

# HANGKIVETŐ FŐNEVEK ANALÓGIÁS KERETBEN

RUNG ANDRÁS

## 1. Bevezetés

A szabályalapú nyelvtanok (akár hagyományosak, akár generatívak) sokszor jó közelítő leírást adnak az alaktani viselkedésről, azonban képtelenek olyan nyelvi jelenségeket magyarázni, mint például a fokozatosság, nyelvi ingadozás, a gyakoriság hatása a nyelvi változásra (Skousen 1989). Ennek legfőbb oka, hogy egzaktásra, a modellek egyszerűségére törekzenek, de emiatt figyelmen kívül hagyják mindazokat az ismereteket, amelyeket a nyelv valós működéséről megtudtunk pszicholingvisztikai és neurolingvisztikai kutatásoknak köszönhetően, illetve megállapításaikat elméleti konstrukciók alapján és részjelenségek megfigyeléséből teszik, sokszor a valós adatok nagy részét figyelmen kívül hagyva. Másik súlyos hibája a hagyományos és a generatív elméleteknek, hogy a nyelven kívüli hatásokat, például a használati gyakoriságot teljesen kiküszöbölték a nyelvi leírásból, holott számos esetben a nyelven kívüli elemeknek jelentős hatása is lehet magának a rendszernek a formálódására (Rung 2008).

Az említett nyelvi jelenségek leírásában az analógiás nyelvtanok jobban közelítik a pszichológiai realitást, azaz a valós nyelvi működést. Felszíni alakokra építik leírásukat, mint a nyelvből megfogható biztos és létező elemekre, működésüket nem a jelenségeket megragadni próbáló rendezett vagy rendezetlen szabályok sora, hanem az alakok egymásra hatása határozza meg. A nyelv működését lokális, atomi szinten vizsgálják, nem pedig általánosabb szabályokból és azok megszorításaiból vezetik le.

Ha bármely nyelv analógiás nyelvtanát kívánjuk megírni, annak egyik alapfeltétele az, hogy tudjuk, mely fonémák-hangok, alakok (Bybee 2001), konstrukciók (Goldberg 1995) hasonlóak az adott nyelvben, és ezek hasonlósága milyen mértékű, min alapszik. Természetesen ehhez először meg kell határozunk, hogy a hasonlóságot milyen paraméterek mentén mérjük, milyen tulajdonságokban számít egyáltalán a hasonlóság.

Feltételezésem szerint az egyes nyelvi elemek nem elszigetelten léteznek, hanem szoros és állandó interakcióban vannak egymással, amelynek egyik legfontosabb mozgatórugója az analógia (Itkonen 2005). Ha a nyelvben

valahol változás következik be, akkor az az erőviszonyokra azonnal hatással van, és a rendszer egészének változásához vezet. Ezt a legtöbb 20. századi nyelvelmélet el is ismeri. Ennek megfelelően az egyes állapotok leírásával foglalkozik a szinkrón nyelvészet, míg az ezek közti átmenetek vizsgálatával a diakrónia.

Ez a megközelítés azonban kimondva vagy kimondatlanul azt közvetíti, hogy vannak mindig stabil, önmagukban megfigyelhető állapotok. A nyelv változik, de maga a változás nem alapvető minősége. Az analógiás nyelvtanok által megfigyelt tények, jelenségek azonban cáfolják ezt a szigorú és merev szétválasztást (Philips 2006). A változás és az állapot nehéz szétválaszthatóságából következik, hogy a rendszer a maga statikusságában nem létezik, vagy legalábbis olyan absztrakt fogalom, amely a nyelvvel való valós munkára alkalmatlanná teszi.

Ha a nyelvi jelenségeket szorosan összefüggőnek vesszük, és változásukat lényegi elemüknek tekintjük, akkor felvetődik a kérdés, hogy egyáltalán lehet-e és értelmes-e a nyelvnek bármely részjelenségét leírni anélkül, hogy más részleteit ne vennénk figyelembe, hisz az összefüggések feltárása nélkül a jelenség értelmezhetetlen vagy csak részlegesen értelmezhető lesz.

A leírás ebben az esetben valóban nem lesz tökéletes, de mivel a nyelvi változás teljességét az arról való meglehetősen korlátozott és szerény tudásunk miatt semmiképp sem tudjuk megragadni, így mégis kénytelenek vagyunk csak egyes darabjait vizsgálni. A hagyományos megközelítésekkel ellentétben azonban nem állítom, hogy ezeket a részleteket önállóan és pontosan le tudjuk írni, hanem úgy vélem, hogy újabb leírások fényében majd kiegészítésre szorulnak a későbbiekben. Adatainkat folyamatos változásukból kifolyólag sosem tudjuk megragadni, de ez nem is gond, hisz a nyelvészet célja nem feltétlenül a leírás, hanem a leírást meghatározó nyelvi folyamatok megértése és pszichológiailag reális feltárása.

A szavak összehasonlításában azok felszíni szerkezetét és alakjaik gyakoriságát veszem alapul, mivel számos, megfelelő mértékben meggyőző kutatás a mögöttes szerkezeteket inkább gátló, semmint hasznos elméleti konstrukcióknak tartja (Bybee 2000, 2001, 2002, 2010, Kálmán 2008, Rebrus–Törkenczy 2008, Benua 1995, 1997, Steriade 2000), így a felszíni alakok közt nem mögöttes alakok alapján keresnek kapcsolatokat és megfeleléseket.

Csak a felszíni alakokra hagyatkozva az ingadozással és fokozatossággal szorosan összefüggő analógiás kiegyenlítődés is jól megragadható jelenséggé válik (Kraska-Szlenk 2007). Feltételezésem szerint az analógiás alapú

változásokat több további szempont is meghatározza (használati mód, jelentés stb.), de ezekből a legfontosabb a hangtani/fonológiai hasonlóság (Skousen és mtsai. 2002). Elemzésemben a nyelvi ingadozást mint a változás folyamatának legkézzelfoghatóbb jelét vizsgálom, nem pedig mint a szociolingvisztikai értelemben vett variabilitás letükröződését. Az adatok vizsgálata során láthatjuk, hogy az ingadozás a szavak belső szerveződésének és viszonyainak, illetve a nyelvhasználat általános hatásainak köszönhető, nem pedig a beszélők sokféleségének és egyedi nyelvhasználatbeli eltéréseiknek.

Írásomban a hangkivető tövek elemzése több célt is hivatott szolgálni. Elsődlegesen a hasonlósági viszonyok és a gyakoriság kapcsolatának vizsgálatával meg kívánom mutatni, hogy egy példányalapú, analógiás megközelítés alkalmas, sőt a szabályalapúaknál lényegesen alkalmasabb ezen tövek elemzésére. Elemzésemet korábbiaknál részletesebben bemutatott nyelvi adatokkal támasztom alá, amelyek más kutatásokhoz is alapot jelenthetnek.

A viszonyok értelmezése azonban sokszor csak az egyedi alakok alapos vizsgálatával hajtható végre. Sajnos az egyes vizsgálatra érdemes kisebb csoportok még nagy korpusz segítségével való elemzés esetén is alkalmatlanok a statisztikai alapú feldolgozásra, és csak akár egyedinek is tekinthető jelenségek és tendenciák felismerésére szorítkozhatunk ezekben az esetekben.

## 2. Források és adatok

Korpuszalapú vizsgálataimban a Szószablya webkorpusz (Halácsy és mtsai. 2003) gyakorisági adatait használtam fel, amelyet azért választottam, mert jelenleg ez a legnagyobb, mintegy 19,1 millió szóalakot tartalmazó korpusz, amely 3,493 millió magyar weboldal és 1,486 milliárd szó alapján készült. Így az ezt méretben követő 187,6 millió szavas Magyar Nemzeti Szövegtárnál ([http://corpus.nytud.hu/mnsz/index\\_hun.html](http://corpus.nytud.hu/mnsz/index_hun.html)) nyolcszor több adatot tartalmaz. A Szövegtár mintegy 90%-a szerkesztett (sajtó, szépirodalom, tudományos és hivatalos) szövegeket tartalmaz, amelyek egy része még a 20. században keletkezett, illetve nem a magyar köznyelvet reprezentálja, hanem annak valamely határon túli változatát (22,9 millió szövegszó). Csak 17,8 millió szövegszóból áll az a része, amely alkalmas lenne a magyar köznyelv változásának és az ingadozás jelenségeinek tanulmányozására. Mivel kutatásomban cél volt, hogy minél több nyelvi adattal dolgozzak, és hogy inkább a hétköznapi, szerkesztetlen szövegek kerüljenek előtérbe, ahol a változás jelenségei jobban tetten érhetők, a Szószablya webkorpusz erre a vizsgálatra sokkal alkalmasabbnak tűnt.

A korpusz tövek helyett szóalakokat, elnagyolt gyakorisági kategóriák helyett pedig pontos gyakorisági számokat tartalmaz, amely alkalmassá teszi nyelvészeti és nyelvtechnológiai kutatásokra. A Szószablya webkorpusz nagy erénye, hogy válogatatlan, sokszor a beszélt nyelvhez közelebb nyelvhasználatot és írásmódot rögzítő anyagokat (fórumok, blogok) is nagy arányban tartalmaz, így az ez alapján tett megállapításaim is jobban közelíthetik a magyar nyelvi valóságot. A Szószablya webkorpusz egy szóalakhhoz 4 gyakorisági számot ad meg, aszerint, hogy a szöveg (weboldal), amelyből származik, mennyire jól követi a magyar helyesírás szabályait a szavak szintjén. Az osztályozásba a szónál (szóköztől szóközig terjedő egység) magasabb szintű helyesírási szabályokat nem vették figyelembe. A legszerkesztettebb szövegek 589 millió szóalakot tartalmaznak, ezekhez a korpusz tartalmaz a hunmorph (Trón és mtsai. 2006) program által készített morfológiai elemzéseket is. Minden esetben a nagyobb, 1,486 milliárd szóalakra vonatkozó gyakorisági számokat használom, mivel a nagyobb korpusz több a beszélt nyelvhez közeli alakot tartalmaz, amelyeknek szerzői a magyar helyesírást, mint kötelező érvényű normát kevésbé tartják fontosnak.

A korpusz használata során a következő tényeket tartottam szem előtt. A korpusz dominánsan magyar szövegek alapján készült, amelyeknek azonosítása gépi úton történt, így még akár a jobb helyesírású szövegekben is maradhattak más nyelvi elemek, amelyek a gyakorisági adatokat is befolyásolták. A nem magyar nyelvi elemek kiszűrése alapvetően könnyű és hibamentes folyamat, de ha egy idegen nyelvű szó azonos egy magyar szóalakkal, akkor azok gyakorisága összeadódik (pl. angol *must* 'kell' v. magyar *must* 'szőlőital'), így az ilyen adatokkal „óvatosan” kell bánni. Ilyen jellegű problémák ékezetellen magyar szövegek során is előfordulhatnak, mint arra majd még a későbbiekben utalni is fogok.

Habár a magyar web már 2003-ban is elég jól és egyenletesen reprezentálta a közbeszéd összes témáját és ezáltal azok szókincsét, az informatikához és az internethez kötődő szavak, mint *rendszergazda* vagy *kernel* vélhetően jóval nagyobb arányban vannak reprezentálva, mint ahogy egy átlagos hétköznapi szövegben elvárhatnánk.

Vizsgált korpuszom közel áll a beszélt nyelvhez, de nem beszélt nyelvi korpusz. A szövegek nem beszédben elhangzott szövegek lejegyzései, ezért pontosan nem is tükrözik annak sajátosságait, ugyanakkor behoznak csak az írott szövegekre jellemző, számunkra érdektelen problémákat (elgépelés, ékezetellen írás stb.) Kutatásom céljaival egy nagyméretű beszélt korpusz

alaján való vizsgálódás állna leginkább összhangban, de erre sem most, sem vélhetőleg a közeljövőben nem lesz mód, főleg, ha azt is figyelembe vesszük, hogy a méreteiben hatalmas Szószablya webkorpusz is kicsiny minden magyar nyelvi jelenség magában foglalásához.

Vizsgálatomhoz a hangkivető főneveket választottam. Választásom több okból esett rájuk. A magyar hangkivető tövek, habár zárt osztályt alkotnak, meglehetősen nagy számúak, így a viselkedésükből levonható következtetések nem szórványos, egyedi és ritka adatokon alapszanak. A hangkivető főnevek viselkedésére jellemző a fokozatos ingadozás, amelyről még a szabályalapú megközelítések engedékenyebb változatai sem tudnak teljesen számot adni.

A hangkivető szavak kimerítő jellemzése mind a mai napig nem történt meg. Generatív vagy hagyományos nyelvészeti keretrendszerekben sok leírás keletkezett (pl. Papp 1975, Vago 1980, Elekfi 1994), de ezek elméleti korlátaikból kifolyólag a hangkivető tövek több sajátosságát sem tudták megragadni. A hangkivető töveket kiindulási pontnak az is kiválóan alkalmassá teszi, hogy korábban alapos leírásokat készített viselkedésükről Törkenczy és Siptár (2000), valamint Rebrus (2000), amelyek már jó viszonyítási alapot képeznek vizsgálódásaimhoz. Ezt Rebrus és Törkenczy (2008) az enyémhez hasonló elméleti keretben újra megvizsgálta. Leírásuk sokban egybecseng eredményeimmel, de kutatásom számos új eredményt is hozott megállapításaihoz képest, mivel a hasonlóság fennállását nem csak teljes szekvenciaazonossághoz kötöm (igaz, az ettől való eltérés elméletbeli lehetőségét ők is elismerik, de leírásukban csak szórványosan alkalmazzák), és a nagy mennyiségű adat beépítése is sok nyelvi tényt helyez más megvilágításba.

A szavak alapos fonológiai ismerete nélkül lehetetlen meghatározni egymáshoz köthető viszonyaikat és különbségüket a többi főnévtől. Amennyiben azt feltételezzük, hogy a szavak hangkivetősége nem véletlenszerűen alakult ki, akkor szükségszerűen ezeknek a szavaknak valamilyen szempontból hasonló jegyeket kell felmutatniuk. Ez a hasonlóság kézzelfoghatóvá is válik, amikor a hangkivető tövek egyes nyelvváltozatokban bevonzanak körükbe hozzájuk nagyon hasonló szavakat (pl. *motor*, *bútor*).

A Szószablya webkorpuszon túl vizsgálatomban még a BME MOKK morphdb.hu szótárára támaszkodom, amely jelenleg a legnagyobb ingyenesen is hozzáférhető nyelvi adatbázis (130 ezer tő, Trón és mtsai. 2006), ahonnan összesen 1212 hangkivető főnévi tövet választottam ki. Ezekből összesen 1097 volt a szótárban hangkivetőként megjelölve, amelyekből kivettem a *kelet*, *sportberkek*, *sodor*, *terem* szavakat. A *kelet* szó egyértelműen a *kelte* szóalak

miatt került be hangkivetőként rögzítve, ahelyett, hogy már ragozott főnévként vették volna fel a szótárba. A *sportberkek* már szóalak, és nem tő, helyesen a szótárban *sportberék*-ként kellene szerepelnie, amelynek hiányos a paradigmája. A *sodor* és a *terem* szavak valóban hangkivető főnevek, de mivel alanyesetük és számos további ragozott alakjuk egybeesik a náluk sokkal gyakoribb *sodor* és *terem* igei tövek alakjaival, ezért célszerűbbnek tartottam ezek kihagyását a vizsgálatból. Úgy véltem, hogy elegendő nyelvi adat birtokában ezek elhagyása nem vezet az eredmények jelentős módosulásához.

A szótárban hangkivetőként megadott töveken túl további 115 tövet választottam ki, amelyek valóban hangkivetők, és szerepelnek is a szótárban, de nincsenek hangkivetőként megjelölve. Ezek a szavak a hangkivetőként megjelölt 1097 szótóbból létrehozott összetett szavak, amelyekben a hangkivető tő az összetétel jobb oldali záró tagját adja. Az adatbázis mindösszesen 4 tövet ad meg ingadozónak (hangkivetés a megfelelő ragok előtt a Szószablya webkorpusz alapján *bajusz*: 35%, *főkabajusz*: nincs adat, *harcsabajusz*: 57%, *macskabajusz*: 25%), amelyeket jól azonosít.

Az egyes tövek ingadozását az összes rag és jel jellegű toldalék előtt megvizsgáltam, amelyek hangkivetést válthatnak ki: tárgyeset, szuperesszívusz, többes szám, birtokos személyragok. A Szószablya webkorpusz alapján megállapítható, hogy a *pityer*, *szlalom*, *vicikvacak* szavak már nem hangkivetők, míg további 116 tő tekinthető ingadozónak, mert ezeknél az esetek legalább 1%-ában a hangkivetést kiváltó toldalékok előtt nem történik meg a hangkivetés. 45 tónél ez az arány meghaladja a 10%-ot, 14-nél pedig több, mint 50%. Célszerű lenne ezeknek az adatoknak az alapján szótárunkat frissíteni, hisz 2010-es *Google*-lekérdezések alapján látható, hogy az analógiás kiegyenlítődes tovább folytatódik (pl. a *fátyolt* aránya 2003-ban 33,31%, 2010-ben már 49,95% a *fátylat* helyett). Természetesen ezen folyamatok pontos leírása és értékelése külön vizsgálatot érdemel az összes alak figyelembevételével. Adataim közt található a metatézises és a többseji magánhangzó-rövidüléssel alakok is, amelyek egyben hangkivetőként is viselkednek (*-kehely*, *-pehely*, *-teher*, *-boholy*, *-lélek* végűek).

Adataimat vizsgálataimhoz átalakítottam egy olyan írásrendszerbe, ahol egy fonémának egy betű felel meg. Az egyes alakokban itt a szóbelseji zöngésedési folyamatoknak megfelelő hangot tüntetem fel, amelyeket eredetileg az íráskép nem rögzít, így lesz a *virágcsokorból virákckor*, az adatokat azonban a magyar helyesírást követő írásmódban mutatom be. A szóalakok mögött zárójelben szereplő adatok azt mutatják, amennyiben nem jelzek mást, hogy az adott alak hányszor fordul elő a Szószablya webkorpuszban.

### 3. Hangkivető szavak jellemzése végük alapján

#### 3.1. Általános jellemzők

A hangkivető szavakat elsősorban végük alapján vizsgáltam meg, mivel úgy vélem, hogy a szavak jobb oldala jelentős hatással van arra, hogy mely szavakkal viselkednek morfofonológiailag azonos vagy hasonló módon. Ez megegyezik a legtöbb nemzetközi és magyar nyelvészeti elmélet elképzelésével, ezért külön ennek bizonyítására és indoklására nem térek ki írásomban.

A hangkivető szavakat csak a többi nem hangkivető főnévhez hasonlítottam, bár nem kizárt, hogy az analógiás kiegyenlítésben más szófajú elemek is szerephez juthatnak. Egyes esetekben külön hivatkozok arra, ha úgy tűnik, hogy az ingadozás kiváltásában egy nem főnévi szófajú szónak is szerepe lehet. Különösen sok *-el*, *-ol*, *-ez*, *-oz* végű ingadozó tő van, amelyekre hathat a nem hangkivető *-z* és *-l* képzős igék tömege, főleg, ha az alaki egyezés teljes (pl. *tegez* (főnév) vagy *tegez* (ige)). A vizsgálatban összesen 1075 hangkivető tövet vetettem össze 48464 nem hangkivető főnévvel, mert a Szószablya korpuszban csak ennyi tőhöz lehetett gyakorisági adatot találni.

A szövégek gyakoriságát, arányait a hangkivető tövek egyes szám, alany esetű, ragozatlan alakjainak gyakorisága alapján számoltam, míg a szavak gyakoriságával kapcsolatos megállapításaimat összes ragozott alakjuk alapján teszem. Ez némileg eltérhet a szótó ragozatlan alakjának gyakoriságától, mégis alkalmazhatjuk ezeket a számokat együttesen. A hangkivető szavaknak mind ragozatlan alakjairól, mind összes előforduló alakjukról pontos adataim vannak a Szószablya korpusz alapján, és a kétféle módon számított gyakoriság közt igen magas a korreláció ( $r(1202) = 0,76$ ,  $t = 40,3$ ,  $p < 0,001$ ).

Az összehasonlítást egy python nyelven megírt program segítségével végeztem el, amelynek eredményét egy módosított trie (Knuth 1997) struktúrában tároltam el. Az eredeti tárolási módot kiegészítettem azzal, hogy az egyes csomópontok alatt eltároltam az egyes szekvenciák és részszekvenciák előfordulási számát. Egy-egy új szó esetében a benne előforduló részszekvenciák számát mindig annnyival növeltem, amekkora az adott szó gyakorisága volt. Így a *fátyol* szó esetében 2047-tel (ez a szótó gyakorisága) növeltem a *lotyáf*, a *lotyá-*, a *loty-*, *lo-*, *l-* szekvenciák számát is. Ezt az összehasonlítást végrehajtottam kizárólagosan a hangkivető tövek magánhangzó- és mássalhangzó-szekvenciáit is figyelembe véve. Ez utóbbit Rebrus és Törkenczy (2008) is elvégezte. Igaz, csak az utolsó két mássalhangzóra, és eredményeiket nem súlyozták a tövek gyakoriságával.





Hangkivető főnevek analógiás keretben

fonéma	előfordulás	arány
-m	1412628	65,39%
-g	378878	17,54%
-k	183660	8,50%
-r	105764	4,90%
-ny	23048	1,07%
-l	22492	1,04%
-n	15644	0,72%
-j	12011	0,56%
-cs	4704	0,22%
-sz	1136	0,05%
-z	471	0,02%

1. táblázat

Hangkivető szavak utolsó fonémái tőgyakorisággal súlyozva

fonéma	hányszor gyakoribb a hangkivetőknél
-m	21,25
-k	2,47
-g	1,97
-cs	0,63
-r	0,49
-ny	0,30
-j	0,27
-n	0,27
-l	0,23
-sz	0,02
-z	0,02

2. táblázat

Hangkivető szavak végei a nem hangkivető főnevekhez viszonyítva tőgyakorisággal súlyozva

főnévi vég	előfordulás	arány	hangkivető vég	előfordulás	arány
-s	16465935	15,46%	-m	1412628	65,39%
-t	12153409	11,41%	-om	807301	37,37%
-r	9689570	9,10%	-lom	769148	35,60%
-g	8714415	8,18%	-alom	738340	34,18%
-a	7752718	7,28%	-em	601185	27,83%
-ó	7328275	6,88%	-lem	534034	24,72%
-ő	6970727	6,54%	-g	378878	17,54%
-ás	6166755	5,79%	-og	352272	16,31%
-és	4666957	4,38%	-log	348955	16,15%
-l	4395598	4,13%	-dalom	344027	15,92%
-i	3951793	3,71%	-delem	211143	9,77%
-et	3817903	3,58%	-adalom	203362	9,41%
-er	3539645	3,32%	-k	183660	8,50%
-ny	3531121	3,31%	-galom	163452	7,57%
-k	3365993	3,16%	-sadalom	144575	6,69%
-ég	3336174	3,13%	-talom	129228	5,98%
-m	3009250	2,82%	-odalom	113418	5,25%
-at	2819341	2,65%	-relem	109994	5,09%
-ség	2690891	2,53%	-rodalom	106663	4,94%
-n	2663468	2,50%	-r	105764	4,90%

## 3. táblázat:

Hangkivető és nem hangkivető főnevek 20 legjellemzőbb vége és azok aránya az adott szócsoporthoz tartozó töveinek összes előfordulásához tögyakorisággal súlyozva

Az 1-es és a 2-es táblázat alapján láthatjuk, hogy a hangkivető tövek legjellemzőbb utolsó fonémája kiugróan az *-m*, valamint a *-k* és a *-g* mondható még tipikusnak. Az összes többi zárófonéma nem jellemző a hangkivetőkre. E három hang nem alkot természetes osztályt. A *-g* és a *-k* egymáshoz áll közel (csak zöngességükben térnek el), míg az *-m* a sokkal ritkább *-n*-nel, *-ny*-nyel (csak képzéshelybeli eltérés) alkothatna csoportot. Legfontosabb közös jegyük az, hogy nem koronális hangok, amelyekhez az ingadozással egyik leginkább együttjáró *-t* tárgyrag kötőhang nélkül fonotaktikai okokból

nem vagy nem jellemző módon kapcsolódik. Ez is lehet egyik oka, hogy a jelenlegi szinkrón állapotban dominanciájuk megerősödött, vagy még erősebb, hisz más végű szavak könnyebben vettek, vesznek részt az analógiás kiegyenlítődésben.

Ezek az adatok azonban nem csak önmagukban, önmagukért érdekesek. Ha megnézzük alaposabban a hangkivetők egyes csoportjait, akkor azt tapasztalhatjuk, hogy az ingadozó szavak végei kevésbé jellemzőek a hangkivető tövekre. A következő részekben szisztematikusan áttekintem az ingadozás, a szavak végei és gyakorisága közti összefüggéseket.

3.2. Erősen ingadozó szavak végei (50,01-100%-ban nem hangkivető sémát követnek)

Fonéma	előfordulás	arány
-sz	5021	60,51%
-r	1105	13,32%
-l	1062	12,80%
-z	447	5,39%
-m	328	3,95%
-j	222	2,68%
-k	113	1,36%
	8298	100%

4. táblázat

Erősen ingadozó szavak (17 db) végei túgyakorisággal súlyozva.

A 4. táblázatból látható, hogy a szavak 94,69%-a olyan végekkel bír, amelyek nem jellemzőek a hangkivetőkre (nem *-k* és nem *-m* végű). Azaz az erősen ingadozó hangkivetők viselkedésükben közelítenek azokhoz a nem hangkivető szavakhoz, amelyekre végeik szerint is jobban hasonlítanak. Az *-m* és a *-k* végeket a *szlalom* (328), *ászok* (110) és *vicikvacak* (3) szavakban találhatjuk. Mind a három szó lényegesen ritkább, mint az a hangkivetőkre átlagosan jellemző, mivel az átlagos példánygyakoriság az összes ragozott alakot figyelembe véve 3890 előfordulás.

A *szlalom* jobb oldala meglehetősen hasonlít az *-alom*-ra, amely kizárólagosan jellemző a hangkivető tövekre. A nem hangkivető főnevekkel azonos viselkedésének oka az lehet, hogy a hangkivető szavak alapvetően [-idegen] jeggyel bírnak, a *szlalom* viszont jelentése és a magyarrá nem jellemző

*szl-* szókezdete miatt [+idegen] jeggyel bír. A [+idegen] jeggyel bíró szavak nem hangkivetők. A *szlalom* szó ezekre e szavakra törekszik jobban hasonlítani. (Az idegen jegy bináris megkülönböztetési jegyként való kezelése mellett elméleti szempontból nem vagyok elkötelezett, elképzelhető, hogy a jelentések is asszociatív módon ragadhatók meg. Itt csupán csak egy ilyen tulajdonság valamilyen formában lévő létezésére kívántam utalni).

Az *ászok* felépítésben az *-áo* záró magánhangzó szekvenciája mindösszesen 3,15-szer gyakoribb, mint a többi főnévnel (A minták átlaga 8,75). A lehetséges utolsó két magánhangzók közül a ritkábbak közé tartozik (13-ból a 9-ik), ezt mutatja meg a 5. táblázat.

sorrend	magánhangzók	hányszor gyakoribb a hangkivetőknél
1.	-ao	28,14
2.	-uo	22,76
3.	-oo	15,52
4.	-üö	6,47
5.	-óo	6,35
6.	-öö	4,33
7.	-io	3,87
8.	-ee	3,85
9.	-áo	3,15
10.	-ée	1,79
11.	-au	0,18
12.	-aa	0,07
13.	-ie	0,07

5. táblázat

A hangkivetők utolsó két magánhangzójának aránya a többi főnévvel összevetve tőgyakorisággal súlyozva.

A *vicikvacak* szónál a hangkivető séma elhagyásának oka a tő ritkasága, és az, hogy a hangkivetőkre egyáltalán nem jellemző *-aa* magánhangzó szekvencia utolsó helyzetben. A többi főnévre ez a szekvencia 14-szer jellemzőbb. Igaz, ezekbe beleszámítódnak azok a szavak is, amelyek *-a*-ra végződnek. A szó *-ak* vége sem tipikus, 1,08-szor gyakoribb nem hangkivető főneveknél, azaz lényegileg úgy vehető, hogy ugyanannyira jellemző a hangkivetőkre, mint a főnevekre.

3.3. 10,01-50%-ban nem hangkivető sémát követő szavak

fonéma	előfordulás	arány
-r	16450	46,45%
-l	15171	42,84%
-k	3277	9,25%
-m	252	0,71%
-z	145	0,41%
-sz	111	0,31%
-n	9	0,03%
	35415	

6. táblázat

10,01-50%-ban ingadozó hangkivető szavak (27 db) végei  
tőgyakorisággal súlyozva.

A 6. táblázat alapján láthatjuk, hogy a hangkivetőkre jellemző végek (-k, és -m) száma 9,96%-ra emelkedett, és a nagyon atipikus -sz, -z végek 0,72%-ra estek vissza. Ebből egyértelműen látható, hogy a 10,01-50%-ban ingadozó tövek közt továbbra is a nem tipikus végek a jellemzőek, de kisebb arányban, mint a hangkivetőnek sorolt, de már dominánsan nem hangkivetőként viselkedő töveknél.

Az e csoportban található tövek is ritkák. Egyedül a *sátor* (15731) és a *kebel* (10189) ugrik ki nagyobb gyakoriságával. Végeik és magánhangzó-felépítésük azonban nem tipikus a hangkivetők szempontjából, csupán az -or végszekvencia és a záró magánhangzó-szekvenciák jellemzőek kicsit jobban a hangkivetőkre, de ez nem elegendő ahhoz, hogy stabilan a hangkivetők közt maradjanak ezek a szavak. Némileg alacsonyabb gyakoriságával és végének kevésbé a hangkivetőkre jellemző hangalakjával a *kebel* kevésbé felel meg a hangkivető sémának, mint a *sátor*. Ezzel áll összhangban az a tény, hogy a *kebel* 67,3%-ban, a *sátor* azonban még 81,5%-ban viselkedik a hangkivető séma szerint.

szóvég	többi főnévhez viszonyított gyakoriság
-r	0,49
-l	0,23
-or	2,39
-el	0,17
-tor	1,02
-bel	0,79
-áo	3,14
-ee	3,85

7. táblázat  
A sátor és a kebel végei

A további tipikus hangkivetőnek mondható szavak nagyon alacsony gyakoriságúak, így esetükben nem meglepő, hogy elindultak az analógiás kiegyenlítődés útján: *gyalom* (5), *korteher* (7), *murok* (14), *bűrök* (32), *sulyom* (40), *előterem* (48). Az *előterem* szó vélhetőleg kisebb gyakorisággal fordul elő, és kevésbé is ingadozik, de alakjai nem minden esetben választathatók szét az *előteremt* ige alakjaitól (pl. *előteremt*, *előteremek*).

A további 3 tipikus végfonémára végződő gyakoribb szó is ritkább, mint a hangkivetők átlagos gyakorisága: *űröm* (159), *gyilok* (213), *vacak* (3018). A *vacak* hangalakját már elemeztük, elvárásaimmal összhangban van az a tény, hogy a nagyobb gyakoriságú *-vacak* végű szó kevésbé ingadozik (*vacak* 24,26%, *vicikvacak* 100%). Habár a *gyilok* gyakorisága önmagában sem túl magas, ingadozása annak köszönhető, hogy a magyar szlengben az eredeti 'kétélű tőr, várfal koronáján végigfutó, mellvédes (és fedett) (védő) folyosó' jelentése helyett 'gyilkolászás, mészárlás, elnyúzás' értelemben használatos. Ezt a jelentését vélhetőleg a *gyilkolás* szóhoz való hasonlósága miatt kaphatta. Természetesen ingadozó alakok előfordulhatnak a közel eredeti 'rövid pengés gyilkolóeszköz' jelentésben is: „A szamurálykard (sic!), machete és egyéb pengés gyilokok tartása leginkább egy bizonyos népcsoportra jellemző.” (<http://tinyurl.com/6zb8pgs>) Az *űröm* szó kezdődőben lévő analógiás kiegyenlítődése az *űröm* településnévnek köszönhető, amelyet (más tulajdonnévként használt rendhagyó alakokhoz hasonlóan) nem hangkivető módon szokás ragozni.

3.4. 1,01-10%-ban nem hangkivető sémát követő szavak

fonéma	előfordulás	arány
-k	84753	45,73%
-m	64708	34,91%
-l	15183	8,19%
-n	8738	4,71%
-j	4991	2,69%
-r	3830	2,07%
-ny	2823	1,52%
-g	190	0,10%
-cs	128	0,07%
	185344	

8. táblázat

1,1-10%-ban ingadozó tövek (75 db) végei tógyakorissággal súlyozva

Jól látható, hogy ebben a csoportban már a hangkivetőkre markánsan jellemző tövek szerepelnek. A *-k*, *-g*, *-m* végfonémák már összesen az alakok 80,68%-át fedik le. Azonban meg kell jegyezni, hogy ezek közt a némileg már ingadozó tövek közt a *-g*, *-k* nagyobb részt hasít ki, mint a legtipikusabbnak mondható *-m* vég. A hangkivetőkre nem jellemző *-sz* és *-z* végek már nem képviselik magukat azon szavak közt, amelyek 1,01-10%-ban nem a hangkivető sémát követik.

Még ebben a csoportban is alapvetően alacsony gyakoriságú szavak találhatók. Az átlagot gyakoriságával az *orom* (3888), *hurok* (4271), *Szentlélek* (4453), *pokol* (4962), *lepel* (5920), *berek* (5929), *vászon* (8587), *ajak* (25039), *küzdelem* (33752), *telek* (43541) szavak haladják illetve közelítik meg. A *vászon*, *pokol*, *lepel* szavak szereplése ebben a csoportban jól igazolható azzal, hogy fonológiai felépítésük alapján közelítenek a nem hangkivetőkhez, de nagy gyakoriságuk lassítja az analógiás kiegyenlítődést.

Legnehezebben a *küzdelem* szó megjelenése igazolható, ami alapvetően a *küzdelemet* szóalak magas előfordulásának köszönhető. A 360 előfordulás a 9864 *küzdelmet* alakra jut, aránya jóval meghaladja az 1% körüli küszöböt,

amellyel esetlegesen elgépelésnek vehetnénk. A nagyobb ingadozásért más, az analógiában szerepet játszó, de nem a gyakoriságon vagy a fonológiai hasonlóságon alapuló tényezők adhatnak számot.

A *berek*, *telek*, *Szentlélek* szavaknál elmondható, hogy az *-ee* és az *-ée* végszekvenciák, habár csak enyhén jellemzőbbek a hangkivetőkre, mint a többi főnévre. Ugyanez igaz az *-rk* mássalhangzó szekvenciára, igaz, az *-lk* már jellemzőnek mondható (11,38). A *Szentlélek* analógiás kiegyenlítődéshöz tulajdonnévi jellege, és az is hozzájárulhat, hogy a *-lélek* végűek a hangkivetőkön belül is egy rendhagyó, gyengébb alsémát alkotnak. Habár a *telek* egyes elemeit tekintve elég tipikus hangkivető, mégis számos olyan szó van, amely nem hangkivető, és markánsan hasonlít hozzá. Ingadozására minden bizonnyal hatással volt a sok nem hangkivetőként viselkedő vagy legalábbis ingadozó *-telek* végű településnév (pl. *Kisteleket* (10) : *Kistelket* (1), *Lakiteleket* (2) : *Lakitelket* (4), *Csanyteleket* (2) : *Csanytelket* (0), *Aggteleket* (9) : *Aggtelket* (0)). A szavak végeinek egymáshoz való hasonlóságát mérő python nyelven implementált algoritmusom szerint a *telekhez* 100 leginkább hasonlító szóból csak 28 hangkivető. A *teleknek* csupán tárgyesete és szuperesszívusza ingadozik, ami összhangban van az ingadozás korai szakaszára jellemző viselkedéssel.

A *hurok* felépítésében a *berekhez* hasonló, igaz, magánhangzói inkább jellemzőek a hangkivetőkre. Az *-uo* magánhangzó végszekvencia a hangkivetőkénél 22,76-szor gyakoribb, mint az átlagos főneveknél. Az analógiás kiegyenlítődéshöz szerepet kaphatnak a *hurokja*- kezdetű alakok, amelyek a jelentésbeli kontrasztot kísérelik meg fenntartani a *hurka* ‘húsétel’ szóval. Az *orom* szónak viszonylag kevés alakját használjuk csak. Ezekből a gyakoriak (többes szám, birtokos alakok) stabilan hangkivetők, csak az *oromot* (66), *oromom* (48) alakok tűnnek analógiásan kiegyenlítettnek, de itt feltűnő, hogy nincsen hangkivető párjuk. A szokatlan eloszlás oka az lehet, hogy ezek tulajdonképpen az *örömöt*, *örömöm* szavak ékezettelen lejegyzései lehetnek. Erre utal, hogy a biztosan az *öröm*-höz köthető alakok is nagyjából ilyen gyakoriságúak (*oromomre* (44), *oromomben* (16)), ami megerősíti ezt a hipotézist, valamint, ha a .hu domainek alatt ezekre a kulcsszavakra keresünk a *Google*-ban, akkor kizárólag ilyen ékezettelen változatokat találunk. Az *ajak* ingadozása hasonló okokra vezethető vissza, mint a *vicikvacak* esetében.



3.5. 0-1%-ban nem hangkivető sémát követő szavak

fonéma	előfordulás	arány
-m	2535228	63,07%
-g	598196	14,88%
-k	464053	11,54%
-r	219618	5,46%
-cs	107067	2,66%
-n	54241	1,35%
-ny	21233	0,53%
-j	16262	0,40%
-l	4388	0,11%
	4019836	100%

9. táblázat

0-1%-ban ingadozó tövek (973 db) végei tőgyakorissággal súlyozva.

Ezek azok a szavak, amelyek szótári besorolásuknak és a hagyományos nyelvtani leírásnak megfelelően viselkednek. 89,49%-uk a hangkivetőkre kifejezetten jellemző tövekből kerül ki. A következő csoport (-cs, -r) 8,12%-ot tesz ki. Igaz, ezek sem túlzottan jellemzőek már a hangkivetőkre. A hangkivetőkre egyáltalán nem jellemző -n, -ny, -j, -l végek csupán csak az esetek 2,39%-áért felelősek. Mivel ezek a tövek viselkednek látszólag elvárásaimmal ellentétesen, ezért ezeket vizsgálom meg alaposabban.

Az -l végű *hólepel* (4), *bélpokol* (4), *birkaakol* (5) szavaknál az ingadozás hiánya elsősorban annak tudható be, hogy nagyon kevés adat szerepel róluk a korpuszban. 2010 február 12-én .hu domain alatt végrehajtott Google-lekérdezés eredményeképpen tárgyrageiaikról ezeket az adatokat kaptam, amelyek alapfeltételezéseimet megerősítik:

*hóleplet* (258 találat) : *hólepelt* (419 találat)

*bélpoklot* (38 találat) : *bélpokolt* (1 találat)

*birkaaklot* (39 találat) : *birkaakolt* (8 találat)

Egyedül a *bélpokol* nem ingadozik, amely csak nagyon kevés alakban fordul elő. A hangkivetők csoportjában a nagyon gyakori, hangkivető formából képzett *bélpoklos* (23 300 találat) tarthatja analógiás alapon.

Az *ököl* (4125) szót viszonylagosan nagy gyakorisága tarthatja ebben a csoportban. A *tengeröböl* (250) ingadozásának hiánya véletlenszerű adathiány lehet. 2010 február 13-án .hu domain alatt végrehajtott Google-lekérdezésre már mutat ingadozást: *tengeröblöt* (1600 találat), *tengeröbölt* (95), ami már viszonylag jelentősebb, 5,9%-os ingadozás, bár még így is elmarad a legtöbb *-öböl* végű szótól.

-ly végű szavak, amelyek nem ingadoznak: *szappanpehely* (4), *burgonya-pehely* (15), *kukoricapehely* (119), *zabpehely* (348), *hópehely* (895), *pehely* (985); *misekehely* (3), *virágkehely* (94); *boholy* (189); *státusfogoly* (14), *államfogoly* (20), *hadifogoly* (2207), *fogoly* (11261); *lángbagoly* (1), *macskabagoly* (25), *gyöngybagoly* (39), *hóbagoly* (43).

A *-pehely*, *-kehely*, *-boholy* végű szavak azért nem ingadoznak, mert habár a metatézissel létrehozott alakok típusgyakorisága igen alacsony (3 szabad alakot: *teher*, *pehely*, *kehely* és egy kötött alakot: *boholy* és ezek származékait számlálják), a *teher* gyakorisága (55913) olyan magas, hogy a mintát életben tartja. Egyedül a *kehely* szó némileg erősebb ingadozása (1,3%) lehet rendellenes ebben a helyzetben, ami a *kehelyt* alak gyakoribb előfordulásának köszönhető. Ez a szóalak gyakran fordul elő archaikus, irodalmi közegben, ahol sajátos hangulatot közvetít, másrészt hatással lehet létezésére a protestáns liturgia mondata úrvacsorakor: „isszátok e pohárt”, amikor valójában kelyhet tartanak fel, és a *pohár* szót sem a köznyelvi *poharat* alakjában használják, ezzel áttételesen erősítik a *kehelyt* alakot.

A kisebb gyakoriságú *-fogoly* végű alakok a nagy gyakoriságú *fogoly* mellett összetartó csoportot alkotnak. Stabil viselkedésük nem meglepő, főleg, hogy az *-oo* (15,52-szer gyakoribb, mint a többi főnévnel), és a *-gj* (8,22-szer gyakoribb) végszekvencia meglehetősen jellemző a hangkivetőkre. A hasonló felépítés segíthet a *-bagoly* végű származékoknak is hangkivető voltuk megőrzésében, amelyeket a nagyon jellemző *-ao* (28,13) végszekvencia még inkább e csoportban tart. Magyarázatra szorul azonban, hogy a *bagoly* (7,3%-ban nem követi a hangkivető sémát) és a *fülesbagoly* (2,7%-ban nem hangkivető) miért ingadoznak ezeknél nagyobb mértékben, különösen, hogy a *bagoly* a legnagyobb gyakoriságú közülük. A *bagoly* esetében ennek az az oka, hogy a *bagoly* szóalakjai közt a jobban ingadozó tárgyeset nagyobb arányban van képviselve, mint a többi alakra leginkább jellemző többes szám:

szótó	többes szám	tárgyeset	arány
<i>bagoly</i>	898	440	2,04
<i>fülesbagoly</i>	62	12	5,17
<i>macskabagoly</i>	14	11	1,27
<i>gyöngybagoly</i>	34	4	8,50
<i>hóbagoly</i>	28	9	3,11

10. táblázat

*-bagoly* végű szavak többes számának és tárgyesetének aránya

A *lángbagoly* egyetlen alakja, a *lángbaglyunk* várható módon hangkivető alakú. Mint láthatjuk, az elvárt tendenciából kilóg a *fülesbagoly*, amelynek tárgyragos alakja kevésszer fordul elő, mégis jobban ingadozik. Igaz, csak elenyésző mértékben. A *macskabagoly*-nak pedig hiába van sok tárgyas alakja, mégsem ingadozik. Ez a tendencia azonban változóban van, bármi is lehetett az oka (Google-lekérdezés .hu domain alatt 2010 febr. 13.: *macskabagolyt* (805): *macskabagolyt* (57)).

Elképzeltető, hogy a *bagoly* alak azért is nagyobb mértékben ingadozó, mert a többi *-bagoly* végű összetett szó inkább a művelt, természettudományos párbeszédre jellemző, amelynek folytatói vélhetőleg a műveltebb rétegekből származnak. Általánosságban elmondható, hogy ennek a társadalmi csoportnak a tagjai hajlamosabbak a rendhagyó sémák „tiszteletére”.

Az *-n* végű alakok közt a leggyakoribb *haszon* (50002) elvárásaimnak megfelelő módon nem ingadozik. A belőle létrehozott összetett szavak az ő megtartó hatása miatt nem ingadoznak. A további *-n* végű szavak közt *-vászon* végű szavak találhatók, habár maga a leggyakoribb *-vászon* végű szó, a *vászon* (8587) 2%-ban ingadozik már. Néhány *-vászon* végű összetett szó pedig ennél is jobban ingadozik: *zsákvászon* (22,22%), *csiszolóvászon* (7,69%), *lenvászon* (2,98%), *gyöngyvászon* (2,22%). Az 1%-nál kevésbé ingadozó szavak az alábbiak:

*szitavászon* (1), *kartonvászon* (1), *házivászon* (2), *lepedővászon* (3), *sátorvászon* (3), *kendervászon* (7), *pamutvászon* (12), *viaszosvászon* (32), *festővászon* (44), *vetítővászon* (503), *mozivászon* (1260), *filmvászon* (2331).

A két csoport tagjai közt egyértelmű különbség nincs, de az elmondható, hogy a második csoport leggyakoribb szavaihoz, a *filmvászonhoz* (2331) és a

*mozivászonhoz* hasonlókat találhatók ebben a csoportban. Inkább műveltségi és inkább hosszabb szavak, míg a *vászon* körül a munkásélethez kötődő rövid kezdőtagú összetett szavak csoportosulnak. Ez alól csak a *csiszolóvászon* a kivétel, de annak kötődése a kétkezi munkához még szorosabb, mint a többi ide tartozó szóé. A műveltségi *vászon* szavak a *-bagoly* végű szavak használatához hasonló okokból kifolyólag ingadozhatnak kevésbé.

Az *-ny* végű szavak egy kivétellel a nagy gyakoriságú *torony-ból* (16922) létrehozott összetett szavak. Viselkedésük elméleti keretrendszerrel teljesen összhangban van, hisz a nagy gyakoriság miatt a *torony* szónak stabilan hangkivetőnek kell maradnia, a hozzá nagyon hasonló szavak számára pedig ő a minta. Az egyetlen más eredetű szó, az *ékhorony* (1) pedig elegendő mértékben hasonlít ezekre, hogy megőrizze hangkivető viselkedését. Meg kell jegyezni, hogy vannak olyan *-torony* végű szavak, amelyek jobban ingadoznak 1%-nál, ez arra utalhat, hogy a *torony* nagy analógiás erővel vonzza magához a belőle képzett összetett szavakat, de ereje már nem abszolút érvényű. Ezek a szavak: *hústorony* (4,55%), *órátorony* (4,35%), *hűtőtorony* (1,69%), *víztorony* (1,2%). A hangkivető sémát stabilan követő *-torony* végű alakok az alábbiak:

*őrtorony, harangtorony, lövegtorony, adótorony, aknatorony, antennatorony, bástyatorony, elefántcsonttorony, erkélytorony, fiatorony, figyelőtorony, fűrőtorony, huszártorony, kilátótorony, kutatótorony, lakótorony, lőportorony, öregtorony, páncéltorony, rádiótorony, silótorony, televíziótorony, templomtorony, tetőtorony, ugrótorony, várostorony, vartorony, világitótorony.*

Nehéz egyértelműen besorolni e szavakat jelentésük alapján, de a kevésbé ingadozó szavak jelentésükben és formájukban közelebb állnak a ‘torony’ fogalom prototípusához. Legtöbbjük történelmi, hadi vagy hírközlési kontextusból való. Az eredeti szótól legjobban az emberre alkalmazható *hústorony* tér el. Ennek megfelelően ez is ingadozik a legjobban.

#### 4. Konklúzió

A hangkivető tövek végeinek alapos vizsgálata során azt tapasztaltuk, hogy az ingadozás szorosan összefügg a gyakorisággal és a formával. Ha eltávolodunk a részletektől, és az adatok egészére tekintünk, akkor ez az összefüggés még markánsabban megmutatkozik, az ingadozás mértéke monoton csökken az átlagos gyakorisággal. Ezt az összefüggést mérsékelt

(a korreláció mértéke 0,16), akkor is ki tudjuk mutatni, ha az egyes szavakat gyakoriságát és hangkivetésének mértékét vetjük össze ( $t = 2.27$ ,  $df = 204$ ,  $p < 0,05$ )<sup>1</sup> az összes alapszó vonatkozásában.

Nem hangkivető viselkedés	Átlagos gyakoriság
50,01-100%	488
10,01-50%	1311
1,01-10%	2469
0-1%	4131

11. táblázat

Hangkivető tövek átlagos gyakorisága ingadozásuk mértéke szerint

Ha csak az egyedi tövek gyakoriságát és ingadozását nézzük, a nagyobb csoportokat figyelmen kívül hagyva, akkor nem tapasztalható korreláció. Ez azzal lehet összefüggésben, hogy a hangkivetők fonológiai felépítésüket tekintve meglehetősen heterogén csoportot alkotnak. A gyakoriság önmagában nem felelős az ingadozásért. Azt azonban láthattuk, hogy ha a formai hasonlóságot is figyelembe vesszük, akkor a gyakorisági hatások már markánsabban megjelennek.

Elemzésemben mindvégig kizárólagosan csak felszíni alakokra hivatkoztam alapvetéseimnek megfelelően, azonban ezek is elégségesnek bizonyultak, hisz a hangkivető tövek viselkedése jól leírható pusztán hasonlósági és felszíni alapon. A szövegeknek nagy szerepe van abban, hogy egy-egy szó miképp viselkedik, sőt használati körülményei, jelentései is meghatározzák ezt. Habár tendenciák felfedezhetők a szövegek és a viselkedés közt, a gyakoriságot és a változatos hasonlósági tényezőket figyelmen kívül hagyva ezek a viselkedési mintázatok leírhatatlanok lennének.

Az egyes hangkivető és nem hangkivető szavak viselkedését az összehasonlítást jelentősen megkönnyítő trie struktúrákkal vettem össze, amelyeknek köszönhetően olyan összefüggéseket ismerhettünk fel, amelyek a korábbi

<sup>1</sup> A számításban az egyes alapszavak szóbokraikkal együtt szerepelnek, így a *pocok* szó ingadozását és gyakoriságát 564-nek és 3,5%-nak vettem, amely összegzi a *pocok* és a *pézsma-pocok* előfordulásait.

leírásokban nem jelentek meg, vagy azok kategorikus voltából kifolyólag túlzott szerephez jutottak. Az összehasonlításokból kiderült, hogy a hangkivető tövek kiegyenlítődésében és vélhetőleg a nyelvi változásban is a csoporthoz való hasonlóságnak, csoportthatásnak (gang effect) különösen nagy jelentősége van.

2010-es *Google*-lekérdezéseket is bevonva a vizsgálatba azt tapasztaltuk, hogy a változatok, az ingadozás leírása statikusan lehetetlen, a szinkrón állapot megragadhatóságának elképzelése csak tudományos illúzió, amelyet feladva sokkal érdekesebb az általános nyelvi összefüggésekre, a mögöttes mozgatórugókra figyelni. Leírásaimból az is jól látható, hogy a nyelvészet adatközpontú megközelítése jelentősen átrajzolhatja a most általános nyelvi elméleteket, de a nyelvi elemek egyszerű deskriptív leírásai is jelentősen módosulhatnak ezáltal.

## 5. Szakirodalom

- Benua, Laura 1995. Identity effects in morphological truncation. In J. N. Beckman – L. Walsh Dickey – S. Urbanczyk (szerk.) *Papers in Optimality Theory*. Amherst MA, GLSA. 77–136.
- Benua, Laura 1997. *Transderivational Identity: Phonological Relations between Words*. PhD-disszertáció. UMass at Amherst.
- Bybee, Joan L. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In Michael Barlow – Suzanne Kemmer (szerk.) *Usage-based Models of Language*. Stanford CA, CSLI. 65–85.
- Bybee, Joan L. 2001. *Phonology and Language Use*. Cambridge, Cambridge University Press.
- Bybee, Joan L. 2002. Phonological evidence for exemplar storage in multiword sequences. *Studies in Second Language Acquisition* 24: 215–221.
- Bybee, Joan L. 2010. *Language, Usage and Cognition*. Cambridge, Cambridge University Press.
- Elekfi László 1994. *Magyar ragozási szótár. Dictionary of Hungarian Inflections*. Budapest, MTA Nyelvtudományi Intézet.
- Goldberg, Adele 1995 *Constructions. A Construction Grammar approach to argument structure*. Chicago, University of Chicago Press.
- Halácsy Péter – Kornai András – Németh László – Rung András – Szakadát István – Trón Viktor 2003. A Szószablya projekt. In *MSZNY 2003*. 299.

- Itkonen, Esa 2005. *Analogy As Structure And Process*. Amsterdam. John Benjamins Publishing Co.
- Kálmán László 2008. *A mögöttes és ami mögötte van*.  
 [http://www.szv.hu/cikkek/a-mogottes-es-ami-mogotte-van]
- Knuth, Donald 1997. *The Art of Computer Programming. Vol 3: Sorting and Searching*. Second Edition. Addison-Wesley. Section 6.3: Digital Searchin, page 492.
- Kraska-Szlenk, Iwona 2007. *Analogy. The Relation between Lexicon and Grammar*. München, Lincom.
- Papp Ferenc 1975. *A magyar főnév paradigmatis rendszere*. Budapest, Akadémiai Kiadó.
- Philips, Betty S. 2006. *Word Frequency and Lexical Diffusion*. New York, Palgarve Macmillan.
- Rebrus Péter 2000. Morfofonológiai jelenségek. In Kiefer Ferenc (szerk.) *Strukturális Magyar Nyelvtan 3. Morfológia*. Budapest, Akadémiai kiadó. 763–947.
- Rebrus Péter – Törkenczy Miklós 2008. Morfofonológia és a lexikon. In Kiefer Ferenc (szerk.) *Strukturális Magyar Nyelvtan 4. A szótár szerkezete*. Budapest, Akadémiai kiadó. 683–786.
- Rung, András 2008. Determining word similarity in the Hungarian language. In *Papers from the Mókus Conference*. Budapest, Tinta Kiadó. 112–118.
- Skousen, Royal 1989. *Analogical Modeling of Language*. Kluwer Academic Publisher.
- Skousen, Royal – Deryle Lonsdale – Dilworth B. Parkinson 2002. *Analogical Modeling*. Amsterdam/Philadelphia, John Benjamin Publishing Company.
- Steriade, Donca 2000. Paradigm uniformity and the phonetics/phonology boundary. In J. Pierrehumbert – M. Broe (szerk.) *Papers in Laboratory Phonology 6*. Cambridge, Cambridge University Press,
- Törkenczy Miklós – Siptár Péter 2000. Magánhangzó~semmi váltakozások a magyarban. *Nyelvtudományi Közlemények 97*: 64–130.
- Trón, Viktor – Péter Halácsy – Péter Rebrus – András Rung – Péter Vajda – Eszter Simon 2006. Morphdb.hu: Hungarian lexical database and morphological grammar. In *Proceedings of 5th International Conference on Language Resources and Evaluation*. ELRA. 1670–1673.
- Vago, Robert M. 1980. *The Sound Pattern of Hungarian*. Washington, Georgetown University Press.