

Märgendajate kasutamine

- Märgendajate kasutusalaad:
 - Infootsing
 - Tekstitöötlus (korrektorid)
 - Kõnetöötlus (süntees ja tuvastamine)
 - Käekirja tuvastamine
 - Masintõlkimine
- Märgendatud korpuste kasutusalaad:
 - Sõnaraamatud ja grammatikad
 - Keeleõppevahendid
 - NLP tarkvara arendamine

Märgendisüsteemide standard

- EAGLES 1993 – Indo-Euroopa keeltele Euroopas
- Põhikategooriad: N, V, AJ, PD, AT, AV, AP, C, NU, I, U, R, PU
- N: tüüp (com, prop), sugu, arv, kääne (nom, gen, dat, acc, voc)
- V: isik, sugu, arv, pöördelisus, kõneviis/vorm (indic, subj, imper, cond, inf, partic, ger, sup)
- AJ: võrre, sugu, arv, kääne
- PD: isik, sugu, arv, poss (sg, pl), kääne, kategooria (p, d), tüüp (dem, indef, poss jne)
- U: neg partikkel, inf marker
- R: tüüp (võõrsõna, valem, sümbol, akronüüm, lühend, muu)

Eesti keele morfoloogilise ühestaja märgendid

- Sõnaliik: S, V, A, (G), P, D, K, J, N, I, X, Y, Z
- Nimisõnad e Substantiivid S
 - Tavaline (com) või nimi (prop)
 - Arv (sg, pl)
 - Kääne (nom, gen, part, adit, ill, in, el, all, ad, acl, tr, term, es, abes, kom)
- Omadussõnad e Adjektiivid A
 - Võrre (pos, com, super)
 - Arv
 - Kääne

Märgendid -2

- Verbid e tegusõnad V
 - Liik: main, aux, mod
 - Kõneviis või vorm: indic, imper, cond, quot, inf, partic, ger, sup
 - Aeg: pres, past, impf
 - Tegumood: ps,imps
 - Isik: ps1, ps2, ps3
 - Arv: sg, pl
 - Kõne: af, neg
 - Kääne: ill, in, el, tr, ab
 - Reksioon jms: .Intr, .Part, .NGP-P, .FinV

Märgendid -3

- Pronoomen e asesõna P
 - Liik: pers,dem, indef, pos, inter rel, refl, rec, det
 - Isik: ps1, ps2, ps3
 - Arv: sg, pl
 - Kääne

Märgendid -4

- Adpositsioonid e kaassõnad K
 - Liik: pre, post
 - Reksioonikääne: .part, .gen, .abes, .kom, .term, .all, .el, .nom
- Sidesõna J
 - Liik: crd, sub
- Arvsõna N
 - Liik: card, ord, digit, roman
- Hüüdsõnad I
- Muu X

Mitmesused

- Sagedasemad mitmesed sõnavormid:
 - Ta, ei, oli, on, kui, aga, ma, oma, siis, mis, nagu, või, veel, ära, kes, pole, selle, küll, kõik, polnud – veerand kõigist mitmestest sõnadest
- Morfoloogilise info põhjal:
 - D & J (aga, kui, siis, nagu)
 - D & V aux neg (ei)
 - Partitsiibid
 - P pers ps3 sg gen & P pers ps3 sg nom (ta/tema)
 - V main & V aux (oli, on)
 - S sg gen & S sg nom & S sg part (maja, kala jne)
 - D & A pos sg abl (kindlalt, tugevalt)
 - S adit & S part & S gen (jaama, pealkirja)
 - P inter rel pl nom & P inter rel sg nom (kes, mis)
 - D & K (vastu üle, all)
 - 1200 tüüpi