

ADJEKTIIVIDE POLÜSEEMIA KORPUSES JA SÕNARAAMATUS

Maria Tuulik

Ülevaade. Artikli eesmärk on uurida eesti omadussõnade semantikat ja eelkõige polüseemiat koondkorpuse 100 sagedama adjektiivi näitel. Esimeses pooles vaadeldakse valimi jagunemist mitmetähenduslikkuse aspektist ja selleks, et selgitada polüseemia ja struktuuri suhteid, analüüsitakse ka adjektiivide morfoloogilise struktuuri varieerumist. Teises pooles keskendutakse semantilistele tüüpidele, mille abil saaks mitmetähenduslike adjektiivide tähendusnihkeid paremini kirjeldada. Tutvustatakse semantiliste tüüpide rakendusvõimalusi sõnaraamatute koostamisel ja pakutakse välja eesti adjektiivi peamised semantilised tüübid.

Võtmesõnad: semantika, leksikoloogia, leksikograafia, sõnamoodustus, semantilised tüübid, eesti keel

1. Sissejuhatus

Adjektiive pole traditsioonilises leksikaalsemantikas nii ulatuslikult uuritud kui nimisõnu või verbe, ometi on nad semantiliselt vähemalt sama mitmekesised (I. Peters, W. Peters 2000). Näiteks võib adjektiivi tähendus muutuda vastavalt sellele, millist konkreetset substantiivi tähenduskomponenti ta rõhutab ja samuti vastavalt sellele, mis tüüpi substantiivi täiendina ta esineb.

Artiklis käsitletakse eesti keele adjektiivide polüseemiat korpusandmete ja leksikograafiliselt koostatud kirjelduste põhjal. Sõnaraamatuesituste uurimisel kasutatakse peamise allikana praegu koostatavat ühekõitelist eesti keele seletavat sõnaraamatut (Langemets jt 2010), kus polüseemia ühtlasema esituse huvides märgitakse sõnatähenduste juurde semantilisi tüüpe. Spetsiaalselt adjektiivide jaoks pole eesti keeles semantilisi tüüpe veel välja töötatud, vaid kasutatakse substantiivide analüüsil tuvastatud tüüpe (Langemets 2010). Kuna adjektiivide semantilised tüübid sobituvad substantiivi omadega vaid osaliselt, siis arutletakse, milliseid oleks veel otstarbekas lisada.

Et leksikograafias toetab semantilise tüübi lisamine sõnatähendustele sama tähendusvälja sõnade sarnasemat kirjeldust, siis on adjektiivide semantiliste tüüpide väljatöötamine ka oluline etapp Eesti Keele Instituudi leksikograafi tööriista – EELexi sõnastikusüsteemi – arendamisel. Artikli rakenduslik väljund ongi eesti adjektiivide semantiliste tüüpide loend, mida saab sõnaraamatutöös kasutada.

Lisaks aitab semantiliste tüüpide selgitamine tulevikus tuvastada eesti adjektiivide süstemaatilise polüsemia malle. Margit Langemets (2010: 252) toob välja, et süstemaatilisi malle kodeerides on võimalik ühtlustada semantilise info esitust keelevaras, näidata sõnatähenduste vahelisi loogilisi seoseid ja vähendada seeläbi polüsemia loetlevat esitust, mis on omane traditsioonilisele sõnaraamatule.

2. Materjali valik

Korpuse töötlemise programmi Sketch Engine¹ abil on Tartu Ülikooli koondkorpusest tehtud adjektiivide sagedusloend (100 adjektiivide) ja puhastatud valim automaatse morfoloogilise analüüsi vigadest, et sisse jääksid ainult nii-öelda tõelised omadussõnad. Kuid millised on tõelised adjektiivid ja mida tuleks sõnade adjektiiviks määratlemisel arvesse võtta?

“Eesti keele grammatika” (EKG 1995: 20) defineerib omadussõnu kui tüüpjuhul käändes, arvus ja võrdlusastmetes muutuvaid sõnu, mis väljendavad (asja) omadust ning võivad lauses esineda ühilduva atribuudi või teatud liiki verbide laiendina ka predikatiivi ja seisundimääruse positsioonis. Kuid kõik semantilised adjektiivid pole keeles grammatiliselt adjektiivid. Eesti keeles on keeruline piiri tõmmata näiteks partitsiipide või käändumatute adjektiivide puhul. Käändumatud adjektiivid, mis automaatse analüüsiga esile tulid, jäeti semantika ja süntaksi põhjal valimisse sisse, nt *eri*. Samas partitsiipide iga juhtumit vaadati eraldi, kuna sagedusloendis leidis läbiseegi reaalselt adjektiividena kasutusel partitsiipe ning puhtaid verbivorme, mis olid korpuses samamoodi adjektiiviks märgendatud.

Sagedusloendi 100 esimese adjektiivide seas oli 17 sõna, milles võis kahelda. Need olid korpusetextides ainult verbina kasutusel olevad sõnad (nt *olev*, *pandud*, *hakanud*, *kuuluv*), peamiselt verbina kasutusel olevad sõnad (nt *läinud*) ja sõnad, kus adjektiivne kasutus esines verbivormidega pooleks (nt *tulev*, *kehtiv*). Kuigi sõna *läinud* puhul on võimalik ka adjektiivne kasutus tähenduses ‘eelmine’ (nt *läinud aasta*), siis korpust ja sõnavisandeid uurides ilmnes, et valdav osa *läinud* kasutustest olid siiski verbivormid (*hästi läinud*, *tööle läinud*, *meelest läinud* jne). Tuli ette ka keerulisemaid juhtumeid, kus adjektiivne kasutus paistis korpuses verbivormidega pooleks esinevat, nt sõnad *tulev* ja *kehtiv*. Sõnavisanditest võis näha, et üsna palju on adjektiivset kasutust, nt *tuleval aastal*, *nädalal*, *hooajal*, aga samuti leidis hulgaliselt verbivorme, nt *välismaalt tulev*, *arutusele tulev*.

Sellistel juhtudel hinnati, kas partitsiibi adjektiivse kasutuse tähendus on nihkunud, see tähendab, et ei kattu enam verbi omaga (nt *elav* tähenduses ‘liikuv, elurõõmus’). Siin võis mängida rolli ka partitsiibi semantiline läbipaistmatus. Seejärel vaadati, kas võrdlusastmed on võimalikud, näiteks saab võrrelda *särav*, *säravam*, *säravaim*, ent mitte *kuuluv*, *kuuluvam*, *kuuluvaim*. Ja viimaks prooviti, kas partitsiibile on võimalik adverbi ette panna, näiteks kõlab täiesti loomulikult

¹ Sketch Engine on tekstikorpuste töötlemisprogramm, mille abil saab koostada ja töödelda konkordantse, teha sagedusloendeid, sõnavisandeid jm (Kilgarriff jt 2004). Selle eestikeelse mooduli kohta vt Kallas 2013.

väga erinev, kuid mitte *väga toimuv*. Kokkuvõtvalt võib öelda, et lähtuti leksikaal-semantilisest vaatepunktist ja valimisse jäeti sisse leksikaliseerunud, sisulises või vormilises mõttes oma paradigmast väljunud partitsiibid.

Automaatse analüüsiga sattus adjektiivide sagedusloendisse veel sõna *kogu*, millel EKSS-is adjektiivset tähendust polegi. Kokku tõsteti kõrvale 17 sõna ja võeti sama palju juurde n-õ õigeid adjektiive, mis sagedusloendis järgmiseks tulid. Eesti keele kõigi sõnade 100 sagedama sõna hulka mahtus adjektiividest üksnes kolm: *uus*, *suur* ja *viimane*.

3. Ülevaاتlikult valimist

Koondkorpuse põhjal on kõige sagedamad adjektiivid eesti keeles *uus*, *suur*, *viimane*, *hea* ja *võimalik*. Võrdlusena võiks tuua Oxfordi korpuse² põhjal inglise keele sagedamad adjektiivid *good* 'hea', *new* 'uus', *first* 'esimene', *last* 'viimane' ja *long* 'pikk'. Kolm viiest sagedamast adjektiivist (*uus*, *hea* ja *viimane*) kattuvad ning ka sõna *esimene* sobiks oma sageduse poolest loetelusse, kuid on koondkorpuses märgendatud numeraalina ja seetõttu siinses eesti adjektiiviloendis esile ei tule.

Võrreldes ülejäänud eesti ja inglise keele 100 sagedamat omadussõna, leidis kinnitust asjaolu, et valdav osa sagedatest adjektiividest on universaalsed. Samas paistab eesti omadussõnade tagant korpuse iseloom. Kuna suur osa koondkorpuse tekstidest on ajakirjanduslikud, siis on valimis sees sõnad nagu *majanduslik*, *sotsiaalne*, *rahvuslik*, *juriidiline* ja *erapooletu*, kuid esile ei tule näiteks meeleolu märkivad adjektiivid nagu *rõõmus*, *kurb* või *vihane*, mis mõjuvad märksa igapäevasemana.

Värvidest kuuluvad eesti 100 sagedama adjektiivi hulka *must*, *valge*, *punane* ja *roheline*. See langeb kokku Berliini ja Kay (1969) hierarhiaga, mille järgi on *must* ja *valge* põhivärvinimed ehk kõige primaarsemad. Kui keeles on kolmas värvisõna, siis on see *punane*, ja kui neljas, siis on see *roheline* või *kollane*. Valimist nähtub, et samas järjekorras tulevad värvisõnad ka sageduse põhjal esile. Kuid miks tuleb eesti keeles esimesena, peaaegu poole suurema sagedusega välja *roheline*, mitte *kollane*? Korpuse lauseid analüüsides selgub, et *rohelise* sagedusele aitab kaasa keskkonnanakaitsjate tähendus ja keskkonnakaitsesega seotud väljendite rohkus (nt *roheline mõtteviis*, *liikumine*, *energia* jne). Ka *kollasel* on sagedaid kollokatsioone teises tähenduses 'põnevust, sensatsiooni taotlev', nt *kollane ajakirjandus*, *ajaleht*, aga üldiselt ei tundu *kollase* teine tähendus nii produktiivne ja semantiliselt avatud kui *rohelise* oma.

4. Morfoloogiline struktuur ja polüseemia

Morfoloogilise struktuuri järgi jaotub valim järgmiselt: juuradjektiive on 34, tuletisi 57, liitsõnu 3 ja *v*-kesksõnu 6 (vt tabel 1). Sõnade juuradjektiivideks, tuletisteks ja liitsõnadeks jagamisel tugineti Silvi Vare sõnapere sõnaraamatule (2012).

² Vt <http://oxforddictionaries.com/words/the-oxford-english-corpus> (15.10.2013).

Tabel 1. Valimi jaotumine morfoloogilise struktuuri järgi

Valim	Kokku	Näited
juuradjektiivid	34	<i>uus, suur, hea, vana, noor, vaba, väike</i>
tuletised	57	<i>tõeline, vajalik, järgmine, edukas, ilus</i>
liitsõnad	3	<i>rahvusvaheline, käesolev, lugupeetud</i>
v-kesksõnad	6	<i>erinev, huvitav, sobiv, vastav, tulev, elav</i>

Reet Kasiku (2004: 140) järgi on oluline osa eesti keele adjektiividest *lik-*, *line-* või *ne-*tuletised ja juuradjektiive on suhteliselt vähe. Vaadeldud omadussõnade loendis moodustasid juuradjektiivid kolmandiku ja sageduse poolest olid need esil, esimese 20 seas oli tüvisõnu isegi üle poole, kinnitades seega Zipfi (1949) vähima pingutuse seadust, mille järgi kõige sagedamad sõnad on ühtlasi ka kõige lühemad. Eesti sagedamad juuradjektiivid on enamasti “kulunud” 3-4-täheliseks.

Tuletised jagunevad järgnevalt: 28 *ne-*tuletist, 13 *lik-*tuletist, 10 *line-*tuletist ja 6 muud, mille hulgas omakorda leidus *tu-*liitelisi, *s-*liitelisi ja *ik-*liitelise tuletis. Tabel 2 esitab tuletusliidete järgi jagatud tuletiste esinemisarvud koos näidetega uuritud valimist.

Tabel 2. Tuletised liidete järgi

Tuletised	Kokku	Näited
<i>ne-</i> tuletised	28	<i>tänane, lihtne, tõsine, endine</i>
<i>lik-</i> tuletised	13	<i>ohtlik, võimalik, kohalik, vajalik</i>
<i>line-</i> tuletised	10	<i>tavaline, poliitiline, roheline, tõeline</i>
muud (<i>tu-</i> , <i>ik-</i> , <i>s-</i> tuletised)	6	<i>edukas, ilus, tähtis, võimatu, üksik, erapooletu</i>

Adjektiivide tähenduste arv pärineb koostamisel oleva ühekõitelise eesti keele sõnaraamatu (Langemets jt 2010) tööbaasist. Võrreldes suure seletava sõnaraamatuga (EKSS 2009), kus tähendusi loetleti paljuski lineaarselt, esitatakse siin polüseemiat prototüüpsete tuumtähenduste kaupa, millest lähtuvad teised, tuletatud tähendused. See tähendab EKSS-i peente tähenduseristuste üldistamist, ümberrühmitamist ja sageli kokkuvõtmist (Langemets jt 2010: 799–780). Praegu on märksõnastiku ligikaudu 9000 adjektiivist koostatud umbes pooled.

Polüseemseid sõnu, st selliseid, kus sõnaraamatus on eristatud kaks või enam põhitähendust, oli 100 sagedama adjektiivi hulgas kokku 80, sealhulgas ka üks homonüüm (*täis*). Monoseemseid adjektiive, millel sõnaraamatus on üks põhitähendus, oli valimis 20. Polüseemsete adjektiivide suur ülekaal on ootuspärane, kinnitades samuti juba Zipfilt pärinevat väidet, et sageli esinevad sõnad kipuvad olema polüseemsemad (Zipf 1945: 251–256).

Uuritud 100 adjektiivil oli põhitähendusi kokku 341. Tabel 3 annab ülevaate sellest, kuidas sagedamad adjektiivid jagunevad vastavalt tähenduste arvule.

Tabel 3. Valimi jagunemine põhitähenduste arvu põhjal

Tähenduste arv	1	2	3	4	5	6	7	8
Sõnu vastava tähenduste arvuga	20	31	8	10	9	9	8	5

Substantiivide analüüsil on Langemets (2010: 88) välja toonud, et kõige lühemad sõnad ehk tüvisõnad on kõige polüseemsemad, seetõttu huvitas uurijat, kas sama tendents kehtib ka adjektiivide juures. Sel eesmärgil arvutati eraldi juuradjektiivide, tuletiste, liitsõnade ja partitsiipide polüseemia indeks (tabel 4).

Tabel 4. Adjektiivide polüseemia indeks ehk keskmine põhitähenduste arv sõna kohta

Valim	Kokku	Polüseemia indeks
juuradjektiivid	34	5
tuletised	57	2,6
liitsõnad	3	1,3
v-kesksõnad	6	2,3

Tabelist 4 nähtub, et juuradjektiivide polüseemia indeks on 5, tuletistel aga 2,6 ehk ligikaudu poole väiksem. Järgnevad v-kesksõnad (2,3) ja liitsõnad (1,3). Kõik 6-tähenduslikud ja valdav osa 7-tähenduslikest sõnadest olid juuradjektiivid. Monoseemseid juuradjektiive leidis 100 adjektiivi seas ainult kaks, *eri* ja *kuulus*. Kõige polüseemsemad tüvisõnad ja tuletised valimis olid sellegipoolest tähenduste arvu järgi võrdsed: *hea* ja *vana* 8 tähendusega ning *viimane* ja *positiivne*, samuti 8 tähendusega. Kõige vähem polüseemsed olid liitsõnad, mis taas toetab Zipfi (1945, 1949) väidet, et sagedamad sõnad on lühemad. Seega võib arvata, et liitsõnad (“tänu” oma pikkusele) kalduvad olema pigem monoseemsed.

5. Semantilised tüübid ja nende rakendamisvõimalused

Semantilised tüübid on abstraktsed abivahendid tähendusüksuse juures, mille (kombinatsioonide) abil on võimalik võrrelda omavahel semantiliselt seotud sõnu, kirjeldada sõna sisemisi tähendusmalle. Näiteks Langemetsa järgi (2010: 198–200) võib isikuomaduste sõnade puhul üsna regulaarselt esilduda suhe OMADUS–INIMENE, mis tähendab, et üks ja sama sõna tähistab nii (füüsilist) omadust kui ka isikut, keda selle omaduse järgi tuntakse või iseloomustatakse. Semantilised tüübid on vastavalt OMADUS ja INIMENE_OMADUS (näide 1).

- (1) OMADUS–INIMENE
 Kõlas mahe bass. (OMADUS)
 Esines kuulus bass. (INIMENE_OMADUS)

Semantilised tüübid võivad olla uurijale abiks polüseemsete sõnade analüüsil, eriti süstemaatilise polüseemia selgitamisel. Ühtlasi võivad semantilised tüübid sõnaraamatutöös olla ka praktiliseks abivahendiks sõna semantika kodeerimisel. Kui teatud sõnade tähenduste vahelduses ilmnevad regulaarsused on juba kindlaks tehtud, siis võib vastavaid malle kodeerides sõnaraamatus semantilise info esitust korrastada ning näidata ja selgitada sõnatähenduste vahelisi loogilisi seoseid (Langemets 2010: 252).

Peale selle võimaldab semantiliste tüüpide kasutamine sõnaraamatutöös samasse tähendusvälja kuuluvaid sõnu ühtlasemalt kirjeldada. Näiteks koostati

eesti keele põhisõnavara sõnastiku (PSV 2014) substantiivid Langemetsa tüpoloogialusel mõistegruppide kaupa (linnud, loomad, kehaosad, kuud jne), mitte tähestiku järjekorras, nagu on sõnastik ise. Sellises koostamismudelil on võimalik sama tüüpi sõnade puhul kasutada sarnast leksikograafilist malli (nt sarnase ülesehitusega definitsioone), mille tulemuseks on semantilise info ühtlasem esitus sõnaraamatus.

Teine väga oluline semantiliste tüüpide rakendusvaldkond on korpuste semantiline märgendamine, mida Eestis pole seni veel tehtud. Semantilise märgendamise peamine eesmärk on sõnatähenduste ja sõnaühendite tähenduste ühestamine, mis on vajalik näiteks masintõlkeprogrammide loomisel. Sisuliselt tähendab see polüsemse sõna tähenduste eristamist tekstis. Semantiline tüüp aitab märgendada sõnade tähenduslikku kuuluvust: korpuses esinevale sõnale või sõnaühendile lisatakse märges vastavalt selle semantilisele kategooriale. Sel moel saaksid semantilised tüübid toetada automaatset semantilist keeleanalüüsi ja sõnatähenduste ühestamist.

6. Eesti adjektiivide semantilised tüübid

Semantiliste tüüpide väljatöötamisel analüüsiti ühekõitelise seletava sõnaraamatu põhjal 100 sagedama adjektiivide 341 põhitähendust. Tähdused jagati semantikast lähtudes rühmadesse ja uuriti definitsioone, et vaadata, kas teatud tüüpi adjektiivid on ka sarnasel viisil defineeritud ja ehk annab see, millist tähenduse elementi või suhet definitsioonis välja tuuakse, vihje semantilise tüübi määramiseks.

Ühekõitelises eesti keele seletavas sõnaraamatus märgitakse adjektiivide juurde substantiivide analüüsil tuvastatud tüüpe (Langemets 2010). Kuna adjektiivide semantilised tüübid substantiivide semantiliste tüüpidega päriselt ei sobitu, siis ongi käesoleva artikli üks eesmärk pakkuda välja tüübigi just spetsiaalselt omadussõnadele.

Adjektiivide semantiliste tüüpide arv võib erinevates käsitlustes olla väga erinev. Uurimuses kasutati eeskujuna ja võrdlusmaterjalina Dixoni (1982) seitset olulisemat adjektiivide semantilist tüüpi, leksikaalse andmebaasi SIMPLE koostamisel väljatöötatud 18 tüüpi (Lenci jt 2000: 30–33) ning Hundsnurderi ja Spletti (1982) uurimuse põhjal loodud GermaNeti 15 adjektiivide semantilist klassi, millel on omakorda 78 alaliiki (vt ka Orav 2000).

Tuntud adjektiivide uurija Robert Dixon on välja toonud 7 olulisemat adjektiivide semantilist tüüpi: DIMENSIOON (nt *suur, pikk, lai*); VANUS (nt *uus, vana*), HINNANG (nt *hea, halb, veider*), VÄRVUS (nt *valge, must, punane*), FÜÜSILISED OMADUSED (nt *kõva, tugev, puhas*), KIIRUS (nt *kiire, aeglane, äkiline*) ja INIMESEGA SEOTUD OMADUSED (nt *vihane, rõõmus, armukade*). Neli esimest (DIMENSIOON, VANUS, HINNANG ja VÄRVUS) on Dixoni järgi kõige enam seotud adjektiivide kategooriaga ja moodustavad omadussõnade semantilise prototüübi, see tähendab, et keeltes, kus on piiratud hulk omadussõnu, kuuluvad need tõenäoliselt just nendes tüüpidesse (Dixon 1982: 46).

Semantilise tüübigi koostamisel katsuti silmas pidada ka rakenduslikku eesmärki, et seda oleks võimalik sõnaraamatutöös päriselt kasutada. See tähendab, et tüüpe ei tohiks olla nii palju, et need muutuksid asjaks iseeneses, samuti ei tohiks tekkida olukorda, kus leksikograafil on iga tähenduse juures kõvasti peamurdmist,

millise tüübi kasuks otsustada. Seetõttu üritati ka tüüpide nimetused hoida võimalikult läbipaistvad ja arusaadavad. Samas oli analüüsi huvides jagada tüübid peeneteralisemaks, kui need hetkel ühekõitelises seletavas sõnaraamatus on. Kuna selleks, et oleks võimalik järgmise sammuna selgitada adjektiivide regulaarseid tähendusvaheldusi ehk avada süstemaatilise polüseemia malle, peavad nende semantilised tüübid olema senistest (vt tabel 5) märksa kitsamad.

Tabel 5. Ühekõitelise seletava sõnaraamatu omadussõnade semantilised tüübid

Semantiline tüüp	Kirjeldus
OMADUS_FÜÜS	füüsikaline omadus (mass, kõrgus, kiirus, heli, värv jne)
OMADUS_FÜÜS_VÄRV	värv, värvus
OMADUS_KVAL	hinnanguline omadus
OMADUS_MAX	intensiivsust, suurust rõhutav omadus
OMADUS_MIN	väiksust, märkamatus rõhutav omadus
OMADUS_PSÜHH	psüühikaga (tunnetega, tajuga) seotud omadus

Praegu on ühekõitelises seletavas sõnaraamatus adjektiivide jaoks kasutusel kuus semantilist tüüpi (lisaks veel üldine tüüp OMADUS). Sellise tüübigruppiga ei tule tähendusvaheldused alati esile, näiteks *vana* tähenduses 'kunagine', mis praegu on märgitud lihtsalt omaduseks, võiks olla OMADUS_AEG, *vana* tähenduses 'teatava vanusega', mis on samuti omaduseks märgitud, võiks olla OMADUS_VANUS jne.

Tabelis olevatest tüüpidest jäid uude tüpoloogiasse samal kujul OMADUS_FÜÜS ja OMADUS_PSÜHH, OMADUS_MAX ja OMADUS_MIN sulatati kokku tüübiks OMADUS_RÕHUTAMINE, kuna materjali uurides torkas silma, et muidu võis tüübi nime ja selle sisu vahel kohati vastuolu tekkida, näiteks *viimase* tähendus 'muude seas vähima väärtusega, kõige madalamal asuv, kõige viletsam' (näitelause: *Kas mina olen siis see viimane, kellega tantsidagi ei kõlba?*) on märgitud tüübiks OMADUS_MAX ja on arusaadav, et mõeldakse 'maksimaalselt vähim, viletsaim', kuid siiski võiks seda tähendust semantika põhjal tõlgendada ka tüübina OMADUS_MIN ehk 'kõige vähem väärtuslik'.

OMADUS_FÜÜS_VÄRV ja OMADUS_KVAL jäid samuti uude tüübigruppi sisse, kuid need nimetati ümber tüüpideks OMADUS_VÄRV ja OMADUS_HINNANG. Esimene seetõttu, et sobituks vormi poolest paremini ülejäänud tüüpidega ja teine lihtsalt nimetuse läbipaistvuse huvides.

Valimi analüüsist koorus 15 omadussõna semantilist tüüpi, mis on esitatud tabelis 6 koos seletuse ja näidetega. Näited esindavad vastavate adjektiivide EKSS-i esimest põhitähendust (mitte kõiki). Kui silmas peetakse mitteprimaarset tähendust, on näite taha lisatud selgitus. Tärniga (*) on tähistatud näited, mis valimisse ei sattunud.

Ideaalis võiksid kõik semantilised tüübid üksteist välistada, kuid tegelikkuses on sellist tulemust väga raske – kui mitte võimatu – saavutada. Tabelis 6 toodud tüüpidest on mõni laiem kui teised, näiteks OMADUS_FÜÜS hõlmab ka tüüpe OMADUS_MASS, OMADUS_VÄRV, OMADUS_MÕÕDE jms. OMADUS_FÜÜS on mõeldud kasutamiseks teise valikuna ehk siis, kui adjektiivi tähendust ei saa paigutada kitsamasse tüüpi.

Tabel 6. Adjektiivide semantilised tüübid

Semantiline tüüp	Kirjeldus	Näited valimist
OMADUS_FÜÜS	füüsikalised omadused, mis ei kuulu kitsamatesse tüüpidesse	<i>tugev, nõrk, puhas</i>
OMADUS_MÕÕDE	ruumilised omadused	<i>pikk, lai, lühike, kõrge</i>
OMADUS_LIIKUMINE	liikumise ja tempoga seotud omadused	<i>kiire, aeglane</i>
OMADUS_VÄRV	värvusega seotud omadused	<i>must, valge, punane, roheline</i>
OMADUS_MASS	massiga seotud omadused	<i>raske, kerge</i>
OMADUS_TEMPERatuur	temperatuuriga seotud omadused	<i>külm, soe</i>
OMADUS_AEG	ajaga seotud omadused	<i>eilne, tänane, käesolev, praegune, endine</i>
OMADUS_KOHT	mingi kohaga seotud omadused	<i>sealne, kohalik</i>
OMADUS_VANUS	kellegi või millegi eaga, olemasolu ajaga seotud omadused	<i>vana (ka tähenduses 'kulunud'), noor, uus, värske</i>
OMADUS_KOGUS	hulgaga seotud omadused	<i>suur (tähenduses 'arukas'), vähene</i>
OMADUS_ÜHISKOND	sotsiaalsete suhetega seotud omadused	<i>poliitiline, rahvuslik, sotsiaalne, avalik, riiklik</i>
OMADUS_RÖHUTAMINE	intensiivsust, suurust või väiksust rõhutavad omadused	<i>täielik, viimane (nt käitub nagu viimane narr)</i>
OMADUS_HINNANG	hinnangu ja sealhulgas modaalsusega seotud omadused	<i>õige, huvitav, ohtlik, sobiv, halb, võimalik</i>
OMADUS_PSÜHH	meeleolu, iseloomu ja vaimsete võimetega seotud omadused	<i>tõsine, tark*, rõõmus*</i>
OMADUS_SUHTELISUS	omadused, mida defineerib võrdlus või suhe millegi muuga	<i>sarnane, keskmine, viimane</i>

Näites (2) on lühendatud kujul esitatud EKSS-i artikkel *endine*, mille tähendusi on varasematesse semantilistesse tüüpidesse keeruline liigitada. Nende järgi klassifitseeruksid mõlemad tähendused üldsõnalise tüübi OMADUS alla, kuid uusi tüüpe kasutades võiks esimene tähendus olla OMADUS_AEG ja teine tähendus OMADUS_SUHTELISUS.

(2) **endine** <adj> <-se 5 või -se 4>

1 [OMADUS_AEG] varem kellekski või millekski olnud, varem kasutusel olnud, kunagine. *Endine õpetaja, direktor.*

■ varasemasse, möödunud aega kuuluv, muistne. *Endised ajad, inimesed.*

2 [OMADUS_SUHTELISUS] selline nagu varem, muutumatu. *Elab endist viisi.*

Loodud tüübigraafik ei ole kindlasti ammendav ja võib edasise analüüsi käigus täiendada. Kuid loodetavasti saab siiski ka praegu asendada sõnastikubaasis lihtsalt omaduseks märgitud tähendused täpsema tüübiga. Ja leksikograafi vaatepunktist võib polüsemse sõna puhul juba semantiliste tüüpide lisamine aidata mõtteid organiseerida ja paremini tähendusjaotust struktureerida.

7. Kokkuvõte ja tulevik

Artiklis kirjeldati ja analüüsiti eesti 100 sagedamat adjektiivii nii polüseemia kui morfoloogilise struktuuri seisukohast. Eraldi vaadeldi juuradjektiivide, tuletiste, liitsõnade ja partitsiipide polüseemia indeksit ehk keskmist tähenduste arvu sõna kohta. Kõige polüseemsemad on eesti keeles juuradjektiivid (polüseemia indeks 5). Tuletiste polüseemia indeks on 2,6 ehk ligikaudu poole väiksem. Kõige monoseemsemad olid 100 sagedama adjektiivii hulgas liitsõnad (polüseemia indeks 1,3).

Artikli teises pooles tutvustati semantiliste tüüpide rakendusvõimalusi, nii teoreetilise abivahendina polüseemsete sõnade analüüsil kui ka praktilise abivahendina sõnasemantika kodeerimisel sõnaraamatutöös. Analüüsides ühekõitelise seletava sõnaraamatu põhjal 100 sagedama adjektiivii põhitähendusi, pakuti välja 15 semantilist tüüpi spetsiaalselt eesti keele (polüseemsete) adjektiivide kirjeldamiseks.

Järgmise sammuna on plaanis tuvastada eesti adjektiivide süstemaatilise polüseemia mallid, mida nüüd, tänu olemasolevatele semantilistele tüüpidele, on võimalik teha. Selle töö eesmärk omakorda oleks uurida eesti adjektiivii polüseemiat ja luua ka teoreetiline alus, mille põhjal saaks ühtlustada adjektiivide polüseemia esitust sõnaraamatus.

Viidatud kirjandus

- Berlin, Brent; Kay, Paul 1969. Basic Color Terms. Their Universality and Evolution. Los Angeles.
- Dixon, Robert M. W. 1982. Where have all the adjectives gone? And other essays in syntax and semantics. *Janua linguarum. Series Maior*, 107. Berlin: Walter de Gruyter. <http://dx.doi.org/10.1515/9783110822939>
- EKG 1995= Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1995. Eesti keele grammatika I. Eesti Teaduste Akadeemia, Eesti Keele Instituut. Tallinn.
- EKSS = Eesti keele seletav sõnaraamat 1–6. (“Eesti kirjakeele seletussõnaraamatu” 2., täiendatud ja parandatud trükk.) [The Explanatory Dictionary of Estonian.] Toimetanud Margit Langemets, Mai Tiits, Tiia Valdre, Leidi Veskis, Ülle Viks, Piret Voll. Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus, 2009.
- Hundsnurher, Franz; Splett, Jochen 1982. Semantik der Adjektive im Deutschen: Analyse der semantischen Relationen. Westdeutscher Verlag. <http://dx.doi.org/10.1007/978-3-322-87624-9>
- Kallas, Jelena 2013. Eesti keele sisusõnade süntagmaatilised suhted korpus- ja õppeleksikograafias. [Syntagmatic Relationships Of Estonian Content Words In Corpus And Pedagogical Lexicography.] Tallinna Ülikooli humanitaarteaduste dissertatsioonid 32. Tallinn: Tallinna Ülikool.
- Kasik, Reet 2004. Eesti keele sõnatuletus. [Derivation in Estonian.] Tartu: Tartu Ülikooli Kirjastus.
- Kilgarriff, Adam; Rychly, Pavel; Smrž, Pavel; Tugwell, David 2004. The Sketch Engine. – Proceedings of the 11th EURALEX International Congress (Lorient, 6–10 July 2004). Lorient: Université de Bretagne-Sud, 105–117.
- Langemets, Margit 2010. Nimisõna süstemaatiline polüseemia eesti keeles ja selle esitus keelevaras. Tallinn: Eesti Keele Sihtasutus.
- Langemets, Margit; Tiits, Mai; Valdre, Tiia; Voll, Piret 2010. *In spe*: ühekõiteline eesti keele sõnaraamat. – Keel ja Kirjandus, 11, 793–810.

- Lenci, Alessandro; Bel, Nuria; Busa, Federica; Calzolari, Nicoletta; Gola, Elisabetta; Monachini, Monica; Ogonowski, Antoine; Peters, Yvonne; Peters, Wim; Ruimy, Nilda; Villegas, Marta; Zampolli, Antonio 2000. SIMPLE: A General Framework for the Development of Multilingual Lexicons. – *International Journal of Lexicography*, 13 (4), 249–263. <http://dx.doi.org/10.1093/ijl/13.4.249>
- Orav, Heili 2000. Adjektiivid kui semantiline probleem: wordnet-tüüpi tesaauruste koostamise kogemused. – Tiit Hennoste (Toim.). *Arvutilingvistikalt inimesele*. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 1. Tartu: Tartu Ülikooli Kirjastus, 153–166.
- Peters, Ivonne; Peters, Wim 2000. The Treatment of Adjectives in Simple: Theoretical Observations. – *Proceedings of LREC 2000: 2nd International Conference on Language Resources and Evaluation*.
- PSV 2014 = Eesti keele põhisojavara sõnastik. [The basic dictionary of Estonian.] Toimetanud Jelena Kallas, Mai Tiits, Maria Tuulik. Tallinn: Eesti Keele Sihtasutus, 2014.
- Zipf, George Kingsley 1945. The meaning-frequency relationship of words. – *Journal of General Psychology*, 3, 251–256. <http://dx.doi.org/10.1080/00221309.1945.10544509>
- Zipf, George Kingsley 1949. *Human behavior and the principle of least effort*. Addison-Wesley Press.
- The Oxford English Corpus. <http://oxforddictionaries.com/words/the-oxford-english-corpus> (15.10.2013).
- Vainik, Ene 2002. Millest on tehtud eestlaste emotsioonisõjavara? – *Keel ja Kirjandus*, 8, 537–553.
- Vare, Silvi 2012. *Eesti keele sõnavered*. Tänapäeva eesti keele sõjavara struktuurianalüüs I–II. Tallinn: Eesti Keele Sihtasutus.

Maria Tuulik (Eesti Keele Instituut, Tartu Ülikool) on eesti keele põhisojavara sõnastiku töörühma liige ja kirjutab doktoritööd eesti adjektiivide polüseemiast. Põhilised uurimisvaldkonnad: leksikaalne semantika, leksikograafia, tekstilingvistika.
 Roosikrantsi 6, 10119 Tallinn, Estonia
maria.tuulik@eki.ee

ADJECTIVAL POLYSEMY IN CORPORA AND DICTIONARIES

Maria Tuulik

Institute of the Estonian Language

The article describes and analyzes the 100 most frequent Estonian adjectives with regard to both polysemy and morphological structure. Also examined is the polysemy index, i.e. the average number of meanings per word, for root adjectives, derivatives, compound words and participles. The most polysemous of these are root adjectives, with a polysemy index of 5. The polysemy index of derivatives is roughly half that, 2.6. Among the 100 most frequent adjectives, the least polysemous are compound words (polysemy index 1.3).

The second half of the article discusses the application of semantic types, both as a theoretical aid in the analysis of polysemous words and as a practical tool in the coding of semantic information in dictionaries. Based on the primary meanings of the 100 most frequent adjectives as given in the single-volume explanatory dictionary of Estonian, 15 semantic types are proposed in order to describe Estonian (polysemous) adjectives.

The next planned step, made possible by the existence of the semantic types, is to introduce the patterns of systematic adjectival polysemy in Estonian. The aim of this would be to study Estonian adjectival polysemy and establish a theoretical framework on the basis of which the presentation of adjectival polysemy in dictionaries could be harmonized.

Keywords: semantics, lexicology, lexicography, word formation, semantic types, Estonian