

Elmerekprezentációon alapuló számítógépes nyelvészet „felszíni szintaxissal”

Kleiber Judit és Alberti Gábor

A számítógépes és kísérletes nyelvészet új útjai és korszerű eszközei

Pécs, 2019. május 2.

Miről lesz szó?

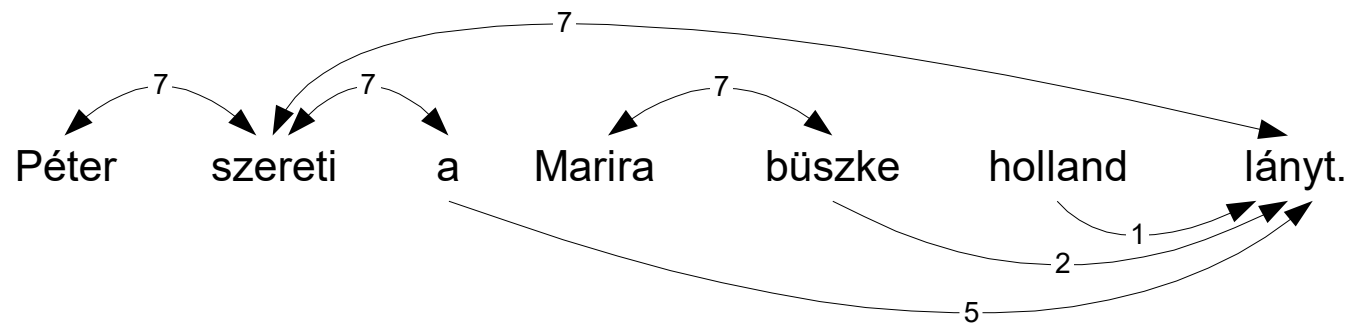
- GeLexi-projekt (kb. 2000–2008)
 - Elemző és gépi fordító (magyar, angol), „kopredikációs hálózat”
 - „Játékgrammatika” (párszáz szavas adatbázis), PROLOG nyelv
- ReALIS
 - Elmereprezentáció, világról és egymásról való tudás, BDI (A, E)
 - Élethossziglani építkezés: óriási DRS
 - Kopredikációs hálózat kiterjesztése
 - Élethossziglaniságból adódó kulturális–enciklopédikus és interperszonális hálózat
- Cél (kérdésfeltevés)
 - Elméletben működő elmereprezentációs hálózat implementálása
 - GeLexi-ben alkalmazott megközelítés kiterjesztése elmereprezentációra

A GeLexi-projekt

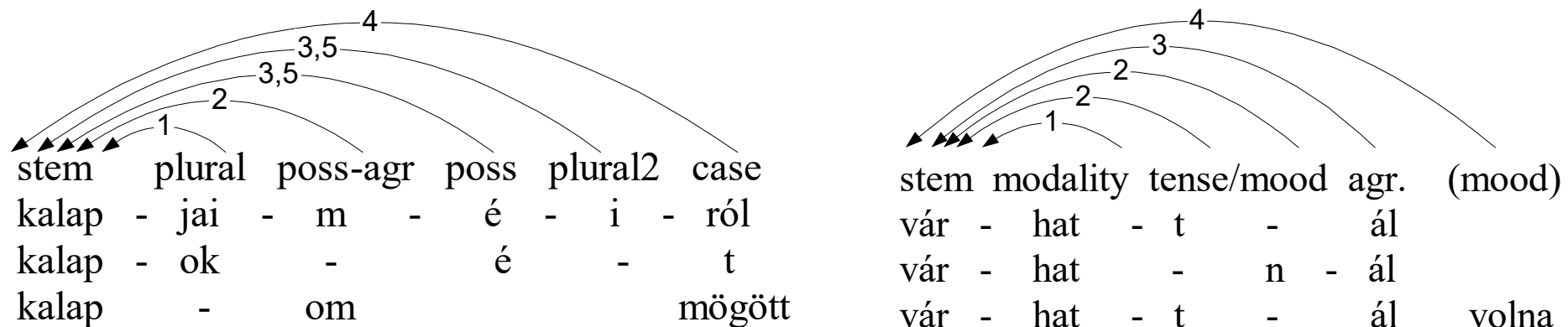
- Elméleti háttér: GASG (Alberti 1999)
 - Generative/Generalized Argument Structure Grammar
 - Unifikációs kategoriális nyelvtan
 - Radikális lexikalizmus (Karttunen 1986)
 - Frázis-struktúra kiküszöbölhető
 - Lexikai egységek komplex leírása, minden információ egyszerre
 - Morfoszintaxis (elemsorrend)
 - Szemantika (DRS)
 - Elemek keresik egymást
 - A mondatban szomszédosak akarnak lenni, különféle ranggal
 - Gyengébb rang: elszakadhatnak egymástól, morfológiai jelöltség árán
- Elméleti cél: egzaktságának bizonyítása implementációval

Elemrend: rangparaméterek

- Szavak közti keresés (Alberti et al. 2002)

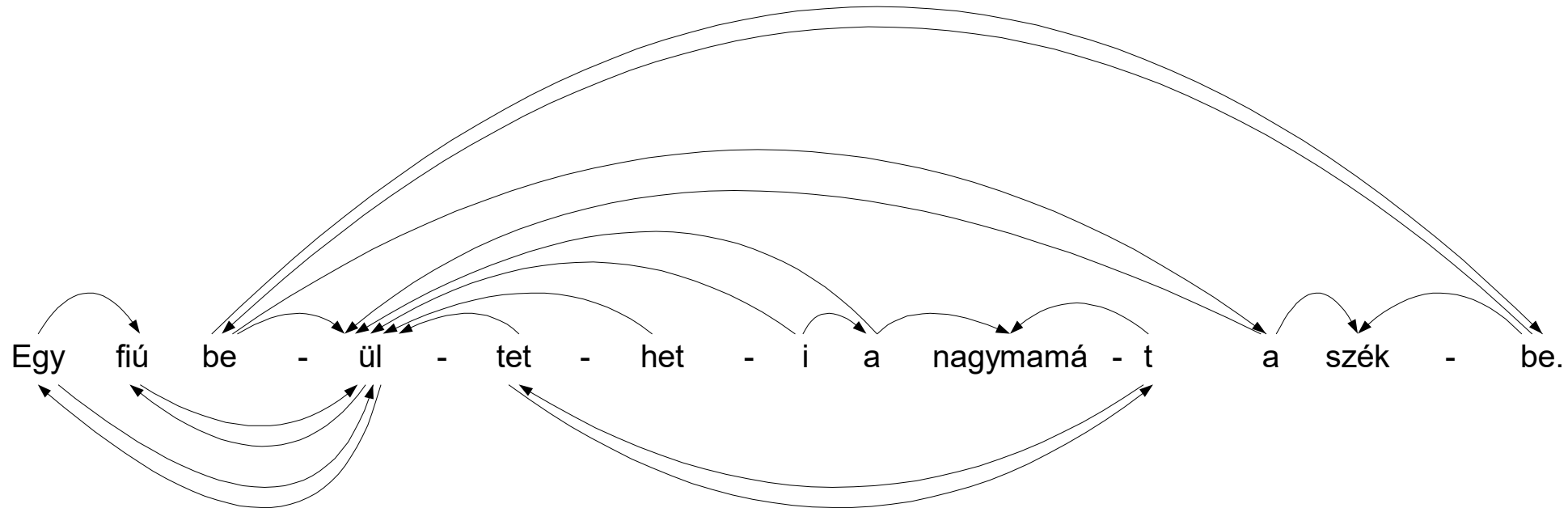


- Szón belüli keresés (főnév, ige) (Alberti et al. 2005)



Elemrend: rangparaméterek

- Morfémák közti keresés, szón belül és kívül → univerzalitás



A GeLexi-projekt eredményei

- Gyakorlati cél
 - Párszáz szavas adatbázison működő program („játékgrammatika”)
 - PROLOG programnyelv
 - Generatív alapfeladat (grammatikalitás)
 - Agrammatikus szósorok felismerése (*könyveihoz; kalapjaimrólé; holland okos lány*)
 - Elemző (Alberti et al. 2003; 2004)
 - Kimenete: lexikai egységek, szintaktikai viszonyok, szemantikai leírás (DRS)
 - Többértelműség (*dobom; Péter keresi a lányt a kutyával*)
 - Gépi fordító (Alberti–Kleiber 2004)
 - Magyar és angol nyelv között (→ angol nyelvű elemző is)
 - Kopredikációs hálózaton keresztül (ugyanarról a referensről tesznek állítást)

A program kimenete

- `gramm("Péter keresteti Marit a rendőrséggel.")`.
- **LEXIKAI EGYSÉGEK:**
 - Péter:** `n(1,1,li(m("", "Péter", ""), labstem("Peter", phonfst(1,2,0,2), 1, [])))`
 - keres:** `n(2,1,li(m("", "keres", ""), labstem("look-for", phonfst(1,2,2,2), 2, [{"NOM"}, {"ACC"}])))`
 - tet:** `n(2,2,li(m("t", "A", "t"), labder("cause", phonfsu(2,2,0.2,2), 2, ac(-1,0,1))))`
 - i:** `n(2,3,li(m("", "i", ""), labsuff("sg3obj+def", phonfsu(1,3,1,3), 2, 3)))`
 - Mari:** `n(3,1,li(m("", "Mari", ""), labstem("Mary", phonfst(2,2,0,2), 1, [])))`
 - t:** `n(3,2,li(m("V", "t", ""), labsuff("ACC", phonfsu(1,1,1,3), 1, 4)))`
 - a:** `n(4,1,li(m("", "a", "Z"), labstem("the", phonfst(1,3,3,3), 3, [])))`
 - rendőrség:** `n(5,1,li(m("", "rendőrség", ""), labstem("police", phonfst(1,2,3,2), 1, [])))`
 - gel:** `n(5,2,li(m("S", "A", "l"), labsuff("INSTR", phonfsu(1,2,2,3), 1, 4)))`
- **SZINTAXIS:**
 - ...
 - regent-noun-subj:** **keresteti-Péter**
 - regent-det-subj: keresteti-Péter
 - regent-noun-obj:** **keresteti-Marit**
 - regent-det-obj: keresteti-Marit
 - regent-noun-obl:** **keresteti-rendőrséggel**
 - regent-det-obl: keresteti-a
 - det-noun:** **a-rendőrséggel**

A program kimenete

- `gramm("Péter keresteti Marit a rendőrséggel.")` .

- **SZEMANTIKA:**

```
provref("fixpoint", [e(2,2,1)])
provref("old", [r(1,1,1)])
pred("Peter", 1, [r(1,1,1)])
provref("new", [e(2,1,1)])
pred("look-for", 2, [e(2,1,1), r(4,1,1), r(3,1,1)])
provref("new", [e(2,2,1)])
provref("=", [e(2,2,1), e(2,1,1)])
pred("cause", 2, [e(2,2,1), r(1,1,1), e(2,1,1)])
provref("old", [r(3,1,1)])
pred("Mary", 3, [r(3,1,1)])
provref("old", [r(4,1,1)])
provref("<or=", [r(4,1,1), e(2,2,1)])
pred("police", 5, [r(4,1,1)])
```

- **KOPREDIKÁCIÓS VISZONYOK:**

```
copr("look-for", 2, 1, "Peter", 1, 1, 1, 1, "arg")
copr("look-for", 2, 1, "Peter", 1, 1, 1, 0, "arg")
copr("look-for", 2, 1, "Mary", 3, 1, 2, 1, "arg")
copr("look-for", 2, 1, "Mary", 3, 1, 2, 0, "arg")
copr("cause", 2, 2, "Peter", 1, 1, 1, 1, "arg")
copr("cause", 2, 2, "Peter", 1, 1, 1, 0, "arg")
copr("cause", 2, 2, "look-for", 2, 1, 2, 0, "arg")
copr("the", 4, 1, "police", 5, 1, 0, 1, "free")
```

r1	r3	r4	e21	e22
Péter(r1)				
Mari(r3)				
rendőrség(r4)				
e22: okoz(r1,e21)				
e21: keres(r4,r3)				
...				

Gépi fordítás

- `translate_Hun_Eng_print("Szeretlek") .`

...

SZEMANTIKA:

```
provref("fixpoint", [e(1,1,1)])
```

```
provref("new", [e(1,1,1)])
```

```
pred("love", 1, [e(1,1,1), r(0,1,1), r(0,1,2)])
```

In English: I love you.

...

- `translate_Eng_Hun("I love you."), fail.`

In Hungarian: Szeretlek.

In Hungarian: Szeretlek téged.

In Hungarian: Szeretlek titeket.

In Hungarian: Én szeretlek.

In Hungarian: Én szeretlek téged.

In Hungarian: Én szeretlek titeket.

no

- `translate_Hun_Eng("A Pétert váró lányra büszke fiú énekel.") .`

In English: The boy proud of the girl waiting for Peter sings.

yes

ReALIS – a rendszer

- ReALIS: Reciprocal and Lifelong Interpretation System (Alberti 2011)
- Elmereprezentáció
- Világról és egymásról való tudás (BDI, A, E)
 - A szereplők: én (i), te (u), ő (other)
 - A modalitások: **vélekedés** (Belief), **vágy** (Desire), **szándék** (Intention), **autoritás** (Authority), **tapasztalat** (Experience)
 - Időparaméter (régén, most, majd) → eseményszerkezet megragadása
 - Intenzitás → bizonytalanság kifejezése; diskurzusjelölők, pl. *ugye, szerintem*
- Az interpretáló pillanatnyi elmeállapotának modellálása
- Matematikai definíció
 - Relációk: formula (kondíciósor) megalkotása (σ); referens-azonosítás, horgonyzás (α); dobozhierarchia (λ); szaliens elemek kijelölése, kurzor (κ)
- Élethossziglani építkezés: óriási DRS

ReALIS – mire jó a hálózat?

- Elméleti cél: makacs nyelvészeti problémák kezelése
- Az élethossziglaniságból adódó kulturális–enciklopédikus és interperszonális hálózat alkalmazása
 - Implicit antecedensek keresése (vö. Kálmán 1990)
 - *Péter tegnap megnősült. A pap hosszan beszélt. /#A kutya hangosan ugatott.*
 - Az esküvőhöz asszociált szereplők.
 - Áthidalás vagy lexikális elburjánzás (vö. Pustejovsky 1995)
 - *Vörös lány* (haja? bőre? ...); *pécsi vonat* (Pécsről jön? Pécsre megy? Ott gyártották?)
 - Kopredikációs hálózat: vörös (x), lány (x) → először segédváltozóval, de „letárolható”
 - Szabálytalan melléknevek (vö. Kiefer 2000)
 - *Az egykori / állítólagos kém* (nem igaz, hogy kém(x) és egykori/állítólagos (x))
 - Eseményszerkezet, időparaméter / másik „világocska” (λ -szint), hiedelemvilág (iB=0; iBoB=+5)

ReALIS – mire jó a hálózat?

- Elméleti cél: makacs nyelvészeti problémák kezelése
- Az élethossziglaniságból adódó kulturális–enciklopédikus és interperszonális hálózat alkalmazása
 - Eventuális vonzat helyettesítése főnévivel (vö. Pustejovsky 1995)
 - *Befejeztem a szendvicset.* (Evését? Készítését? Lefestését? ...)
 - Minden értelmes lexikai egység egy hálózatot aktivál (használati hálózat)
 - Birtoklás (vö. Alberti 1995)
 - *Péter háza* (lakó, tulajdonos, építész, „kukázó”, ...)
 - Amit lehet vele tenni, az alapján lehet birtokolni.
 - Vonzatöröklés képzés során (vö. Laczkó 2000)
 - *Péter megoperálása* (csak páciens) / *operációja* (orvos is); további szereplők
 - Komplex esemény vs. eredmény; név árulkodó lehet (*Kovács doktor / bácsi*)

ReALIS – mire jó a hálózat?

- Elméleti cél: makacs nyelvészeti problémák kezelése
- Az élethossziglaniságból adódó kulturális–enciklopédikus és interperszonális hálózat alkalmazása
 - Különleges birtokosok, elidegeníthetetlenség (vö. Kiefer 2000; Farkas–Alberti 2017)
 - *A növény virága / Mari virágja; Péter anyja / a világ legjobb anyája*
 - Hálózatban a szereplők hierarchiában: elsődleges birtokos (elidegeníthetetlen)
 - Poliszém rendszerek (vö. Kiefer 2000)
 - *Megvárlak az iskola előtt / iskola után; Az egész iskola kirándul / egyetért; ...*
 - Igei vonzatszatszerkezet-családok (vö. Alberti 2006)
 - *Ás* ige: centrális: vki földet mozgat.
 - További ágentív és nem ágentív szerepek (műveltetés, cél, funkció...)
 - *A kalózvezér elásatta a kincset a kalózokkal.*
 - Az ige a teljes (részbenrendezett) hálózatot meghívja (megjósolja az eseteket).

ReALIS – gyakorlatban?

- Gyakorlati cél: az elmélet implementálása
- Kérdés: az elméletben működő rendszer alkalmazható lehet-e intelligens nyelvtechnológiai célokra?
 - Pl. gépi fordítás minőségének javítása
- A Gelexi-projektben alkalmazott megközelítés
 - Kopredikációs hálózat kiterjesztése az elmereprezentációs hálózatra
 - Elmében tárolt hálózatokra való hivatkozás elemzés közben
 - Hogyan rögzíthető ez a tudás?
 - Hogyan vehető elő?
 - Létező nyelvtechnológiai eszközök az implementáláshoz?
 - A felvázolt rendszer megközelítésének alkalmazása létező rendszerekben?

Hivatkozások

- Alberti, Gábor (1995): Role assignment in Hungarian possessive constructions. In István Kenesei, Csaba Pléh (eds.): *Approaches to Hungarian 5. Levels and Structures*. Szeged: JATE. 13–28.
- Alberti, Gábor (1999): GASG: The Grammar of Total Lexicalism. In: *Working Papers in the Theory of Grammar 6/1*, Theoretical Linguistics Programme, Budapest University and Research Institute for Linguistics, Hungarian Academy of Sciences.
- Alberti Gábor 2006. A szóképzéssel együtt járó vonzatszerkezet-változások rendszere. *Nyelvtudományi Közlemények* 103: 75–105.
- Alberti Gábor (2011): *ReALIS: Interpretálók a világban, világok az interpretálóban*. Budapest: Akadémiai Kiadó.
- Alberti Gábor, Balogh Kata, Kleiber Judit, Viszket Anita (2002): A totális lexikalizmus elve és a GASG nyelvtan-modell. In: Maleczki Márta (szerk.): *A mai magyar nyelv leírásának újabb módszerei V.*, SZTE, Szeged, 193–218.
- Alberti, Gábor, Kata Balogh, Judit Kleiber, Anita Viszket (2003): Total Lexicalism and GASGrammars: A Direct Way to Semantics. In: Alexander Gelbukh (eds.): *Proceedings of CICLing2003*, Springer-Verlag, 37–48.
- Alberti, Gábor, Kata Balogh, Judit Kleiber, Anita Viszket (2005): Towards a Totally Lexicalist Morphology. In: István Kenesei, Christopher Piñón (eds.): *Approaches to Hungarian 9*, 9–33.
- Alberti, Gábor, Judit Kleiber (2004): The GeLexi MT Project. In: John Hutchins (ed.): *Proceedings of EAMT 2004 Workshop*, Valletta: Univ. of Malta, 1–10.
- Alberti, Gábor, Judit Kleiber, Anita Viszket (2004): GeLexi project: Sentence Parsing Based on a GEnerative LEXICon. In: *Acta Cybernetica* 16, 587–600.
- Farkas Judit, Alberti Gábor (2017): Általánosítások a potenciális magyar birtokszóváltozatokra vonatkozóan. *Jelentés és Nyelvhasználat* 4: 59–79.
- Kálmán, László (1990): Deferred Information: The Semantics of Commitment. In: László Kálmán, László Pólos (eds.): *Papers from the Second Symposium on Logic and Language*. Budapest: Akadémiai Kiadó. 125–157.
- Karttunen, Lauri (1986): *Radical Lexicalism*, Report No. CSLI 86–68, Stanford.
- Kiefer Ferenc (2000): *Jelentéselmélet*. Budapest: Corvina.
- Laczkó Tibor (2000): Az ige argumentumszerkezetét megőrző főnévképzés. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 3. Morfológia*. Budapest: Akadémiai Kiadó. 293–407.
- Pustejovsky, James (1995): *The Generative Lexicon*. MIT Press, Cambridge, MA.