

Mis on leksikograafia

Piret Voll

eesti keele instituudi leksikograaf

Sõna *leksikograafia* tuleneb kreeka keelest: *lexikon* sõnaraamat + *graphō* kirjutun. Leksikograafia tähendab nii teooriat — sõnaraamatute koostamise õpetust kui ka praktilist poolt — sõnastike tegemist ennast. Leksikograaf on inimene, kes on kursis leksikograafia teoreetiliste suundadega ja kasutab neid teadmisi sõnastike loomisel. Leksikograafia oli tähtis teadusharu juba Vana-Kreekas, kus koostati vananenud või haruldastest sõnadest käsikirjalisi sõnastikke.

Sõnastik

on kindla eesmärgiga korrastatud sõnavarakogum, mis võib olla vormistatud raamatuna, raamatu osana, kaartidel, magnetlindil või elektroonilisel kujul veebis, kettal vm. Sõnastik koosneb sõnaartiklitest. **Sõnaartikkel** jaguneb mitmeks osaks, millest tavalisemad on märksõna ja kirjeldus. Mingis mõttes on ka märksõna ise sõna kirjeldus, andes teada, milline on sõna õigekiri. Olenevalt sõnastiku otstarbest võib sõnaartikkel sisaldada infot grammatika, sõna päritolu, sõnavaraliste seoste, esinemissageduse vms kohta. **Märksõna** on sõna, mis on mingil kindlal eesmärgil sõnastikku valitud. Näiteks võõrsõnade leksikoni märksõnavaliku kriteerium välistab omasõna sattumise sõnastikku ja slängisõnaraamatusse pääsemiseks peab sõna esindama mingit slängi. Tänapäeva sõnaraamatutes esitatakse märksõnad enamasti tähestikulises järjekorras, harvemini sisu (kirjeldust) arvestavalt. Tähestikulises sõnastikus järjestatakse sõnaartiklid tavaliselt algustähtede järgi, ainult pöördõnaraamat annab sõnad nende lõpptähtede järjekorras. Kirjeldust arvestav sõnastik on näiteks sagedussõnaraamat, mis näitab sõnade sagedust ühe keele tekstides ning sõnad on sellises sõnastikus järjestatud harilikult sagedasemast alustades. Mõisteline sõnaraamat koondab sisult kokkukuuluvad, aga vormilt erinevad sõnad ühte artiklisse (näiteks *taime* juures *leiam* *puu*, *aedvilja*, *aianduse*, *kapsa*,

istiku, lehe, kasvama jne), mis tähendab, et sellise sõnastiku juurde peab kindlasti kuuluma põhjalik tähestikuline register. Sõnaartikli kirjeldus sõltub sellest, kelle jaoks ja mis eesmärgil on sõnastik tehtud.

Järgnevad näited sõna *hiir* esitusest erinevates sõnastikes osutavad võimaluste paljususele, olenevalt sellest, keda sõnastikku luues on silmas peetud. Et kõik näited pärinevad eesti viimase aja sõnastikest, siis annavad nad päris hea ülevaate sellest, kui erinevalt võib üht ja sedasama sõna vastavalt vajadusele kirjeldada. Samuti näeme, kuidas leksikograafid juba ilmunud sõnastikke oma töös kasutavad (võrdle näiteks ÕS ja „Arvutikasutaja sõnastik”).

- „Eesti kirjakeele seletussõnaraamat” (parandatud ja täiendatud versioon, ilmumas):

hiir, -e, -t 35 s. 1. väike, saleda keha ja pika sabaga näriline. *Väike, suur hiir. Hall, valge hiir. Hiir kröbistab, piiksub. Hiired rikuvad ja hävitavad toiduaineid. Nagu hiir lõksus. Laps oli nurgas vait kui hiir. Hiilib vaikselt, tasa kui hiir. Kui kassi kodus pole, on hiirtel pidu. Vs.*

Ls. jutttselg-, kaelus-, kodu-, maja-, mets-, pisihiir.

2. esineb liitsõna järellosana väikeste hiirt (1. täh.) meenutavate loomade nimetustes.

Ls. nahkhiir.

3. *info* arvutihhiir. *Optiline, juhtmeta, traadita hiir. Vajuta hiire parempoolset klahvi.*

Ls. arvutihhiir.

[kellegagi] kassi ja hiirt ~ kassi-hiirt ~ nagu kass hiirega mängima vt. kass.

- **ÕS 2006:**

hiir <30: hiire, hiirt> (*näriline*); INFO *laual libistatav nuppudega osutusseadis*. ZOOLOG: jutttselg-, kaelus-, leet-, mets-, kodu-, pisi-, uru-, villakhiir *tšintsilja*; nahkhiir. Hiireauk, -kuningas NALJ *kass*, -lõks, -poeg, -tipsi <-tipsi> LASTEK *väike hiir*, -karva, -viu ZOOLOG. Hiireklahv, -padi (*arvutihhiirel*). Hiirvaikne

- „Arvutikasutaja sõnastik” (Keelevara):

hiir *laual libistatav nuppudega osutusseadis*.

- „Eesti etümoloogiline teatmik”, 1982:

hiir *smugri: ung. Egér*

- „Väike murdesõnastik”, 1982:

hiir *Jõe Kuu Hlj VNg Lüg Jõh IisR Vai Jäm Khk Mus Krj Pha Vll Muh Emm Käi Rei Phl LNg Rid Mar Kul Mär Vig Kir Lih Kse Han Var Tõs Khn Aud Pjg Tor Hää Saa Ris Kei Hag Rap Juu Kos Jür Amb JMd Jjn Ann Tür Koe VMr Kad*

VJg Iis Trm Kod MMg Äks Lai Plt Pil KJn Vil Trv Hls Krk Hel Ran Puh TMr
Ote San Urv Krl Har Rõu Plv Vas Ráp Se Lei Lut ± loom

- „Eesti esimene slängisõnaraamat”, 1991:
hiir, ka hiireke noor ja rumal, vist päris nägus. *Teema*: tüdruk
- „Poetilised sünonüümid” (Keelevara):
hiir
Teema: tütar
Siit on nirki nikistänüd, / siit on kassi kallal käinüd, / siit on iiri ammustanud./
Sa olid tarka, mõissid kosta:/ Sie sinu oma iirekene, / sie sinu oma nirgikene,
/ kahejalgnõ kassikene./ RKM 10, 46 (4) Khn. Vt. ka *eide hiirekene, memme hiir*.
memme hiir
Teema: tütar
Mina ügzi memme hiire, / Kahejalgne taadi kassi:/ H II 56, 352 (17) MMg;
Vt. ka *hiir*.
- „Eesti-inglise sõnaraamat”, J. Silvet 1980:
hiir [hiire] mouse (*pl.* mice); vaikne kui ~ (as) quiet as a mouse



Esimesed sõnaraamatud

Maailma esimese sõnaraamatu ilmumisaeg ja -koht ei ole täpselt teada, aga juba muistsetest aegadest – III ja II aastatuhandest eKr – pärinevad hiina, india, sumeri ja akadi sõnastikud, samuti egiptuse, araabia, kreeka ja ladina sõnaraamatud. Õieti oli tegu sõnaloenditega, milles puudusid tähenduse seletused ja näitelauseid. Üles tähendati vanad või haruldased sõnad. Need sõnaloendid olid koostatud sisust, mitte tähestikulisest järjekorrast lähtudes. Seega esitasid maailma vanimad ükskeelsed sõnaraamatud eelkõige kirjakuju, mille abil ühele või teisele asjale osutada. Suurte seletussõnaraamatuteni jõuti alles 17. sajandil, kui Prantsusmaal hakati looma akadeemilist, seletustega varustatud sõnaraamatut, mis pidi kirjeldama n-ö õiget keelt. Kohe pörkuti tüüpiliste hädadega, mis saadavad iga suursõnaraamatu koostajaid ajast aega: tegijate erinevast keeletajust johtuv ebaühtlus, koostamiseks kuluv pikk aeg, pidevalt muutuv keel jne. Ühe keele esimese seletussõnaraamatu valmimine on nii saksa, inglise kui ka prantsuse keele puhul aega võtnud 50–70 aastat, leedu keele sõnaraamatu viimane köide ilmus siis, kui koostamise algusest oli möödunud koguni 100 aastat. „Eesti kirjakeele seletussõnaraamatu” tegemine võttis aega umbes 50 aastat.

Suure seletussõnaraamatu alustamiseks vajatakse korralikku tekstikogu. Sajandeid koguti keeleainest n-ö käsitsi, sõnu ja väljendeid sedelitele kirjutades. Näiteks eesti kirjakeele seletussõnaraamatu eeltööna valmis viieaastase töö tulemusel umbes miljoni sedeliga kartoteek. Miljon sedelit aga on sõnaraamatu tegemiseks ilmselgelt liiga vähe. Läks aastaid, enne kui alles 1980. aastatel küündis sedelite arv üle 4 miljoni. Sedelkartoteegi puuduseks on aga see, et ehkki kartoteek sisaldab suure hulga sõnu, on need sõnad kättesaadavad üksnes ühe sõna kaupa. Seega pole imestada, et suurte seletussõnaraamatute esimeste versioonide loomisaeg kestab aastakümneid.

Kuidas tehakse sõnastikke tänapäeval

Tänapäeva leksikograafiaraamatutes jagatakse sõnaraamatu tegemise õpetust hoopis teisel moel. Alustatakse äriplaanist. Loodav sõnaraamatutüüp otsustatakse pärast turu-uuringuid: millist sõnaraamatut kasutaja vajab ja/või milline sõnaraamat puudub raamatukaupluse riiulilt. Kui kirjastaja on kokku pannud eelarve, algab sisuline töö. Koostajal-toimetajal võib olla sadu häid ideid, aga lõpliku sõna ütleb kirjastaja, kelle käes on rahakott. Seesama rahakott dikteerib ka

meeskonna suuruse ja sõnaraamatu valmimise aja, seetõttu võivad nn kommertssõnastikud maksta lõivu põhjalikkusele ja kvaliteedile. Õnneks ei mõjuta turusituatsioon (vähemalt mitte otseselt) teaduslike sõnaraamatute väljaandmist, ehkki oma probleemid on siingi.

Sõnastiku loomise sisuline töö koosneb tänapäeval õige paljudest etappidest. Alustuseks töötatakse välja kasutaja profiil. See tähendab, et üritatakse võimalikult täpselt selgeks teha, kellele ja mis otstarbel sõnastikku koostama hakatakse. Täiskasvanu või laps, keeleõppija või keelt emakeelena kõnelev, tavakasutaja või kitsama ala asjatundja – kõigil neil on sõnaraamatule erinevad nõudmised. Põhjalikult analüüsitakse, mida tulevane kasutaja sõnastikust otsida tahab ja milline oleks parim moodus selle esitamiseks. Seejärel otsustatakse ligikaudne maht – väga oluline on algusest peale teada, kas koostatakse väikest, 5000 märksõnaga piirduvat, või hoopis suurt, 100 000 märksõnani küündivat sõnastikku. Selles etapis lülitub töösse IT-meeskond, kelle ülesandeks on luua (või osta, otsida jne) sobilik sõnastikupõhi ning seda vastavalt leksikograafide vajadustele pidevalt täiendada ja muuta. IT-inimeste abi on vaja ka korpuste koostamisel-leidmisel. Tänapäeval saab keeletehnoloogia vahendeid kasutades lühikese ajaga tekitada mitmekümne miljoni sõnaga tekstikorpuse. Arvataksegi, et umbes 20 miljoni sõnaga korpus võiks olla keeleuurimiseks piisav. Sellisest korpusest saab otsida mis tahes sõnu ja sõnavorme, samuti jälgida sõnade esinemisümbrust, -sagedust ja konteksti. Suur materjalihulk annab muidugi rohkem võimalusi, samas on raskem teha valikuid, mida võtta, mida jätta. Üks põhimõte on toetuda esinemissagedusele, aga sealgi võib korpus vales suunas mõjutada. Üha enam räägitakse tasakaalustatud korpusest, kus erinevad tekstiliigid – ajalehe-, ilukirjandus- ja teadustekstid – oleksid võrdselt esindatud. Seega on infotehnoloogial tänapäeva leksikograafiamaailmas aina suurem roll. Tavaliselt saab iga leksikograaf kindla tööloigu, kas tähestikupõhiselt, mõisterühma järgi või hoopis mõne läbiva ülesande (näiteks grammatika või fraseoloogia esitus vms). Abiks on prooviartiklite koostamine, kus võivad esile tulla kitsaskohad. Et jõuda parima lahenduseni, arutatakse kõik suuremad probleemid ühiselt läbi. Oluline on ka see, et meeskonnatööna tehtav sõnastik ei oleks ülemäära iga üksiku tegija nägu. See võib viia ebaühtluseni; igal inimesel on ju teistest vähem või rohkem erinev keeletaju ja -kogemus. Suurte sõnastike puhul kasutatakse mitmeastmelist toimetamist, kus esimene lüli koostab, teine toimetab ning viimase sõna ütleb peatoimetaja.

Seega peaks leksikograaf teadma keelest justkui kõike: ühes suures ükskeelses sõnastikus võime näiteks leida nii grammatika, häälduse, etümoloogia, seletuse kui ka leksikaalseid suhteid kajastava info. Kaks- või mitmekeelsete sõnastike puhul peab valdama lisaks emakeelele väga heal tasemel teisi keeli. Tegelikult muidugi üliinimesi pole ja alati saab konsulteerida asjatundjatega. Seletussõnaraamatu tegijatel on tihtipeale vaja täpsustada just kitsama eriala termineid, puudutagu need geoloogiat, füüsikat, sõjandust, majandust vms. Abi leiab ka erialateatmikest, käsiraamatutest, samuti võrguväljaannetest.

Kui leksikograafid on oma töö teinud, jääb üle sõnastik välja anda – kas raamatuna, laserplaadil, internetikeskkonnas või mitmel kujul korraga.

Eesti leksikograafia tähtteoseid

- 1637 Heinrich Stahl „Anführung zu der Ehstnischen Sprach”
- 1732 Anton Thor Helle „Kurtzgefaszte Anweisung zur Ehstnischen Sprache”
- 1780 August Wilhelm Hupel „Ehstnische Sprachlehre für beide Hauptdialekte”
- 1869 Ferdinand Johann Wiedemann „Eesti-saksa sõnaraamat”
- 1925–1937 Johannes Voldemar Veski, Elmar Muuk „Eesti õigekeelsuse-sõnaraamat”
- 1933 Elmar Muuk „Väike õigekeelsus-sõnaraamat”
- 1939 Johannes Silvet „Inglise-eesti sõnaraamat”
- 1958–1968 Andrus Saareste „Eesti keele mõisteline sõnaraamat”
- 1960 „Õigekeelsuse sõnaraamat”(KKI)
- 1961 Richard Kleis, Johannes Silvet, Eduard Väari „Võõrsõnade leksikon”
- 1965 Johannes Silvet „Eesti-inglise sõnaraamat”
- 1976 „Õigekeelsussõnaraamat”(KKI)
- 1982 Paul Saagpakk „Eesti-inglise sõnaraamat”
- 1991 Asta Õim „Sünonüümisõnastik”
- 1992 Paul Saagpakk „Sünonüümisõnastik”
- 1993 Asta Õim „Fraseoloogiasõnaraamat”
- 1995 Asta Õim „Antonüümisõnastik”
- 1999 „Eesti keele sõnaraamat ÕS 1999” (EKI)
- 2006 „Eesti õigekeelsussõnaraamat ÕS 2006” (EKI)
- 1988–2007 „Eesti kirjakeele seletussõnaraamat” (EKI)
- 1994– „Eesti murrete sõnaraamat” (EKI)