támpontot a felbontott szóelemeknek a ténylegesekhez mért növekedése is: török: 173,3% (15 tényleges elemhez képest 26 plusz); szláv: 100,0% (18-hoz képest 18); latin: 33,3% (18-hoz képest 6); — ugyanez a százalékarányok viszonylatában: török: +9,6%; szláv: +0,3%; latin: —10,0%. Nyilván több a véletlennél, s bizonyos valóságos helyzetet közelíthet meg az, hogy a három kategória közül az adatok tanúsága szerint a viszonylag legrégebb és sok fontos szót tartalmazó török réteg vett részt legjobban a belső magyar szóalakítás alapjaiban, vagyis ez csökkent vissza legjobban az alaktani felbontás útján nyerhető plusz szóelemek szintjére; ugyanakkor — a szláv réteg mintegy közbülső helyzete mellett — az időben fiatalabb és aránylag sok periferikus szót tartalmazó latin réteg meglehetősen kis szerepet játszik a belső magyar továbbfejlődés terén.

7. Utolsó vizsgálati szempontként a szóelemek alaktani szerkezete szolgált. Itt — mint említettem — csak az előbbi vizsgálati mód alapján belső keletkezésűeknek minősült szavakat vettem számításba; tudniillik egynéhány, még a rokon nyelvi együttélés során való keletkezésre, azaz ősi eredetre mutató képzésen vagy elhomályosult összetételen kívül — melyeknek aránya a megvizsgált szövegben igen minimális — e belső keletkezésű kategória az, amelyik a benne lefolyt különböző szóalkotások révén az alaktani felbonthatóság állapotában van. A továbbiakban tehát a HB. 59 tényleges és 64 felbontott, a Born. 82 tényleges és 108 felbontott, a Szem. 135 tényleges és 211 felbontott, a Ped. 172 tényleges és 296 felbontott viszonylatú belső keletkezésű szavát vizsgálom.

A megvizsgált szavakat négy alaktani csoportba soroltam: 1. képzett, 2. „ragozott”, 3. összetett, 4. elvont szavak. „Ragozott”-on itt természetesen a ragszilárdulással keletkezett, többnyire határozószói, kisebb részben névmási vagy egyéb szófajiságú elemeket kell érteni. Az összetett szavak kategó-

riájába kerültek az igekötős igék is, ezekről azonban még külön is közlök majd néhány adatot. A szóalkotás kisebb, szórványosabb módjait (vegyülés, hangrendi átcsapás, játszai alakítás stb.) nem soroltam még gyűjtőkategóriába sem, ezeknek biztos esete ugyanis a megvizsgált szövegekben nem merült föl. A kifejező jellegű, onomatopoeitikus eredetű szavakat — melyek száma itt szintén nem túlságosan nagy — a képzőrendszerbe való analógiás beilleszkedésű szövegződéseiknél fogva a képzett elemekhez soroltam. — Ismeretes, hogy e kategóriák közül az összetett szavaknak mint ilyeneknek a felismerése, vagyis az egyszerű szókapcsolatoktól való megkülönböztetése milyen nehéz dolog, főként a régebbi időszakokra nézve. E tekintetben a képzetváltás mozzanatára, az egybe- vagy különírás tényére és a mai formákkal való egybevetésre támaszkodtam, mint főbb besorolási kritériumokra; mégis e kategória adatainál maradt még éppen eléggé nagy bizonytalansági hányados.

A tényleges és a felbontott viszonylat kettőssége e vizsgálatban konkrétan azt jelenti, hogy például a HB. *bírságnap* szava az első viszonylatban egy, mégpedig összetett elemnek, a második viszonylatban azonban két elemnek: *bírságnap* összetett és *bírság* képzett elemnek számít; ez utóbbi azért nem négynek, mert itt a *bír* és a *nap*, mint tovább már nem bontható elemek *k i r e k e s z t e t t e k* a *v i z s g á l a t b ó l* (az ilyen abszolút alapszavak, tövek száma az előző statisztikák adataiból könnyen kiszámítható). A felbontott szóelemek statisztikája természetesen még ilyen korlátozott jelleggel is érdekes eredményekre vezet: például a tényleges statisztika adataihoz mért növekedése megmutatja, hogy a képzett, „ragozott”, összetett és elvont kategória maga is milyen mértékben vett részt a belső magyar szóalkotás kiinduló, „alap”-rétegében.

A számszerű adatok tényleges viszonylatban a következők:

	Képzett	Ragozott	Összetett	Elvont
HB. ....	42	16	1	—
Born. ....	37	27	18	—
Szem. ....	57	34	43	1
Ped. ....	98	26	45	3

A felbontott viszonylat adatai pedig ezek:

	Képzett	Ragozott	Összetett	Elvont
HB. ....	45	18	1	—
Born. ....	53	35	20	—
Szem. ....	106	53	51	1
Ped. ....	169	54	66	7

## A tényleges viszonylat szerinti vizsgálat eredménye százalékszámokban:

	Képzett	Ragozott	Összetett	Elvont
HB. ....	71,2	27,1	1,7	—
Born. ....	45,1	32,9	22,0	—
Szem. ....	42,2	25,2	31,9	0,7
Ped. ....	57,0	15,1	26,2	1,7

## A felbontott viszonylat szerinti vizsgálat eredménye százalékszámokban:

	Képzett	Ragozott	Összetett	Elvont
HB. ....	70,3	28,1	1,6	—
Born. ....	49,1	32,4	18,5	—
Szem. ....	50,3	25,1	24,1	0,5
Ped. ....	57,1	18,2	22,3	2,4

A közölt adatok a nyelvtörténeti fejlődés szempontjából két általánosabb jellegzetességet mutatnak. Egyrészt az egyes kategóriák egymáshoz és önmaguk diakrón fejlődéséhez viszonyított aránybeli eltérései — talán a képzett szóelemek határozottabb kivételével — nem túlságosan nagyok, noha bizonyos fejlődési irányzatok itt-ott kirajzolódnak bennük. Másrészt — de az előbbivel összefüggésben — majd mindegyik kategória időbeli fejlődésében némiképp az ingadozás: emelkedés és visszacsökkenés váltakozása tapasztalható. Az összefoglaló kép tehát igen hasonló ahhoz, amit a szófaji vizsgálat tárt elénk. Ez a hasonlóság nem is egészen véletlen, hiszen a szóalkotási módok érvényesülésének vannak bizonyos szófaji kötöttségei: például a képzett szavak túlnyomó többsége az igei és főnévi kategóriához kapcsolódik, a „ragozott” alakok a határozószókban és a névmások egy részében jelentkeznek stb., ennél fogva ezek diakrón fejlődésének sajátosságaitól nem szakadhatnak el nagyobb mértékben. Azonkívül a megvizsgált szövegek viszonylag nem nagy terjedelme folytán a nagyobb intenzitású, határozottabb irányban való fejlődési kilendülés hiánya magától értődően számszerű ingadozásokon keresztül jelentkezik csupán. Sőt a szófajoknál érintett — és itt ismételnél fölösleges — stilisztikai jellegzetességek közvetlenebbül vagy a szófajokon át közvetve szintén hatottak némileg az alaktani kategóriák számadataira.

Ami mármint a részleteket illeti, mind tényleges, mind felbontott viszonylatban, azonkívül a nyelvtörténet minden szakaszában a képzett szavak mutatkoznak a legkiterjedtebb kategóriának. A nyelvtörténeti

átlagot véve mindkét viszonylatban 50% körül mozog arányuk, — az egy HB. kivételével, ahol azonban magasabb arányszámukat az magyarázza, hogy az „ellenlábás” kategóriák közül csak a „ragozott” típus van számottevően képviselve, a később tekintélyes arányokra emelkedett összetett szói típus viszont még majdnem teljesen hiányzik.

A ragszilárdulással keletkezett szókatégória egészében lassú visszacsökkenést mutat az idők folyamán. Úgy gondolom, hogy az a kép, amit a megvizsgált szövegben e kategória enyhe visszahúzó-dásáról kapunk, aligha állhat túlságosan messze a valóságos helyzettől. A megkövesült ragos alakulatok igen jelentős része ugyanis korábbi állapot hagyományozódása, relictuma nyelvünkben, újabb időkben való keletkezésük viszonylag gyérebbe volt miatt „utánpótlásuk” egyre gyengült, s így nem tudtak eléggé versenyt tartani más, sokkal hatásosabb, intenzívebb szóalkotási módokkal, elsősorban a képzéssel és az összetétellel.

A z ö s s z e t e t t s z a v a k diakrón szempontból való aránynövekedése még akkor is kétségtelen, ha beleszámítjuk az itt felmerülő súlyos kategorizálási nehézségeket, valamint ha a Ped. szövege — alighanem csak a viszonylag kis terjedelem miatti véletlenségek, alkalmoszerűségek folytán — a növekedési folyamatot nem is látszik mindenben igazolni. A HB. egyetlen bizonyosra vehető összetett szava (*bírságnap*) után a Born. szövege már ugrásszerű emelkedést mutat; könnyen lehetséges azonban, hogy a növekedés az e két emlék közti időben a valóságban nem volt ilyen erőteljes, tudniillik a Born.-ban jónéhány olyan többé-kevésbé összetételnek látszó szókapcsolattal találkozunk, amelynek ilyenként való megítélése — nem utolsósorban a szövegnek a ma nyelvétől való aránylag nagyobb időbeli távolsága következtében — korántsem kétségtelen; tehát az összetételek valóságos száma itt még alatta maradhat a táblázatokban megadottnál. Bár jól tudjuk, hogy az egyes szókapcsolatok összetétel vagy nem összetétel voltának az eldöntése még a ma nyelvében is mennyire nehéz vagy éppen reménytelen, a két utolsó szövegnek (Szem. és Ped.) e kategóriába sorolható példái már szilárdabb vázattal alkotnak a Born.-féle összetételeknél, biztosabb jelei e kategória létének és erősödésének. A HB. gyakorlatilag alig számottevő arányaihoz képest mindenesetre a Szem. 30%-ot is meghaladó tényleges viszonylatú összetett szói adatai reális különbségnek látszanak, és jól szemléltetik azt a nagy utat, amelyet ez a szókatégória a diakrón kiterelvényesedés tekintetében megtett. — Egyébként hogy ne csak a magyarban keletkezett szavak körén belüli arányoknál maradjunk, az sem egészen érdektelen, hogy a Szem. 192 ténylegesen számításba vehető szóadatából 149 az egyszerű, 43 az összetett szó, a Ped. ugyanilyen viszonylatú arányai pedig: 213 tényleges szóból 168 az egyszerű, 45 pedig az összetett szó. Ugyanez százalékarányban kifejezve azt jelenti, hogy az egyszeri előfordulás elve alapján számbavett tényleges szóelemeknek a Szem.-ben 22,4%-a, a Ped.-ban pedig 21,1%-a összetett szó. Mindez ad bizonyos támpontot arra nézve, milyen fontos helyet foglal el az újabbkori magyar nyelv szókinszrendszerében az összetett szavak kategóriája.

A z e l v o n t s z ó a l a k o k néhányából nem volna helyes bármily aránybeli következtetést is levonni. Az azonban nyilván nem véletlen, hogy e pár adat is az újabb szövegekben jön elő.

A három nagyobb alaktani kategóriának, a képzett, a „ragozott” és az összetett szavaknak a magyar szóalkotás „alapjai”-ban való részvételére a belső keletkezésű szavak keretén belül konkrét számokban kifejezhető tám-

pontokat kapunk, ha e kategóriák felbontott viszonylatú adatainak a tényleges viszonylatúakhoz mért növekedést tekintjük. Ez a képzett szavaknál a következő: HB.: 7,1% (42 tényleges elemhez képest 3 plusz); Born.: 43,2% (37-hez képest 16); Szem.: 86,0% (57-hez képest 49); Ped.: 72,4% (98-hoz képest 71). Ugyanez a százalékos táblázat adatainak tükrében: HB.: -0,9%; Born.: +4,0%; Szem.: +8,1%; Ped.: +0,1%. — A „ragozott” szavak adatai ezek: HB.: 12,5% (16 tényleges elemhez képest 2 plusz); Born.: 29,6% (27-hez képest 8); Szem.: 55,9% (34-hez képest 19); Ped.: 107,7% (26-hoz képest 28). Ugyanez a százalékos táblázat arányaiban: HB.: +1,0%; Born.: -0,5%; Szem.: -0,1%; Ped.: +3,1%. — Az összetett szavak adatai ezek: HB.: 0,0% (1 tényleges elemhez képest 0 plusz); Born.: 11,1% (18-hoz képest 2); Szem.: 18,6% (43-hoz képest 8); Ped.: 46,6% (45-höz képest 21). Ugyanez a százalékos táblázat tükrözött arányokban: HB.: -0,1%; Born.: -1,5%; Szem.: -7,8%; Ped.: -3,9%.

Az ily szempontból felsorakoztatott adatokból is látható, hogy a belső magyar szóalkotáshoz a HB.-ben még mennyire csekély számban szolgálnak alapul a már maguk is magyarban keletkezett szavak; később azonban „alap”-ként való szereplésük arányai egyre fokozódnak, és eléggé jelentős mértékűvé növekednek. E növekedés arányai legnagyobbak a „ragozott” szavaknál, legkisebbek az összetett szavaknál. Másik oldalról nézve a dolgot ez annyit jelent, hogy a belső keletkezésű szavak nagy csoportján belül a „ragozott” szavak azok, amelyek a szóelemek tényleges előfordulásának szintjéről a legnagyobb mértékben süllyedtek le a csak alakítani bontással kihámozható szóelemek szintjére, az összetett szavak viszont azok, amelyek a tényleges szinten a legjobban tartani tudták magukat. Ez érthető is, mivel az összetétel egy némely tulajdonsága (viszonylag újabb keletkezés, szerkezeti komplikáltság, hosszú hangalak stb.) nemigen kedvezett e szótípus egyedei alakítani továbbfejlődésének; az összetétel egy biztos értelemben mintegy befejező, végső formája már az alakítani való szógyarapításnak.

Az összetett szavakon belül az igekötős összetételek arányai a következők (elől az összetett szavak együttes adatszám, hátul az igekötős adatok száma): Tényleges viszonylatban: HB.: 1 : 0; Born.: 18 : 8; Szem.: 43 : 10; Ped.: 45 : 7. Felbontott viszonylatban: HB.: 1 : 0; Born.: 20 : 14; Szem.: 51 : 14; Ped.: 66 : 26. A HB.-ben tehát nincs biztosan igekötőnek fölfogható szóelem. A Born.-ban az igekötők viszonylag nagy számához hozzá kell venni azt a bizonytalansági tényezőt, amelyet e szóvegyes összetett szavainak a megítélésével kapcsolatban már említettem, ez ugyanis az igekötős igékre is vonatkozik. A Ped. igekötős igéinek jelentős része a felbontott szóelemek plusz részlegébe, a szóalkotási „alap” szintjére szorult le. — E problémák értékelésében mindenesetre tekintetbe kell venni, hogy a viszonylag kis kategóriára való lebontás miatt az adatok száma itt már föltétlenül kevesebb annál, hogysem a magyar igekötők fejlődésének a megítélésében hitelesebb erővel támaszkodhassunk rá.

**8.** A szókincs-szerkezet statisztikai vizsgálatának itt bemutatott — de a szempontok és a részletek tekintetében még távolról sem kimerített — formái nemcsak újszerűbbek, hanem kétségtelenül szövevényesebbek, nehezebben átláthatók is annál, amilyen vizsgálatokat a szokványos nyelvészeti munkálkodásban eddig alkalmaztunk. A szempontok bővítése, a részletekbe való nagyobb behatolás, a nyelvnek a szókincsen kívüli aspektusaira való kiter-

jeszkedés és nem utolsósorban a jóval nagyobb szöveganyagon kívánatos munkálkodás az efféle vizsgálatok bonyolultságát nagymértékben fokozza, azonkívül még az itt vázolt jelleggel és mértékben is rendkívül munkaigényes voltukat is meghatványozza. Mindez azonban — úgy gondolom — nem lehet akadálya annak, hogy a nyelvi rendszer statisztikai vizsgálatában a következő időkben ne haladjunk lényegesebben előre. Az ilyen természetű vizsgálat nehézségei, eredményeinek bonyolultsága magának a nyelvi rendszernek a hallatlan sokrétűségéből, szövevényességéből fakad, és egyúttal azt tükrözi. Általa azonban olyan eredményekhez jutunk, amelyek egyrészt konkrét, megfogható, számszerűen kifejeződő jelleggel mutatják meg a nyelv fejlődésének és állapotának azokat a nagyon lényeges jellegzetességeit, amelyekről csak nagy általánosságban, inkább sejtelemszerűen volt eddig tudomásunk, másrészt olyan fontos problémákra vetnek világot, amelyeket eddigi kutatásaink során nem tudtunk felszínre hozni. Ha az efféle vizsgálatokban rejlő lehetőségek igen széles skáláját ez a cikk nem is volt képes átfogni, talán adott belőle némi ízelítőt.

BENKŐ LORÁND