

Nyelvtechnológia 1

BME, 2007. november 6.

Prószéky Gábor



A számítógépes nyelvészet történelme

- ☐ Általános tapasztalat: a nyelv változik
- ☐ Ezért: a nyelvészet a 20. századig = történeti nyelvészet
- ☐ A deskriptív nyelvészet (és a „preskriptív” nyelvészet)
- ☐ A világháborúk körül:
 - 1947-49: USA/UK Booth és Weaver: kriptográfia
- ☐ Amerikai kezdemények
 - 1951: Bar-Hillel (MIT) - később (1959) megkérdőjelezi a TAGF-t
 - 1954: Georgetown/IBM - az USA kormánya támogatta
- ☐ Szovjetunió és Kelet-Európa: matematikai nyelvészet
- ☐ Generatív grammatika: Chomsky „Syntactic Structures”-e épp 50 éve (!)
- ☐ Transzformációk a „csúcson”: az „Aspects” (1965)
- ☐ A GF-korszak vége: ALPAC Report (1966)
- ☐ Woods (1969): Lunar (holdprogram!), ATN
- ☐ Winograd (1972): SHRDLU
- ☐ GF-túlélők: Systran, Logos (vietnami háború!) és Metal
- ☐ Az USÁ-n kívül:
 - Kanada: METEO
 - Európai közösség: EC-Systran + közvetítőnyelves GF (Eurotra, DLT)
 - Japán: 5. generációs számítógép (!)

A számítógépes nyelvészet történelme 2.

- 1970-es évek vége: megjelennek a „Bay Area” nyelvtanok
- 1980-as évek: tudás-alapú GF-rendszerek (CMU)
- 1980-as évek: Rosetta(Montague-szemantika mint interlingva)
- 1983: Language as a Cognitive Process (Winograd)
- 1983: kétszintes morfológia (Koskenniemi)
- 1980-1990-as évek: megjelennek a piaci alkalmazások
 - Apple Macintosh, IBM PC: helyesírás, elválasztás stb.
 - Logos
 - Siemens Metal > Langenscheidt T1
 - PC-alapú rendszerek: Globalink, Tovna, Kielikone, ProMT
- 1980-as évek vége: IBM - statisztikai módszerek
- 1990-es évek: megszületőben a korpusznyelvészet
- 1990-es évek vége: a Lernout & Hauspie tündöklése és bukása, majd a ScanSoft (amit 2005-től a Nuance) felvásárolja
- 2000-es évek (?): a beszédfordító rendszerek ígérete
- ... és minden egyre inkább statisztikai módszerekkel