

Kálmán László

MTA Nyelvtudományi intézet, Budapest

kalman.laszlo@nytud.mta.hu

Az összetett szavak és a nyelvelmélet

Kivonat

Az írás célja, hogy a szóösszetételek példáján mutassa be, hogy mi a pontos eltérés a szabályokon és megszorításokon alapuló (hagyományos és generatív) nyelvfelfogás és a mintázatok hasonlóságain alapuló (analógiás) megközelítés között. Az írás fő állítása az, hogy a szóösszetétel fogalma sajátos átmenetet képez a memorizált szókincs (lexikon) jellemzői és a produktív nyelvi eszközök között, ezért általános vele kapcsolatban a tanácstalanság a szabályközpontú elméletekben. A kategória átmeneti jellegét az írás a „N + V + -ás/-és” szerkezetű magyar összetételeken mutatja be. A kiinduló kérdésre pedig azt válaszolja, hogy az analógiás elméletek éppen abban különböznek a szabályközpontúaktól, hogy tartalmas viszonyt tételeznek fel a mintázatok produktivitása és a memorizált tapasztalatok szerkezete között, ami a szabályalapú leírásokban nem lehetséges.

Kulcsszavak

nyelvelmélet, analógia, összetett szavak

1. Bevezetés

Ebben az írásban az összetett szavak példáján mutatom be, hogy mi is a pontos különbség napjaink két nagy nyelvelméleti paradigmája, a szabályközpontú és az analógia-alapú között. Először röviden tisztázom, hogy mi a jellegzetessége a kétféle megközelítésnek. Nyitva hagyom azonban azt a kérdést, hogy van-e valóban elvi különbség a kétféle megközelítés között, vagy inkább csak heurisztikus, szemléleti vagy ideológikus az eltérés. A következő pontban arról lesz szó, hogy mi okozhatja a hagyományos és modern nyelvtanok jól érezhető tanácstalanságát az összetett szavak kezelésével kapcsolatban. Arra a következtetésre jutok, hogy a problémát a szóösszetétel kettős természete okozza: egyrészt a szókincs (lexikon) gazdagításának eszköze, tehát a nyelvi változás egyik formája, másrészt viszont többé-kevésbé produktív szerkezetalkotási eszköz is. Ezután egy konkrét magyar összetételi típus, a „N + V + -ás/-és” felépítésű összetételek néhány fajtáját veszem sorra, és megmutatom, hogy milyen leírási nehézségeket okoznak ezek a szabályokkal dolgozó nyelvreírásban. Rámutatok, hogy mind a formai, mind a jelentéstani nehézségeket a memorizáltság különböző fokozatai okozzák, és ezek összefüggése a produktivitás különböző mértékével. Ezzel eljutok az írás végső következtetéséhez: Az analógiás megközelítés nemcsak szemléletében különbözik a szabályalapútól, hanem konkrét nyelvfelfogásában és következményeiben is. A különbség lényege az, hogy az analógiás felfogás szoros összefüggést feltételez a szabályok érvényesülése (illetve a kategóriák éllessége) és a szókincsben felfedezhető mintázatok között, míg a szabályalapú elméletek ilyen összefüggést még akkor sem tudnának kifejezni, ha egyébként lehetővé teszik a sztochasztikusan (csak bizonyos valószínűséggel) érvényesülő szabályok és az elmosódó határokkal rendelkező kategóriák létezését.

2. Szabályok és mintázatok

2.1. A kérdés rövid története

A 20. század elejéig a nyelv szabályközpontú és analógiás megközelítése nem került ellentmondásba egymással. Végülis a görög *analógia* szó nem jelent mást, mint 'szabályszerűséget'. Az emberiség nagyon korán felfigyelt arra, hogy a nyelvi megnyilatkozások és azok részei szabályszerűen épülnek fel, szabályszerűségeket mutatnak. Legfeljebb némi hangsúlykülönbséget jelentett, hogy ezt azzal magyarázták-e, hogy a beszélők szabályokat sajátítanak el, és azokat követik, vagy azzal, hogy a korábbi nyelvi tapasztalataikból kibontakozó mintázatok befolyásolják a viselkedésüket, azok hasonlatosságára alkotják meg és értik meg az új megnyilatkozásokat. A szabályközpontú felfogás uralkodott például a klasszikus iskolai grammatikákban, amelyeknek döntően normatív volt a megközelítése: a cél az, hogy a diák megtanulja és kövesse a klasszikus auktorok műveiből kikövetkeztethető szabályokat. Viszont az analógiás felfogás dominált a 19. század pszichológizáló gondolkodásában, mint amilyen Paul (1886) nyelvtörténeti munkájában található.

Napjainkban azonban a kétféle felfogás élesen szembenálló táborokat alkot, méghozzá azért, mert amikor formálisan, precízen próbáljuk megragadni a nyelvi szabályszerűségek természetét, akkor meg kell hoznunk bizonyos döntéseket, amelyeket a korábbi, informális elméleteknek nem kellett. A szabályközpontú felfogást képviselik a már említett generativista irányzatok, amelyek szerint a nyelvtan célja, hogy megadja a nyelvnek mint (potenciális) megnyilatkozások osztályának az összetételét, és ez természetesen csak egy (rekurzív) szabályrendszerrel lehetséges. (Ennek az irányzatnak a kezdetét Chomsky 1957 jelentette.) Ezzel szemben az analógiás felfogás hívei, akik nagyon sok apró iskolát képviselnek, azt hangsúlyozzák, hogy a nyelv nem fogható fel megnyilatkozások osztályaként, mert egy ilyen statikus osztály feltételezése nem plauzibilis. (Ennek az irányzatnak néhány jelentős munkája Skousen 1989, Eddington 2006 és Blevins és Blevins 2009.) Mindezek az iskolák a nyelvi szokások és kategóriák graduális jellegét hangsúlyozzák. Szerintük a nyelvi szabályszerűségeket nem azzal kell magyarázni, hogy a beszélők interiorizált szabályokat követnek, hanem azzal, hogy a korábbi nyelvi tapasztalataikból kibontakozó mintázatokkal van tele az emlékezetük, és ezek dinamikus befolyásolják a döntéseiket, amikor új megnyilatkozásokat értenek meg vagy produkálnak.

2.2. A kétféle irányzat megkülönböztetésének problémája

A kétféle megközelítés nyilván nagyon eltérő módszerekkel és filozófiai alapállással jár együtt. A szabályközpontú felfogás követői szabály- vagy megszorítás-rendszerekben gondolkoznak, míg az analógiás felfogás hívei nagy adathalmazok (korpuszok) statisztikai jellegzetességeiben és a megnyilatkozások különböző szempontú hasonlóságának, valamint ezek gyakoriságának mérésében. Ugyanakkor eddig nem találok olyan okfejtéssel, amelyik meggyőzően feltárja, hogy miért elvileg összeférhetetlen a kétfajta iskola, és hogy az ideológiai jellegű érveken túl mi hozható fel valamelyik mellett vagy ellen. A szabályközpontú modellekbe is beilleszthető ugyanis a gradualitás, a bizonytalanság, a „szürke zónák” létezése. Elég, ha sztochasztikussá, valószínűségi alapúvá tesszük a szavak szófaji és jelentéstani besorolását, a szabályok alkalmazhatóságát (a teljesen kötelezőtől a teljesen választhatóig), és így tovább. Van-e az analógiás felfogásnak a gradualitáson kívül más ellenérve a generativista elképzeléssel szemben? Ezt fogom megvizsgálni a szóösszetétel példáján.

3. Az összetétel közelebbről

3.1. A nyelvtanítók tanácstalansága

Legkorszerűbb hagyományos (iskolai) grammatikánkban a következő rejtélyes mondatot olvassuk a szóösszetételről (Keszler 2000, a továbbiakban: MG, 321. o.):

„Az összetett szavak a mondatok szókapcsolataiból vagy egyéb mondatrészletekből származnak, illetőleg azokra vezethetők vissza.”

A forrásokból nyilvánvaló, hogy itt sem a *származás*, sem a *visszavezethetőség* szavak nem etimológiailag értendők, de hogy akkor hogyan, arra nem derül fény.

Valójában a nemzetközi szakirodalomban sincs sok okosság az összetett szavak mibenlétéről. Annyit szoktak írni, hogy olyan szavakról van szó, amikben egyetlen tartalmas tő helyett kettő vagy több van egymás után. És a legtöbb idegen nyelvű munkában is írnak ahhoz hasonlókat, mint amiket a MG a 322. oldalon:

„Az összetett szavak és a szintagmák közötti különbségtevés legtöbb esetben szemantikai alapú, például *franciatanár* és *francia tanár* [...]”

A „legtöbb esetben” kifejezést itt nagyon merésznek tartom, hiszen becslésem szerint az összetett szavak döntő többségénél a két vagy több tő egyáltalán nem fordulhatnának elő külön szóként, szó szerkezetet alkotva (például van *buszmegálló*, de nincs **busz megálló*). És olyan butaságokat, mint ami az ezt követő mondatban áll, ritkán találunk máshol a szakirodalomban:

„Vannak azonban olyan esetek, amikor csupán a helyesírás szabja meg, hogy a szerkezetet egyetlen szónak, vagy szintagmának tekintsük-e; vö. *nejlonzacskó*, *papírzacskó*, *műanyag zacskó*, *autóverseny* és *kötélhúzó verseny*, *be-benéz* és *ide-oda néz*.”

Az a gyanúm, hogy a szóösszetételekkel kapcsolatos általános semmitmondás nem más, mint a tehetetlenséggel járó frusztráció külső megnyilvánulása. A nyelvtanírók nem tudnak mást mondani az összetett szavakról, mint hogy milyen (nem szóösszetételnek minősülő) parafrázisokkal lehet hasonló tartalmakat kifejezni. Például *buszmegálló* 'olyan hely, ahol buszok szoktak megállni'. Ugyanakkor a nyelvtannak egyetlen más területén sem tartják szükségesnek, hogy egy-egy szerkezetet valamilyen más szerkezetből „származtassanak”, még az igekötők esetében sem, pedig az igekötős ige tulajdonképpen a szóösszetétel alfaja. Vajon miért?

3.2. A jelentéstani szempont

A magyar és a nemzetközi szakirodalom csak annyit mond az összetett szavak szemantikájáról, hogy azok valamiféle „jelentéstöbbletet” hordoznak a nekik megfelelő szó szerkezetekhez képest. Nem is csak az a baj, hogy ez teljesen semmitmondó, hanem inkább az, hogy nem igaz. Például a *buszmegálló* mennyiben hordoz „jelentéstöbbletet” ahhoz képest, hogy a *busz megállója*? Vagy hogy egy igazi úgynevezett „jelentéssűrítő” összetétellel példálózunk: a *franciatanár* milyen „jelentéstöbbletet” tartalmaz a *francia tanár*-hoz képest?

A hagyományos és modern nyelvtanok alapfeltevése az, hogy a *francia* szó „alapjelentése” az, hogy 'francia születésű vagy illetőségű', ezért a *francia tanár* szó szerkezetben ez a szó a „szó szerinti jelentésében” fordul elő. Ezért lehet azt mondani, hogy a *franciatanár* szóban, mivel ebben úgy értelmezzük, hogy 'a francia nyelvet, irodalmat, kultúrát oktató', nyilván valami extra mechanizmusnak kell működnie.

Valójában azonban semmiből sem következik, hogy a *francia* szónak pont azt az értelmezését kellene „alapjelentésnek” tekinteni, ami jelzős szerkezetekben fordul elő, és nem azt, ami például összetett szavakban, mondjuk a *franciatanár*-ban, esetleg valami harmadik dolgot. Valójában a jelzős szerkezet esetében is van egy „rejtett” tartalom (születési hely, származás, lakhely stb.), amit az értelmezéskor a hallgatóságnak fel kell fejtenie, akárcsak a *franciatanár* esetében azt, hogy a francia nyelvvel, kultúrával stb. van kapcsolatban. Sőt, ha úgy tetszik, az utóbbi még egyértelműbb is, mint a

jelzős szerkezet.

A „jelentéstöbblet” mint az összetett szavak különleges sajátossága különösen abszurd az olyan esetekben, mint a *franciaágy*, aminek az értelmezése egyáltalán nem „többletet” tartalmaz a *francia ágy* értelmezéséhez képest. Lehet, hogy ez az ágyfajta francia eredetű (bár még valószínűbb, hogy nem), de az értelmezéséhez semmiképp sincs szükség a *francia* szónak sem az „alapjelentésére”, akármit jelentsen is az, sem valamilyen rejtett tartalom felfejtésére. Egyszerűen arra van szükség, hogy ismerjük ezt az összetett szót.

3.3. A tanácstalanság forrása

Vajon mitől van az, hogy a nyelvtanok semmit sem tudnak kezdeni az összetételekkel, és ezért csak csupa haszontalanságot olvashatunk róluk? Mind a nyelvtani, mind a jelentéstani szempont esetében azt láttuk, hogy a tanácstalanság a megszilárdultság és a produktivitás kettősségével kapcsolatos. A nyelvtani szempontú vizsgálat állandóan ingadozik a szóösszetételek létrejöttének történeti magyarázata és szinkrón nyelvtani leírása között, mintha nem tudná eldönteni, hogy a nyelvi változás egy fajtájáról, vagy nyelvtani szabályokról lenne-e szó. Valóban a szóösszetétel sok fajtája a szinkrón nyelvi rendszerben is produktív lexéma-alkotási módszer, ezért állandóan létrejönnek újabb összetett szavak, így az összetétel nemcsak a nyelvi változás egyik tényezője, hanem a nyelv rendeltetésszerű használatának alapvetően a szinkrón rendszerhez tartozó eszköze. Ezért például a MG siet különbséget tenni a megszilárdult és a produktív összetételek között. Utóbbiakat szerinte

„az adott grammatikai szabályok alapján, az adott jelentéskörben korlátlan vagy csaknem korlátlan számban megalkothatjuk.” (322. o.)

Hasonló dilemmával találkoztunk a jelentéstani szempont vizsgálatánál is. A „jelentéstöbblet” nesze-semmi fogalmával szemben azt találjuk, hogy az összetett szavak egyes típusainak az értelmezése a felépítésüknek megfelelő, viszonylag rendszeres mintát követ (elég a *franciatanár*-hoz hasonló összetételek tömegére gondolni, mint amilyen a *számítógépguru* vagy a *nyugdíjszakember*), bár ezek között is vannak gyakoribbak és *ad hoc* jellegű alkotások, míg másoknak, mint amilyen a *franciaágy*, az értelmezésében nem sokat segít a részeik megértése, és csak akkor tudjuk őket értelmezni, ha már találkoztunk velük, és tudjuk, mire használatosak.

A nagy kérdés az, hogy valóban létező, feltételezhető és hasznos különbségtétel-e a produktív szóalkotás és a megszilárdult alakulat. Ez az első látásra kézenfekvő különbségtétel ugyanis mind a hagyományos (iskolai), mind pedig a modern (generativista) nyelvleírásnak óhatatlan velejárója, ugyanakkor annyira alapvető problémát rejt magában, amitől a nyelvnek magának a mibenlétére, természetének lényegére vonatkozó nézeteink függenek.

Ha ugyanis kiderülne, hogy ez a megkülönböztetés nem egyértelmű, vagy egyáltalán nem is értelmes, vagy akár csak nem kategorikus, hanem graduális, akkor bizonytalanná válik, hogy pontosan miket írhatunk bele egy nyelv grammatikájában a szabályok közé, és az mind a hagyományos, mind a generatív nyelvten számára katasztrofális lenne, hiszen mindkét irányzat feltételezi a szabályoknak és a szókincnek (az építőelemek tárának) megkülönböztetését. Ha az összetett szavak egy részéről nem lehet eldönteni, hogy a szókincs részét képezik-e, tehát valamilyen nyelvi változás révén a megszilárdult építőkövek közé kerültek, vagy pedig nyelvtani szabályok alkalmazásának az eredménye, akárcsak egy mondatszerkezet, akkor tarthatatlan az, ahogy a szabályközpontú felfogás elképzei a nyelvtant. A hagyományos és a generatív nyelvtenban ugyanis nincsen köztes régió a szókincs és a szabályrendszer között, így nem utalhatjuk az összetételek egy részét ilyen régióba, legfeljebb informálisan mondhatnánk, hogy „megszilárdulóban van”, de nem tudnánk ezt pontosan értelmezni a rendszeren belül.

3.4. Egy konkrét vizsgálat

3.4.1. Vannak-e csoportok?

Egyetlen konkrét összetettség-típussal foglalkoztam, méghozzá a „legszabályosabbak” egyikével, az „N + V + -ás/-és” típusú összetételekkel, mint amilyen a *kulcskeresés*. Ezek nagy részében az N-t a V tárgyaként értelmezzük (a kulcs az, amit keresünk). Így ilyenkor az igének természetesen tárgyassá kell lennie. Ezt a tényt akár szabályként is felírhatjuk. Ugyanakkor vannak az ugyanilyen felépítésű összetett szavaknak érdekes altípusai, amelyek – maradványként a szabályközpontú megközelítésnél – kivételes al-szabályokat igényelnek. Ilyen például az az eset, amikor az ige (V) tárgyatlan, és a főnevet (N) az alanyként kell értelmezni. Ez azonban csak bizonyos esetekben figyelhető meg, például ha az ige hangadást fejez ki: *madáracsicsérgés, kutyaugatás* stb.¹ További alszabályokra is szükség lenne azonban azokra az esetekre, amikor az ige nem hangadást fejez ki, például *hajhullás, vízesés, torokgyulladás, fogolyszökés, cseresznyevirágzás, atomrobbanás, gyerekbűnözés* stb.

A probléma az, hogy ennek az utóbbi sorozatnak a tagjai között nem sok közöset találunk. Szinte mindegyik más-más esetet képvisel, más-más természetűek az igék és a főnevek. Létezhetnek ugyan közöttük népesebb alcsoportok, amelyek szabállyal leírható mintázatokat alkotnak, például ilyen csoport lehet a *torokgyulladás, szívbénulás, inszalagszakadás* és társaik. (Orvosi körökben például nem kelt meglepetést, ha az ember mondjuk *bőrnedvedzésről* beszél, pedig ez az interneten a hibahatáron belül, kétszer fordul elő.) Azt viszont már nehéz megmondani, hogy pontosan mely szavak minősülnek ezek „társainak”: csak azok, amelyekben az előtag egy szervre, az utótag pedig annak meghibásodására utal? Vagy ezek valójában abba a tágabb körbe tartoznak, amelybe a *maghasadás* is? Esetleg ide sorolhatjuk a *fogolyszökés*-t és a *gyerekbűnözés*-t is, mint amelyek szintén az előtagban megnevezettek rendellenes viselkedésére, diszfunkciójára utalnak? És akkor itt ugyanarról a mintázatról van szó, mint a földrengés, a *földcsuszamlás, a filmszakadás, a tengelytörés, a bányaomlás* vagy a *csőrepedés* esetében?

Nem tudom elképzelni, hogy létezne olyan elméletileg megalapozható módszertani eljárás, amivel egy szabályalapú leírásban el lehetne dönteni, hogy pontosan mely eseteket kell egyetlen szabály segítségével leírni, mint önálló (al)osztályt, és hol kell az osztályok között határokat húzni. Ez a fajta kényszerű önkényesség viszont mindenképpen a szabályközpontú megközelítés ellen szól. Azt mondhatnánk, hogy minél népesebb egy-egy ilyen csoport, annál produktívabb szabályt feltételezhetünk, és annál kevésbé mondhatjuk, hogy a tagjai már megszilárdult, a szókinccs részét alkotó szavak. Csakhogy ehhez először is tisztáznunk kellene, hogy egyáltalán milyen csoportok léteznek, és – mint említettem – már ez is lehetetlennek látszik. Csak annyit mondhatunk biztosan, hogy vannak egyedinek látszó, csoporthoz nehezen sorolható szavak, mint a *vízesés*, amelyeket nyilván olyan memorizálандó, megszilárdult szónak kell tekintenünk, mint amilyen a korábban említett *franciaágy*. De semmi sem szól amellett, hogy ezt a típust élesen el tudnánk választani a csak néhány elemet tartalmazó csoportoktól (ha egyáltalán elfogadjuk, hogy csoportot alkotnak), mint amilyen a *fogolyszökés* és a *gyerekbűnözés*.

3.4.2. Az összetételek funkciója

Az összetett szavak esete arra hívta fel a figyelmünket, hogy a szabályalapú nyelvi leírás fő problémája, támadható pontja az, hogy kategorikus különbséget kell tennie a megszilárdult, memorizálандó alakok és a szabállyal magyarázható, nem feltétlenül memorizálандók között. Lehetséges sztochasztikus szabályokat és sztochasztikus kategóriákat feltételezni, de a szabályközpontú irányzatban belül nehéz elképzelni, hogy a megszilárdultság, a memorizálандóság is sztochasztikus legyen. Márpedig az összetett

1 Szilágyi N. Sándor hívta fel a figyelmemet arra, hogy ehhez hasonló sajátos alosztály azok között az összetételek között is van, ahol az ige tárgyas, ti. az, ahol a főnevet ennek ellenére az ige alanyként kell értelmezni, viszont az igének olyasmit kell jelentenie, ami az embernek szenvedést okoz, pl. *darázscsípés, kutyaharapás*.

szavak esetében éppen ezt tapasztaljuk. Vannak a skála egyik végén olyanok, mint a *franciaágy* vagy a *vízesés*, és a másik végén olyanok, mint a *kulcskeresés*. Az utóbbit talán sosem hallottam, de ha hallottam is, nem kell rá emlékezmem, hogy megértsem. A skála két vége között pedig olyanok vannak, amelyekről több vagy kevesebb biztonsággal tudom csak megmondani, hogy hallottam-e már, és amelyeket több vagy kevesebb biztonsággal tudom csak a megalkotójuk szándéka szerint értelmezni.

Az összetett szavak éppen a természetük miatt alkalmasak ennek a megmutatására. Ezek ugyanis éppen olyan tartalmak kifejezésére szolgálnak, amelyekre akár egyetlen tőszó is létezhetne. Tudom, hogy ez nagyon nehezen ellenőrizhető, nem mérhető szempont, inkább intuitív meghatározásnak szánom. Mindenesetre nagyon gyakori, hogy amit az egyik nyelvben összetett szóval fejezünk ki, annak a megjelölésére egy másik nyelvben önálló tőszó létezik, például azt, amit magyarul a *szabálysértés* szóval fejezünk ki, angolul tőszóval (*offence*) fejezik ki, a magyar *vízesés* pedig franciául *chute* vagy *cascade* (mindkettő tőszó).

Hogy miért van az, hogy bizonyos célokra plauzibilisnek találjuk egyszerű, egy tőből álló szavak használatát, más célokra pedig nem? Van egyfajta ökonómiai racionalitás abban, hogy a nyelvek olyan jelenségek jelölésére használnak tőszavakat, amelyek gyakran fordulnak elő, és bizonyos jelentőségük van az ember és a társadalom életében. A legegyszerűbben ezt a két motívumot úgy tudnám együtt megnevezni, hogy „intézmény-jellegű” jelenségek megnevezésére szoktunk tőszavakat használni. Igaz, hogy az *intézmény* szóval általában társadalmi intézményekre szoktuk utalni, de hasonlóan gyakori, visszatérő, jelentős és tartós jelenségek nemcsak a társadalom életében vannak, hanem például a természetben is.

Például a *kulcsbeszorulás* vagy *kulcsszorulás* kiválóan alkalmas lehetne arra, hogy a beszélő érzékeltesse ennek az eseménynek az intézményesült (gyakori, visszatérő, jelentős, tartós) előfordulását. Ennek ellenére az elsöre egyáltalán nem találtam példát; a második az interneten legalább egyszer előfordul. Nyilván inkább lakatos-körökben várnánk e szavak megjelenését; az interneten talált egyetlen találat is gépkocsilakatos szövegből származik. Egy másik példa: az anyagneves összetételek talán minden korlátozás nélkül produktívnak nevezhetők, a legtöbb esetben mégis félrevezető lenne például *szénembernek* nevezni egy szénből kifaragott emberfigurát, mert azt sugallná, hogy szénből rendszeresen faragnak figurákat (sőt, talán hogy ennek valami speciális, például kultikus célja is van).

Valójában nagyon kevés olyan eset van, amikor bizonyosan állíthatnánk, hogy nem lenne plauzibilis az egytagú szó létezése. Olyankor lehetne ezt elmondani, amikor a szándékolt értelmezés egy epizodikus, egyszeri eseményre vagy jelenségre utalna. Például teljesen félrevezető lenne a *babaválogató* szóval utalni valakire, aki éppen babák között válogat a játékboltban. Azt sugallnánk vele, olyan irányba terelnénk a hallgatóság asszociációit, hogy az illetőnek ez rendszeres elfoglaltsága, sőt, hogy ez többeknek a rendszeres elfoglaltsága, ha nem egyenesen a szakmája.

Akár ikonikus eszköznek is tekinthetjük tehát a szóösszetételt. Mivel az összetett szó hangsúlyozásában a tőszavakra emlékeztet, az összetételek alkalmazásával arra utalunk, hogy az illető dolog akár tőszóval is kifejezhető lenne, mert a beszélő (a fenti értelemben) intézményesnek tekinti azt, amire utal vele. Ezt a nehezen formalizálható, de intuitíve jól érthető szempontot megvizsgáltam egy nagy számítógépes korpuszon is.

Az „N + V + -ás/-és” felépítésű összetett szavakat válogattam ki a *Szószablya Webkorpusz* (Halácsy és mtsai. 2004, Kornai és mtsai. 2006) szavai közül. Ez összesen 2718 tövet jelentett, amelyeknek előfordulási gyakorisága körülbelül 3.5 millió. Ezek között a tövek között összesen 13 olyat találtam (vagyis az öt ezreléküket), amelyek nem feleltek meg az előző pontban ismertetett intuitív kritériumnak, vagyis nem „intézményesült” jelenségre utaltak. (Ezek előfordulási gyakoriságukat tekintve 317 darabot jelölnek, ami csak azért fontos, mert ez átlagosan kb. 24 előfordulást jelent, vagyis ezek nem tekinthetők nemlétezőnek. Minden más szempontból a típusgyakoriság a fontos ebben az

esetben.)

A vizsgált összetett szavak közül például a *búcsúkihallgatás* vagy az *embersportfogadás* szavakat minősítettem „kivételesnek”, de ennek az is lehet az oka, hogy olyan szakterületekhez tartoznak, amelyeken nem vagyok járatos, és azokon a bizonyos szakterületeken olyan terminus technicusnak minősülnek, amelyeket akár tőszóval is jelölhetnének.

4. Általános tanulságok

Az első tanulság, amelyet az előző részben már említettem, hogy az analógiás megközelítésnek mindenképpen előnye az, hogy nem kell állást foglalnia arról, hogy melyik részmintázat (esetünkben összetételi altípus) felel meg valamilyen „szabálynak”, és pontosan melyik összetétel melyik részmintázatot képviseli. Ezekről a szabályalapú elméletekben önkényesen kell döntenie a nyelvésznek, és az ilyen döntést empirikusan lehetetlen alátámasztani.

A második tanulság az összetételek funkcionális, jelentéstani sajátosságaira vonatkozik. Mivel az analógiás felfogás szerint csak a hasonlóságokkal lehet magyarázni egy bizonyos szerkezet(típus) használatát, természetesen adódik, hogy egyazon összetettség-típus különböző alcsoportjainak igen heterogén lehet a használata, aszerint, hogy különböző történeti véletlenek és egyéb tényezők révén milyen memorizált elemek kerültek bele az egyes alcsoportokba. Az egyazon részmintázatba, alcsoportba tartozó összetételek használata között az analógiás megközelítés szerint csak családi hasonlóságnak kell érvényesülnie, ami a különböző használatok mindenféle átfedését is magába foglalja, míg a szabályalapú megközelítés vagy egységes jelentést, vagy egymástól elkülönülő alcsoportok létezését jósolja.

Az összetett szavak jellegzetességeinek vizsgálata után visszatérhetünk kiinduló kérdésünkhöz: Van-e valóságos különbség a szabályszerűségek szabályokkal, illetve mintázatokkal való magyarázata között? Mindkettő visszatérő elrendezések leírására szolgál, de míg a szabályközpontú megközelítések azt feltételezik, hogy léteznek alapelemek, amelyekből a beszélők összetett jeleket alkotnak, az analógiás felfogás szerint a listázható, „szótározható” elemek és a belőlük kibontakozó mintázatok egymással dinamikus viszonyban vannak.

Ezt a dinamikus viszonyt különösen jól illusztrálják az összetett szavak: a listázásuk gyakorlatilag és elméletileg is lehetetlen, hiszen az összetétel a szóalkotás aktívan alkalmazott, produktív eszköze. Ugyanakor a produktivitása a különböző típusokban eltérő nagyságú, sok feltétlenül listázandó összetétel létezik, és sok olyan, amely akár listázva is lehetne (amit akár memorizálhatnak is a beszélők), ugyanakkor a nyelv beszélői számára felbontható, mert sok vele formailag és szemantikailag párhuzamos összetétel létezik abban az értelemben, hogy többé-kevésbé ténylegesen memorizálva van, vagy egyesével, vagy mint valamennyire általános mintázat.

Ha mindez így tartható, akkor a kiinduló kérdésre a következő a válasz. A szabályalapú és a mintázat-alapú (analógiás) megközelítés között az a legfontosabb eltérés, hogy az analógiás elméletek közvetlen kapcsolatot tételeznek fel a „felbonthatóság” (motiváltság, ha úgy tetszik: szabállyal való magyarázhatóság) mértéke és a memorizált tapasztalatokból kibontakozó mintázatok között. Ez az a kapcsolat, amit még egy sztochasztikus szabályokkal és sztochasztikus kategóriákkal operáló szabályrendszer sem képes megragadni. Olyan szabályalapú elmélet ugyanis nem létezik (illetve ha létezne, akkor már analógiás elméletnek minősülne), amelyben a szabályok alkalmazási valószínűsége (vagy annak a valószínűsége, hogy egy adott elem egy bizonyos kategóriába tartozik-e) attól függ, hogy milyen a memorizált, listázott elemek összetétele.

Hivatkozások

- Blevins, J.P. és J. Blevins (szerk.) 2009. *Analogy in grammar: Form and acquisition*. Oxford University Press, Oxford.
- Chomsky, Noam 1957. *Syntactic Structures*. Mouton, Hága.
- Eddington, D. 2006. Paradigm uniformity and analogy: The capitalistic versus militaristic debate. *International Journal of English Studies* 6: 1–18.
- Halácsy Péter, Kornai András, Németh László, Rung András, Szakadát István, Trón Viktor 2004. Creating open language resources for Hungarian. In: *Proceedings of the 4th international conference on Language Resources and Evaluation (LREC2004)*.
- Kornai András, Halácsy Péter, Nagy Viktor, Oravecz Csaba, Trón Viktor és Varga Dániel 2006. Web-based frequency dictionaries for medium density languages. In: Adam Kilgarriff és Marco Baroni (szerk.): *Proceedings of the 2nd International Workshop on Web as Corpus*. ACL-06, 1–9.
- Keszler Borbála (szerk.) 2000. *Magyar Grammatika*. Nemzeti Tankönyvkiadó, Budapest.
- Paul, Hermann 1886. *Principien der Sprachgeschichte*. Max Niemeyer, Halle.
- Skousen, R. 1989. *Analogical modeling of language*. Kluwer Academic Press, Dordrecht.

Angol nyelvű cím:

Compounds and theory of language

Angol nyelvű összefoglaló (max. 20 sor)

The paper intends to answer the question what exactly distinguishes analogy-based approaches to language form rule- and constraint-based ones. Using the example of compounding, the paper argues that one problem area in which the difference is relevant is the gradual character of the boundary between memorized vocabulary and regularly formed expressions. A brief survey of a type of Hungarian compounds (consisting of a noun, a verb plus a nominalizing ending) reveals why rule- and constraint-based theories are unable to explain the degrees of productivity and transparenance within such a large family of expressions. The conclusion is that analogy-based theories, but not constraint- and rule-based ones, posit a strict and meaningful relationship between the stochastic parameters of rules and categories, on the one hand, and the patterns arising from memorized experience, on the other.