

Választás és véletlen a nyelvben

Megjegyzések Gustav Herdan „Language as Choice and Chance” című könyvéhez

„A nyelv mint a közlés eszköze három fő szempontból vizsgálható: a fonéma-rendszer, a szókincs és a grammatika szempontjából. Valamely nyelvközösségre azok a személyek tartoznak, akik — a különböző variációs lehetőségektől eltekintve — ugyanazt a fonéma-rendszert, szókincset és grammatikát használják. Kvantitatív vizsgálatok megmutatták azonban, hogy nemcsak ezekben egyeznek meg valamely nyelvközösség tagjai, hanem abban is, hogy milyen gyakorisággal használnak bizonyos fonémákat, szavakat, grammatikai formákat és szerkezeteket. Hasonlóság tehát nemcsak abban van, hogy mit használnak, hanem abban is, hogy milyen gyakran.” G. HERDANNak ez a megállapítása a nyelvnek lényeges aspektusát tárja fel, melyet a nyelvészek már régóta felismertek. Az a tény, hogy a nyelv nemcsak minőség, hanem mennyiség is, képezi alapját a nyelv statisztikai modelljének. Régóta tudjuk, hogy statisztikai módszerek alkalmazhatók a nyelv mennyiségi jellemzőinek leírására, tehát például a szógyakoriság, fonéma-gyakoriság vizsgálatára. Amikor azonban modellről beszélünk, akkor ennél többet kívánunk. A modellbe, akárhogy vesszük is, bele kell férnie az egész nyelvnek. Igaz, hogy általában a nyelv egyetlen aspektusát tárja fel, de a koncepciót igyekszik az egész nyelvre kiterjeszteni. A valószínűségszámítás, illetőleg a statisztika alkalmazása nyelvi jelenségek vizsgálatára része lehet ugyan a nyelv statisztikai modelljének, de önmagában még csak alkalmazásról van szó, mégpedig a valószínűségszámítás, illetőleg a statisztika sok alkalmazási területe közül az egyikről. A nyelv statisztikai modelljével újban többet foglalkoznak (HERDAN, MANDELBRÖT, SIMON stb.),¹ jelen ismertetésben HERDAN gondolatmenetét szeretném röviden vázolni.

A valószínűségszámítás viszonylag fiatal tudomány, ismertetésével a nyelvészeti alkalmazások szempontjából többen foglalkoztak már hazai viszonylatban is.² A következőkben csak a valószínűségi törvények lényegét foglalom össze röviden.

A valószínűségszámítás véletlen tömegjelenségek vizsgálatával foglalkozik. Mit értünk azonban véletlenen? Vannak a természetben olyan jelenségek, amelyek a következő székémba foglalhatók: ha a körülményeknek, feltételeknek bizonyos K komplexuma fennáll, akkor szükségképpen bekövetkezik egy bizonyos A esemény. Az ilyen székémet k a u z á l i s s z k é m á k nak nevezzük. Például ha desztillált vizet 100° C-ra hevítünk, minden esetben forni kezd 760 torr. nyomáson. De vannak olyan jelenségek is, amelyek a következő székémával írhatók le: a körülmények (feltételek) bizonyos K komplexumának fennállása esetén az A esemény be is következhetik, meg nem is. Az ilyen eseményeket nevezzük **v é l e t l e n e s e m é n y e k** nek, az ilyen székémet pedig **s t o c h a s z t i k u s s z k é m á k** nak. Így például, ha egyetlen rádiumatomot egy bizonyos időn keresztül figyelünk, előfordulhat, hogy ez az atom ez alatt az idő alatt elbomlik, de az is előfordulhat, hogy nem bomlik el. A véletlen eseményeknek természetesen szintén megvannak az okai, de a feltételek közt szereplő okok nem elegendők ahhoz, hogy az A esemény szükségszerűen bekövetkezzék.

A valószínűségszámítás azonban nem foglalkozhat egyszeri véletlen eseményekkel, hanem csak véletlen tömegjelenségekkel. Vagyis, a véletlen események közt meg kell különböztetnünk azokat, amelyek tömegesen fordulnak elő azonos körülmények közt és egyszeri véletlen eseményeket, amelyeknél a körülmények és így maga az esemény sem ismétlődik meg, illetőleg nem ismételtető meg. A véletlen tömegjelenségek olyan véletlen események, amelyek bekövetkezése vagy be nem következése azonos körülmények között igen sokszor figyelhető meg. Valójában tehát a valószínűségszámítás nem egyes

¹Vö. főleg G. HERDAN, *Language as Choice and Chance* (P. Noordhoff N. V. - 1956-Groningen); G. HERDAN, *Type-token Mathematics* (Mouton & Co. 'S-Gravenhage, 1960); B. MANDELBRÖT, *Théorie informationnelle (sémilogique) de la structure statistique des langues* (Technical Report, Massachusetts Institute of Technology, Research Laboratory of Electronics, 1953).

²Vö. ANTAL LÁSZLÓ, *Megjegyzések a szóállomány statisztikai vizsgálatához*: Nyr. LXXXIII., 307—11; FODOR ISTVÁN, *A statisztikai módszer alkalmazásának néhány kérdése*: Nyr. LXXXIV., 196—203; FÓNAGY IVÁN, *Beszéd és valószínűség*: Nyr. LXXXVI., 308—20; PAPP FERENC, *A matematikai módszerek alkalmazásáról nyelvtudományunkban*: I. OK. XVII., 289—306; TARNÓCZY TAMÁS, *A jelölés a hírtartalom nyelveket meghatározó tulajdonságairól*: NyK. LXIII., 161—178; VÉRTÉS EDIT, *Adalékok a magyar nyelv hangtani szerkezetéhez (Statisztikai vizsgálódások)*: NyK. LIV—LVI., hogy csak a legfontosabbakat említem.

eseményekkel, hanem egymással megegyező, tömegesen előforduló események összességével, eseménykategoriókkal foglalkozik.

Valószínűségi összefüggések igen nagy szerepet játszanak ma a legkülönbözőbb tudományágakban. Ha a valószínűségi relációk a klasszikus oksági törvényekhez való viszonyát akarnánk meghatározni, azt kellene mondanunk, hogy a klasszikus törvények mintegy „határértékei” a statisztikai törvényeknek, amennyiben azok ugyanis igen nagyszámú mikroszkopikus eseményre vonatkoznak. Ezt a „kettősséget” a nyelvben is tapasztalhatjuk. A klasszikus szemlélet szerint egy szó használatát különböző körülmények határozzák meg. Amikor ki akarunk fejezni valamit, akkor a kifejezendő tartalomnak megfelelő szót (szóalakot) kell keresnünk. Más szóval ez azt jelenti, hogy egy szó használata és a beszélő vagy az író szándéka közt oksági viszony áll fenn. Az író, illetőleg beszélő a nyelvi formák mennyiségi összetevőit illetően viszont valószínűségi törvényeknek van alávetve. Oksági viszonyt illetően meghatározottságot tulajdonképpen akkor is látunk, ha a nyelvi jelenségek összességét tekintjük például egy adott szövegben. Ha a nyelvről teljes képet akarunk alkotni, mind a mennyiségi, mind a minőségi oldalát figyelembe kell venni. Választás (oksági törvények) és véletlen (statisztikai törvények) ugyanannak a nyelvi valóságnak különböző aspektusai.

Ugyanakkor a nyelv minőségi jellemzőinek, mint például a stílusnak is vannak „mennyiségi” oldalai. Mennyiség és minőség tehát nem választható el egymástól. Hiszen két író stílusának különbözősége abban is megnyilvánul, hogy például bizonyos nyelvi formákat különböző gyakorisággal használ, vagy hogy például különböző nagyságú a szókincsük. Így lesz a nyelv statisztikai modelljének egyik legfontosabb területe a stílus-statisztika.

A szókincs vizsgálatánál különböző arányokat szokás kiszámítani. Ezek közül a stílus statisztikai vizsgálata szempontjából különösen fontos a következő három arány: speciális szókincs : teljes szókincs; speciális szavak (szóalakok) előfordulásának száma : összes szavak előfordulásának száma; speciális szókincs : összes szavak (szóalakok) száma. Ezek az arányok azonban mind függenek a minta nagyságától, mégis felhasználhatók szűk keretek közt a stílus jellemzésére. Ilyen arányok segítségével lehet valószínűsíteni valamely mű ismeretlen szerzőségét. (Ezt tették például Chaucer „The Equatorie” című műve esetében.³)

Az említett arányokat természetesen csak akkor tudjuk megállapítani, ha meg tudjuk számolni a szavakat. Amíg arányokról van szó, kvalitatív statisztikáról beszélünk, amikor azonban eloszlásokat és sűrűségeket vizsgálunk, a kvantitatív statisztika módszereit alkalmazzuk. A mintavételnél két módszert alkalmazhatunk.

1. V a k p r ó b a alapján veszünk mintát, ami azt jelenti, hogy a) a kiválasztás alapja az a tény, hogy a sokaság minden egyes eleme kiválasztásának ugyanaz a valószínűsége; b) két egymás után következő kiválasztás független egymástól. Ha például egy szótárban találhatók szavak számát akarjuk meghatározni, meg kell állapítanunk, hogy egy oldalra általában hány szó jut. A vakpróba-hoz különböző, úgynevezett vakpróba-számok használhatók (például a Tippett-féle táblázatok).⁴ 2. Szavaknál ez a módszer azonban nem minden esetben használható, hiszen általában összefüggő szöveget kell megvizsgálunk, és az ilyen szövegek bizonyos fajta szavakat több példányban tartalmazhatnak, amelyek egy-egy gondolat vagy esemény leírására szolgálnak; ezért, ha a fenti módszert használjuk, helytelen következtetésre jutnánk. A mintavételt tehát valamilyen rendszer szerint az egész szövegre ki kell terjeszteni. Elfogadhatunk például olyan megoldást is, hogy minden oldalon bizonyos sorokat veszünk figyelembe, vagy hogy a vakpróba alapján kiválasztott mintákban megvizsgáljuk az n első vagy utolsó sort.

A stílus statisztikai vizsgálatához azonban nem elég a számlálás. HERDAN olyan mutatót (paramétert) keres, amely a minta nagyságától független, és ezért egyértelműen felhasználható a stílus statisztikai jellemzésére. HERDAN paraméterének, melyet v_m mel jelöl, állandóságát matematikai módszerekkel is be lehet bizonyítani. HERDAN megállapításai szerint v_m állandó értéket mutat, bármekkora mintát vegyünk is ugyanannak a szerzőnek a műveiből. A kérdés most már az, hogy állandó-e a v_m egy adott korszakon belül is, hiszen írói stíluson túl korstílusról is beszélünk. A detroit-i egyetem

³ Vö. G. HERDAN, Chaucer's authorship of the Equatorie of the Planetis: Language 32 (1956), 254.

⁴ Vö. L. H. C. TIPPETT, Random Sampling Numbers, Vol. XV., Tracts for Computers, Cambridge, 1927. A vakpróba-számok tulajdonképpen véletlen számok. Ha például egy könyvből véletlenszerűen akarunk elolvasni néhány oldalt, akkor kezünkbe vesszük a Tippett-féle táblázatot és felütjük a könyvet a táblázatban megadott oldalakon.

szláv intézete szó-statisztikai vizsgálatokat végzett orosz szövegek alapján.⁵ A szövegeket 25 százalékban a XIX. századból, 25 százalékban az 1800—1918-as évekből, és 50 százalékban pedig 1918-tól napjainkig terjedő időszakból válogatták össze. Minden periódusra meghatározták v_m értékét. Úgy találták, hogy v_m még inkább jellemző egy egész korszakra, mint valamely író stílusára. Tekintettel arra, hogy a mintákat a legkülönbözőbb írók műveiből vették, a v_m stabilitása a közös szókincsnek köszönhető, amely 1830-tól napjainkig körülbelül változatlan maradt.

A v_m most már mint egyfajta „langue—parole”-viszony jelzője (indexe) fogható fel, vagyis v_m megmutatja, hogy a teljes szókincs (la langue) és az előfordulások (la parole) aránya egy eléggé tekintélyes korszakon belül állandó marad. Mivel az előfordulások számának megváltozásával szemben v_m invariáns, változás a v_m értékében így csak akkor következhet be, ha a szókincs (akár a részleges, akár a teljes) megváltozik. Ez utóbbit a vizsgálatok messzemenően igazolták. A változások a tapasztalatok szerint csökkenő tendenciát mutatnak. A nyelv bizonyos vonatkozásokban rokonságot mutat tehát a fizikai rendszerekkel és azok jellemzőjével, az entrópiával. Tudniillik ez a csökkenő tendencia más szóval azt jelenti, hogy a szavak használatának gyakorisága a legtöbb szó esetében egy középtérték köré csoportosul, a gyakoriságok az idő múlásával pedig mind kisebb variációs eltérést mutatnak. Amint a szógyakoriság relatív fluktuációja egy minimum felé közeledik, az „állapot” valószínűsége maximális lesz.

A Herdan-féle v_m paraméter YULE K-jellemzőjéhez hasonlít, és lényegében csak abban különbözik tőle, hogy nem függ attól a feltételtől, amelytől K függ. A K -ról YULE ugyanis feltételezte, hogy a szógyakoriságok Poisson-féle eloszlást mutatnak.⁶ A v_m egyetlen hiányossága, hogy a YULE által javasolt kontroll-lehetőségeket HERDAN nem használta ki a v_m használatánál, és így eredményei nem tekinthetők 100 százalékosaknak. HERDAN e hiányosságot részben kiküszöböli második könyvében (Type-Token Mathematics), ahol a stílus jellemzésére újabb paramétert talál.⁷

Nyelvszék körében felmerülhet az az ellenvetés — és ezzel HERDAN is tisztában van —, hogy bár ilyen paraméterek használata bizonyos kényelmet jelent, de ezek nem mondanak semmi újat és érdemlegeset a jelenség nyelvészeti oldaláról. HERDAN azonban utal arra, hogy ha olyan, a nyelvészetben egyébként jól ismert tény, mint hogy a szókészlet arányosan növekszik a szöveg hosszának növekedésével, kvantitatíve le tudunk írni, az jóval több mint kényelmi szempont. De a gyakorlatban is nagy jelentőségű két stílus összehasonlításánál az a tény, hogy ugyanarra az eredményre jutunk, ha statisztikai paraméterekkel dolgozunk, éppúgy mint amikor az eddig szokásos módszereket alkalmazzuk. Tulajdonképpen amikor a statisztikai módszereket alkalmazzuk a nyelvben, a következő hipotézisből indulunk ki: ha egy adott szöveg különböző részeiből veszünk mintákat, amelyek mindegyike elegendő szót tartalmaz, akkor a szavak gyakorisági eloszlása az előfordulások gyakoriságának megfelelően nem fog lényegesen különbözni egymástól a különböző minták esetén, még akkor sem, ha mintákat az egész szöveget vesszük. Más szóval ez azt jelenti, hogy bármely megfelelő nagyságú szövegrészt úgy tekinthetünk a szókészlet gyakorisági eloszlását illetően, mintha a teljes szövegből vakpróba alapján vettünk volna mintát. Ezért a szóelőfordulás statisztikai hipotézise nem azt igyekszik megmagyarázni, hogy miért használunk egy bizonyos szót, vagy miért használunk egy bizonyos grammatikai formát. Függetlenül attól, hogy valamely összefüggő szövegben milyen sajtószerű szavak fordulnak elő, vagy hogy azokat hogyan fűzik egybe a nyelvtani szabályoknak megfelelően, egy elegendően nagy szövegrész vakpróba alapján vett mintának tekinthető a szógyakoriság, a fonémák, a szóhosszúság szempontjából.

Természetesen már magában is fontos tény az, hogy bizonyos nyelvi formák statisztikai törvényeket követnek, amely semmi esetre sem „a priori” tény. Ha azonban a legkisebb nyelvi egységekből — a betűkből, fonémákból — indulunk kis, és így jutunk a nagyobb egységekhez: a morfémákhoz, szótagokhoz, mondatokhoz, akkor a tárgy

⁵ Vö. H. H. JOSSELSOON, The Russian Word Count (para. 5 by B. Epstein, Detroit 1953).

⁶ Annak feltételezése, hogy a szógyakoriság egy határozott valószínűségi törvényt követ, nem indokolható kellőképpen. Ezért HERDAN paraméterének matematikai levezetésében a fenti feltételt kizárja.

⁷ HERDAN egyik leghasználhatóbb statisztikai paramétere e tekintetben a logaritmus jelpéldány (type-token) arány, ahol a jel a nyelvi forma (fonéma, morféma stb.), a példány az illető jel előfordulási száma. Ez a paraméter a minta nagyságától független (legalábbis különböző nagyságú minták esetén az eltérések elhanyagolhatók), és ezért jól használható a stílus statisztikai vizsgálatára.

(tartalom) és az egyéni stílus egyre nagyobb hatást fog gyakorolni a statisztikai paraméterekre; a különbözőségek, a változások a statisztikai paraméterekben pedig igen fontosak lehetnek mind a tartalom, mind a stílus megváltozásában. Ez az oka annak, hogy a statisztikai nyelvészet, többek között mint a stílus statisztikája, oly fontos szerepet kap a nyelvészet különböző ágaiban (glottokronológia, stílusstatisztika, nyelvtörténet, gépi fordítás, nyelvtanítás stb.). Ha tehát fentiek figyelembevételével nem is becsülhetjük túl HERDAN eredményeit, annyi bizonyos, hogy bizonytalan becslések, szubjektív értékelések helyébe egyértelmű jellemzőket kapunk a matematika módszereinek felhasználásával.

HERDAN könyvének fontos részét képezik azok a vizsgálatok, amelyekben több szerző műveinek stilisztikai összehasonlítását végzi statisztikai módszerekkel. A v_m ilyen esetekben nem alkalmazható, bár kiszámításával kapunk valamilyen adatot a stílus egyhangúságát vagy változatosságát illetően, de tekintettel arra, hogy a v_m változik a szókészlettel, és két szerző szókészlete általában nem egyezik meg, összehasonlításra alkalmasabb módszert kell keresni. Két szerző összehasonlítása a közösen használt szavak szempontjából a korrelációs módszer segítségével válik lehetővé. A korrelációs számítás segítségével megállapítható különböző írók stílusának hasonlósága, illetőleg különbözősége. De ebben az esetben is a filológus feladata, hogy eldöntse, mennyiben tudja ezeket az adatokat felhasználni.

HERDAN könyvében a nyelv statisztikai modelljén kívül az információelméleti modellel is foglalkozik. Sok esetben csak kísérletekről, félkész megállapításokról beszélhetünk ugyan, de így is ízelítőt kaphatunk a statisztikai modell lényegéről és a matematikai statisztika, korrelációs számítás, információelmélet különböző nyelvészeti alkalmazási lehetőségeiről. A könyv áttanulmányozása után (minden kétséget kizáróan) megbizonyosodik előtűnk, hogy itt egyrészt a nyelv egy új koncepciójáról van szó, másrészt arról, hogy részleteiben is milyen nagy mértékben segíthetik az említett matematikai diszciplínák a filológust kutatásaiban. Sok esetben ugyan HERDAN eredményeinek nyelvészeti interpretációja nehézségekbe ütközik, talán nem is minden állja meg a helyét megállapításai közt, könyve minden kétséget kizáróan igen fontos mű, és elsősorban azok számára az, akik a matematikai nyelvészet különböző területeibe szeretnének bepillantást nyerni.

KIEFFER FERENC

Sprachwissenschaftliches Wörterbuch

Hrsg. von Johann Knobloch. Heidelberg, 1961. Indogermanische Bibliothek. II. Reihe. Wörterbücher. Lief. 1. 80 l.

A nyelvtudományi terminológiának J. MAROUZEAU ismert könyvénel (Lexique de la terminologie linguistique. 3^e éd. Paris, 1951) részletesebb, pontosabb összefoglalása a mai nyelvészetnek egyik legsürgősebb feladata. Éppen a terminológia tisztázatlansága miatt már van olyan nyelvész, mint J. R. FIRTH, aki — a jelen munkának sajnos nagyon szűkre szabott előszava szerint is — valamiféle „increasing Babel in linguistics” veszélyére figyelmeztet. E feladattal most J. KNOBLOCH, az innsbrucki egyetem professzora próbál megbirkózni, s munkájában olyan jeles erők támogatják, mint W. PORZIG, M. REGULA, H. SEILER, L. WEISBERGER és W. WISSMANN. A munkának egyelőre nincs azonban semmiféle bibliográfiája, s ezért nehéz megállapítanunk, mennyiben lesz az új terminológiai szótár valóban nemzetközi jellegű, mennyiben öleli fel a nyelvtudomány tárgya és nyelv szerint annyira különböző területeit. Mindezen kérdések persze súlyosan vetődnek fel történeti vonatkozásban is. Annyi máris feltűnő, hogy például az orosz, illetve a szovjet nyelvtudománynak, sőt általában a szlavisztikának egyelőre nagyon kevés nyoma akad az új összeállításban,¹ amelynek egyébként nagy előnye, hogy rendszeresen idézi a ritkább műszavak lelőhelyét. *Abbreviatursprache* alatt Wundt, *Abgeschlossenheit des Satzes* alatt E. Lerch, *Abglitt* alatt A. J. Ellis, E. Sievers és M. Grammont nevét találjuk; A. J. Ellis ó-angol fonetikájának (1869) említése talán megközelítő korhatárt jelent a feldolgozott művek felső korhatára szempontjából is. Nagyjából tehát az utolsó 100 évnek elsősorban német nyelvészeti bibliográfiáját

¹Vö. azonban az *Ablautreihe* cikkel, ahol utalás van Leskien ó-bolgár nyelvtanára. *Aktionsart* alatt viszont szláv anyag egyáltalában nincs Fr. Celakovsky ismert tanulmányán kívül (1853). *Akambe* alatt egyetlen orosz szerző neve sem szerepel; az új orosz akadémiai szótárban adatok Dal' óta.

várhatja KNOBLOCHTÓL az olvasó. Ha itt-ott régebbi korokra történik utalás (vö. *Ableitung* [seit Schottel 1641], akkor is elsősorban a német nyelvészet múltjáról van szó. Az ókori grammatikai hagyomány felé csak szórványosan történik kitekintés (vö. ugyan-csak *Ableitung* Dionysios Thrax, Platon és Apollonios Dyskolos említésével).

A kezünkben levő I. füzet 5 iven a (phonetische) *Akustik* szócikkig jut el (sajnos, az utóbbinak szövege mindössze 4 sor!). Ha a címszóanyagot a Marouzeauval összevetjük, a haladás kétségtelenül ugrásszerűnek tűnik: Marouzeaunál még hiányzott — hogy csak néhány szót említsünk — *a mobile, a-Mode, Abbildetendenzen, Abengung, Abessiv(um), Abessiv(us), Abfolge der Lauten, Abgrenzung, Abhängigkeitskompositum* stb. Mindezzel MAROUZEAU a megvizsgált részben csupán 1—2 „pluszt” tud szembeállítani: ilyen például *Abhängig*, illetve *Abhängiger Casus* (= cas oblique). A felvett anyag tehát kétségtelenül gazdagabb és változatosabb. A finnugor nyelvészek és a kaukázológusok annak is örvendhetnek, hogy Knoblochnál megvannak olyan nem indoeurópai esetnevek, mint pl. a Marouzeau-tól kihagyott *Adessiv(us)*.

További jelentős mozzanat a legfontosabb szócikkek tartalmi bővülése. Marouzeau-nál az *Ablaut, Apophonie* és *Alternance phonétique* cikkek együttesen sem tesznek ki annyit, mint Knoblochnál az *Ablaut* cikk, a szó összetételeivel együtt.² Azt persze kevésbé helyeselhetjük, hogy az *Ablaut* fogalma váratlanul a finnugor fokváltakozással keveredik (12); a két jelenséget másutt — pl. az *Alternationen* szócikkben — jobban lehetett volna rokonítani. Gombocz Zoltán az *Ablauthoz* hasonló jelenségnek csak az olyan *a ~ á ~ o* váltakozást tartotta, aminőt az *al/szik ~ ál/om ~ ol/i* esetében tapasztalunk. Még alaposabb és mélyrehatóbb a gyarapodás az *Adjektivum* cikkben, ahol Knobloch szemantikai osztályozása sokkal tagoltabb, mint a konvenciókhoz jobban ragaszkodó Marouzeau-é. Meglepő azonban a melléknévnek mint szófajnak a nem fogalmához való kötése több szerző, elsősorban G. Scarpat nyomán.³ Említést érdemelt volna, legalább utalás formájában, a melléknév helyének kérdése is a jelzős szintagmában.

Problematikusknak látjuk némely szócikk túlságos széttagoltságát; az *Akkusativ* „főszócikk” után tipográfiailag egyenrangú részként következik *Akkusativ, absoluter; Akkusativ, adnominaler*, stb. E vitatható tagolás ellenére örvendünk azonban annak, hogy Knoblochnál megvan többek közt a latinból is jól ismert *Accusativus interiectionis* (vagy *exclamationis*, pl. Vergiliusnál: „O fortunatos nimium . . . agricolas”), amelyről korábban Marouzeau, klasszika-filológus létére, legalábbis Lexique-jában megfeledezett.

A szótár további füzetait nagy érdeklődéssel várjuk; fontos pozitív vonásai joggal fokozzák fel vele szemben támasztott igényünket.

GÁLDI LÁSZLÓ

Ф. П. Филин: Проект «Словаря русских народных говоров»

Москва—Ленинград, 1961. Издательство Академии Наук СССР. 199 I.

Az első — és ez ideig egyetlen — általános orosz tájszótárt 1852-ben adta ki az Orosz Tudományos Akadémia (Опыт областного великорусского словаря, изданный Вторым отделением Имп. Академии Наук. СПб. 1852.). A hat évvel később megjelent pótkötettel együtt (Дополнение к Опыт областного великорусского словаря. Издание Второго отделения Имп. Академии Наук. СПб. 1858.) e szótár címszóanyaga megközelítette a negyvenezretet. Tulajdonképpen azért készült, hogy az orosz nyelvjárások szókincsének bemutatásával kiegészítse az orosz irodalmi nyelvnek 1847-ben közzétett akadémiai szótárát (Словарь церковно-славянского и русского языка, составленный Вторым отделением Имп. Академии Наук. I—IV. СПб. 1847.). Számos fogyatékosága ellenére jelentős szerepe volt az orosz nyelvjárások szókészlete iránti érdeklődés fokozásában és a gyűjtőmunka élénkebbi tételében. A mai kutató azonban a legtöbb esetben könnyen nélkülözheti munkájában. Az 1852—1858. évi tájszótár címszavainak jó háromnegyed része megtalálható ugyanis V. I. Dal' szótárában, amely a XIX. századi orosz nyelv szókincsének eddig legteljesebb gyűjteménye В. И. ДАЛЬ, Толковый

² E ponton KNOBLOCH már támaszkodhatott E. ISING tudománytörténeti alapvetésére: Die Begriffe 'Umlaut' und 'Ablaut' in der Terminologie der frühen deutschsprachigen Grammatik. Sb Berlin, 1955, 3.

³ „G. Scarpat, in АНТИΩΡΩΝ Hugoni Henrico Pauli oblatum, Genua 1956 erkennt das wesentliche Merkmal des Adjektivs darin, dass es kein eigenes Genus besitzt, wie jedes Substantivum, sondern in dieser Hinsicht dem determinierten Substantivum übereingestimmt ist” (38).

словарь живого великорусского языка. I—IV. Москва 1863—1866. СПб. — Москва 1880—1882, 1903—1909, 1912—1914.). Dal' szótárának tájszavai között nagy számmal vannak olyanok is, melyek Dal'-nak vagy másoknak a saját gyűjtéséből származnak és ekkor kerültek először szótározásra. Minden szempontból a legmegbízhatóbb és a tájszavak tekintetében is a leggazdagabb e szótárnak a III. és IV. kiadása, melyet J. Baudouin de Courtenay szerkesztett. A benne feldolgozott szavak száma a 220 000 körül jár. Ennek a hatalmas szómennyiségnek mintegy fele népnyelvi jellegű (tájszavak, mesterségszavak, a városi népnyelv szavai stb.). DAL' szótárát számos kisebb-nagyobb regionális tájszótár követte. Majdnem mindegyik egy-egy északorosz nyelvjárásnak a tájszóanyagát tartalmazza. A legjelentősebb északorosz tájszótárak sorában meg kell említeni A. I. PODVYOCKIJ arhangelszki (1885.), P. I. JAKUSKIN jaroszlavlai (1896.), G. I. KULIKOVSKIJ olonyeci (1898.), I. T. SMIRNOV kasini (1901.), V. VOLOCKOJ rosztovi (1902.), A. GRANDILEVSKIJ holmgori (1907.), N. M. VASNECOV vjatkai (1907.), M. K. GERASIMOV cserepoveci (1910.) szótárát. A délorosz nyelvjárások tájszótárai közül A. V. MIRTOV doni szótára (1929.) a legszámottevőbb. A fehér-orosz nyelvterülettel érintkező orosz nyelvjárások szempontjából fontos V. N. DOBROVOL'SKIJ szmolenszki szótára (1914.). A kaleidoszkópszerűen változatos szibériai orosz nyelvjárásoknak a szókincsébe, melyeknek a legrégebb rétege XVI—XVIII. századi északorosz telepésekre vall, PATKANOV és ZOBININ tobolszki szójegyzéke (1899.), valamint V. G. BOGORAZ kolimai tájszótára (1901.) nyújthat betekintést. Ha az orosz regionális tájszótáraknak ehhez a korántsem teljes felsorolásához hozzávesszük a folyóiratok, évkönyvek stb. közleményeiben szétszórt óriási tájszóanyagot és a kéziratban fekvő gyűjtéseket, elképzelhetjük, milyen fáradságos munkát kell végeznie annak a kutatónak, aki az orosz nyelvjárások szókincsével foglalkozik. A külföldi kutatóknak még a nyomtatásban megjelent nagyobb szótárak sem állnak mindig rendelkezésére, pedig finnugrisztikai, turkológiai stb. célokból is égető szükség volna rájuk. Mindenképpen megérett tehát a helyzet arra, hogy új általános orosz tájszótár foglalja össze és tegye hozzáférhetővé a XIX. és XX. századi orosz tájszótárak és kéziratok gyűjtéseinek anyagát.

Ez lesz a feladata annak az új orosz tájszótárnak, amelynek elkészítése a Szovjet Tudományos Akadémia Orosz Nyelvtudományi Intézetének a tervmunkái között szerepel. A vállalkozás részletes Tervezetét F. P. FILIN, a kiváló leningrádi dialektológus és nyelvtörténész dolgozta ki. A Tervezet tüzetesen tárgyalja az új szótár feladatát, címszóanyagát, kronológiai kereteit, forrásait, a kartotékcédulák megírásának és rendezésének gyakorlati kérdéseit, továbbá a szócikkek összeállításának lexikográfiai elveit. 41 lapnyi terjedelemben próbaszócikkeket is közöl.

A Tervezet szerzője szerint az új orosz tájszótárnak a lehetőséghez képest teljes képet kell adnia a XIX. és XX. századi orosz tájszókincséről, és szervesen kell illeszkednie ahhoz a nagyszabású lexikográfiai komplexumhoz, amelyet a készülő vagy tervbe vett orosz nyelvtörténeti szótárak alkotnak: 1. a XI—XIV. századi órosz nyelv szótára; 2. a XV—XVII. századi orosz nyelv szótára; 3. a XVIII. századi orosz nyelv szótára; 4. a XIX—XX. századi orosz irodalmi nyelv történeti szótára. Összegeznie kell a következő fontosabb forráscsoportokat: 1. az Orosz Nyelvtudományi Intézet lexikális archívumában található, mintegy 150—200 ezerre becsült tájszóanyagot; 2. a nyomtatásban közzétett tájszótárak, tájszójegyzékek stb. anyagát; 3. a szerkesztőség számára elérhető helyen őrzött kéziratok tájszógyűjtések anyagát; 4. a dialektológiai tárgyi kandidátusi disszertációkhoz csatolt szótári mellékletek anyagát. Helyszíni gyűjtést a szótár szerkesztősége nem szándékozik végezni. A szótárban felölelendő tájszóanyag alsó időhatárát 1800-ban állapítja meg a Tervezet. Ezt az időpontot az teszi indokolttá, hogy a tájnyelvi gyűjtések csupán a XIX. század elején kezdtek rendszeresebbé válni. A szótár felső időhatára a szerkesztésnek, illetőleg a megjelenésnek az idejével lesz azonos. Elvileg valamennyi orosz nyelvjárás tájszóanyagának bele kell kerülnie a szótárba, de bizonyos egyenetlenségek előreláthatók. Az északorosz nyelvjárások tájszóanyaga sokkal alaposabban fel van gyűjtve, mint a délorosz nyelvjárásoké. A legkevésbé a szibériai és a távolkeleti orosz nyelvjárások szókincse ismeretes.

Az új orosz tájszótár differenciális dialektológiai szótár lesz, tehát olyan, amelyik csupán a tájszavakat, azaz a nyelvjárásoknak az irodalmi és köznyelv szókincsétől eltérő lexikális elemeit veszi figyelembe. Minthogy pedig a tájszónak a nyelvészeti irodalomban is szokásos, közkeletű meghatározása — „olyan szó, amely csak egy vagy több nyelvjárásban használatos” — annyira általános, hogy a lexikográfiai gyakorlatban félreértéseket, zavarokat okozhat, a tájszó fogalmának új, szabatos és egyértelmű definíciójára van a szerkesztőségnek szüksége. F. P. FILIN felfogása szerint tájszóról akkor beszélhetünk, amikor az illető szó az alábbi két kritérium mindegyikével ren-

delkezik 1. a nyelvterületen belül izoglosszája van, azaz a használata területileg körülhatárolható; 2. nem tartozik bele a mai irodalmi és köznyelv szókészletébe. Ezt az utóbbi megszorítást F. P. FILIN azért tartja szükségesnek, mert az egyes nyelvjárásoknak nem kevés olyan, sajátos szavuk van, amelyek az idők folyamán utat találtak az irodalmi és köznyelv szókincsébe is. Az irodalmi nyelvben meghonosodott *лѣтъ* 'ugat' például bizonyos nyelvjárásokban él csupán; sokfelé a *брѣхѣтъ* 'ua.' igét használják helyette. Hasonló a helyzet az irodalmi nyelv *ржѣтъ* 'nyerít' (vö. nyj. *гогопѣтъ* 'ua. '), *бѣчь* 'nagyon' (vö. nyj. *дѣже, горѣдо, порѣто* 'ua. '), *навѣз* 'trágya' (vö. nyj. *пожѣм, нажѣм, ожѣм* 'ua. ') stb. szavaival is. A *лѣтъ* típusú szavak ki fognak maradni a tervezett orosz tájszótárból. Ez feltétlenül sajnálatos abból a szempontból, hogy a szótárból nem lehet majd választ kapni arra a rendkívül érdekes kérdésre, milyen arányban vett részt ennek vagy annak a nyelvjárásnak a szókincse az orosz nemzeti nyelv legfontosabb és legértékesebb változatának, az irodalmi nyelvnek a kialakításában. Be kell azonban látni, hogy az izoglosszával rendelkező irodalmi és köznyelvi szavak felvétele a szótárba nagymértékben megnehezítené a szerkesztők munkáját és szerfölött megnövelné a szótár terjedelmét. Az efféle „tegnapi tájszavak” izoglosszáinak megállapításához nyelvatlaszt kell használni: hátha szerepeltek a kérdőívben.

A tájszavakat F. P. FILIN a következőképpen osztályozza: I. az irodalmi és köznyelv szavaitól morfológiai és fonetikai felépítésükben, azaz anyagi megformálásukban eltérő szavak, melyeket lexikális tájszavaknak lehetne nevezni; II. a megfelelő irodalmi és köznyelvi szavaktól jelentésükben eltérő szavak, melyek szemantikai tájszavaknak nevezhetők. Ez az osztályozás csupán annyiban különbözik a magyar nyelvtudományban SZINNYEI JÓZSEF Magyar Tájszótára óta alkalmazott hármass felosztástól — 1. tulajdonképpeni tájszavak; 2. jelentésbeli tájszavak; 3. alak szerinti tájszavak —, hogy a lexikális tájszó elnevezés alá vonja mindazt, amit mi „tulajdonképpeni tájszó” és „alak szerinti tájszó” névvel szoktunk jelölni. A lexikális tájszavak kategóriáját F. P. Filin így tagolja tovább: 1. olyan töveket tartalmazó szavak, amelyek hiányzanak az irodalmi nyelvből (pl. *вѣлна* 'gyarjú', *лѣни* 'tavaly', *кѣдлѣй* 'csenevész, satnya'); 2. az irodalmi nyelvben is meglevő tövekből alakult, különleges jelentésű szavak (pl. *сузѣм* 'mocsaras erdőség', *нѣсерѣка* 'harag, rosszindulat', *рѣстѣтся* '〈tyúk〉 tojni készül'); 3. olyan szavak, melyeknek a tövük és jelentésük egyezik az irodalmi nyelv megfelelő szavaival, az affixumuk vagy affixumaik viszont mások (pl. *чернѣца, чернѣга* 'fekete áfonya', vö. *чернѣка* 'ua. '); *зѣгородѣ, зѣгородѣка, изгорѣдина* 'sövény', vö. *изгородѣ* 'ua. '); 4. olyan szavak, amelyek szabályokba nem foglalható, fonematikus jellegű hangtani és alaktani eltéréseket mutatnak az irodalmi nyelv azonos tövű, affixumú és jelentésű megfelelőihez képest (pl. *иржѣ* 'rozsa', vö. *ржѣ* 'ua. '); *ловѣнѣй* 'ön-, önből készült', vö. *оловѣнный* 'ua. '); *пѣхмурный* 'borús', vö. *пѣсмурный* 'ua. ').

Az új orosz tájszótár címszavai között az eredeti orosz szavak mellett a nyelvjárások jövevényszavai szintén helyet kapnak, még akkor is, ha csupán szűk területen élnek és idegen nyelvi származásuk nyilvánvaló. Vonatkozik ez a nemrégiben eloroszosodott etnikai csoportok tájszavaira is. Ilyenek például az uráli Pelim (Пѣлым) folyó mentén beszélt nyelvjárásoknak a vogul jövevényszavai, amelyeket A. K. Matvejev jegyzett föl 1955—1957-ben: *вѣт* 'áradáskor víz alá kerülő rét', vö. vogul *оѣт, оѣм* 'ua. '; *шош, сош* 'patak', vö. vogul *сѣс* 'patak, nedves horhos'; *лѣпча* 'erdei folyóba zuhant fákból keletkezett torlasz', vö. vogul *лѣпци* 'ua. '; *ѣбхѣла* 'napon szárított sózott hal', vö. vogul *ѣхѣлы* 'ua. ' (Филин, Проект 139).

Tulajdonnevek a tervezett szótárban nem lehetnek címszavak. Az egyes szócikk végén azonban kiegészítésképpen fel fogják sorolni azokat a helyneveket, amelyek a tárgyaltnál köznévből alakultak. A szócikk végén szükség esetén történeti, néprajzi és egyéb felvilágosítás, jövevényszavak esetében pedig etimológiai magyarázat is lesz.

A tervezett szótár szócikkeinek a felépítése nagyjából az újabb orosz értelmező szótárakéra emlékeztet azzal a különbséggel, hogy a címszó egyes jelentéseit a földrajzi elterjedés megadása követi, feltüntetve a forrást és évszámot is. A szóhasználatot a forrásokból vett példamondatok szemléltethetik. A szócikkékből megtudhatók a címszavak nyelvtani, hangsúlyozási stb. sajátosságai, valamint a velük alkotható állandósult szókapcsolatok.

A gondosan kidolgozott próbaszócikk, melyek között 5—8 lapnyi terjedelműek is vannak, arról győzik meg az olvasót, hogy az új orosz tájszótár ügye tudományos felelősségérzettel áthatott és magas színvonalbeli követelményeket támasztó szakemberek kezében van. A *нѣдѣ* 'hajófajta' szócikke a magyar szókincs szláv elemeinek a vizsgálatához is becses adalékokat szolgáltat. A homonimák szétválasztására azonban néhol jobban kell ügyelni. A sokjelentésű *лѣда* 'irtványföld; stb.' 11—15. jelentéseként

besorolt *лѧда* 'padlás- vagy pinceajtó', *лѧды* 'spaletta, ablaktábla', *лѧда* 'tolózár, retesz', *лѧда* 'bolti pult', *лѧда* 'a szövőszék alkatrésze' etimológiailag nem az ősi szláv *лѧда* 'irtványföld' szóval, hanem a következő, német eredetű szavakkal függ össze: ukrán *лѧда* 'bolti pult felhajtható táblája; nyílás-záró fedő; padlás- vagy pinceajtó; a szövőszék alkatrésze; stb.', lengyel *lada* 'láda, pult; stb.', román *ladă* 'láda, fiók', magyar *láda* stb. (vö. BERNEKER, SLEtWb. I, 682). Ugyancsak nem a *лѧда* 'irtványföld' szóhoz tartozik a 16. jelentésként nyilvántartott *лѧда* 'lusta ember' sem. Minden bizonnyal annak az ukrán *лѧда* 'lusta ember' szónak az átvételével van itt dolgunk, amelyik a 'lusta ember; semmirekellő' jelentésű *лѧдѧй*, *лѧдѧка*, *лѧдарь* főnévvaltozatok valamelyikéből vonódhatott el (vö. *Преображенский*, ЭтСлов. I, 463; VASMER, REW. II, 25). A *лѧда* szócikke tehát végeredményben három, önálló szócikkbe kívánczozó homonimára bontandó föl.

A tájszótárak szerkesztésének elméleti és gyakorlati kérdéseivel foglalkozó szakirodalom sem nálunk, sem külföldön nem mondható különösen gazdagnak. Már csak ezért is örömmel fogadjuk F. P. FILINnek a Tervezetét, amely sok elvi jelentőségű kérdést tárgyal és számos gyakorlati útbaigazítást is nyújt. A legmélyebb benyomást azonban az teszi a Tervezet olvasójára, hogy az előkészítő munkálatok bonyolult részlet-problémáival való megismerkedés során egy óriási területen beszélt nyelv impozáns méretű és nagy tudományos értékű tájszótárának a körvonalai bontakoznak ki előtte, egy olyan régen várt szótár, amely most már belátható időn belül meg fog valósulni. Őszintén kívánjuk, hogy az új orosz tájszótár minél előbb napvilágot lásson.

KISS LAJOS

Historia osadnictwa i organizacja społeczna pasterstwa oraz słownictwo pasterskie Tatr Polskich i Podhala

Opracowali S. Górzynski, J. Kolowca, B. Kopezyńska-Jaworska, Z. Korosadowicz, Z. Radwańska-Paryska, A. Szyfer, J. Zborowski, Z. i S. Zwolińscy.

Wrocław—Warszawa—Kraków, 1962. Zakład narodowy im. Ossolińskich, wydawnictwo Polskiej Akademii Nauk. — Polska Akademia Nauk. Wydział I — nauk społecznych. Pasterstwo Tatr Polskich i Podhala. Redaktor Włodzimierz Antoniewicz. Tom IV. 272 l.

A lengyel Táttra és a tőle északra, Nowy Targ körül elterülő medence, Podhale (tkp. 'a havasi legelők alja', vö. lengyel *hala* 'havasi legelő') néprajzi és nyelvészeti tekintetben egyaránt számos sajátossággal tűnik ki. Lakóinak, a csekély lélekszámú lengyel guráloknak (vö. lengyel *góral*, tkp. 'hegyi lakó') olyannyira jellegzetes, egyéni areulaturak van a lengvelség etnikai csoportjai között, mint — mondjuk — a székelyeknek a magyarság néprajzi csoportjai között. Mindenképpen érthető tehát az a fokozott figyelem, amely a táttrai gurálok iránt nemcsak a lengyel közönség, hanem a lengyel művészet és tudomány részéről is megnyilvánul. A lengyel irodalomnak nem egy jeles képviselője írt gurál tárgyú műveket, a lengyel tudományosság pedig régóta foglalkoztatja a gurálok története és jelenkori élete. A hegyvidékek és lakóinak a tanulmányozásához önálló folyóirat (Wierchy. Kraków, 1921—.) áll a kutatók rendelkezésére. Nemrégiben a táttrai és podhalei pásztorkodásnak a tüzetes feldolgozására indított vállalkozást a Lengyel Tudományos Akadémia I. Osztálya. A 10 kötetes tervezett sorozat részmonográfiák alakjában mutatja be azt a havasi pásztorkodást, amely jó ideje tartó általános hanyatlás után napjainkban végképp megszűnik. A nyájakkal nem kevés erdei kárt okozó havasi pásztorkodás és az erdők megóvását szorgalmazó erdészeknek a szakadatlan viszálykodása ugyanis a Táttrai Nemzeti Park létrehozásával lezárult: a természetvédelmi területen tilos a legeltetés. Ezért hát a kutatónak sietnie kell. A táttrai pásztorkodás fölött már megkondult a lélekarhang. Eltűnőben és átalakulóban van az az évszázados táttrai állattartási mód, amely szoros magyar kapcsolatai révén a magyar szakemberek érdeklődésére is számot tarthatott (vö. KNEIENZA ISTVÁN, A tót és lengyel költőzködő pásztorkodás magyar kapcsolatai. Ethn. XLV, 62—73).

A nagyszabású sorozat első kötetei a táttrai és podhalei pásztorkodásnak néhány földrajzi és néprajzi problémáját tárgyalják. A kötetek címe a következő: I. Fizjografia i geografia pasterstwa Tatr Polskich i Podhala. 1959. — II. Pastwiska podgórskie i górskie Tatr Polskich i Podhala. Terazniejszość i przyszłość. 1960. — III. Hodowla owiec i bydła w Tatrach Polskich i na Podhalu. Terazniejszość i przyszłość. 1961. —

A legutóbb megjelent és most ismertető IV. kötet történeti, néprajzi és nyelvészeti tárgykörű tanulmányokat foglal magában. Minden tanulmány végén francia nyelvű összefoglalás van.

Folyóiratunk jellegének megfelelően csupán a két utolsó, nyelvészeti tárgyú dolgozathoz fűzök egypár megjegyzést, kiemelve a magyar vonatkozásokat. ANNA SZYFER tanulmánya a tátrai és podhalei pástorkodás szókészletének nyelvföldrajzi módszerű leírását adja. A 75 kérdésből álló kérdőívére kapott feleleteket a szerző szöveges formában és — részben — térképre vetítve tárja az olvasók elé. Az összegyűjtött anyagot etimológiai szempontból is rendezti. Feltűnő, hogy szlovák eredetű réteg teljesen hiányzik ebből az összeállításból, holott nyilvánvaló, hogy a szerzőtől cseh eredetűnek tartott szavak tulajdonképpen nem a csehből, hanem a szlovákból valók. Az etimológiai elemzés a dolgozatnak egyébként is gyenge oldala. A szerző egyáltalán nem használta — többek között — V. Machek cseh és szlovák etimológiai szótárát, pedig a *redyk* 'tavaszi nyájkijajtás a havasi legelőre', *kierdel* 'a havasi legelőre hajtott nyáj' és számos más szó magyarázatához érdemes lett volna. A dolgozatban a következő szavak fordulnak elő mint magyar származásúak, illetőleg mint magyar alapszóból képzettek: *baca* 'juhászgazda, bacsó', *juhás* 'juhász', *juhásic* 'juhot legeltet, juhászkodik', *bojtár* 'bojtár', *szalas* 'szállás', *buňkos* 'bunkós pástorkot'. A *baca* azonban minden bizonnyal nem a magyarból, hanem a szlovák *bača*-ból való (KNEZSA: *Studia Slavica* VIII, 435). Ugyancsak a szlovákból származik a szerző által ismeretlen eredetűnek minősített *habarka* 'a megoltott tej szétkavarására használatos eszköz', maga a szlovák *habarka* 'kavarófa' viszont egy magyar *habaró*-ból alakulhatott (MACHEK, *EtSl.* 119).

Z. RADWAŃSKA-PARYSKA tanulmánya a Táttra és Podhale gurál növényneveinek a szótárát tartalmazza. Az 508 növénynevből 184 a szerző saját helyszíni gyűjtéséből, a többi pedig mások gyűjtéséből és a korábbi szakirodalomból származik. A legrégebbi adatokat a Vörös Kolostorban élt Ciprián barátunk az 1760-as évek táján készült herbáriumában olvasható jegyzetek szolgáltatják. A magyar vonatkozású gurál növénynevek közül az 'áfonya' jelentésű *jafer*-re szeretném felhívni a figyelmet. E név úgy függhet össze a magyar *áfonya*-val, hogy végső forrásuk azonos. (A kérdés újabb irodalmából lásd MACHEK, *EtSl.* 120; E. BELCARZOWA: *Język Polski* XL, 284; CIORANESCU, *DiccEt-Rum.* 10). A gurál *nedregula* 'nadrágulya' a szerző szerint a szepesi német *Nadragulle* közvetítésével a magyar *nadrágulya*-ra megy vissza. Nem lehet azonban számitáson kívül hagyni a keleti szlovák *nadragul'a* közvetítő szerepét sem.

A fényképekkel és térképvázlatokkal gazdagon illusztrált, reprezentatív kiállítású sorozat a lengyel Táttra és Podhale komplex módszerrel végzett kutatásának az eredményeit összegezi. Köteteit számos tudományág művelői fogják haszonnal forgatni. A nyelvész számára a IV. kötet nyújtja a legbővebb anyagot, de értékes adalékok a többi kötetben is találhatóak.

KISS LAJOS

Jaroslav Jelínek—Josef V. Bečka—Marie Těšitelová: *Frekvence slov, slovních druhů a tvarů v českém jazyce*

Praha, 1961. Státní pedagogické nakladatelství. 587 l.

Az orosz nyelv gyakoriságszótára¹ után újabb szláv nyelv szókincsének ilyenfajta feldolgozását üdvözölhetik örömmel a szakemberek. A cseh nyelv gyakoriság-szótárának munkálatait 1940-ben kezdték meg, s 1953-ban fejezték be. A szótár beilleik azon nyelvészeti munkák sorába, melyek a „matematikai nyelvészet” címszó alá tartoznak, s amely matematikai nyelvészetet így vagy úgy, de általában a strukturalizmus fogalmával kapcsolnak egybe. (Bár a matematikai, közelebből statisztikai módszerek alkalmazása nyelvészeti munkákban ismeretes volt már a strukturalizmus megjelenése előtt is.²) Az utóbbi években különösen a lexikológiában, de a nyelv más síkjain is egyre inkább

¹ JOSSELSON, HARRY H., *The Russian word count*. Detroit, 1953.

² Андреев, Н. Д.—Зиндер, Л. Д., *Основные проблемы прикладной лингвистики* В. Я. 1959/4. 19. l.

érvényesítik a matematikai módszereket.³ A jelen munka kiadása is összefüggésben áll ezzel a tendenciával, s arra mutat, hogy a prágai strukturalisták (funkcionális nyelvészek) feltámadásának 1957-es deklarációját⁴ nagy jelentőségű munkák is követik. Éppen ebben a cikkben olvashatjuk, hogy „для полного познания языковой действительности следует сочетать качественный анализ элементов языка с количественным (статистическим) анализом.” (52. l.).

A cseh gyakoriság-szótár első fejezete (A) ismerteti a munka megindulásának előzményeit és a munka lefolyását, majd tömör értékes összefoglalót ad az eddig megjelent gyakoriság-szótárakról (M. TĚŠITELOVÁ). JAR. JELÍNEK leírja azokat az elveket, melyek a szójegyzék összeállítását irányították. Összesen 75 műből 1 623 527 szót dolgoztak fel. A feldolgozott műveket 8 csoportból merítették: A) széppróza, B) költészet, C) ifjúsági irodalom, D) dráma, E) szakirodalom, F) újság, G) tudományos irodalom, H) beszédek. Két szójegyzéket is közöl a szótár. Az első jegyzék (99—159. l.) a leggyakoribb 10 000 szót tartalmazza, gyakorisági sorrendben. A második jegyzék (163—583. l.) a 3 vagy 3-nál nagyobb gyakoriságú szavak betűrendes jegyzéke. Bemutatok egy szót a két jegyzékben:

co 7907—8 — 75

Az első szám azt jelenti, hányszor fordult elő összesen a szó. A második, hogy hány műfaji csoportban szerepelt, s a harmadik, hogy hány műben. A betűrendes jegyzékben még bővebb az egyes szavak adatolása:

		A	B	C	D
co	7907 — 8 — 75:	2889/15	381/10	1375/10	1571/10
		E	F	G	H
		343/9	235/7	553/10	554/4

Az újabb számok jelentése: hányszor fordult elő a szó az egyes műfaj-csoportokban, és ezeken belül hány műben volt fellelhető.

A szótár második (B) fejezete tartalmazza az itt elmondottakat, és ugyanakkor foglalkozik a szó megállapított gyakoriságának értékelésével. Helyesen állapítja meg a fejezet szerzője, JAR. JELÍNEK, hogy az értékelésnél mind a három számadatot figyelembe kell venni. Azonos gyakoriságú szónál annak lesz magasabb a besorolása, amelyik több műven szerepelt (3. szám), s ha ez a szám is azonos, akkor a 2. szám dönt: az a szó a magasabb rangsorú, amely több műfaj-csoportban szerepelt. A harmadik (C) fejezet (JAR. JELÍNEK) az ún. szóismétlési indexet és az egyes szófajok gyakoriságát közli. A szóismétlési index a felhasznált összes szó (jel) és a felhasznált különböző szavak (típus) hányadosa. Ez a szám tehát bizonyos mértékig jelzi a mű nyelvének gazdagságát. A szótár közli az egyes műfaj-csoportok, majd ezeken belül az egyes művek szóismétlési indexét (60—64. l.). JAR. JELÍNEK érdekes statisztikát közöl az egyes szófajokról, azoknak jel-gyakoriságáról, típus-gyakoriságáról és ismétlési indexéről az egyes műfaj-

³ Az ismertetett könyv végén gazdag irodalom-jegyzék van, amely tartalmazza egyrészt a szótár elkészítéséhez felhasznált anyagot, másrészt a matematikai módszer alkalmazását hozó legfontosabb irodalmat. Ezenkívül megemlítjük még a következő cikkeket, amelyek a legújabb eredményeket ismertetik és egyúttal új bibliográfiát is közölnek: a SaS 1959/1-ben megjelent két nagyszerű cikk: PETR SGALL, Nové otázky matematických metod v jazykovědě. 44—55. l. és JIŘÍ KRÁMSKÝ, Teorie sdělné promluvy 55—66. l. Ugyancsak e folyóirat hasábjain jelent meg JITKA ŠTINDLOVÁ két cikke: Stroje na zpracování informací a jejich význam pro jazykovědu. SaS 1961/3. 208—215. l. és Uplatňování metod mechanizace a automatizace v lexikologické a lexikografické práci v zahraničí. SaS 1962/1. 67—71. l. Az utóbbi főként az Olaszországban és Franciaországban elektronikus számológép segítségével folyó munkálatokat ismerteti. A szovjet nyelvészek e kérdéssel foglalkozó cikkei közül a 2. alatt említetten kívül mindenképp fel kell hívni a figyelmet P. M. Фрумчина cikkeire: Применение статистических методов в языкознании ВЯ. 1960/4. 129—133. l. és Статистическая структура лексики Пушкина. ВЯ. 1960/3. 78—81. l. A matematikai s ezen belül a statisztikai módszert ismertető magyar cikkek közül a következőket említem meg: FODOR ISTVÁN, A statisztikai módszer alkalmazásának néhány kérdése: Nyr. 84: 196—205. l.; PAPP FERENC, Új irányzatok a szovjet nyelvtudományban: NyK. LXI, 392 kk.; PAPP FERENC, A matematikai módszerek alkalmazásáról a nyelvtudományban: I. OK. XVII, 289—306. l. A statisztikai módszer stilisztikai felhasználására példa PAPP FERENC szép dolgozata: A stíluselmzés egy mennyiségi mutatókra épített módszere: Filológiai Közöny, 1961. 1—2. sz. 69—85. l.

csoportokon belül (32—59. l.). Általános lexikológiai szempontból is különleges érdeklődésre tarthatnak számot azok az adatok, melyek az egyes szófajok kölcsönös viszonyát világítják meg: ez a viszony jellemzi a különböző műfaj-csoportokat is (65—82. l.). Az utolsó (D) rész JOSEF V. BEČKA munkája (83—96. l.). Itt a szóalakok gyakoriságáról és viszonyatairól kapunk képet. Az ismertetett bevezető rész után (9—98. l.) következik a már említett két szójegyzék.

Egy gyakoriság-szótár milyenségét két, egymásba kapcsolódó tényező határozza meg: 1. hogyan sikerült az anyagot kiválasztani, 2. milyen a munka felhasználhatósága. Az első követelménnyel kapcsolatban azt jegyezém meg, hogy én a címet túl tágnak érzem. Tágnak egyrészt azért, mert a szótár csak a cseh *irodalmi* nyelv anyagát dolgozza fel. Ugyanis a beszélt nyelvi anyag még akkor is kicsiny lenne, ha a drámai művek nyelvét is ideszámítom ($96\ 031 + 140\ 215 = 236\ 246$ szó), de ez az eredmény is hamis, hisz nem azonos a köznyelvvél a dráma nyelve (s elsősorban szógyakorisági szempontból nem azonos), s még kevésbé a nyomtatásban is megjelent, előre megfogalmazott beszédek nyelve. Ugyanakkor jó lenne a munka címében meghatározni az időszakot is, amelyet a feldolgozott művek felölelnek, hisz éppen ez a munka, melyet több mint húsz éve kezdtek el, bizonyította be a legfényesebben, hogy mennyire nem időtálló egy ilyen szótár egésze. (Mint kirívó példát a „pán” szót hozza fel Jaromír Belič a szótár előszavában e tétel igazolására.)

A gyakoriság-szótár megbízhatóságának egyik alapja a szerzők és műveik helyes megválogatása.⁵ Itt egyedül talán Vladislav Vančura művének a beválogatása kifogásolható. Nem hiszem, hogy Vančura annyira egyéni szókincse mérvadó lehet egy gyakoriság-vizsgálatkor.

A szótár megítélésének másik és fő szempontja: felhasználhatósága. Az ismertetett mű e szempontból már eddig is fényesen kiállta a próbát. Marie Těšitelová szerint a gyakoriság szótárak három kérdés-kör⁶ megoldásához járulhatnak hozzá:

1. A nyelv lexikai oldalának megismerése.
2. Stilisztikai elemzés.
3. Nyelvtani kérdések megoldása és egyéb kérdések (pl. didaktika).

Az első ponttal kapcsolatban már eddig is jelentek meg erre a szótárra alapozott munkák. Elsősorban MAC VEY értékes megjegyzései csillantották fel az ilyen irányú kutatás komoly távlatait.⁷ Csak egy érdekességet említek meg: eddigi optimista elképzeléseinkkel szemben a cseh gyakoriság-szótár azt bizonyítja, hogy a szókincs 90%-ának megértéséhez 8000 szó szükséges!

A másik két ponttal kapcsolatban a szótár szerzői tettek figyelemre méltó észrevételeket. Tekintve, hogy az egyes művek cédulái külön vannak elraktározva, ez az anyag az egyes írók szókincsének stilisztikai elemzésekor igen becses lehet. A gyakoriság-szótárak didaktikai szerepéről nem is kell beszélnem.

Az elmondottakat összegezve: A cseh gyakoriság-szótár megjelenését örömmel üdvözölhetjük. Haszna, értéke már eddig is megmutatkozott. Meggyőződésem, hogy e munkát kézikönyvként fogják használni mind a cseh szókincssel foglalkozók, mind a szókincs törvényszerűségeit kutatók általában.

BOJTÁR ENDRE

Hibaigazítás

„XVII. századi közlés a Volga menti magyarokról” című közleményem (NyK. LXV, 411—2) 2. bekezdésének egész 7. sora az eredeti szövegben kurzív; a 3. bekezdés 8—9. sorában az *impohterum* szó helyes elválasztása: *impohterum*; ugyane bekezdés 9. sorában az első álló szedésű szó helyesen: *cæterum*, 11. sorában a *fi* után az eredetiben kötőjel van.

V. O. A.

⁴ Б. Трнка и др.: К дискуссии по вопросам структурализма. ВЯ. 1957/3.

⁵ A JOSSELSON-féle orosz gyakoriság-szótárral kapcsolatban többek között éppen e téren merültek fel jogos kifogások. Л. А. Новак—Р. Г. Пуотровский, Н. Н. Josselson: The Russian word count and frequency analysis of grammatical categories of standard literary Russian. Detroit, Wayne unio press, 1953. ВЯ, 1958/3. 134. l.

⁶ TĚŠITELOVÁ, MARIE, K. statistickému výzkumu slovní zásoby. SaS 1961/3. 171—181. l.

⁷ Az ismertetett szótár előszavában fel vannak sorolva M. VEY cikkeinek bibliográfiai adatai.

TARTALOM

RÉDEI KÁROLY: Obi-ugor jövevényszók a zürjén nyelvben	3
VÁSZOLYI ERIK: Zürjének a Kanyin-félszigeten	17
MIKOLA TIBOR: Szamojéd nyelvtanulmányok I.	35
MOÓR ELEMÉR: Szláv jövevényszavaink kvantitási anomáliái	43
KNIEZSA ISTVÁN: Szláv jövevényszavaink magánhangzó-kvantitása	59
KÁROLY SÁNDOR: A mondatfajták vizsgálata a funkció és a forma szempontjából	67

KISEBB KÖZLEMÉNYEK

GÜLYA JÁNOS: A vahi osztják főnevek tipológiai osztályozása a fonémaeloszlás alapján	89
FOKOS DÁVID: Etimológiai, jelentéstani és szintaktikai adalékok	95
RÉDEI KÁROLY: Szófejtések	103
GÜLYA JÁNOS: Etimológiai jegyzetek	111
BERECZKI GÁBOR: Etimológiai megjegyzések	117
ERDŐDI JÓZSEF: A mari U. <i>melna</i> , K. <i>mele-nä</i> szó etimológiájához	123
M. VELENYÁK ZSÓFIA: Négykézláb	125
LŐRINCZ JENŐ: Az <i>ős</i> rokon nyelvi megfelelőinek szelkup tagja	127
PAPP FERENC: A magyar szószerkezet-rendszer néhány sajátosságáról	129
BENKŐ LÁSZLÓ: A szavak stilisztikai értékeléséről	141
SZÁNTÓ ÉVA: Megjegyzések egy beszédtagolással foglalkozó elmélethez	151

ISMERTETÉSEK — SZEMLE

T. LOVAS RÓZSA: Melich János	157
Temesi Mihály: Klemm Imre Antal	160
Lakó György: Beke Ödön	162
JAKOBSON, ROMAN: A IX. Nemzetközi Nyelvészkongresszus záróelőadása	165
SZ. KISPÁL MAGDOLNA: Finnugor szak a magyar egyetemeken	173
SZABÓ LÁSZLÓ: Az Észk SzSzK TA Nyelv- és Irodalomtudományi Intézetének mun- kájáról	175
LÁSZLÓ GYULA: Az uráli őshaza kérdéséről	179
KISS JENŐ: <i>Béla Szent-Iványi</i> , Finnisch-ugrische Sprachwissenschaft und Ungarn- kunde an der Berliner Universität	181
FABRICIUS-KOVÁCS FERENC: <i>Hajdú Péter</i> , The Samoyed Peoples and Languages	182
HAUHIA, ULLA: <i>Sulo Haltsonen</i> , Suomalaista kaunokirjallisuutta vierailta kielillä	184
K. SAL ÉVA: Virittäjä	186
BERECZKI GÁBOR: <i>Paula Palmeos</i> , Karjala Valdai murrak	188
ERDÉLYI ISTVÁN: Очерки мордовских диалектов	189
LŐRINCZ JENŐ: Современный марийский язык. Фонетика	192
FOKOS DÁVID: Коми-пермяцкий язык. Введение, фонетика, лексика и морфология.	196
ERDÉLYI F. István A zürjénföldi régészeti irodalom új terméséből	204
SZATHMÁRI ISTVÁN: <i>Bárczi Géza</i> , A magyar nyelv életrajza	207
FODOR ISTVÁN: <i>Papp László</i> , Nyelvjárástörténet és nyelvi statisztika	210
FARKAS VILMOS: <i>Rácz Endre</i> , A magyar nyelv következményes mondatai	214
FÁBIÁN PÁL: <i>Molnár József</i> , A könyvnyomtatás hatása a magyar irodalmi nyelv kialakulására 1527—1576 között	216
FÓNAGY IVÁN: <i>Elekfi László</i> , Vizsgálatok a hanglejtés megfigyelésének módjaihoz	217

FÓNAGY IVÁN: <i>László Zsigmond</i> , Ritmus és dallam. A magyar vers és ének prozódiaja	219
FÓNAGY IVÁN: <i>André Abraham Moles</i> , Théorie de l'information et perception esthétique	222
FÓNAGY IVÁN: <i>Karel Sedláček—Antonín Sychra</i> , Hudba a slovo z experimentálních hlediska	225
VÉRTES O. ANDRÁS: Proceedings of the Fourth International Congress of Phonetic Sciences	226
HAZAI GYÖRGY: <i>Henri & Renée Kahane—Andreas Tietze</i> , The Lingua Franca in the Levant	228
FODOR ISTVÁN: Psycholinguistics: A Book of Readings	230
PAPP FERENC: <i>И. И. Резун</i> , Модели языка	234
KIEFER FERENC: Választás és véletlen a nyelvben	239
GÁLDI LÁSZLÓ: Sprachwissenschaftliches Wörterbuch	242
KISS LAJOS: <i>Ф. П. Филин</i> , Проект „Словаря русских народных говоров”	243
KISS LAJOS: Historia osadnictwa i organizacja społeczna pasterstwa oraz słownictwo pasterskie Tatr Polskich i Podhala	246
BOJTÁR ENDRE: <i>Jaroslav Jelínek—Josef V. Bečka—Marie Těšitelová</i> , Frekvence slov, slovních druhů a tvarů v českém jazyce	247
Hibaigazítás	249

A kiadásért felelős az Akadémiai Kiadó igazgatója

Műszaki szerkesztő: Vidosa László

A kézirat nyomdába érkezett: 1964. I. 28 — Példányszám: 650 — Terjedelem: 22 (A/5) ív

64.58362 Akadémiai Nyomda, Budapest — Felelős vezető: Bernát György

MUNKATÁRSAINKHOZ!

Kérjük munkatársainkat, hogy a szerkesztőség és a nyomda munkájának megkönnyítése érdekében vegyék figyelembe a következőket:

- 1. A kéziratokat szabványos papírlap egyik oldalára gépeljék.*
- 2. Egy lapra legfeljebb 30 sort, soronként legfeljebb 60 betűt írjanak.*
- 3. A kézirat szövegében a következő jelzéseket használják: —— (nyelvi adatok jelölésre), ritkított (értelmi kiemelésre), = = félkövér (alcímek betűtípusaként), ~ ~ kapitálchen a szerzők nevének kiemelésére).*
- 4. A kézirat géppel írott szövegében az utólagos javításokat, a fonetikai jeleket tintával írják.*
- 5. A folyóiratok és sorozatok kötetszámát római számmal írják. A kötet-szám és lapszám közé vesszőt tegyenek.*

Ára: 16,— Ft

Előfizetési ára egy évre: 28,— Ft

INDEX: 25.666



AKADÉMIAI KIADÓ BUDAPEST