

## A magyar szókészlet finnugor elemeinek statisztikája

0. Az alábbi adatokat azzal a céllal közlöm, hogy a statisztika tükrében mérhessük fel a finnugor etimológiai kutatások helyzetét és fejlődését a hetvenes évek közepéig. Ekkor zárták le a szerzők (pontosan 1975. november 30-án) A magyar szókészlet finnugor elemei című etimológiai szótár harmadik kötetének kéziratát (I–III. Főszerkesztő: LAKÓ GYÖRGY, szerkesztő: RÉDEI KÁROLY és K. SAL ÉVA, munkatársak: ERDÉLYI ISTVÁN, FABRICIUS-KOVÁCS FERENC, GULYA JÁNOS, K. SAL ÉVA, VÉRTES EDIT. Akadémiai Kiadó, Budapest 1967, 1971, 1978. 727 l.).

Az első kötet megjelenésekor már készítettem hasonló statisztikát (NyK 70:392–396) és úgy vélem, hogy a teljes mű birtokában mindenképpen érdemes kiegészíteni az akkor közölt adatokat. A most közlésre kerülő statisztika készítésében segítségemre voltak az ELTE finnugor szakos hallgatói: BERNÁTH ÉVA, KLIMA LÁSZLÓ, MAROSÁN LAJOS, NAGY JÓZSEF, ROMETSCH HAIK és VARGA JUDIT. Ők végezték az anyaggyűjtést és részben az adatok összesítését, rám hárult munkájuk teljes ellenőrzése, a táblázatok összeállítása és a szöveges részek megírása. Ezúton köszönöm meg a felsorolt hallgatók lelkiismeretes munkáját.

Itt említem még meg, hogy JANURIK TAMÁS már korábban elkészítette az MSzFE (akkor részben még kézirat) anyaga alapján az uráli, finnugor és az ugor alapnyelv hangstatisztikáját (NéprNytud. 15–16:57–76). Bár a megjelent mű alapján készített hangstatisztika bizonyára némileg eltért volna Janurikétól (ő 691 szócikk alapján végezte számításait, a megjelent műben viszont 694 szócikk van), mégsem tartottam szükségesnek ennek elkészítését, főleg azért nem, mert a remélhetőleg hamarosan megjelenő Uralisches Etymologisches Wörterbuch alapján teljesebb képet kaphatunk majd az alapnyelv hangállományáról.

Az alább következő táblázatok értékeléséhez — mintegy használati utasításként — a következőket ajánlom az olvasó figyelmébe: A nyelvészetben a statisztikát sem túlbecsülni, sem lebecsülni nem lenne helyes. Az itt közölt számadatok nem hasonlíthatók fizikai vagy kémiai mérések eredményeihez, mert elsősorban a szótár szerzőinek véleményét tükrözik. Ezenkívül egy-egy új rokon nyelvi adat előkerülése módosíthatja az etimológia korát, minősítését stb. Az említett bizonytalansági tényezők figyelembevételével azonban mégis hasznos és tendenciájukat tekintve objektív információkat kaphatunk a finnugor eredetű szavak szerepéről nyelvünkben, a finnugor nyelvek rokonsági viszonyairól, az alapnyelvről stb. Az ilyen és hasonló statisztikák készítése tehát mindenképpen helyes és szükséges, még akkor is, ha az „errare humanum est” jegyében eleve feltételezzük, hogy a statisztika

összeállítója is követ el hibákat és ha az olvasó ellenőrző számításai nem mindig esnek egybe a cikkben közöltekkel.

1. Az MSzFE három kötetében 694 szócikk van, tehát nyelvünkben ennyi uráli, finnugor, illetve ugor eredetű tőszó található. Ez a szám természetesen magába foglalja a bizonytalan és vitatott eredetű szavakat is. Bár pontos adatok egyelőre nem állnak rendelkezésünkre, úgy tűnik, hogy a többi finnugor nyelv ősi szókészlete is hasonló nagyságrendű. Azt is tudjuk, hogy a végső elemzésben finnugor (uráli, ugor) eredetű szavak száma a képzések és összetételek jóvoltából szótári szókészletünkben sokkal nagyobb, mint a fenti szám és hogy az ilyen szavak használati aránya még ennél is nagyobb. Feltétlenül szükséges és időszerű lenne már ezekről modern statisztikák készítése az MSzFE, a TESz. és a készülő gyakorisági szótár alapján.

A címszavak kezdőbetűi szerinti statisztika a következő képet mutatja:

<i>a, á</i>	41	<i>n</i>	14
<i>b</i>	11	<i>ny</i>	27
<i>cs</i>	28	<i>o, ó</i>	15
<i>d</i>	6	<i>ö, ő</i>	14
<i>e, é</i>	42	<i>r</i>	21
<i>f</i>	74	<i>s</i>	20
<i>gy</i>	5	<i>sz</i>	47
<i>h</i>	79	<i>t</i>	41
<i>i, í</i>	23	<i>u, ú</i>	7
<i>j</i>	8	<i>ü, ü</i>	5
<i>k</i>	49	<i>v</i>	33
<i>l</i>	33	<i>z</i>	4
<i>ly</i>	1	<i>zs</i>	1
<i>m</i>	45		

Mint látható, nyelvünkben nincsenek *c, g, p, ty* kezdetű, finnugor eredetű szavak. Ennek hangtörténeti okai vannak, amelyekre most nem térhetek ki.

2. A címszavak döntő többsége köznyelvi szó (618 címszó, 89%), kisebb része népnyelvi, nyelvjárási (63 címszó, 9,1%), elenyésző része pedig kihalt, csak a régiségben előforduló szó (13 címszó, 1,9%). Ez utóbbi szám azt mutatja, hogy a magyar írásbeliség évszázadai alatt finnugor eredetű tőszavainknak csak meglepően kis hányada veszett ki nyelvünkéből. Bár itt a véletlen szerepét is figyelembe kell venni. Hiszen legtöbbször a véletlenen múlt, hogy egyik-másik szavunkat leírták-e a régebbi korokban. Így az is lehetséges, hogy egy-egy finnugor eredetű szó kiveszett nyelvünkéből, anélkül, hogy egyszer is leírták volna. Itt jegyzem még meg, hogy a szótárban népnyelviként megjelölt szavak közt több olyan is van, amely széles körű elterjedtsége miatt szinte már köznyelvinek tekinthető, pl. *csobolyó, csuhé, fan, hagymáz, iafia, ipa, napa, tér 'fér', zaj 'jégzajlás'* stb. Ez még inkább alátámasztja azt a megállapítást, hogy finnugor eredetű szavaink legnagyobb része élő, a mai magyar köznyelvben használt szó. Talán említésre méltó, hogy a népnyelvi szavak aránya a *cs* kezdetűek között a legnagyobb (28-ból 11), kérdés, hogy ez csak véletlen jelenség vagy valamilyen tendenciát takar.

3. A címszavak szófaji megoszlása (több szófajhoz tartozó szavaknál csak az elsőként megadott szófajt vettem figyelembe):

ige	232	33,5%
főnév	359	51,7%
melléknév	43	6,2%
számnév	19	2,7%
névmás	10	1,4%
határozószó	29	4,2%
egyéb	2	0,3%
összesen	694	100 %

A táblázathoz a következő megjegyzéseket fűzhetjük:

1. Az igék jelentős száma azt mutatja, hogy az alapnyelv sokrétűen tudta megjelölni a különböző cselekvéseket.

2. A melléknévek kis száma azzal magyarázható, hogy (mint a mai finnugor nyelvekben is) az alapnyelvben a főnév és melléknév kategóriája alig volt differenciálva, tehát a főneveket lehetett melléknévi funkcióban használni és viszont.

3. Tőszámneveink — kevés kivételtől eltekintve — finnugor eredetűek. (A mai magyarban mintegy két tucat tőszónak tekinthető számnév van, ide számítva a *negyven*, *ötven* féle alakokat is.)

4. Az előbbi megállapítás érvényes a névmásokra is.

5. Az igekötőkön és névutókon kívül az önálló szóból keletkezett ragoakat is a határozószók közé soroltam.

6. Viszonylag kevés (12) finnugor eredetű nomen-verbum található a szótárban. Ezek között is négy ikes ige van, négy szó pedig — nomen-verbumként — csak a régiségben vagy a nyelvjáráásokban fordul elő.

4. A szótár jelekkel, illetve szöveggel minősíti is az egyes etimológiákat. A „~” jel (tilda) áll a biztos egyeztetések előtt, 396 szó (57,1%) tartozik ide. A kevésbé biztos egyeztetéseket „?~” jelzi. Mindössze 9 ilyen szó van (1,3%). A következő kategória minősítése: egyeztethető (90 szó, 13%). A talán egyeztethető szavak csoportjába 95 szó (13,7%) tartozik. Vitatott eredetű 85 szó (12,2%). Ezeknél a szótár a valószínűség sorrendjében két-három származtatási lehetőséget is megad. A fennmaradó 19 szóból (2,7%) 12 finnugor elemekből alakult elhomályosult összetétel, 4 finnugor eredetű többlől képzett szó, 3 pedig az őspérmiből származó átvétel.

A vitatott etimológiáknál 179 lehetőséget ad meg a szótár. Ezek megoszlása: egyeztethető 36, talán egyeztethető 70, képzés finnugor többlől 32, összetétel finnugor elemekből 4, onomatopoeitikus 9, belső keletkezésű 3, jövevényszó 25. A lehetőségek döntő többsége (142) itt is finnugor eredetre utal.

A „~”, „?~” és „egyeztethető” (tehát biztos) etimológiák száma 495, a szótár etimológiáinak 71,3%-a. Ez a szám azt mutatja, hogy finnugor eredetűnek tartott szavaink nagy részénél a származtatás helyességét megfelelően dokumentálni tudjuk.

Anélkül, hogy értékelni akarnám a szótárban kialakított minősítési rendszert, szeretnék megemlíteni néhány szempontot, amelyek a statisztika

készítésekor merültek fel. Szembetűnő, hogy milyen ritka a „?~” minősítés. Ez is arra mutat, hogy felesleges volt ilyen kategóriát felvenni. Mint az előző bekezdésből kitűnik, az egyeztethető minősítésű etimológiákat is biztosnak tekintettem (ugyanígy járt el JANURIK TAMÁS is említett statisztikájában, l. i. m. 58). Tettem ezt azért is, mert nem tudtam megállapítani, hogy mi ezeknek az egyeztetéseknek a jellemző vonása a többi csoporthoz képest. Ha helyes ez a felfogás, akkor viszont nem lehetne a vitatott eredetű szavak magyarázási lehetőségei között az egyeztethető terminust használni. Ha ugyanis egy szó eredete vitatott, illetve vitatható, akkor az eleve nem lehet biztosan (egyeztethetően) finnugor. Talán jobb lett volna, ha a szótár csak biztos és bizonytalan (ezen belül vitatott) etimológiájú szavakat különböztetett volna meg.

5. A szótár megjegyzéseiből, illetve a rokonyelvi megfelelőkből arra is következtethetünk, hogy az illető szó uráli, finnugor vagy ugor eredetű-e:

uráli	139 szó	20%
finnugor (?uráli)	29 szó	
finnugor	325 szó	
finnugor összesen	354 szó	51%
ugor (?uráli)	8 szó	
ugor (?finnugor)	18 szó	
ugor	153 szó	
ugor összesen	179 szó	26%
nem minősíthető	22 szó	3%

A kapott eredmény nagyjából megfelel várakozásunknak. Megjegyzendő, hogy már egy új rokon nyelvi megfelelő felbukkanása révén átkerülhet a szó egy másik alapnyelvi rétegbe, tehát ez a táblázat eléggé labilis. Mégis valószínű, hogy nagy vonalakban a fenti megoszlás a későbbiekben is érvényben marad. A fenti számítás szerint az ugor korban az alapnyelv szókészlete jelentős mértékben növekedett (mintegy 36%-kal), míg a hang- és alaktanban ilyen jelentős különbséget nem tudunk prezentálni. (Bár az ugor nyelveknek vannak határozottan jellemző vonásai, mint azt legutóbb HONTI LÁSZLÓ kimutatta, vö. NyK 81:225—245.) Úgy vélem, ez a tény indokolttá teszi ezen a téren a további kutatásokat.

6. A szótár kilenc finnugor és öt szamojéd nyelvből közöl megfelelőket, s ezen kívül néhány adatot idéz kisebb finnségi, illetve kihalt szamojéd nyelvekből. A megfelelőlek számát nyelvenként az alábbi táblázat tartalmazza. Zárójelben „?” után áll a bizonytalan megfelelők száma, a %-os arányokba ezeket is beszámítottam. Az első %-oszlop azt fejezi ki, hogy a szótár 694 címszava közül hány %-nak van megfelelője az illető rokon nyelvben, a második pedig azt, hogy az összes rokon nyelvi megfelelőnek (3115) az illető nyelv szavai hány %-át alkotják.

Nyelv	A megfelelések száma	A címszavak %-ában	Az összes megfelelék %-ában
vogul	390 (?18)	58,8	13,1
osztják	365 (?11)	54,2	12,1
zürjén	253 (?41)	42,4	9,4
votják	228 (?32)	37,5	8,3
cseremisiz	181 (?30)	30,4	6,8
mordvin	180 (?17)	28,4	6,3
finn	236 (?37)	39,3	8,8
észt	214 (?20)	33,7	7,5
lapp	192 (?42)	33,7	7,5
jurák	103 (?31)	19,3	4,3
jeniszjei	72 (?19)	13,1	2,9
tavgi	72 (?14)	12,4	2,8
szelkup	107 (?22)	18,6	5,1
kamassz	72 (?17)	12,8	2,9
egyéb finnségi	65	9,4	2,1
egyéb szamojéd	34	4,9	1,1

Az összes megfelelék száma (a bizonytalanokkal együtt) 3115. Egy-egy címszóra tehát átlagosan 4,5 megfelelő jut (ez a számba vett nyelveknek harmadrésze). Ez azt mutatja, hogy az etimológiák megbízhatóságát megfelelő számú adat támasztja alá. A megfelelék számának alakulása egyrészt a magyar és az illető nyelv közötti rokonság fokát, másrészt az illető nyelv tanulmányozottságát tükrözi. Megfigyelhetjük, hogy közelebbi rokonság, illetve jobban ismert nyelv esetén a megfelelék száma elelkezik.

7. A szótár minden szócikkéhez részletes irodalmat közöl. A finnugor lexikológiában alapvető fontosságú négy művet választottam ki (MUSz. 1873—188; NyH<sup>7</sup> 1927; FUV 1955; SKES 1955-1978) és megvizsgáltam, hogyan foglal állást az MSzFE az említett művekben közölt olyan etimológiákkal kapcsolatban, amelyek ez utóbbi szótárnak is címszavai. A táblázat tehát azt mutatja, hogy az MSzFE szócikkei közül hány van meg az illető műben, és mennyiben helyes vagy helytelen az ott közölt etimológia.

	Helyes	Kevesebb nyelvből közöl adatot	Részben téves	Téves	Összesen	Ebből helyes
MUSz.	65	69	262	98	494	13,2%
NyH <sup>7</sup>	219	86	59	18	382	57,3%
FUV	212	78	76	5	371	57,1%
SKES	164	36	28	3	231	71%

A táblázatból első látásra is jól látható a MUSz. és a NyH<sup>7</sup> megjelenése közötti óriási fejlődés. A SKES-ben kisebb az MSzFE címszavainak száma, mint a többi műben. Ez azzal magyarázható, hogy a SKES csak a finn tagú etimológiákat közli és a kézirat lezárásáig utolsó két kötete még nem jelent meg. A helyes egyeztetések magas aránya (71%) és a teljesen téves egyeztetések elenyészően kis száma (3 etimológia) a SKES-ben azt mutatja, hogy a vezető finn és magyar finnugristák a finnugor szó- és hangtörténet alapvető kérdéseiben lényegében egyetértenek. Ez a tény is arra utal, hogy az MSzFE minden tekintetben méltó reprezentánsa és összegzője a finnugor etimológiai kutatásoknak.

CSÚCS SÁNDOR

### A Statistics of the Finno-Ugric Elements of the Hungarian Word-Stock

by SÁNDOR CSÚCS

The statistical data published here by the author are based on the etymological dictionary *A magyar szókészlet finnugor elemei* [The Finno-Ugric Elements of the Hungarian Word-Stock]. The author's main points are the following:

1. There are 694 entries in the dictionary of which 618 (89%) are standard, 63 (9,1%) dialectal and 13 (1,9%) are obsolete forms.
2. The number of attested Finno-Ugric etymologies is high (495 words, 71,3%).
3. Of all the entries 139 (20%) are of Uralic, 354 (51%) of Finno-Ugric and 179 (26%) of Ugric origin.
4. The number of correspondences found in the cognate languages, on the one hand, depends on the degree of relationship between Hungarian and the language in question and, on the other hand, on the volume of research done on the language in question.
5. In the case of etymologies the opinion of the co-workers of the Hungarian and the Finnish etymological dictionaries largely agrees with or nears to one another.