

# A (PRE)MODÁLIS TARTOMÁNYOKKAL ÖSSZEFÜGGŐ FŐNÉVI IGENEVES MINTÁZATOK KLASZTEREZÉSÉNEK LEHETŐSÉGE A TÖRTÉNETISÉG PERSPEKTÍVÁJÁBAN<sup>1</sup>

BAJZÁT TÍMEA BORBÁLA


## 1. Bevezetés

A tanulmány célkitűzése egy olyan klaszterezési eljárás mód tapasztalatainak rögzítése történeti nyelvi anyagok tükrében, amely már előzetesen mai magyar nyelvi korpuszadatokon [a Magyar Nemzeti Szövegtár (MNSZ v2.0.5)] alkalmazásra került (vö. Bajzát 2023). A dolgozatban bemutatott vizsgálat a premodális tartományokkal összefüggő főnévi igenes szerkezetek elemzésére és mintázatalapú csoportosítására fókuszál számítógépes nyelvészeti eszközökkel, elméleti keretét tekintve pedig a funkcionális kognitív nyelvészeti (Langacker 1987, 1991, 2008; Tolcsvai Nagy 2017c) kiindulópontot érvényesíti. A funkcionális kognitív szemléletmód választása több aspektusból is motívált. Egyfelől támogatja a nyelvi adatok használat alapú, holisztikus megközelítését, másfelől pedig elutasítja a szinkronia és a diakronia éles elválaszthatóságát (Tolcsvai Nagy 2017a: 52). A kutatásnak ugyan a segédige/predikatív melléknév + főnévi igenév szerkezetek vizsgálata áll a fókuszában, hangsúlyozandó az, hogy a dinamikus jelentéskonstruálás (Tolcsvai Nagy 2017b: 123–126) során az absztrakció magasabb fokán kezelhető mintázatsémák a nyelvi környezetek komplex, kvalitatív vizsgálatával válhatnak hozzáférhetővé.

A kutatás a jelen szakaszában a Történeti magánéleti korpuszból (TMK <http://tmk.nyttud.hu/> vö. Dömötör–Gugán–Novák–Varga 2017; Novák–Gugán–Varga–Dömötör 2018) és a Magyar Nemzeti Szövegtár 2-ből (MNSZ2 v2.0.5. <http://clara.nyttud.hu/mnsz2-dev/> vö. Oravecz – Váradi – Sass 2014) vételezett minták klaszteranalízissel való feldolgozására fókuszál, s enged betekintést az eredmények összehasonlító elemzésébe. A vizsgálatban a manuális elemzéssel rögzített tulajdonság-együttállások (szemantikai és morfoszintaktikai jegyek rögzítésével) automatizált klaszterezése a PAST szoftverben (Hammer–Harper–Ryan 2001) valósult meg, azért, hogy a példányok által reprezentált konstrukciós mintázatok elérhetővé váljanak, és visszacsatolást kapjunk arra vonatkozóan, hogy az eljárás mód alkalmazása hogyan hasznosítható a diakrón irányú vizsgálatokban.

A mintavételezésnél törekedtem az intuíción minimalizálására. Ebből kifolyólag a vizsgálandó szerkezetek meghatározása korpuszalapon történt, indirekt eljárási módszerrel (vö. Bajzát 2023) az MNSZ2 főnévi igenévi kifejezéseinek gyakorisági listája alapján (lásd 2. fejezet). A mintavételezéskor olyan főnévi igenes szerkezetekhez kértem le

---

<sup>1</sup>„Az Innovációs és Technológiai Minisztérium ÚNKP-21-3 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.” 

konkordanciákat, amelyek a dinamikus jelentésfeldolgozás során valamilyen, az elsődleges figura/megnyilatkozó belső, mentális állapotait (pl. TUDÁS, VÁGY, HAJLANDÓSÁG, AKARAT) teszik hozzáférhetővé. A tanulmányban ezentúl bemutatom azt a szempontrendszert, amelyet mind a mai magyar nyelvi adatok elemzésében (vö. Bajzát 2023), mind pedig a diakrón korpuszokból (TMK) vett konkordanciák feldolgozásában alkalmaztam (néhány szükséges módosítással). Az elemzési szempontok rögzítésénél törekedtem arra, hogy a szemantikai jegyek (ontológiai csoportosítások) vizsgálata együttesen történjen meg a morfoszintaktikai és szintaktikai jellemzők azonosításával, hiszen feltételezhető, hogy ezek együttesen dolgozzák ki, s teszik értelmezhetővé a nyelvilleg aktivált szereplők és azok körülményeinek viszonyrendszerét a premodális–modális tengelyen. A detektált jellemző-együttállásokat adattáblába felvéve, majd a klaszteranalízis módszerét alkalmazva rajtuk, olyan módon lehet osztályozni az adatok csoportjait, hogy azok valószínűsítést tegyenek a funkcionálisan összetartozó kifejezőmódra, s a vizsgált nyelvallapotok közötti eltérést is tükrözzék a csoportok típusaiban. Egyfelől az előbb részletezett módszerrel körülhatárolt mintázatsoportokat van lehetőségünk azonosítani a premodális–modális határsáv szimbolikus viszonyai szerint, másfelől lehetőség nyílt arra, hogy a vizsgált szerkezetek nyelvi profiljába, mintázatokba való szerveződésük változásaiba kaphassunk betekintést.

A tanulmányban bemutatott megközelítésmód és módszertani eljárás reményeim szerint hozzájárul ahhoz, hogy a modalitás kifejezése konstrukciós mintázatainak alakulástörténetére vonatkozóan újabb adalékok kerüljenek felszínre a nyelvészeti diskurzusban.

## **2. A dolgozat által érvényesített elméleti kiindulópontok, összefüggésben a modalitással**

A jelen tanulmány nem vállalkozik arra, hogy a teljesség igényével rögzítse a nemzetközi és a hazai tudományos diskurzusokban megjelenő szemléletmódokat a modális szemantika területéről, azonban igyekszik röviden betekintést nyújtani abba, hogy a jelen részletezett vizsgálat milyen elméleti alapfelvetésekkel közelít a kutatási célkitűzéseinek megválaszolásához. Kugler (2017a: 469, 480) alapján a modalitás olyan mondatsemantikai jelenségeként határozható meg, amellyel a diskurzusszituációban a megnyilatkozó fél konstruáló szubjektumként kifejezésre juttatja a saját mentális műveleteit, ezáltal jelöltté teszi azt, hogy az általa nyelvilleg megkonstruált jelenet egy lehetséges szituáció.

A modalitás vizsgálatában elterjedt nézőpont, hogy a szükségszerűség és a lehetőségesség fogalmi dichotómiájának tengelyéről teszik értelmezhetővé a modális jelentéstartalmakat a kategorizációs kísérletekben (vö. van der Auwera–Plungian 1998: 80; Kiefer 2000, 2005). A szereplő szempontjából a belső lehetőségesség/szükségesség azt jelenti, hogy a szereplő tulajdonságaitól függ a potenciális esemény megvalósulásának lehetősége, ez tehát a diszpozicionális kategóriának felel meg. A szereplő részéről külső hatóerő mind a cirkumsztanciális, mind a deontikus tartománynak jellemzője. Ilyenkor a szereplő erődinamikai viszonyba lép a fizikai és a társas környezettel. A jelenetből kiindulva külső lehetőségesség és szükségesség az episztemikus tartománnyal kapcsolódik össze, és nem szereplői, hanem konceptualizálói (megnyilatkozóí, ill. beszélői) nézőpontot feltételez. Fontos kiemelni, hogy van der Auwera és Plungian (1998: 84) nem tárgyalja modális szemantikai tér részeként a volitív (3) (volitív modalitás lásd Palmer: 1986:

115, 116–119, 152–153) és a buletikus (buletikus vö. Kiefer 2000: 298) tartományt, viszont elismerik annak szoros kapcsolódását a modális szemantikai térhez (van der Auwera–Plungian 1998: 86). A modalitásfajták a beszédhátterüket tekintve mutatnak egymástól eltérést. A deontikus, a cirkumsztanciális, a buletikus és a diszpozicionális modális beszédháttereket jellemzően dinamikus modalitásnak nevezi a szakirodalom (Kiefer 2000: 298), ezzel ezeket a szereplőorientált modalitásokat (vö. van der Auwera–Plungian 1998) kategorikusan megkülönböztetve az episztemikus modalitástól (vö. Kiefer 2000: 298), amely beszélőorientált. A dinamikus modalitásnak azonban ennél létezik egy szűkebb megközelítésmódja is, amely szerint csak a diszpozicionális, cirkumsztanciális és részben a buletikus tartományok azok, amelyek ebben a kategóriában kezelhetők (Van linden 2010). Kiefer fenntartja az objektív-szubjektív kettősséget is (2005: 74). Ennek alternatíváját a funkcionális kognitív megközelítés a konstruálás módjában látja (lásd Langacker 1987: 128–132, 2008: 77, 260–264).

A premodális (vö. Van linden 2010: 9–10) és az alapmodális (dinamikus modalitás és a deontikus modalitás egyszerre) szemantikai tartományok közötti határvonalat nem lehet élesen megállapítani, a prototípuselvű és kontinuumelv szerint való szerveződés lehetőségét fenn kell tartani. A premodális tartomány(ok) esetében fontos jellemző az, hogy a megnyilatkozás a jelentésképzés során olyan állapotokhoz kapcsolódik, amely önmagában hordozza egy folyamat megvalósulásához kötődő feltételeket, ezért implikálja a folyamat bekövetkeztének lehetőségét (lásd (1a)–(1b)).

(1a) Peti tud zongorázni.

(1b) Peti képes zongorázni.

Az (1a)–(1b) jelentésének a része, hogy Peti rendelkezik olyan TUDÁS-sal, KÉPESÉG-gel, amely elengedhetetlen a ZONGORÁZÁS-hoz (Kiefer 2005: 96–111, 2000: 337–339), ugyanakkor a (7a) és a (7b) mondatok magát a TUDÁS-t, KÉPESSÉG-et figyeltetik meg. A jelentésnek a fenti példákban azonban nem egyformán része az, hogy az elsődleges figura zongorázik (szokott zongorázni). A két példában maga a TUDÁS és a KÉPESSÉG van a figyelem előterében, nem a Petiben rejlő képesség megnyilvánulása, a zongorázás. Ugyanakkor a TUDÁS és a KÉPESSÉG megkonstruálása alapján következtetünk arra, hogy megvan a lehetősége annak, hogy Peti használja is a megszerzett tudást.

A modális szemantikai térbe való bejutásnak a premodális tartományból az az alapfeltétele, hogy a nyelvilleg kidolgozott elemi jelenetben a folyamat és azok megkonstruált erőviszonyai emergens kapcsolatot hozzanak létre magával a potencialitással. A premodális tartományban – a világról való tudásunk alapján – az adott konstelláció inherens részének tartjuk a szükségességet (Nuyts 2006: 3; Van linden 2010: 4). A modális tartományba való átlépés akkor történik, amikor a folyamat potencialitása már nem az így értelmezett inherens viszony, hanem 'lehetséges' vagy 'szükséges' jelentés a jelenet erődinamikai konstrukciójának az értelmezésében jön létre. Langacker (1991: 272) ezt a potencialitást modellálja kiterjesztett realitásként, és a kiterjesztés műveletében ragadja meg a modális konstruálás lényegét.

### 3. A mintavétel

A koncepció a szerkezetek kiválasztásánál az volt, hogy olyan főnévi igeneves kompozitumokat vizsgáljak, amelyek a jelentésképzés művelete során valamilyen mentális, az elsődleges figura szempontjából belső állapottal asszociálódnak a jelentésképzésben (pl. VÁGY, AKARAT, SZÁNDÉK, KÉPESSÉG). Az ehhez a kategóriákhoz kapcsolódó konstrukciók meghatározásánál szükséges volt egy olyan fogalmi készlet összeállítása, amely elemeinek környezetében előfordulhat olyan segédige, illetve olyan melléknév, amely az elsődleges figura mentális állapotát jelöli (a fogalmi készlet összeállításáról vö. Bajzát 2023).

Az alábbi felsorolás ismerteti azokat a komponenseket, amelyek a keresőkifejezésekkel kompozitumot alkottak, és megfeleltek a választott szemantikai kritériumoknak, tehát az elsődleges figura nézőpontjából belső állapotot reprezentálnak, a zárójel az MNSZ2-ben való előfordulási értéküket mutatja: *tud* (1 317 457); *akar* (612 684); *szeret* (487 039); *kíván* (192 733); *képes* (135 704); *mer* (63 905); *képtelen* (48 840); *hajlandó* (48 449); *bír* (22 218); *szándékozik* (9023); *remél* (6580); *óhajt* (5508); *imád* (5135); *hajlamos* (4552); *vágyik* (1681); *utál* (1474); *vonakodik* (1167); *gyűlöl* (191); *vágyakozik* (35) (Bajzát 2023).

A jelen vizsgálatba a gyakoriság szerinti első négy szerkezettagot vontam be (*tud*, *akar*, *szeret* és *kíván*). Két korpuszból vételeztem a mintáimat: a Történeti magánéleti korpuszból és a Magyar Nemzeti Szövegtárból. A Történeti magánéleti korpuszban az alábbi kifejezéssel kerestem a vizsgált segédige + főnévi igenév szerkezetekre: L~*tud*@C~Vfin .. Inf; L~*akar*@C~Vfin .. Inf; L~*szeret*@C~Vfin .. Inf; L~*kíván*@C~Vfin .. Inf. Az MNSZ2-ben először a részletes keresés funkcióval a korpuszban előforduló összes főnévi igenév előfordulásra indítottam szűrést, majd az így kapott konkordanciát tovább szűrtem a fent megadott segédigei komponensekre a nódusztól 2L–2R keresési ablakban.

A mintavételezés után manuálisan távolítottam el a konkordanciákból azokat az eseteket, amikor az adatok nem az általam keresett szerkezetet tükrözték, illetve a duplumokat. Az így kapott listákból reprezentatív elemszámú (Krejcie–Morgan 1970 alapján) mintát vettem (1. táblázat), az Excel táblázatkezelő program =VÉL() függvényét alkalmaztam erre a feladatra a TMK esetében. A táblázat 3. oszlopa tartalmazza a vizsgálat során elemzett minták méretét.

Kompozitumszerkezet	Teljes konkordancia az adatok kézi ellenőrzése után (db)	Az elemzéshez előkészített minta mérete (db)
<i>tud</i> + NI	1383	302
<i>akar</i> + NI	1570	310
<i>szeret</i> + NI	28	28
<i>kíván</i> + NI	278	162
<b>Összesen</b>	<b>3259</b>	<b>802</b>

1. táblázat. Mintavétel a TMK-ból

A Magyar Nemzeti Szövegtárból eleve a reprezentatív mintavételnek megfelelő elemszámú konkordanciát nyertem ki a „Véletlen minta” funkció alkalmazásával, azonban az adatok kézi ellenőrzését így sem tudtam elkerülni (2. táblázat). A táblázat 3. oszlopa tartalmazza a vizsgálat során elemzett minták méretét.

Kompozitumszerkezet	A konkordancia mérete (db)	Elemzéshez előkészített minta mérete (db)
<i>tud + NI</i>	500	488
<i>akar + NI</i>	384	378
<i>szeret + NI</i>	384	359
<i>kíván + NI</i>	384	370
<b>Összesen</b>	<b>1652</b>	<b>1595</b>

2. táblázat. Mintavétel az MNSZ2-ből

#### 4. Az elemzési módszer ismertetése

##### 4.1. Az elemzéshez alkalmazott szempontrendszer

Az elemzési metódus során a klaszteranalízis módszerét alkalmaztam azért, hogy a tulajdonság-együttállások, relációs mintázatok hozzáférhetővé váljanak a vizsgálatban. Ahhoz, hogy a jellemző-együttállások szisztematikusan feldolgozhatók legyenek az automatikus klaszterezésben, egy olyan elemző szempontrendszer kialakítása volt szükséges, amellyel a szemantikai szempontok és a morfoszintaktikai jellemzők együttesen kezelhetővé válnak, továbbá (szintén egyidejűleg) az elemzés a nyelvi aktívált és kidolgozott szereplők (Paradis 2005), valamint a körülmények kategorizációjával is kiegészüljön (lásd. 3. táblázat) (klaszteranalízis alkalmazása vö. Bajzát 2023).

Szempontok	(A)kategóriák
1. Mondattípus	1. Direkt kérdő mondat
	2. Független kérdő mondat
	3. Direkt felszólító mondat
	4. Független felszólító mondat
	5. Egyéb főmondat
	6. Egyéb mellékmondat
2. <i>Ha</i> kötőszó	7. <i>Ha</i>
3. Az elsődleges figura ontológiája (Paradis 2005 alapján) élőség szerint	
3.1. Élő elsődleges figura	8. Humán, konkrét
	9. Nem humán, de élő, konkrét
	10. Metonimikusan konstruált
3.2. Nem élő elsődleges figura	11. Konkrétum
	12. Absztraktum
	13. Topikfolytonosság

4. Az elsődleges figura nyelvi kidolgozása abban az esetben, ha az nem főnévvel történik	14. Névmási koreferencia
	15. Sematikus névmási kidolgozás személyes névmással
	16. Határozatlan névmási kidolgozás
	17. Kidolgozás általános névmással
5. A beszédpartnerhez való lehorgonyzódás	18. Beszédpartnerhez való lehorgonyzódás
6. A kopolán vagy a segédigén kidolgozott inflexiói toldalek szám- és személybeli lehorgonyzása	19. E/1.
	20. E/2.
	21. E/3.
	22. T/1.
	23. T/2.
7. A segédigén vagy a kopolán jelölt grammatikai mód	24. T/3.
	25. Kijelentő mód
	26. Feltételes mód
	27. Konjunktívusz
8. A segédigén vagy a kopolán jelölt grammatikai idő	28. Felszólító mód
	29. Múlt idő
9. A főnévi igenév által aktivált folyamattípus kategorizációja a WordNet ontológia alapján (vö. Fellbaum 1998; Miller 1995)	30. Jelen idő
	31. <body> (TESTI FOLYAMATOK)
	32. <change> (VÁLTOZÁS)
	33. <communication> (KOMMUNIKÁCIÓ)
	34. <competition> (VERSENGÉS)
	35. <consumption> (FOGYASZTÁS)
	36. <contact> (ÉRINTKEZÉS, KAPCSOLAT)
	37. <creation> (ALKOTÁS)
	38. <emotion> (ÉRZELEM)
	39. <motion> (MOZGÁS)
	40. <perception> (ÉRZÉKELÉS)
	41. <social> (TÁRSAS/TÁRSADALMI ESEMÉNYEK)
	42. <possession> (BIRTOKLÁS)
	43. <stative> (ÁLLAPOT)
44. <cognition> (KOGNÍCIÓ)	
45. <weather> (IDŐJÁRÁS)	
10. Azonosíthatóvá válik-e a megnyilatkozó részeről valamilyen beszédcselekvés foratókönyvi sémája?	46. közvetett beszédaktus

**3. táblázat.** A nyelvi elemzésre alkalmazott szempontok (Bajzát 2023)

A mondat típusok szisztematikus megkülönböztetése (1. szempont) azért volt szükséges, mert azok tipikusan összekapcsolódnak a beszédcselekvéssel, nyelvi viselkedést jellemző, konvencionalizálódott sémákkal (Imrényi 2017: 763). A *ha* kötőszó jelölését az indokolja (2. szempont), hogy ez a kötőszó hipotetikus mentális teret képes aktiválni a jelentésfeldolgozás során, prototipikusan a főmondati részben konceptualizált elemi jelenet megvalósulásához kapcsolódó feltételt vezet be (Kugler 2017: 852), tehát keretként képes a figyelem előterébe hozni szerkezet a potencialitást 'ha..., akkor' viszonyban.

A példányok által kidolgozott elsődleges figura ontológiai kategorizációja szintén megtörtént (3. szempont). Ez azért indokolt, mert feltételezhető az, hogy az elsődleges figura a különböző jelentéskiterjesztések (lásd Tolcsvai Nagy 2017b: 255–261; Kövecses–Benczes 2010) létrehozásában is szerepet kapnak az absztrakt segédige/predikatív melléknév + főnévi igenév sémák esetében. Egyfelől a szereplők azonosítása hozzájárul a kiterjesztett jelentés azonosításához, másfelől összefügg a potencialitás megítélésével. Módszertani relevanciával bír az elsődleges figura nyelvi kidolgozásának kategorizációja (4. szempont), mert azokban az esetekben, amikor a vizsgált elemi mondaton belül nem főnévi kidolgozással vált elérhetővé az elemi jelenet elsődleges szereplője, a szövegkörnyezetből következtetéssel végeztem el a szereplő ontológiai kategorizációt, ezáltal szükségesnek találtam rögzíteni ezen eseteket, ahol a döntés nagyobb fokú elemzői intuíciót követelt meg. Az 5. szempont azon példányok jelölésére szolgál, ahol a segédigei komponens a diskurzusból rész vevő és azonosított beszédpartnerhez horgonyzódik le az inflexiók segítségével, ehhez a kategóriához szorosan kapcsolódik a 6. címke, amely a segédigén azonosítható személyjelöléssel függ össze. A segédigén vagy a kopulán jelölt grammatikai mód azonosítása is megtörtént, hiszen a nyelvi kidolgozás során potenciálisan olyan tipikus konstrukcionális mintázatok megvalósulásaival asszociálódhat, amelyek képesek megnyitni a volitív/buletikus modális szemantikai teret (pl. feltételes mód). A deiktikusan jelölt igeidő szintén hatással lehet a modális térbe való belépés feltételeire, méghozzá úgy, hogy a múlt idő (de nem szabályszerűen minden esetben) eliminálhatja a nyelvileg megkonstruált elemi jelenet 'lehetséges' értelmezését (8. szempont).

A vizsgált kompozitumok főnévi igenévi komponense által aktivált eseményszerkezetek kategorizálását a WordNet (Fellbaum 1998; Miller 1995) ontológiája alapján végeztem el (vö. Bajzát 2023). Ez a kategorizációs eljárás 15 szemantikai osztály alapján teszi leírhatóvá a főnévi igenév által aktivált folyamatípust (9. szempont).

A megnyilatkozások által végrehajtott közvetett beszédaktusok (vö. Tátrai 2017: 1019–1022) azonosítását is elvégeztem (46. (al)kategória), mert ezek szintén tipikus konstrukciós mintázatokat képesek aktiválni a közvetett beszédaktusok metonimikus feldolgozásából kiindulva (Thornburg–Panther 1997). Valamint ezek a deontikus vagy volitív/buletikus modális tartományokat aktiválhatják, hiszen részt vesznek a diskurzusból megkonstruálódó társas viszonyok dinamikus alakításában.

#### 4.2. A klaszteranalízis

A klaszteranalízis egy olyan dimenziócsökkentő eljárás, amellyel az általunk vizsgálni kívánt adatok, adatsoportok különböző algoritmusok és eljárások segítségével csoportokba rendezhetők (Divjak–Fieller 2014), a klaszterezés alapulhat a hasonlóságon,

de akár a különbségeken is. A számítógéppel végrehajtott elemzési eljárás előnye, hogy egyszerre nagy mennyiségű adatot vagyunk képesek kezelni, azonban tudnunk kell azt, hogy az általunk választott elemzőszoftver beállításait a kutatási kérdéshez, és anyaghoz optimalizálva kell megtalálnunk (Divjak–Fieller 2014). A klaszteranalízis módszere kevésbé alkalmas hipotézisek tesztelésére, hiszen az eredményként kapott dendrogramokon ugyan látjuk az adatscsoportjainkat, és azt, hogy ezen csoportok a többi csoporttól számítva milyen távolságban helyezkednek el, azonban az analízis önmagában nem kínál fel információt arra vonatkozóan, hogy ezen csoportok miért tartoznak össze. Az elemzői interpretáció bevonása tehát nem kerülhető el.

A klaszteranalízis elkészítéséhez ebben a vizsgálatban (vö. Bajzát 2023) is a PAST szoftvert alkalmaztam (Hammer–Harper–Ryan 2001). Az interpretációkor figyelniünk kell arra, hogy a dendrogramok ágain az egymáshoz közelebb elhelyezkedő tulajdonságscsoportok nem feltétlen határoznak meg önmagukban konstrukciós mintázatokat. A tulajdonságok szoros összekapcsolódását motiváló jelenség lehet az is, hogy bizonyos jegyek nem specifikálnak önmagukban mintázatokat azért, mert túlságosan általánosak (pl. az ige jelen ideje), s ebből következően őket sem specifikálja más formai, illetve szemantikai jellemző.

A PAST szoftverben a Bray–Curtis-féle similarity függvényt beállítva generált a program egy olyan dendrogramot, amely az Y-tengely mentén 0 és 1 közötti értékkel (arányskálán) jelzi azt, hogy az adattáblára felvitt adatok milyen mértékben alkotnak egymással csoportosulásokat az azonosság alapján. Azért esett a Bray–Curtis similarity függvényre a választásom, mert ezen módszer a hasonlóságokat veszi figyelembe a klaszterezési eljárás során; korábbi vizsgálatok alapján úgy tűnik (Bajzát 2023), hogy érzékeny az adatok előfordulásának gyakoriságára, tehát vélhetően azon jellemzők csoportosításakor működik megbízhatóan, amelyek relatív gyakorisága magas volt a vizsgált példányok körében, kevésbé érzékeny a hapaxokra, tehát alkalmasabb lehet az általánosabb mintázataktól tendenciák feltárásában, viszont a kutatás későbbi szakaszában indokoltta válik az, hogy más függvényekkel (pl. koszinusz, Manhattan-távolság) is feldolgozzuk a kézi elemzéssel létrehozott adattáblákat.

### 4.3. A klaszteranalízist kiegészítő kvalitatív szempontok

Az elaboratív műveletek azonosítása mellett szükséges volt egy olyan szintaktikai modellnek az alkalmazására is, amelyben lehetővé válik a szemantikai-művelési szempontok érvényesítése egységes, többdimenziós ábrázolással a választott nyelvelméleti keretbe illeszkedően. Ennek révén a dolgozat Imrényi András viszonyhálózati modelljében (2017: 664–760) kezeli és a modell terminusait használva írja le azokat a műveleteket, amelyek a predikátum jelentését kívülről módosítják.

Az elemzésben a kontextualizációs utasításként működő megnyilatkozásrészek és a kontextualizált hálózatrész azonosításán kívül fontos volt bevonni a funkcionális pragmatika által alkalmazott *kontextus* fogalmát (Tátrai 2017: 927–951) és a kontextualizálás műveletét, hogy a példányok diszkurzív szerveződését és beágyazottságát is kezelni tudjuk.

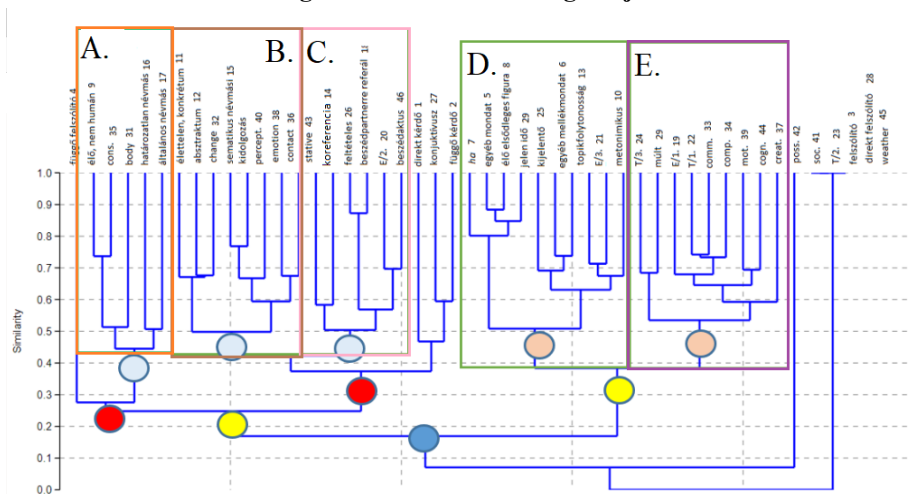
Összegezve megállapítható az, hogy (i) a diskurzus alapjához való személybeli lehorgonyozódás (társas deixis), (ii) az elsődleges figura ontológia, (iii) a főnévi igenév által kidolgozott folyamat típus, (iv) a deiktikusan jelölt igeidő, (v) a vizsgált példányok

magmondatainak szintaktikai környezetének (különösen a felülírás és a kiterjesztés műveletének (Imrényi 2017: 720–725); és a kontextualizáló viszonyok azonosítása kulcs szerepet játszik a vizsgálatban mint szempont.

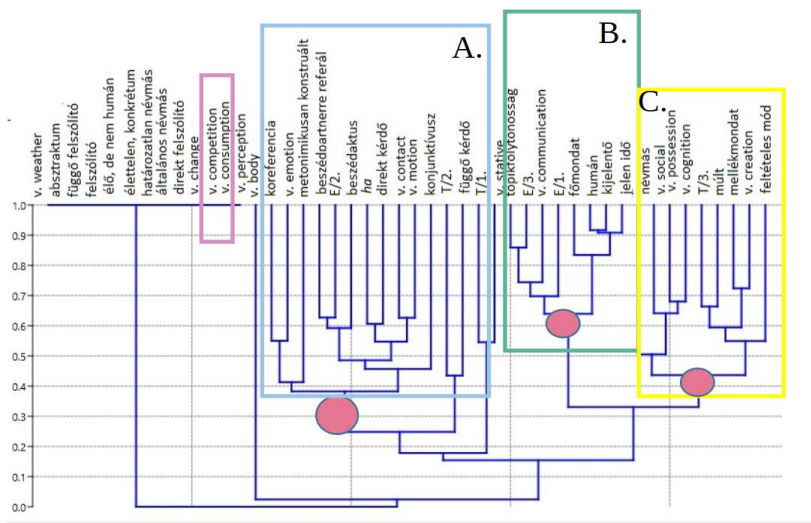
### 5. A módszer és az eredmények bemutatása a *tud* + főnévi igenév és az *akar* + főnévi igenév kompozitszerkezet feldolgozásán keresztül

A jelen dolgozatban azon szerkezetek kvalitatív elemzésének bemutatására vállalkozom, amelyek a premodális–modális viszony és a potencialitás azonosítási módjának szempontjából összefüggésekre mutatnak rá a diakrón adatokat tekintve, összevetve a mai magyar nyelvi adatokkal. Illetve olyan adatok bemutatására, amelyek feltehetően egy körülhatárolható s a bemutatott konstrukció szempontjából releváns mintázati sémák megvalósulását mutatják. A következő ábrákon (1–4.) látható dendrogramok jobb áttekinthetőségét segítik a fejezetben található táblázatok (4–10.) A dendrogramokon igyekeztem a hasonlósági alapon összetartozónak számító tulajdonságcsoportokat elkülöníteni (az ábrákon ezek színekkel és betűkkel válnak jelöltté), ezek a tömbök láthatóak kiterítve a táblázatokban is. A táblázatokban megjelenő sorszámok megegyeznek a 3. táblázatban látható (al)kategóriák sorszámával, így azonosíthatóvá válik az, hogy melyik jellemző melyik szempontcsoporthoz tartozik. Tehát például az 4. táblázat az 1. ábrán látható E ágat tükrözi. A jelen dolgozatban nem törekszem arra, hogy a dendrogramokon látható csoportok teljes körű bemutatását elvégezzem.

#### 5.1. A *tud* + főnévi igenév szerkezet dendrogramjai



1. ábra. A *tud* + főnévi igenév szerkezet dendrogramja az MNSZ2-ből származó példányok alapján (Bajzát 2023)



**2. ábra.** A *tud* + főnévi igenév szerkezet dendrogramja az TMK-ból származó példányok alapján

A *tud* + főnévi igenév szerkezet a KÉPESSÉG forrástartományból léphet be a modalitás szemantikai tartományaiba. A továbbiakban a tulajdonság-együttállásokat vizsgálom a premodális–modális jelentés aspektusából összekapcsolva a két korpuszból vételezett minta közötti eltérések rögzítésével.

Az 1. és a 2. ábra összehasonlításával hozzáférhetővé válik az számunkra, hogy a TMK-ból és az MNSZ2-ből vételezett adatok elemzése után kirajzolódó szemantikai mintázatok eltérnek abban a tekintetben, hogy a TMK-ból vett mintában nem azonosíthatunk olyan főnévi igenévi előfordulásokat a *tud*-dal egy kompozitumon belül, amelyek az MNSZ2-ből vett mintában megtalálhatók voltak (<competition>, <consumption>, <perception>, <change> és <body>). Ezen kívül az elsődleges figura ontológia típusok közül az absztraktum és a nem humán, de élő kategóriákra nem volt adatolható előfordulás a TMK-ból származó listában.

24.	T/3.	34.	<competition>
29.	Múlt idő	39.	<motion>
19.	E/1.	44.	<cognition>
22.	T/1.	37.	<creation>
33.	<communication>		

**4. táblázat.** A *tud* + főnévi igenév dendrogramjának E ágán megjelenő szempontok (az MNSZ2-ből vett példányok alapján)

13.	Topikfolytonosság	5.	Egyéb főmondat
21.	E/3.	39.	Humán elsődleges figura
33.	<communication>	44.	Kijelentő mód
19.	E/1.	37.	Jelen idő

**5. táblázat.** A *tud* + főnévi igenév dendrogramjának B ágán megjelenő szempontok (a TMK-ból vett példányok alapján)

Az tette indokoltta, hogy az MNSZ2 adatai alapján készült dendrogram E ágának és a TMK-ból vételezett minta vizuális reprezentációjának B ágának összehasonlítását emeljem ki elsőként az eredmények ismertetésében, hogy a csoportokhoz kapcsolódó példányok a diskurzusalaphoz képest az MNSZ2 esetében a múlt időben horgonyzódnak le, míg a TMK tekintetében a jelen időben. Azonban a mindkét táblázat által mutatott mintázatokban jellemző volt az E/3.-hoz és az E/1.-hez való lehorgonyzódás (az 4. táblázatban a T/3. is megjelenik). A mintázatokra specifikus folyamattípusoknál azt látjuk, hogy az MNSZ2-adatok sokkal nagyobb variabilitást tükröznek, míg a TMK-ból származó mintában a <communication> volt az egyetlen specifikáló folyamattípus.

(3) [...] *árról és a tervezett darabszámról azonban a cég képviselői még nem tudtak közelebbit mondani.* (MNSZ2, doc#1111)

Az E ágon olyan kirajzolódó mintázatok voltak adatolhatók, amelyek az alapmodális térrel asszociálódhatnak a folyamattípusok és a szereplők viszonyában, azonban a múlt idejű konstruálás révén következtethetünk arra, hogy az elemi mondatban megképződő események már lezárultak, tehát a potencialitás mérlegelése a beszédhelyzetben nem bír relevanciával, így ugyan még azonosíthatók a modális források a múltra vetítve, azonban a beszéd idejében vélhetőleg ezek már nem állnak fenn.

A (3) aktiválja a diszpozicionális tartományt, hiszen az esemény (*mondani*) az elsődleges figura információhiánya révén nem tudott megvalósulni. A szubjektívizáltan jelölt időviszonyok (vö. *még nem*) azonban jelzik azt a megnyilatkozási értelmezést, hogy a megjelenés előtt (de legkésőbb a megjelenéskor) ismertté válnak a pontos adatok (a dolgok normális működése szerint). Az esemény tehát a megnyilatkozás jelenében kizárható, a jövőben azonban biztosan bekövetkezik, egyik sem engedi meg a 'potenciális' értelmezést.

(4) *Az doktor felől mast sem tudok Kdnk mit írnom, mert volt egynihányszer nála, [...]* (TMK)

A B ágon tükröződő jellemző-együttállás egy olyan konstrukciós altípus/sémát határolt körül a TMK-ból vételezett mintákban, amely (4) főmondatként, a megnyilatkozó metapragmatikai tudatosságának jelzéseként (Tátrai 2017: 1046) a társalgási lépés színrevitelével kontextualizálja a mellékmondatban kidolgozódó állítást. Egyfelől ez a metapragmatikai tudatosság jelölőjeként funkcionáló megnyilatkozásrész a közös jelen-

tésfeldolgozás számára explikálja a konstruáló szubjektum vélt TUDÁS-át a diskurzus által tematizált információval kapcsolatban, valamint jelzi a kommunikációs szándékát. A tagadás révén a példány esetében nincs relevanciája a modális térrel való összefüggés mérlegelésének. Azonban érdemes számolnunk az átmenetiség esetével is, a példány(ok) ugyan nem bizonyulnak modálisnak az egészleges feldolgozás során, viszont mégis előtérbe állítják az elmondás vágyát elkülönítve azt az elmondás tényétől, valamint a *mást* mutató névmás explikálja azt, hogy az információátadás részben potenciálisan megvalósulhatott/megvalósulhat az adott diskurzusban.

43.	<stative>	20.	E/2.
14.	Névmási koreferencia	46.	Beszédaktus
26.	Feltételes mód	18.	Beszédpartnerhez való lehorgonyzódás

**6. táblázat.** A *tud* + főnévi igenév dendrogramjának C ágán megjelenő szempontok (az MNSZ2-ből vett példányok alapján)

14.	Névmási koreferencia	46.	Beszédaktus
38.	<emotion>	7.	<i>Ha</i>
10.	Metonimikusan konstruált elsődleges figura	1.	Direkt kérdő mondat
18.	Beszédpartnerhez való lehorgonyzódás	36.	<contact>
20.	E/2.	39.	<motion>
27.	Konjunktívusz	23.	T/2.
2.	Függő kérdő mondat	22.	T/1.

**7. táblázat.** A *tud* + főnévi igenév dendrogramjának A ágán megjelenő szempontok (a TMK-ből vett példányok alapján)

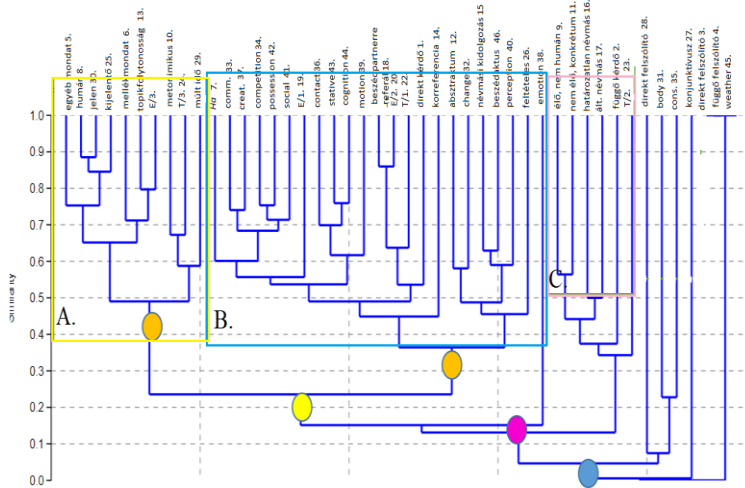
(5) *Akkor inkább ne mondj semmit sem rólam, ha nem tudod azt mondani, hogy én a te kedvesed vagyok.* (MNSZ2, doc#433)

(6) [...] *tude a Tanu valamit réa mondani, hogy cselekedett volna?* (TMK)

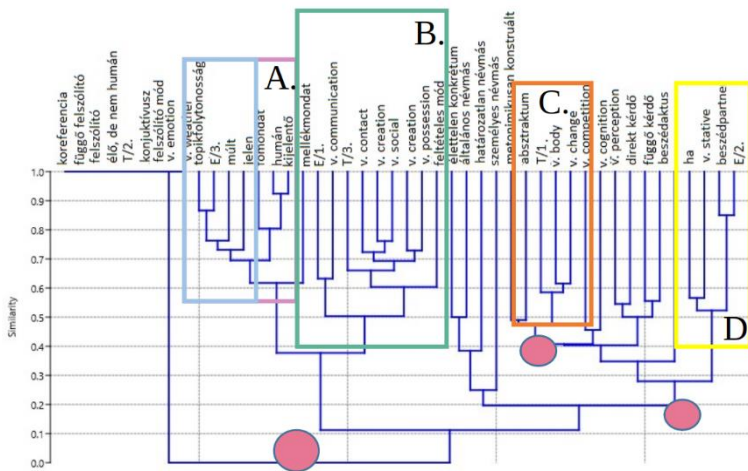
Az (5) és a (6) példányban a partner verbális tevékenységének potencialitásán keresztül juttatja érvényre a felé irányuló elvárását a megnyilatkozó, amely metonimikus figyelemáthelyeződéssel dolgozható fel a jelentésképzés művelete során. A példányban a szereplők a diskurzusszituáció résztvevői, ezzel a diskurzusban működő interperszonális viszonyok közvetlen alakításában is funkcionálnak a megnyilatkozások.

A beszédpartnerhez lehorgonyzott segédígés, 'verbális közlés' jelentésű folyamatot tartalmazó fenti példányok esetében a társalgási implikátúra hozza létre azt, hogy a megkonstruált folyamat 'lehetséges, nem kizárható' eseményként konceptualizálható. Az implikátúra konvencionálódása révén a konstrukció részévé válhat ez a deontikus modális jelentés. Látható, hogy a jellemzően középmagyar kori nyelvi anyagból vételezett konkordanciában is megjelenik a deontikus modális térrel asszociálódó társalgási implikátúra-mintázat.

## 5.2. Az *akar* + főnévi igenév szerkezet dendrogramjai



3. ábra. Az *akar* + főnévi igenév szerkezet dendrogramja az MNSZ2-ből származó példányok alapján (Bajzát 2023)



4. ábra. Az *akar* + főnévi igenév szerkezet dendrogramja az TMK-ból származó példányok alapján

Az *akar* + főnévi igenév szerkezet a VÁGY, AKARAT premodális forrástartományból léphet be a modális szemantikai térbe. Ha a 3. és a 4. ábrát vetjük össze egymással, akkor kirajzolódik az a tendencia, hogy már a TMK adataiban találunk előfordulást szinte az összes folyamattípus-kategóriára (az IDŐJÁRÁSI JELENSÉG-ekkel összekapcsolódó kategória az MNSZ2-ben sem adatolható az *akar* + főnévi igenév esetében), kivéve az ÉRZELMI ÁLLAPOTOK-kal összekapcsolódó csoportra vonatkozóan, tehát kevésbé bizonyul

korlátozottak az, hogy a vizsgált szerkezettag milyen folyamatípus-kategória példányként azonosított főnévi igenévvel alkothat kompozitumot, szemben a *tud* + főnévi igenév középmagyar kori előfordulásaival. A másik feltűnő jelenség a vizualizáción az, hogy az MNSZ2 példányainak összefüggéseit bemutató ábrán az A ág együttállási mintázata erőteljesebben generikus, kevésbé jellemzik a specifikusabb, a sémák körülhatárolására alkalmas tulajdonság-együttállások. A TMK adataihoz tartozó dendrogramon ezek jobban elkülönülnek egymástól (4. ábra).

7.	<i>Ha</i>	41.	<social>
33.	<communication>	19.	E/1.
37.	<creation>	36.	<contact>
34.	<competition>	43.	<stative>
42.	<possession>	1.	Direkt kérdő mondat
44.	<cognition>	14.	Névmási koreferencia
39.	<motion>	12.	Absztraktum
18.	Beszédpartnerhez való lehorgonyzás	32.	<change>
20.	E/2.	15.	Sematikus névmási kidolgozás
22.	T/1.	40.	<perception>
46.	Beszédaktus	38.	<emotion>
26.	Feltételes mód		

**8. táblázat.** Az *akar* + főnévi igenév dendrogramjának B ágán megjelenő szempontok (az MNSZ2-ből vett példányok alapján)

6.	Egyéb mellékmondat	37.	<creation>
19.	E/1.	41.	<social>
33.	<communication>	39.	<motion>
24.	T/3.	42.	<possession>
36.	<contact>	26.	Feltételes mód

**9. táblázat.** Az *akar* + főnévi igenév dendrogramjának B ágán megjelenő szempontok (a TMK-ből vett példányok alapján)

A TMK-ból vételezett mintában az E/1.-hez és a T/3.-hoz lehorgonyzódó segédigei példányok rendeződtek mintázatba a KOMMUNIKÁCIÓ-val összekapcsolódó folyamatípust előhívó főnévi igenévi példányokkal, de ha megfigyeljük a dendrogramot, akkor észrevehetjük, hogy erősebben az E/1.-es jelöltség asszociálódik velük. Ezen tendencia megfigyelhetővé válik az MNSZ2-ből nyert minta esetében is, tehát a két minta ebből az aspektusból szemlélve hasonlóságot mutat egymással. Szintén egyezést mutatnak afelől közelítve, hogy az ALKOTÁS; ÉRINTKEZÉS, KAPCSOLAT; TÁRSAS/TÁRSADALMI ESEMÉNYEK és a BIRTOKLÁS folyamatokkal összefüggő eseménytípusok közelítenek az előbb említett kategóriákhoz.

- (7) *Ha én százalékosan akarom kifejezni az egy 10-20 százalék között lehet, a most bajbajutott embereknek az a rétege.* (MNSZ2, doc#2754)
- (8) *mindazonáltal elébb Disputálni akarnék [...]* (TMK)
- (9) *Félre ne érts, csak poénkodtam, nem akartalak bántani! Bocszsz.* (MNSZ2, doc#2815)
- (10) *Kegyelmednek tudván sok dolgait, terhelni nem akartam velle, [...]* (TMK)

A (7) példánynál megfigyelhető, hogy a megnyilatkozó által nyelviileg jelöltté tett metapragmatikai tudatosság jelzéseként funkcionál az *akar* + főnévi igenév szerkezet megvalósuló példánya, mely a főmondatbeli vélekedést kontextualizálja a konstruáló szubjektum kiindulópontjából.<sup>2</sup> Ebben az esetben a százalékos kifejezés tényleges megvalósulás (a *ha* kötőszós tagmondat nem valódi feltételt közöl), csak az érték becslült, bizonytalan. A mennyiségi arányoknál a százalékos kifejezés a számítással kapcsolatos tudáson alapuló képesség, nem mérlegeljük azt, hogy a példány összefüggne a potencialitás szemantikai terével, tehát nem modális.

A (8) példány a Történeti magánéleti korpuszból származik. Megfigyelhető, hogy a buletikus/volitív modális térrel kapcsolódik össze, hiszen az elemi mondatban nyelviileg kidolgozott elemi jelenet egy lehetséges szituáció. A modális jelentéstartalommal bíró példány mintázatához a feltételes módú konstruálás is tetten érhető a segédigei komponensen. Ez a mintázat nagy arányban jellemezte a csoporthoz tartozó többi példányt is az elemzett konkordanciában, a (7)-hez hasonló, tehát az *akar* + főnévi igeneves szerkezeten keresztül történő explicit metapragmatikai tudatosság jelzéseként funkcionáló előfordulás nem volt azonosítható a KOMMUNIKÁCIÓ-val összekapcsolódó folyamattípus esetében.

A (9) nyelvi adat feldolgozásánál a BOCSÁNATKÉRÉS beszédaktusának forgatókönyvi sémája aktiválódik. A megnyilatkozó a saját kommunikatív szándékaihoz fűződően juttatja érvényre az értékelői reflexióját, azonban a múlt időhöz való episztemikus lehorgonyzás a folyamatot az aktuális beszédhelyzethez képest eltérő időben kontextualizálja ('nem volt bántásra irányuló akarat'). Mindazonáltal a BOCSÁNATKÉRÉS forgatókönyvébe illeszkedően, a jelek szerint a partner bántásként értelmezhetette a poénkodást, a bántás tehát nem kizárható lehetőség. A megnyilatkozás a beszédpartner érzelmi beállítódását igyekszik összehangolni a beszélő/elsődleges figura szándékaival. A (10)-ben szintén egy <contact> folyamattípus-csoportba tartozó példány kerül a segédige + főnévi igenév sematikus szerkezetbe, valamint a megnyilatkozó szintén a múlt időhöz való episztemikus lehorgonyzáson keresztül teszi nyelviileg kidolgozottá az értékelői reflexióját az előzetesen vélt szándékaival összefüggésben. A megnyilatkozásban nem eliminálható annak a potencialitása, hogy a partner mégis terhelőnek tulajdonítja a

<sup>2</sup> A kontextualizáló tagmondat a megnyilatkozó saját nyelvi tevékenységéhez való reflexív viszonyulását fejezi ki, ennek explicit jelölője az E/1.-hez és jelen időhöz lehorgonyzódó performatív 'létrehozás' és 'közlés' jelentésű ige (Tátrai 2017: 1046–1053).

megnyilatkozó által közölni szándékozott információkat. Ennél az esetről ugyan azonosítottuk a múlt időhöz való lehorgonyzódást, viszont a kontextusból következtethetünk arra, hogy a kimondás még nem realizálódott esemény, tehát modálisként azonosíthatjuk a példányt.

12.	Absztraktum	32.	<change>
23.	T/1.	34.	<competition>
31.	<body>		

**10. táblázat.** Az *akar* + főnévi igenév dendrogramjának C ágán megjelenő szempontok (a TMK-ból vett példányok alapján)

Az eddigi elemzésekben láthattuk, hogy olyan szerkezetek kerültek felszínre, amelyek esetében az elsődleges figura ontológiai kategóriája minden esetben humán elsődleges figura. Azonban indokolt lehet megvizsgálni azon eseteket, amelyeknél elkülönbözést adathatunk, mivel a humán kategóriától elváló elsődleges figurák a jelentéskiterjesztések létrehozásában erősebb potenciállal bírnak az absztrakt segédige + főnévi igenév sémák tekintetében, kiemelten azokban az előfordulásokban, amelyek a VÁGY, AKARAT premodális forrástartományból aktiválhatják a modális szemantikai teret, hiszen feltételezzük azt, hogy ezen tartományok alapvetően az emberi tudatossággal bíró entitásokkal kapcsolódnak össze.

(11) *Egyáltalán mindenki feszült, mert hajszojja a boldogság két madarát, ami nem nagyon akar megmutatkozni.* (MNSZ2; doc#2284)

(12) [...] *hogy az Isten úgy akarja maga munkáját végbevinni [...]* (TMK)

A (11) példány az MNSZ2-ből származik. Az absztraktumok csoportjára ezen korpuszban jellemző a variabilitás, illetve a metaforikus jelentéskiterjesztésen alapuló konstruálás, amelyet a példaként idézett adat is szemléltet. A (11)-ben a segédige predikátum polaritásának negatív felülírása eliminálja a modális olvasat lehetőségességét, hiszen a kifejezés feldolgozásánál azonosíthatóvá válik az, hogy a metaforikusan megkonstruált boldogság (absztraktum) megmutatkozásának hiányáról szól a közléység a megszemélyesítésen keresztül. A (12) példány kategorizációja az elsődleges figura ontológia felől kérdéses. Az adatok feldolgozásakor azt az elemzői döntést hoztam, hogy minden természetfeletti erővel összekapcsolható entitást absztraktumként címkézek, azonban megfigyelhető az ágens magas fokú antropomorfizációja az idézett példányban. Ebből fakadóan akár a humán entitások kategóriájához is rendelhetnénk mint periferikus kategóriapéldányt, de ez az antropomorfizáció a kontextus viszonyában dolgozódik fel, tehát indokoltnak láttam azt, hogy megtartsam a primer kategorizációs eljárásomat. A TMK példányaiban megfigyelhető az, hogy az absztraktum csupán ezzel a nyelvi megvalósulással asszociálódik, tehát nem variábilis a példányok köre, amelyre kiterjedhet az MNSZ2 előfordulásaival szemben.

## 6. Összefoglalás, kitekintés

A tanulmány célja egy olyan módszer tapasztalatainak rögzítése volt diakrón perspektívában, amely a mai magyar nyelvi korpuszadatokon már előzetesen alkalmazásra került. A dolgozatban bemutatott vizsgálat a premodális tartományokkal összefüggő főnévi igeneves szerkezetek elemzésére és mintázatalapú csoportosítására fókuszált számítógépes nyelvészeti eszközökkel, funkcionális elméleti keretben. A jelen dolgozatban csupán a feldolgozott adatok töredékének bemutatása valósulhatott meg. A vizsgált szerkezetek mintázatainak alakulástörténetére csak nagyon óvatos következtetéseket tehetünk, a kutatás későbbi szakaszában a dendrogramokon látható összes csoport szisztematikus feldolgozásával és bemutatásával várhatjuk ezeket, valamint érdemes lesz a következőkben bevonni más diakrón korpuszok nyelvi anyagának a feldolgozását is [Ómagyar Korpusz (Simon–Sass 2012), MTSz (Sass 2017)]. A *tud* + főnévi igenév esetében megfigyelhető volt az, hogy bizonyos folyamattípusok (VERSENGÉS, FOGYASZTÁS, ÉRZEKELÉS, VÁLTOZÁS és TESTI FOLYAMATOK) nem fordulnak elő a *tud* szerkezettag környezetében, tehát korlátozottabbak a megvalósulások a főnévi igenév típusokat tekintve. Ez az *akar* + főnévi igenév középmagyar kori előfordulásaiban nem volt kimutatható (kivéve egy kategóriában). Továbbá a *tud* TMK-ból vételezett példányai között kevésbé voltak változatosak az elsődleges figura ontológiai kategóriáinak az előfordulásai is. Megfigyeltük, hogy a középmagyar kori adatoknál is kimutatható volt a deontikus modális tér aktiválása, amely a grammatikalizációs ösvényen való elmozdulást jelöli a dinamikus modalitás kategóriái felől (vö. Van linden 2010). Az *akar* + főnévi igenév TMK-ból vételezett mintájában szintén azonosíthatók voltak a példányok azon csoportjai, amely a társalgási implikátúra konvencionalizálódottabb formáit valósítják meg.

A későbbiekben célkitűzés a bemutatott szempontrendszer optimalizálása (kategóriák számának csökkentése), további függvények tesztelése, illetve a vizsgálat kiterjesztése további szerkezettagokra.

## Hivatkozások

- Bajzát Tímea Borbála 2023: A premodális tartományokkal összekapcsolódó segédige/melléknév + főnévi igenév konstrukciók mondatszintű szemantikai vizsgálata. In Tátrai Szilárd – Tolcsvai Nagy Gábor szerk.: *Magyar mondat és kontextuális környezete*. Megjelenés előtt.
- Divjak, Dagmar – Fieller, Nick 2014: Cluster analysis. Finding structure in linguistic data. In Dylan, Glynn – Robinson, A. Justyna eds.: *Corpus Methods for Semantics. Quantitative studies in polysemy and synonymy*. Amsterdam–Philadelphia: John Benjamins, 405–439.
- Dömötör Adrienne – Gugán Katalin – Novák Attila – Varga Mónika 2017: Kiútkeresés a morfológiai labirintusból – korpuszépítés ó- és középmagyar kori magánéleti szövegekből, *Nyelvtudományi Közlemények* 113, 85–110.
- Fauconnier, Gilles 1985: *Mental spaces: aspects of meaning construction in natural language*. Cambridge–Massachusetts–London: Bradford Book – MIT Press.
- Fellbaum, Christiane ed. 1998: *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

- Hammer, Øyvind – Harper, David A.T. – Ryan, Paul D. 2001: PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* 4/1, 1–9.
- Imrényi András 2017: Az elemi mondat viszonyhálózata. In Tolcsvai Nagy Gábor szerk.: *Nyelvtan*. (A magyar nyelv kézikönyvtára 4.) Budapest: Osiris, 663–761.
- Kiefer Ferenc 2005: *Lehetőség és szükségszerűség*. Budapest: Tinta Könyvkiadó.
- Kiefer Ferenc 2000: *Jelentéelmélet*. Budapest: Corvina Kiadó.
- Krejcie, Robert V. – Morgan, Daryle W. 1970: Determining Sample Size for Research Activities. *Educational and Psychological Measurement* 30, 607–610.
- Kugler Nóra 2017a: Az evidencialitás és a modalitás. In Tolcsvai Nagy Gábor szerk.: *Nyelvtan*. (A magyar nyelv kézikönyvtára 4.) Budapest: Osiris, 467–494.
- Kugler Nóra 2017b: Az összetett mondat. In Tolcsvai Nagy Gábor szerk.: *Nyelvtan* (A magyar nyelv kézikönyvtára 4.) Budapest: Osiris, 806–895.
- Langacker, Ronald W. 1987: *Foundations of Cognitive Grammar. Vol. 1. Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 1991: *Foundations of Cognitive Grammar. Vol. 2. Descriptive application*. Stanford: Stanford University Press.
- Langacker, Ronald W. 2008: *Cognitive Grammar: A basic introduction*. Oxford: Oxford University Press.
- Novák Attila – Gugán Katalin – Varga Mónika – Dömötör Adrienne 2018: Creation of an annotated corpus of Old and Middle Hungarian court records and private correspondence. *Language Resources and Evaluation* 52, 1–28.
- Miller 1995: WordNet: A Lexical Database for English. *Communications of the ACM* 38/11, 39–41.
- Nuyts, Jan. 2006: Modality: Overview and linguistic issues. In William Frawley eds.: *The expression of modality*. (The expression of cognitive categories 1.) Berlin: Mouton de Gruyter, 1–26.
- Oravecz Csaba – Váradi Tamás – Sass Bálint 2014: The Hungarian Gigaword Corpus. In Calzolari, Nicoletta – Choukri, Khalid – Declerck, Thierry – Loftsson, Hrafn – Maegaard, Bente – Mariani, Joseph – Moreno, Asuncion – Odijk, Jan – Piperidis, Stelios eds.: *Proceedings of the ninth international conference on language resources and evaluation (LREC-2014)*. Reykjavik: European Languages Resources Association (ELRA), 1719–1723.
- Palmer, Frank Robert 1986: *Mood and modality*. Cambridge: Cambridge University Press.
- Paradis, Carita 2005: Ontologies and construals in lexical semantics. *Axiomathes* 15: 541–573.
- Sass Bálint 2017. Keresés a korpuszban: a kibővített Magyar történeti szövegtár új keresőfelülete. In Forgács Tamás – Németh Miklós – Sinkovics Balázs szerk.: *A nyelv-történeti kutatások újabb eredményei IX*. Szeged: Szegedi Tudományegyetem Magyar Nyelvészeti Tanszék. 267–277.
- Simon Eszter – Sass Bálint 2012. Nyelvtológia és kulturális örökség, avagy korpusz-építés ómagyar kódexekből. *Általános Nyelvészeti Tanulmányok XXIV*, 243–264.
- Tátrai Szilárd 2011: *Bevezetés a pragmatikába. Funkcionális kognitív megközelítés*. Budapest: Tinta Kiadó.

- Tátrai Szilárd 2017: Pragmatika. In Tolcsvai Nagy Gábor szerk.: *Nyelvtan*. (A magyar nyelv kézikönyvtára 4.) Budapest: Osiris, 899–1057.
- Thornburg, Linda – Panther, Klaus-Uwe 1997: „Speech Act Metonymies.” In Liebert, Wolf Andreas – Redeker, Gisela – Waugh R., Linda eds.: *Discourse and Perspective in Cognitive Linguistics*. Amsterdam–Philadelphia: John Benjamins, 205–219.
- Tolcsvai Nagy Gábor 2017a.: A nyelv emberi lényege. A magyarázat látóköralkotó változatai a nyelvtudományban. In Tolcsvai Nagy Gábor szerk.: *Megértés és megértetés*. Budapest: Gondolat Kiadó, 38–62.
- Tolcsvai Nagy Gábor 2017b.: Jelentéstan. In Tolcsvai Nagy Gábor szerk.: *Nyelvtan*. (A magyar nyelv kézikönyvtára 4.) Budapest: Osiris, 207–466.
- Tolcsvai Nagy Gábor szerk. 2017c: *Nyelvtan*. (A magyar nyelv kézikönyvtára 4.) Budapest: Osiris.
- van der Auwera, Johan – Plungian, Vladimir A. 1998: Modality’s semantic map. *Linguistic Typology* 2, 79–124.
- Van linden, An. 2010: From premodal to modal meaning: Adjectival pathways in English. *Cognitive Linguistics* 21/3, 537–571.