

TARTU ÜLIKOOL
MATEMAATIKA-INFORMAATIKATEADUSKOND
Arvutiteaduste instituut
Informaatika eriala

Maarja-Liisa Natus

**LIITSÕNADE TÄISKUJU LEIDMINE
SÕNAARTIKLITEST**

Bakalaureusetöö (4AP)

Juhendaja: Jaak Vilo, PhD

Autor: “.....“ mai 2006
Juhendaja: “.....“ mai 2006
Õppetooli juhataja: “.....“ 2006

TARTU
2006

Sisukord

SISUKORD	2
SISSEJUHATUS	3
1.SÕNARAAMATUTE ÜLESEHITUS	5
1.1 SÕNAARTIKLI SISU	5
1.1.1 Märksõna	5
1.1.2 Grammatilised andmed.....	6
1.1.3 Seletused.....	7
1.1.4 Pragmaatilise info märgendamine.....	7
1.1.5 Viited.....	8
1.1.6 Näited	9
1.1.7 Sõnaartiklite erinevus sõnaraamatuses	10
2. LIITSÕNADE TÄISKUJU LEIDMINE SÕNAARTIKLITEST	13
2.1 LIITSÕNAD ÕS-IS JA SELETUSSÕNARAAMATUS	14
2.1.1 <i>Liitsõnad algfailis ja HTML kujul</i>	14
3. PROGRAMMI LIITSÕNADE TÄISKUJU GENEREERIMISEKS	17
4. PROGRAMMI TULEMUSE ÕIGSUS	18
4.1 PROGRAMMI KORREKTSED TULEMUSED	18
4.2 PROGRAMMI EBAKORREKTSED TULEMUSED	21
KOKKUVÕTE	25
GENERATING THE FULL FORM OF THE COMPOUND WORDS	27
VIITED	28
URL-ID	29
LISA 1	30
LISA 2	42

Sissejuhatus

Paberkujul olevad sõnastikud peavad olema kompaktsed, sisaldamaks kogu vajalikku informatsiooni. Sageli peab kasulik sõnaraamat ka taskusse mahtuma. Arvutisõnastike eelis trükitud sõnastike ees on see, et nad võivad olla mahukad ning sisaldada sageli ka vähem tähtsat informatsiooni. Tavakasutaja eest peidetakse suurt hulka sõnastikes sisalduvat infot, mis sisaldub sõnastiku algfailis.

Käesolevas bakalaureusetöös käsitletakse sõnastike struktuuri, sõnaartiklites kajastuvat informatsiooni ja liitsõnade lühendamist sõnaartiklites. Töö eesmärgiks on lühendatud liitsõnade täiskuju genereerimine.

Töö teoreetilises pooles tuginetakse Margit Langemets'i väitekirjale „Sõnaraamatu arvutilingvistiline analüüs” (vt. [Lan00]), Kristel Uihoaed'i bakalaureusetööle „Sõnaseletuste genereerimine tesauruse info põhjal”(vt. [Uib05]), Külli Kivi diplomitööle “Sõnastikud struktureeritud tekstiandmebaasina”(vt. [Kiv93]) ning internetis antud teemal ilmunud artiklitele. Selle töö osas antakse ülevaade sõnaartikli ehitusest ning näidete abil seletatakse lahti selle tähtsamad osad.

Praktilises pooles kasutatakse järgmisi Keelevaras asuvaid Eesti keele sõnaraamatuid ÕS 1999 (edaspidi ÕS) (vt. [url:QS1999]), Eesti Keele Seletussõnaraamat (vt. [url:SELETAV]), Õpilase Õigekeelsussõnaraamat (edaspidi Õpilase ÕS) (vt. [url:QPILASE_QS]). Liitsõnade genereerimiseks kasutakse programmeerimiskeelt Perl. Töökeskkonnaks on Windows XP ning programmi kirjutamiseks kasutatakse vabavarana veebis olevaid Perli redaktoreid OptiPerl ja PerlBuilder. Praktilise poole esimeses osas koostatakse programm, millega püütakse lahendada lühendatud liitsõnadest pikkade liitsõnade genereerimise probleem. Bakalaureuse töö praktilise osa eesmärk on kõikidest lühendatud liitsõnadest täispikkade liitsõnade genereerimine. Tavapärase lühenduse viga on selles, et paljusid sõnu, mida on sõnastikus kasutatud lühendatult, on võimatu üles leida. Põhjuseks on see, et lühendamine on mõeldud rohkem paberkujul oleva raamatu jaoks. Kuna aga paberkujust on genereeritud elektrooniline väljaanne, siis on seal säilinud kõik algsed reeglid. Erinevates sõnaraamatutes on genereerimine erinev, kuna sõnaraamatud on erineva raskuse ja struktuuriga. Reegleid on

palju ning iga sõnastiku puhul peab neid kindlalt järgima. Käesoleva töö autori arvates võiks koostatud tööst abi olla ka edaspidi arvutisõnastike täiuslikumaks muutmisel.

1. Sõnaraamatute ülesehitus

Sõnastikud on struktureeritud tekstiandmebaasid, millel on enamasti hierarhiline üldstruktuur ning mis ühendavad endas nii andmebaasisüsteeme kui ka vabas vormis teksti tunnusjooni [Lan00]. Sõnaraamat koosneb erineva pikkusega tähestikulises järjekorras olevate märksõnadega algavatest sõnaartiklitest. Need on omakorda jagatud funktsionaalseteks alamosadeks, mis on varustatud spetsiifiliste märgenditega, moodustades kirjetaolised struktuurid [Kiv93].

1.1 Sõnaartikli sisu

Sõnaartikli esitamise viis sõltub sellest, millise sõnastikuga on tegemist. Informatsiooni saab anda märksõnadega kas tähestikulises järjekorras või mõisteliselt (väljendatud seda moodustavate sünonüümsete sõnade kaudu ning lisatud seletused, kommentaarid vms) [Uib05]. Sõnaartikkel on sõnaraamatu organiseerimise printsiip, mille mõtte on näidata, missugune info (mingitelt alustelt lähtudes) kuulub kokku. Sõnaartikkel võib sisaldada järgmist informatsiooni: märksõna, leksikaalne info (grammatika, süntaks), etümoloogia, näited. Järgnevates alapunktides antakse lühike ülevaade sõnaartikli sisust.

1.1.1 Märksõna

Sõnaartikkel algab märksõnaga, millele järgneb terve hulk erineva märgendiga osi, mis annavad informatsiooni sõna õigekirja ja struktuuri kohta. Leksikaalse info hulk aga sõltub märksõnast endast. Märksõnade iseloomustamisel on püütud esitada seda, mida keeletarvitaja võiks sõnaraamatust otsida [ÕS99]. Märksõnaks võib olla liht- või liitsõna või sõnaühend.

jooksutama < 52 > *joosta laskma, jooksmata panema*

(ÕS)

jooksu/poiss. ◇ Hotelli jooksupoiss

(ÕS)

maha .jooksmata. ◇ Jooksis teise jalust maha. Kolm vagonetti jooksis rööpmelt maha. Kas oled oma sarved lõpuks maha jooksnud *ülele* (noorustulisusest üle saanud)? Maha jooksnud = maha/jooksnud aeg *ülele* tühjal kulunud aeg. Maha/jooks ka tehn aeglustuv liikumine pärast mootori väljalülitamist

(ÕS)

Vaatluse all olevates sõnaraamatutes, kui märksõnaks on liitsõna või liitsõna, on märksõna nimetavas käändes. Verbina esitatud liitsõnad on *ma*-infinitivis. Märksõnas võib olla ära märgitud morfeemipiir (formatiivipiir, liitsõnapiir), ka rõhk, välde jms. Liitsõnapiiri on tähistatud sõnastikes erinevalt.

udu/.sul'g, -sule, -.sulge

(Õpilase ÕS)

termo|püksid <pl. > voodri ja vatiiniga pikad soojad spordipüksid.

(Seletussõnaraamat)

Teistes keeltes olevate sõnaraamatute ülesehitus ja reeglid erinevad autori vaatluse all olevate sõnaraamatute omadest. Näiteks soome keeles on verbid reeglina *da*-infinitivis. Mõne märksõna puhul on ära märgistatud mingi märksõna osa, mis tähendab, et antud sõna saab kasutada nii koos märgistatud osaga või ka ilma. Käesolevas töös uuritud sõnaraamatutes on sõna osa märgitud kõigis erinevalt: ÕS-is on kasutatud nurksulge (nt. märksõna **hellake[ne]**), Seletussõnaraamatus ümarsulge (nt. märksõna **kullake(ne)**) ning Õpilase ÕS-is ainult püstkriipsu (nt. märksõna **kal'like**, -se, -st = **kal'like|ne**, -se, -st).

1.1.2 Grammatilised andmed

Sõnastikus pole grammatiline info esmatähtis, kuid selle esitamine on enamasti siiski vajalik. Grammatiliste andmete maht on sõnastikes erinev. Sõnade grammatiline iseloomustus sõnastikes hõlmab:

- 1) morfoloogilist infot, s.t. sõnamuutumist ehk vormimoodustuslikku külge, lisaks sõnaliike ja muid morfoloogilisi kategooriaid
- 2) süntaktilist infot, s.t keeleelementide kombinatsioone ehk funktsionaalseid erinevusi. [Lan00].

Grammatiline info on enamasti esitatud noolsulgudes. Seal võivad olla andmed häälduse, rektsiooni, muuttüübi numbri ja käänamise või pööramise vajalike vormide kohta. Tüübitabel ilmub nii ÕS-il, Seletussõnaraamatul kui ka Õpilase ÕS-il, kui klikkida hiirega märksõna järel oleva tüübinumbri peale.

kaalium <-i, -i 10> s. keem. element, hõbevalge pehme läikiv leelismetall (K). Looduses leidub kaaliumi üksnes ühenditena.

Liitsõnad: sööbekaalium.

(Seletussõnaraamat)

Noolsulgudes on esitatud omastava ja osastava käände lõpud, muuttüübi number 10. Järgneval real on tüübitabeli sõna nr. 10, mille järgi käändub märksõna „kaalium”.

10 **seminar**, seminari, seminari, seminarisse e. seminari, seminaride, seminare e. seminarisid, seminaridesse e. seminaresse (Seletussõnaraamatu tüübitabel)

1.1.3 Seletused

Seletused on märksõna semantiline info, mis peaks avama märksõna sisu. Seletus peaks olema keerulise sõna lahtiseletamine vähemkeeruliste sõnade abil [Uib05]. Märksõna tähenduse seletus on enamasti kirjeldav, informaalne ning lihtsasti arusaadav tekst. Vaatluse all olevates sõnaraamatutes on märksõnade seletused, sünonüümivasted ja tähendusvihjed antud ainult vajaduse korral. Need on esitatud märksõna järel kaldkirjas. Kui sõnal on mitu tähendust, siis märksõna järel (mõnedel juhtudel ka ees) on kirjutatud homonüümi järjenumbr vm indeks, mille järel on sõna erinevad tähendused.

kulul *postp. [gen.] arvel.*

1. *kellegi rahadega, kellegi ülalpidamisel.* Elas teiste kulul. Poeg juba täismees, aga elab veel vanemate, isa-ema kulul. Vallavaene maeti riigi kulul.

2. *kellegi huve kahjustades, kedagi v. midagi mingis suhtes piirates.* Ajas rahva kulul suure varanduse kokku. Tahab teiste kulul kannuseid teenida, oma nahka päästa.

3. *kedagi v. midagi pilgates, kellegi v. millegi üle v. kohta.* Tema kulul tehti halba nalja. Teravmeelitses oma taipamatu kaaslaste kulul. Nüüd saavad jutumoorid nende kulul lõksutada, hambaid teritada. Suudab oma äparduste kulul naljagi heita. Nad löbustasid end minu hirmu kulul, minu kõrgete kontsadega kingade ja kübara kulul. (*L. Promet*)

(Seletussõnaraamat)

Kui võrrelda ÕS-i, Seletussõnaraamatut ning Õpilase ÕS-i, siis võib täheldada erinevusi märksõnade seletustes. Kui ÕS-i ja Õpilase ÕS-i puhul ei ole taime- ja loomanimedel juures ladinakeelseid nimetusi, siis Seletussõnaraamatu märksõnade seletamisel on need olemas. Lisaks võib leida ka keemiliste elementide sümboleid, mõnede sõnade rahvusvahelisi tähiseid ja lühendeid (nt. mõõtühikud).

1.1.4 Pragmatilise info märgendamine

Sõnastikes kasutatakse teksti märgendamist tähistamiseks pragmaatilist informatsiooni:

- eriala (nt *kunst*), kasutusvaldkonda (nt *bot*).

trell <-i 22> **II s.** *tehn. trellpuur.* Ta tõi töö juurest trelli ning võristas õhtu otsa puurida. (*A. Valton*)

Liitsõnad: aku-, elektri-, käsi-, lööktrell.

(Seletussõnaraamat)

Kui aga märksõna või seletus kirjeldavad juba eriala, siis pole märgendit vaja.

tuuma/füüsika. *aatomituuma käsitlev füüsika osa.*

Eelpool toodud näide ei vaja märgendit *füüs*, kuna see on juba nii märksõna kui seletusega öeldud.

- stiili

nael/kumm. *argi naastrehev auto* (ÕS)

- märksõnade murdelisust

ubin <-a 2> *s. murd. õun. .. kümnekond uibukest ... jänestest kooritud, / nendeski ubinad kõvad ja hapud. (P. Haavaoks)* (Seletussõnaraamat)

Veel saab märgendite kaudu näidata tähenduse ülekandmist, hinnangulist värvingut, sagedust jms. Õpilase ÕS-is puuduvad aga üldse kõik märgendid, kuna selle sõnastiku struktuur on ÕS-i ja Seletussõnaraamatu omast tunduvalt primitiivsem. Selles sõnastikus on enamasti märksõna järgi ainult mõnesõnalised seletused.

1.1.5 Viited

Võttes kätte paberkujul oleva sõnaraamatu, peab kasutaja väga täpselt teadma, kuidas seda kasutada. Siin kohal tuleb välja elektroonilise sõnastiku eelis, mis oleks dünaamilisus ja liigendatus. Viidete abil on hõlbus liigelda omavahel seotud märksõnade vahel. Paljude sõnaartiklite juures juhatatakse kasutaja antud märksõna kohta rohkemat infot otsima ühe või mitme märksõna kirjest. Viidatud märksõnaga koos on tihti täpsemalt kirjas tähenduse number, sõnaliik või on osutatud sünonüümile, antonüümile, joonisele vms [Lan 02].

.väike *vt väikene* (Õpilase ÕS)

päi *vt ka pea* (Seletussõnaraamat)

kumm 1. < 20 kummi, .kummi> *kummiese; kummiaine. ◇ Auto/ argi autorehv, püksi/kumm; närimis/, vaht/kumm. Kummi/kanal (rõival kummipaela jaoks). vt ka kummi* (ÕS)

Tihti peale on viitamist kasutatud ainsuslike või mitmuslike sõnade puhul, kus on viidatud siis vastava sõna mitmusele või ainsusele.

omaksed *vt ka omas* (Seletussõnaraamat)

Viiteid võib leida ka Seletussõnaraamatu sõnaartiklite lõpus oleva fraseoloogia osa alt, kus on viidatud mõnele fraseologismis olevale sõnale.

– **välja pääsema** *mingist paigast, hrl. ruumist välja saama; võimalik olema sealt pääseda.*

Pääses põlevast majast akna kaudu välja. Väravad on suletud, kuidas välja pääseda? Papagoi on puurist välja pääsenud. Vaenlane püüdis piiramisrõngast välja pääseda. Ei siit pääse hobusega välja kaugemale, sest ei ole teed, mis kannaks hobust peal. (*G. Vilbaste*)

• (*mingist olukorrast pääsemise kohta*). Arutati, kuidas halvast olukorrast välja pääseda. .. ta töötab seni, kuni ta nende killast välja pääseb. (*V. Õun*)

Fraseoloogia:

haljale oksale pääsema *vt ka oks*

[kelle(l)gi] keel pääseb lahti ~ **valla** *vt ka keel*

kerge nahaga pääsema *vt ka nahk*

(Seletussõnaraamat)

Kaugemate seoste esitamiseks on käsitluse all olevates sõnaraamatus kasutatud lühendit *vrđ*. Nii võrreldakse õigekirja või sõnade sarnasusi, mis võivad sageli kasutajaid eksitada. Vaadates järgmisi sõnaartikleid Seletussõnaraamatus on näha, et kirjapilti võiks kahe märksõna puhul segamini ajada, kuid nende sisu on teineteisest täiesti erinev, mistõttu antud võrdlus on otstarbekohane.

.alla/tuule, .alla/tuult. pärituult. vastand: vastutuult. ◊ Pööras näo allatuule v allatuult. *vrđ*

all/tuule

(Seletussõnaraamat)

all/tuule. tuulealuses küljes. vastand: pealtuule. ◊ Istub alltuule. Alltuule/külge, /parras *mer.*

vrđ .alla/tuule

(Seletussõnaraamat)

1.1.6 Näited

Näideteks on sõnaühendid või lühilauseid, mis täpsustavad sõna sisu ja kasutuspiirkonda, näitavad sõltumust. Paljud sõnaraamatud ei sisalda üldse näiteid või kasutavad neid vastavalt vajadusele.

Üheks näidete põhiosaks on püsiühendid: vanasõnad, fraseologismid, metafoorsed ühendid, tsitaadid jm. Tsitaadi autor on märgitud sulgudesse.

susi <11> *s. kõnek. pahatahtlik urgitsev, suskiv inimene.* See mutt on igavene susi. .. ega oleks arvatagi võinud, et ta on niisugune, kes jookseb iga sõna ette kandma, susi on, muud midagi. (*E. Krusten*)

(Seletussõnaraamat)

Mõnede sõnaartiklite puhul koosneb näidete osa liitsõnade loetelust, mida märksõnaga moodustada saab. Sageli on mõne artikli sees toimunud liitsõnade loetelus üleminek mõnele teisele märksõnale, mis on eelistatum või kaasaegsem kasutada sõnaühendites.

.aatom < 9 -i> *keemilise elemendi väikseim osake*. ◇ Hapniku aatom = hapniku/aatom. Aatomi/parem: tuuma/: /patarei, /plahvatus, /relv, /relvastus, /rännak, /sõda. Aatomi/katel *vmo* = aatomi/reaktor, parem: tuumareaktor. Aatomi/sisene. Aatomite/vaheline. Aatom(i)/number, parem: laenguarv, järjenumbr. Aatom/kell, /kütus *tuumkütus*. *vt ka* tuum (ÕS)

Liitsõnadest ning nende moodustamisest on pikemalt juttu punktis 2.

1.1.7 Sõnaartiklite erinevus sõnaraamatutes

Sõnaartiklite ehitus võib olla mõne sõnastiku puhul keerulisem kui teise oma. Enamasti koosneb sõnaartikkel väikesest osast võimalikust infost (nt märksõna ja viide), kuid esineb ka mitmeleheküljelisi kirjeid [Kiv93]. Eelnevat selgitab järgnev võrdlus ÕS-i, Õpilase ÕS-i ja Seletussõnaraamatu näitel. Sõnaartikli ehituse näiteid märksõnaga *armas* :

1. ÕS:

armas < 35 .armsa, .armsat ja **armast**>. ◇ Mu **armas**! **Armas** sõber! Ole nii **armas** ja käi poes ära. Väga **armas** sinust, et. **Armas** kuulda, et. Oh sa **armas** aeg! Ime/, üli/, hale/**armas** >>

```
<art nr="2056"><hdr><mrks>armas</mrks> <grm><tyyp_grp><tyyp>35</tyyp> <vorm>.armsa, .armsat
<lyh>ja</lyh> armast</vorm></tyyp_grp></grm>.</hdr> <naited><naide>Mu armas! Armas sõber! Ole nii
armas ja käi poes ära. Väga armas sinust, et .. Armas kuulda, et .. Oh sa armas aeg! Ime/, üli/,
hale/armas</naide></naited> </art>
```

Joonis 1.1 Märksõnaotsing veebis ÕS-ist ja sama kirje väljavõte sõnastiku originaalfailist.

2) Õpilase ÕS:

armas, .armsa, .armsat = **armast**; .armsam, kõige .armsam = .armsaim. Armsate inimeste keskel. >>

```
<art><mrks>armas</mrks>, .armsa, .armsat = armast; .armsam, kõige .armsam = .armsaim. Armsate inimeste
keskel.</art>
```

Joonis 1.2 Märksõnaotsing veebis Õpilase ÕS-ist ja sama kirje väljavõte sõnastiku originaalfailist.

1) Seletussõnaraamat

armas <armsa, -t 18 e. armsat 17>

1. adj. **armas** *tatud, kallis*. **Armas** ema, õde, mees. Minu kõige armsam tädi. Siiski oli see tüdruk Sandrile **armas**. Meile armsaks saanud kool, laul. Tahaksin olla jälle armsas kodus. Meie **armas** kodumaa. Ärge unustage **armast** emakeelt! Jookse, kui sulle elu on **armas**. Mida armsam laps, seda kibedam vits. (*Vanasõna*)



o *iron.* Armsad sugulased hoolitsesid selle eest, et naine kõik teada saaks. Meie **armas** koolivend valetab, nii et suu suitseb.

• (*keskarmas* v. *ülivõrdes*;) *lemmikarmas*. Tema armsaim ajaviide, tegevus, laul. Lapse armsamad mänguasjad. Minu armsaim raamat. Missugune on su kõige armsam lill? Tema armsamaks toiminguks näis olevat võtta napsi.. (*J. Vahtra*)

• (*pöördumisfraasides, näit. kirja v. kõne alguses*). **Armas** sõber! Armsad kuulajad! Aga, **armas** inimene, sa ju lubasid! Sul on õigus, **armas** Jaan. "Armas hing, armu peab nõudma, mitte õigust," õpetas vaimulik. (*A. H. Tammsaare*)

2. adj. *kena, meeldiv, sümpaatne; lahke, armastusväärne*. Küll teil on armsad lapsed. Perenaine oli meie vastu väga **armas**. Ta on suured armsad silmad. Kassipojad on nii armsad. Kui armsad on kevadel esimesed sinililled!

o (*viisakusvormelites*). Väga **armas**, et tulid! **Armas** kuulda, et teil kõik hästi läheb. Leili ja Mari, olge nii armsad ja tehke see ära! Kui **armas** Sinust, et Sa vaeva paljaks ei pidanud. (*E. Krusten*) || *hubane, õdus*. Kodus oli kõik puhas ja **armas**. Aga rõõmsalt lõõmas kasehalgude tuli, mille armsas õhetuses vennaksed end meeltemööda soojendasid. (*F. Tuglas (tlk.)*)

Liitsõnad: hale**armas**, ime**armas**, üli**armas**.

3. s. **armas** *tatu, kallim; lähedane, armastatud inimene*. Ütle, **armas**, millal me jälle näeme! Jääge jumalaga, mu armsad! ..nägi unes oma kallist kodukohta ja oma armsaid, kes teda seal ootavad. (*O. Luts*)

Fraseoloogia:

nii et hing **armas** vt ka **hing**

(oh sa ~ sina) **armas** aeg ~ taevas ~ jumal vt ka **aeg, taevas, jumal** >>

<art> <mrks>armas</mrks><grm>armsa, -t <tyyp>18</tyyp> e. armsat <tyyp>17</tyyp> </grm><nr>1.</nr>
<sl>adj. </sl><sel>armastatud, kallis. </sel><naide>Armas ema, õde, mees. Minu kõige armsam tädi. Siiski oli see tüdruk Sandrile armas. Meile armsaks saanud kool, laul. Tahaksin olla jälle armsas kodus. Meie arms kodumaa. Ärge unustage armast emakeelt! Jookse, kui sulle elu on armas. Mida armsam laps, seda kibedam vits. Vs. </naide><kasutusjuht>o</kasutusjuht> <stiil>iron. </stiil><naide>Armsad sugulased hoolitsesid selle eest, et naine kõik teada saaks. Meie armas koolivend valetab, nii et suu suitseb. </naide><txh_varjund>•</txh_varjund> <sel>(kesk- v. ülivõrdes;) lemmik-. </sel><naide>Tema armsaim ajaviide, tegevus, laul. Lapse armsamad mänguasjad. Minu armsaim raamat. Missugune on su kõige armsam lill? *Tema armsamaks toiminguks näis olevat võtta napsi.. </naide><o>J. Vahtra. </o><txh_varjund>•</txh_varjund> <sel>(pöördumisfraasides, näit. kirja v. kõne alguses). </sel><naide>Armas sõber! Armsad kuulajad! Aga, armas inimene, sa ju lubasid! Sul on õigus, armas Jaan. *"Armas hing, armu peab nõudma, mitte õigust," õpetas vaimulik. </naide><o>A. H. Tammsaare. </o><nr>2.</nr> <sl>adj. </sl><sel>kena, meeldiv, sümpaatne; lahke, armastusväärne. </sel><naide>Küll teil on armsad lapsed. Perenaine oli meie vastu väga armas. Tal on suured armsad silmad. Kassipojad on nii armsad. Kui armsad on kevadel esimesed sinililled! </naide><kasutusjuht>o</kasutusjuht> <sel>(viisakusvormelites). </sel><naide>Väga armas, et tulid! Armas kuulda, et teil kõik hästi läheb. Leili ja Mari, olge nii armsad ja tehke see ära! *Kui armas Sinust, et Sa vaeva paljaks ei pidanud. </naide><o>E. Krusten. </o><i>|| hubane, õdus. </i><naide>Kodus oli kõik puhas ja armas. *Aga rõõmsalt lõõmas kasehalgude tuli, mille armsas õhetuses vennaksed end meeltemööda soojendasid. </naide><o>F. Tuglas (tlk.). </o><l>Ls. hale-, ime-, üliarmas. </l><nr>3.</nr> <sl>s. </sl><sel>armastatu, kallim; lähedane, armastatud inimene. </sel><naide>Ütle, armas, millal me jälle näeme! Jääge jumalaga, mu armsad! *..nägi unes oma kallist kodukohta ja oma armsaid, kes teda seal ootavad. </naide><o>O. Luts. </o><fras>Fraseoloogia</fras> <f> nii et hing armas </f><lyh>vt ka</lyh> <ptr>hing</ptr> <f>(oh sa ~ sina) armas aeg ~ taevas ~ jumal </f><lyh>vt ka</lyh> <ptr>aeg</ptr>, <ptr>taevas</ptr>, <ptr>jumal</ptr> </art>

Joonis 1.3 Märksõnaotsing veebis Seletussõnaraamatust ning sama kirje väljavõte sõnastiku originaalfailist.

Ülaltoodud näited annavad ülevaate ka sellest, millise sõnaraamatuga on tegu. Kui Seletava sõnaraamatu puhul on selge, et on oluline anda võimalikult põhjalik info, eelkõige pikad seletused, siis Õpilase ÕS-i puhul on vajalik minimalistlikkus, st. kõikide märksõnade kohta on antud ainult kõige vajalikum info. Õpilase ÕS-is järgneb enamasti märksõnale ainult grammatiline info ning lühike seletus või ainult viide.

kala/.pilk, -pilgu, -.pilku *ükskõikne külm pilk*

(Õpilase ÕS)

ÕS-is on sageli palju näiteid, kuid seletused on ülevaatlilikud ning lühikesed, grammatikale on aga pööratud rohkem rõhku kui Seletussõnaraamatus, mis on orienteeritud märksõnadele seletuste andmisele ning mitmerealistele fraseologismidele.

Üldiselt võib öelda, et iga arvutisõnastik on unikaalne ning paljude nüanssidega.

2. Liitsõnade täiskuju leidmine sõnaartiklitest

Peaaegu igal paber kandjal olevas sõnaraamatus on sõnaartikli juures liitsõnadel lühendatud korduvaid osi. Samuti on ka arvutisõnastikega. Kui esimeste puhul on eesmärgiks ruumi kokku hoida, siis elektrooniliste puhul on püütud järgida trükitud sõnaraamatu infot, eesmärgiga, et mõlemad sõnastikud näeksid välja sarnased. Senistes veebis üleval olevatest sõnaraamatutest pole kõiki liitsõnu võimalik üles leida. Töö järgmises pooles kirjeldatakse, kuidas on arvutisõnastikes liitsõnu kirjutatud ning kuidas autori koostatud programmi abil on võimalik genereerida lühendatud liitsõnadest pikad. Käesoleva töö autor on püüdnud seda probleemi lahendada järgmiselt:

- 1) vaadeldes läbi erinevate sõnade lühendamiste stiilinäidised nii visuaalselt kui tehniliselt
- 2) kirjutades vastava programmi iga märksõnaga seotud liitsõnade leidmiseks
- 3) kontrollides genereeritud liitsõnu Microsoft Wordi Spell Checker'iga ning pisteliselt eesti keele spelleriga, kasutades Filosofti programmi
- 4) tehes statistikat, kui palju on õigeid liitsõnu, mis reegleid peab kindlasti järgima ning millised on eriti keerulised liitsõnad

Järgmises punktis vaatleb autor erinevaid stiile, kuidas on sõnu lühendatud visuaalselt ning ka tehniliselt.

2.1 Liitsõnad ÕS-is ja Seletussõnaraamatus

2.1.1 Liitsõnad algfailis ja HTML kujul

Internetis on sõnastikud elektroonilisel kujul, kuid selle saamiseks peab eelnevalt töötleva algfaili. Sõnastiku algfail on selline märgendatud tekst, kus enamik infot jääb elektroonilisel kujul olevast sõnaraamatust välja (tagid, kommentaarid). Iga sõnastiku jaoks on temale kohaldatud algtekst, kusjuures eri sõnastike puhul on vaja kasutada erinevas mahus erinevaid märgendeid. Edaspidi käsitletakse ainult seda märgendatud teksti osa, mis on seotud liitsõnadega. Kõigepealt analüüsitakse sõnastiku struktuuri ja reegleid, mida liitsõnad sisaldavad ning seejärel tuuakse välja võimalused, kuidas on liitsõnu lühendatud. Vaadeldakse, kuidas on ÕS-i algdokumendis ning sõnastiku HTML kujul sõnu lühendatud. Osade näidete järel on toodud töö autori tehtud programmi väljund antud liitsõnade kohta. Programmist ja selle põhimõttest on pikemalt juttu töö 3. peatükis.

ÕS-i algfailis asuvad liitsõnad näite tagide (<naide></naide>) vahel. Loeteludena on esitatud lühendatud liitsõnad, milles korduvad osad on asendatud liitsõnamärgiga. Liitsõnamärk võib asetseda loetelus oleva sõna ees või taga, vastavalt sellele, kas märksõna esineb kas põhi- või täiendsõnana. Põhisõna kannab mõiste olulist tunnust, täiendsõna täpsustab, iseloomustab teda [Lin05]. Seega liitsõnapiir võib asetseda lühendatud liitsõnas:

1) sõna järel, mistõttu asub loetelu lõpus pikk liitsõna (näites märgitud kollasega), milles liitsõnapiirist vasakule jääb põhisõna, olles ka kõigi eelnevate põhisõnaks. Seda iseloomustab järgnev näide märksõna *kaev* sõnaartikli osa ÕS-i algfailist.

```
<art nr="11702"><hdr><mrks>kaev</mrks> <grm><tyyp_grp><tyyp>20</tyyp> <vorm>kaevu,  
.kaevu</vorm></tyyp_grp></grm>.</hdr> <naited><naide>Läks kaevule, käis kaevul. Joogivee/, karja/  
küla/kaev. Pursk/, puur/, salv/kaev. Pump/kaev</naide> <term>tehn</term> = <naide>pumba/kaev</naide>  
<stiil>argi.</stiil> <naide>Vinn/kaev = vinnaga kaev. Imb/, kaabli/, kanalisatsiooni/, kogumis/, sette/  
vaatlus/kaev.
```

Nagu käesoleva alapeatüki alguses märgitud, sisaldavad sõnastike algfailid palju rohkem infot, kui seda on näha arvutiekraanil. Alljärgnev näide sama sõnaartikli osast HTML kujul on toodud võrdluseks algfailile, mida on kujutatud ülalpool.

```
kaev < 20 kaevu, .kaevu>. ◇ Läks kaevule, käis kaevul. Joogivee/, karja/, küla/kaev. Pursk/, puur/, salv/kaev.  
Pump/kaev tehn = pumba/kaev argi. Vinn/kaev = vinnaga kaev. Imb/, kaabli/, kanalisatsiooni/, kogumis/, sette/  
vaatlus/kaev.
```

2) sõna ees, mistõttu loetelu algab täispika liitsõnaga, milles liitsõnamärgist paremale jääb täiendsõna. See on ka kõigi loetelus olevate sõnade täiendsõnaks. Selle kohta käib järgmine näide märksõna *kaev* sõnaartikli teisest poolst ÕS-i algfailist.

```
Kaevu/kett <hom>-keti</hom>, /kook <hom>-koogu</hom>, /rake <hom>-rakke</hom> = /salv <hom>-salve</hom>, /rõngas, /toru, /pump, /pang. Kaevu/vinn <hom>-vinna</hom> = /ling = /jung</naide><stiil>murdes.</stiil> <naide>Kaevu/vesi, /meister</naide></naited> </art>
```

Antud sõnaartikli osa elektroonilisel kujul näeb välja järgmine:

```
Kaevu/kett (-keti), /kook (-koogu), /rake (-rakke) = /salv (-salve), /rõngas, /toru, /pump, /pang. Kaevu/vinn (-vinna) = /ling = /jung murdes. Kaevu/vesi, /meister
```

Kuna töö eesmärgiks on programmi abil lühendatud sõnade täiskuju leidmine, siis edaspidi on mõnede näidete juures ka programmi väljund antud sõnaartikli või selle osa kohta. Järgnev näide on programmi väljund märksõna *kaev* kohta, kus rea alguses olev number näitab sõnaartikli järjekorranumbrit, peale võrdusmärki on märksõna, siis pärast semikoolonit tulevad märksõnast moodustatavad liitsõnad täiskujul. Kuidas programm liitsõnade täiskujud leiab, on juttu bakalaureusetöö 3. punktis.

```
11702 marksona = kaev; joogiveekaev, karjakaev, külakaev, purskkaev,
puurkaev, salvkaev, pumpkaev, pumbakaev, vinnkaev, imbkaev, kaablikaev,
kanalisatsioonikaev, kogumiskaev, settekaev, vaatluskaev, kaevukett,
kaevukook, kaevurake, kaevusalv, kaevurõngas, kaevutoru, kaevupump,
kaevupang, kaevuvinn, kaevuling, kaevujung, kaevuvesi, kaevumeister;
```

(Programmi väljund)

Ühesõnaga, märksõna võib esineda liitsõnade loetelus nii põhi- kui täiendsõnana. Täiendsõnana võib liitsõna esineda kas nimetavas, omastavas käändes või konsonanttüveliseks lühenevalt (*ne-*, *lane-*, *line-*, *mine-*liitelised nimisõnad).

```
<art nr="40196"><hdr><mrks>.süütamine</mrks> <grm><tyyp_grp><tyyp>12</tyyp>
<vorm>-mise</vorm></tyyp_grp></grm>.</hdr>
<naited><naide>Süütamis/tung</naide> <term>med</term>
<sel>püromaania</sel></naited> </art>
```

(ÕS algfail)

Seega moodustuvad liitsõnad, kus *ne-*, *lane-*, *line-*, või *mine-* liiteline märksõna on täiendsõnaks, lühenenud tüvega. Seetõttu on programmi abil väga keeruline genereerida täiskujul olevaid liitsõnu. Antud näite programmi väljund näeb välja selline:

Enamasti on sõnaartiklis liitsõnad, mis on saadud antud artikli märksõna kasutades. Harvadel juhtudel aga on sõnaartikli sees üle mindud uuele märksõnale ning edasi liitsõnu välja kirjutatud või lühendatud sellega. Sellised võivad olla olukorrad, kus antud märksõna asemel on kasutusel uuem või rohkem levinud sõna. ÕS-is leidub taolisi näiteid üpris mitmes kirjes. Vaatleme märksõna *kaal* ning temast moodustatavaid liitsõnu (allolevad näited algfailis ning elektroonilisel kujul). Võib kasutada nii liitsõnu *keha-* ja *eluskaal* kui ka *keha-* ja *elusmass*. Teiseks näiteks on liitsõna *aatom-* ja *molekulkaal*, mille asemel soovitab ÕS kasutada sõna *aatom-* ja *molekulmass*. Sellisel moel liitsõnade täiskuju genereerimine programmi abil on üks keeruline protsess, millest tuleb juttu alapunktis 4.1. Näited liitsõna põhisõna muutmiseks märksõnaga *kaal*:

```
<art nr="11452"><hdr><mrks>kaal_</mrks> <grm><tyyp_grp><tyyp>20</tyyp> <vorm>kaalu,
.kaalu</vorm></tyyp_grp></grm></hdr> <tahendus> <sel>(keha) raskus,</sel> <k>oskuskeeles</k>
<aa>mas's;</aa> <sel>kaalumisseade, kaalud; tähtsus, väärtus, mõju.</sel> <stiil>vmo</stiil> <sel>10 puuda
(linade kaaluühik).</sel> </tahendus> <naited><naide>Koorma, kauba kaal. Võtab kaalus <lyh>v</lyh> kaalu
juurde, kaalust maha. Kaaluta olek</naide> <sel>kaalutus.</sel> <naide>Müüb kaaluga marju</naide>
<sel>(kaalukaubana).</sel> <naide>Kaaluga müüdiv kangas = kaalu/kangas. Paneb oma elu, saatuse
kaalule</naide> <sel>(seab ohtu).</sel> <naide>Kaalul on eluõnn</naide> <sel>(otsustamisel v ohus).</sel>
<naide>Sõnal on kaalu</naide> <sel>(väärtust, mõju).</sel> <vrd2><naide>Keha/, elus/kaal,</naide>
<lyh>vrd</lyh> <k>oskuskeeles</k> <aa>keha/, elus/mass.</aa></vrd2> <naide><halb_v>Aatom/,
molekul/kaal</halb_v> <etem_p>aatom/, molekul/mass</etem_p>.</naide> <term>sport:</term>
<naide>helves/, sulg/, kerge/, raske/kaal. Kang/, osuti/, vedru/, selve/kaal(ud). Apteegi/kaal(ud). Tasa/, vastu/,
üle/kaal.</naide> <vrd2><naide>Kaaluiive, /kadu, /ühik,</naide> <lyh>vrd</lyh> <k>oskuskeeles</k>
<aa>massi/iive, /kadu, /ühik.</aa></vrd2> <naide>Kaalukang, /kauss, /keel, /pomm, /viht <hom>-vihi</hom>.
Kaal/iste <hom>-iste</hom></naide> <term>sport</term></naited> </art>
```

(ÕS-i algfail)

kaal < 20 kaalu, .kaalu> (keha) raskus, oskuskeeles mas's; kaalumisseade, kaalud; tähtsus, väärtus, mõju; vmo 10 puuda (linade kaaluühik). ◊ Koorma, kauba kaal. Võtab kaalus v kaalu juurde, kaalust maha. Kaaluta olek kaalutus. Müüb kaaluga marju (kaalukaubana). Kaaluga müüdiv kangas = kaalu/kangas. Paneb oma elu, saatuse kaalule (seab ohtu). Kaalul on eluõnn (otsustamisel v ohus). Sõnal on kaalu (väärtust, mõju). Keha/, elus/kaal, vrd oskuskeeles keha/, elus/mass. {Aatom/, molekul/kaal} -> aatom/, molekul/mass. sport: helves/, sulg/, kerge/, raske/kaal. Kang/, osuti/, vedru/, selve/kaal(ud). Apteegi/kaal(ud). Tasa/, vastu/, üle/kaal. Kaaluiive, /kadu, /ühik, vrd oskuskeeles massi/iive, /kadu, /ühik. Kaalu/kang, /kauss, /keel, /pomm, /viht (-vihi). Kaal/iste (-iste) sport

(ÕS-i elektrooniline kuju)

Programmi väljund näeb välja järgmine:

```
11452 marksona =
```

```
kaal; kaalukangas, kehakaal, eluskaal, aatomkaal, molekulkaal, aatommass, molekulmass, helveskaal, sulgkaal, kergekaal, raskekaal, kangkaal, osutikaal, vedrukaal, selvekaal, apteegikaal, tasakaal, vastukaal, ülekaal, kaaluiive, kaalukadu, kaaluühik, kaalukang, kaalukauss, kaalukeel, kaalupomm, kaaluviht, kaaliste;
```

Erinevatest märksõna kujude ja sellega seotud liitsõnade lühendamise näited on ära toodud ka

Lisas 1.

3. Programm liitsõnade täiskuju genereerimiseks

Sõnaraamatutes lühendatud liitsõnade täiskuju leidmiseks tehti programm (Lisa 2). Alljärgnevalt on toodud programmi algoritmi kirjeldus punktide haaval.

Algoritm:

- 1) Loetakse artikli algfailist.
- 2) Kontrollitakse, kas algfail on lõppenud
- 3) Kui on lõppenud, siis minnakse punkti 11. Kui algfail pole lõppenud, siis jätkatakse punktist 4)
- 4) Omistatakse stringile \$marksonad = ' '.
- 5) Moodustatakse märksõnade massiivi.
- 6) Omistatakse muutujale \$i=0 (esimesel korral algväärtustatud nulliga).
- 7) Kontrollitakse, kas märksõna leidub näite liikmes ühel või teisel pool liitsõnamärki..
- 8) Kui ei leidu kummalgi pool, siis minna punkti 12. Kui leidub jätkatakse punktist 9.
- 9) Kontrollitakse, kas märksõna sisaldub liitsõnapiirist ees- või tagapool.
- 10) Kui märksõna sisaldub liitsõnapiirist eespool, siis jäetakse meelde piirist paremal pool olev täiendsõna ning minnakse punkti 11. Kui sisaldub tagapool, siis minnakse lihtsalt punkti 11.
- 11) Eemaldada liitsõnapiir sõna keskelt ja minna punkti 15.
- 12) Kontrollitakse, kas liitsõnapiir asub lühendatud liitsõnast paremal või vasakul pool.
- 13) Kui piir asub vasakul pool, siis liidetakse täiendsõna korduva osa (liitsõnapiiri) asemele ning minnakse punkti 15.
- 14) Kui paremal pool, siis liidetakse liitsõnapiiri asemele märksõna ning minna punkti 15.
- 15) Kontrollitakse, kas on tegemist massiivi lõpuelemendiga, kui ei ole, minnakse punkti 5) . Kui on, siis loetakse järgmine artikkel algfailist
- 16) Jätkatakse punktist 2)
- 17) Töö lõpp.

4. Programmi tulemuse õigsus

Käsitletavate sõnastike (ÕS, Seletussõnaraamat ja Õpilase ÕS) koostamisel on iga sõnastiku jaoks kasutatud oma reegleid, mis muudab programmi töö keerulisemaks. Seega on raske saavutada iga sõnaraamatu jaoks 100% õige tulemus, st. kõik genereeritud liitsõnad vastavad eesti keele reeglitele. Kui suudetakse programmis ette anda kõik liitsõnade esinemise viisid, siis suudab programm õiged lahendused väljastada. Käesoleva bakalaureuse töö raames tehtud programm ei tule toime kõigi juhtudega, kuna eesti keele sõnade käänamisega seotud infot on antud programmile keeruline ette anda. Seepärast üritab autor saavutada sõnade moodustamisel võimalikult kõrge protsendi õigeid liitsõnu. Järgmistes punktides antakse ülevaade, milliseid olukordi on võimalik genereeritud programmiga lahendada ning millistes kohtades moodustuvad keeleliselt ebakorrektsed sõnad. Keskendatakse põhiliselt ÕS-i tingimustele vastava programmi tööle, kuna Seletussõnaraamatu jaoks on programmi tööpõhimõtte sama.

4.1 Programmi korrektsed tulemused

Koostatud programm (vt Lisa 2) saab edukalt hakkama olukordadega, kui sõnaartiklis on liitsõnu lühendatud nii, et korduv osa on antud märksõna kõikide liitsõnade puhul samas käändes.

1) Kui märksõna esineb liitsõnade loetelus ainult põhisõnana, siis toimetab liitsõnadega järgnev programmilõik, mis kontrollib liitsõnamärgi asukohta. Kui see on lühendatud liitsõna taga, siis asendatakse märk puuduoleva põhisõnaga (märksõna nimetavas käändes).

```
if ( ( length $j ) - 1 == $kaldu )
{
    $j = substr( $j, 0, $kaldu ) . $i;
}
```

Joonis 4.1 Programmilõik, mis tegeleb lühendatud liitsõnadega, kus liitsõnapiir on lühendatud liitsõna taga.

Kui programm saab näiteks sisendiks sõnaartikli märksõnaga *rikas*, siis programmilõigus 4.1 saab muutuja \$j väärtuseks hetkel lühenenud liitsõna koos liitsõnapiiriga ning muutuja \$i

väärtuseks on märksõna *rikas* põhisõnana nimetavas käändes. Kui liitsõnapiiri (antud näites '/') indeks võrdub lühenenud liitsõna viimase sümboliga, siis omistatakse see märksõnale.

```
<art nr="34604"><hdr><mrks><anto_v>rikas</anto_v></mrks>
<grm><tyyp_grp><tyyp>33</tyyp> <vorm>.rikka,
rikast</vorm></tyyp_grp></grm></hdr> <tahendus> <sel>jõukas, rahakas,
varakas; rikkalik; -rohke.</sel> <xx><anto_p>vaene</anto_p>.</xx>
</tahendus> <naited><naide>Püsti, ratsa rikas = väga rikas. On püstijalu,
jalgsi rikkaks läinud</naide> <sel>on kiiresti rikastunud.</sel>
<naide>Rikaste linnaosa. Rikas maak sisaldab palju kasulikke aineid. Seente
poolest rikas <lyh>v</lyh> seentest rikas <lyh>v</lyh> seene/rikas mets.
Puru/rikas. Uus/rikas. Kära/, liigi/, rahva/; laste/, muljete/, ohvrite/;
kuulsus/, ootus/, vastutus/rikas</naide></naited> </art>           (ÕS-i algfail)
```

ÕS-is esineb märksõna *rikas* kõikidest temast moodustavates liitsõnades pärast liitsõnapiiri, seega programmi väljundiks on liitsõnad, kus märksõna *rikas* sisaldub põhisõnana.

```
34604 marksona = rikas; seenerikas, pururikas, uusrikas, kärarikas,
liigirikas, rahvarikas, lasterikas, muljeterikas, ohvriterikas,
kuulsusrikas, ootusrikas, vastutusrikas;           (Programmi väljund)
```

Üldiselt võib väita, et kui liitsõnade täiskujud on vaja moodustada sellisel moel, et sõnaartiklis on lühendatud liitsõnade korduvaks osaks märksõna nimetavas käändes, saab programm sellega korrektselt hakkama. Valikuliselt on tulemusi on kantud ka Lisasse 1.

2) Kui märksõna esineb liitsõnade loetelus ainult täiendsõnana, siis tegeleb lühendatud liitsõnadega järgnev programmilõik. Programm fikseerib liitsõnapiiri indeksi ning kui selleks on lühenenud liitsõna esimese sümboli indeks, asendatakse indeks puuduva täiendsõnaga. Täiendsõnaks on antud juhul loetelu alguses asuva liitsõna koostises olev märksõna mingis käändes.

```
$pos1 = index( $j, "/" );
if ( $pos1 >= 0 ) {
    if ( $pos1 == 0 )          {
        $j = $i . substr( $j, 1 );
    }
}
```

Joonis 4.2 Programmilõik, mis tegeleb lühendatud liitsõnadega, kus liitsõnapiir on lühendatud liitsõna ees.

Analüüsimise programmilõigu tööd märksõnaga *rask*.

```
<art nr="33691"><hdr><mrks>rask/</mrks>.</hdr> <sel>raske;
raskesti.</sel></tahendus> <naited> <term>tehn:</term>
<naide>rask/bensiin</naide> <sel>(kõrge keemistemperatuuriga),</sel>
<naide>/metall</naide> <sel>(suure tihedusega),</sel> <naide>/õli.
Rask/betoon</naide> <term>eh.</term> <naide>Rask/lahustuv, /sulav,
/süttiv</naide></naited> </art>
```

(ÕS-i algfail)

Kui programmi sisendiks on sõnaartikkel märksõnaga *rask*, siis programmilõigu 4.2 alguses on muutuja \$i väärtuseks loetelu ees oleva liitsõna täiendsõna ning \$j väärtuseks lühenenud liitsõna koos liitsõnamärgiga sõna alguses. Programmilõik kontrollib, kas liitsõnapiiri indeks võrdub lühenenud liitsõna esimese sümboliga. Kui on, siis asendatakse liitsõnapiir muutuja \$i väärtuseks oleva õige täiendsõnaga. Programmi väljund on järgmine:

```
33691 marksona = rask; raskbensiin, raskmetall, raskõli, raskbetoon,
rasklahustuv, rasksulav, rasksüttiv;
```

Võib öelda, et antud stiilis olevate liitsõnade täiskuju moodustamine on programmi abil võimalik ning tulemus nagu eelmise punkti 1) korral ka üldiselt kontrollitav. Kui on võimalik programmile ette anda lühenenud liitsõna kuju, suudetakse genereerida õiged liitsõnad.

3) Kui liitsõnade loetelusid on mõlemas stiilis (punktid 1) ja 2)) korraga, siis kontrollib programm iga liitsõna korral, kas on tegemist lühenenud sõnaga.

```
$kaldu=index($j, '/');
if ($kaldu >=0 and length $j>0){
  if (index($j,$i)<0) {
    $pos1=index($j, substr($i,0,3));
```

Joonis 4.3 Programmilõik, mis kontrollib, kas tegu lühendatud liitsõnaga.

Kõigepealt fikseeritakse liitsõnapiiri indeks ning seejärel kontrollitakse, kas ühel või teisel pool liitsõnastringi juba ei ole märksõna, st et pole täiskujul liitsõna juba olemas. Kui muutujas \$1 olevat märksõna ei sisaldu lühenenud liitsõnas, siis minnakse edasi kontrollima punkte 1) ja 2).

```
<art nr="662"><hdr><mrks>aine</mrks> <grm><tyyp_grp><tyyp>31</tyyp>
<vorm>.aine, ainet</vorm></tyyp_grp></grm>.</hdr> <naited><naide>Tahke
aine</naide> <sel>tahkis.</sel> <naide>Ainete ringkäik (looduses) =
aine/ringe. Maitse/, toime/, kütte/aine</naide> <sel>kütus,</sel>
<naide>ründe/aine</naide> <term>sõj,</term> <naide>kleep/, lihv/aine.
Kiud/, värv/, valk/aine</naide> <term>biol,</term> <naide>kõhr/aine</naide>
```

```

<term>anat.</term> <vrd2><naide>Toit/ = toite/aine</naide>, <lyh>vrd</lyh>
<naide>toidu/aine.</naide></vrd2> <term>anat</term> <k>(ajus):</k>
<naide>valge/ ja hall/aine. Aher/aine</naide> <term>mäend.</term>
<naide>Toor/aine</naide> <sel>toore.</sel> <naide>Anti/aine</naide>
<term>füüs.</term> <naide>"Kalevipoja" ainel <lyh>v</lyh> ainetel loodud
maalid. Arutluse aine = arutlus/aine. Jutu/, kõne/, uurimis/aine.
Fakultatiivne (õppe/)aine = fakultatiiv/aine = vaba/aine. Alternatiivne
(õppe/)aine = alternatiiv/aine = valik/aine. Humanitaar/ ja
reaal/ained.</naide> <bb>Om-ga:</bb> <naide>aine/osake; aine/kogu</naide>
<sel>ainestiku kogu;</sel> <naide>aine/tund, /õpetaja,
/test</naide></naited> </art>

```

(ÕS-i väljundfail)

Programmi korrektne väljund näeb välja selline :

662

```

marksona=aine;aineringe,maitseaine,toimeaine,kütteaine,ründeaine,kleepaine,
lihvaine,kiudaine,värvaine,valkaine,kõhraine,toitaine,toiteaine,toiduaine,v
algeaine,hallaine,aheraine,tooraine,antiaaine,arutlusaine,jutuaine,kõneaine,
uurimisaine,õppeaine,fakultatiivaine,vabaaaine,õppeaine,alternatiivaine, vali
kaine,humanitaaraine,reaalaine,aineosake,ainekogu,ainetund,aineõpetaja,aine
test;

```

Programm analüüsib liitsõnade loetelusid eraldi, iga uue loetelu puhul kontrollitakse, kas märksõna peab lühendatud kujul esinema põhi- või täiendsõnana. Juht 3) on samuti enamikel puhkudel võimalik lahendada, kui märksõnaga ei toimu palju tüvemuutusi loetelu sees.

Punktis 4.2 on välja toodud olukorrad, millega programm korrektselt hakkama ei saanud.

4.2 Programmi ebakorrektsed tulemused

Programmi väljundis ei vasta kõik liitsõnad eesti keele reeglitele. Allpool tuuakse olukorrad, millal ei suudetud genereerida korrektset tulemust.

Alapunktis 2.1.1 öeldakse, et liitsõnas võib esineda täiendsõnana lühenenud kujul. Sellisteks sõnadeks on *ne-*, *lane-*, *line-*, *mine-*liitelised nimisõnad, mis lühenevad konsonanttüveliseks. Selliste sõnade puhul osutus probleemiks märksõnas eraldatud lõpp. Vaatleme põgusalt ÕS-i algfailis sõnaartiklit märksõnaga *väike[ne]*.

```

<art nr="48486"><hdr><mrks>.väike[ne]</mrks> <grm><tyyp_grp><tyyp>12</tyyp>
<vorm>.väikese <lyh>ja</lyh> .väikse, .väikest, .väikesesse <lyh>ja</lyh>
.väiksesse <lyh>ja</lyh> .väikesse, .väikeste, .väikesi <lyh>ja</lyh>
.väikseid, .väikestesse <lyh>ja</lyh> .väiksesse <lyh>ja</lyh>
.väikesisse</vorm></tyyp_grp>; <vqrre_kesk>vähem <lyh>ja</lyh> .väikesem
<lyh>ja</lyh> .väiksem</vqrre_kesk>, <vqrre_üli>kõige vähem <lyh>ja</lyh>
kõige .väikesem <lyh>ja</lyh> kõige .väiksem <lyh>ja</lyh> vähim
<lyh>ja</lyh> .väikseim <lyh>ja</lyh> .väikesim</vqrre_üli></grm></hdr>
<tahendus>. </tahendus> <naited><naide>Arv, tulu, saak on väike. Oli
väiksest peast väga hea laps. Kardab väiksest saadik <lyh>v</lyh> väiksest
peale pimedust. Peab ennast väikest viisi <lyh>v</lyh> väikest/viisi
<lyh>v</lyh> ^väikselt kunstnikuks. Kaduv/, ime/, üli/väike.
Väik(e)se/formaadiline, /mahuline, /palgaline, /ruuduline, /tiraapiline.

```

Väiksema/palgaline. Väike/aju</naide> <term>anat.</term>
<naide>Väike/(sõidu)/auto, /buss, /veoauto; /linn, /rahvas, /tööstus,
/trükis.</naide> <term>zool:</term> <naide>väike/-karihiir,
/pütt</naide></naited> </art> (ÕS-i algfail)

Antud sõnaartikli puhul ei tööta programm samamoodi kui eelnevates näidetes, kuna lõpp *ne* on märksõnas eraldatud. Mõne liitsõna puhul tundub, nagu võiks põhisõnana kasutada ka *ne*-lõpulist nimisõna, kuid eesti keeles kasutatakse lühenenud tüvega sõna. Seetõttu ei pea autor väljundit absoluutselt valeks, kuid mõonab, et eesti keele reeglitele tuginedes peaks kasutama lühenenud tüve varianti.

48486 marksona = väike[ne]; väikestviisi, kaduvväike[ne], imeväike[ne], üliväike, väik(e)seformaadiline, väik(e)semahuline, väik(e)sepalgaline, väik(e)seruuduline, väik(e)setiraapiline, väiksemapalgaline, väikeaju, väike(sõidu)auto, väikebuss, väikeveoauto, väikelinn, väikerahvas, väiketööstus, väiketrukis, väike-karihiir, väikepütt; (Programmi väljund)

Lisaks märksõnale *väike[ne]*, millega saab sarnaselt ÕS-is liitsõnu moodustada, on ka sõnad *ainuke[ne]*, *armuke[ne]*, *asjake[ne]*, *ivake[ne]*, *kalake[ne]*, *kannike[ne]*, *karbike[ne]*, *kelluke[ne]*, *kepike[ne]*, *killuke[ne]*, *kivike[ne]*, *kohuke[ne]*, *kooruke[ne]*, *korvike[ne]*, *kübe[me]ke[ne]*, *lapike[ne]*, *lühike[ne]*, *majake[ne]*, *mehike[ne]*, *mutike[ne]*, *märjuke[ne]*, *natuke[ne]*, *nääpsuke[ne]*, *osake[ne]*, *paike[ne]*, *paluke[ne]*, *piisake[ne]*, *pikuke[ne]*, *pilguke[ne]*, *pisike[ne]*, *plaadike[ne]*, *poisike[ne]*, *põieke[ne]*, *päike[ne]*, *pätsike[ne]*, *raasuke[ne]*, *rakuke[ne]*, *salguke[ne]*, *taadike[ne]*, *terake[ne]*, *toruke[ne]*, *tossike[ne]*, *tuluke[ne]*, *tädike[ne]*, *tükike[ne]*, *vanake[ne]*, *vatsake[ne]*, *veidike[ne]*, *õhuke[ne]*, *äike[ne]*.

Autori arvates on selliste märksõnadega kirjed kõige nüansirohkemad ning nõuavad edaspidi arvutisõnastike koostamisel suuremat tähelepanu. Tingimusi, mis käivad liitsõnade täiskuju moodustamise kohta, tuleks *ne*-liiteliste nimisõnade kohta eraldi kirja panna.

Peale *ne*- lõpuliste sõnade osutusid ÕS-is rasketeks veel sellised sõnad, milles ühe näite piires on märksõna nii täiend- kui põhisõnaks. Näitena vaadeldakse märksõna *aade*:

```
<art nr="13"><hdr><mrks>aade</mrks> <grm><tyyp_grp><tyyp>31</tyyp>  
<vorm>.aate, aadet</vorm></tyyp_grp></grm></hdr> <tahendus> <sel>ülev  
mõte.</sel> </tahendus> <naited><naide>Vabaduse, karskuse aade = vabadus/,  
karskus/<aade>. <aate>/kaaslane, /mees</naide> <sel>aatleja</sel></naited>  
</art> (ÕS-i algfail)
```

Antud sõnas liitsõna *aatemees* saamiseks tuleb pärast liitsõna *aatekaaslane* moodustamist jätta meelde täiendsõna omastavas käändes, kuid see tähendab seda, et märksõna teiseneb

nüüd uuele kujule (*aate*). Seega tuli programmis teha täpsustavaid muudatusi nii, et asi jääks kehtima ka enne õigesti moodustatud liitsõnade juures.

Programmi väljund sõnaga *aade* näeb välja järgmine:

```
13 marksona = aade; vabadusaade, karuskusaade, aatekaaslane, aatemees;
```

Lisaks märksõnale *aade* võib analoogiliselt arutelu teha märksõnaga *ingel*.

Keerukaks osutusid ka astmevahelduslikud sõnad. Ühe näitena võib tuua märksõna *mesi*. Kuna sõna käändub kolmes esimeses käändes *mesi, mee, mett*. Nimetavas ja omastavas käändes oleva sõna sisehäälikute muutuse tõttu ilmnesid algul programmi väljundis valed liitsõnad. Programmi algfailis on sõnaartikkel järgmine:

```
<art nr="23451"><hdr><mrks>mesi</mrks> <grm><tyyp_grp><tyyp>26</tyyp>
<vorm>mee, mett</vorm></tyyp_grp></grm>.</hdr> <naited><naide>Magus kui
mesi = mesi/magus. Määrib ülemusele mett mokale, moka peale, keelele
<lyh>v</lyh> keele peale</naide> <stiil>ülek</stiil> <sel>(ütleb meelitusi,
meelitab).</sel> <naide>Lehe/, õie/, mets</naide> <sel>kimalase
mesi,</sel> <naide>kärje/, vurri/, kanarbiku/, tehis/mesi. Mee/lõhn, /purk.
Mee/võtmise = /võtt, /korje, /kärg.</naide> <term>bot:</term>
<naide>mee/huul, /ketas, /nääre.</naide> <term>aiand:</term>
<naide>mee/kurn <hom>-kurna</hom>, /magasin, /raam, /vurr <hom>-
vurri</hom>. Mesi/jook, /leib</naide> <sel>(sisaldab mett).</sel>
<naide>Mesi/puu = /taru</naide> <sel>taru.</sel> <naide>Mesi/neste</naide>
<sel>nektar.</sel> <naide>Mesi/murakas</naide> <stiil>argi</stiil>
<sel>soomurakas</sel> <term>bot.</term> <naide>Mesi/ohakas</naide>
<term>bot</term></naited> </art>
```

Esialgses väljundis olid sellised sõnad nagu *mesimagasin, mesiraam, mesivurr*, kuid nagu üleval olevast algfailist näha, oleks pidanud moodustuma sõnad *meemagasin, meeraam* ning *meevurr*. Hilisemal analüüsil ning testimisel lisati programmi täiendsõna meeldejätmise käsk, mis tõi kaasa korrektse tulemuse kujul:

```
23451 marksona = mesi; mesimagus, lehemesi, õiemesi, metsmesi, kärjemesi,
vurrimesi, kanarbikumesi, tehismesi, meelõhn, meepurk, meevõtmise, meevõtt,
meekorje, meekärg, meehuul, meeketas, meenääre, meekurn, meemagasin,
meeraam, meevurr, mesijook, mesileib, mesipuu, mesitaru, mesineste,
mesimurakas, mesiohakas;
```

Astmevahelduslikest sõnadest raskemad sõnad on ÕS-is veel *sõda, nuga, ladu, ader, aed*. Eelnevalt loetletud sõnad on kajastatud ka Lisas1.

Kontrollides väljundfaili Wordi Spell Checker-iga ilmnesid üksikud esinemised, kus liitsõna pole suudetud genereerida. Peamisi valesti kirjutatud sõnade liike on kahesuguseid:

- märksõnade puhul, mille alamstring sisaldub lühenenud liitsõnas.

Näiteks märksõna *aeg*, mille sõnaartikli väljavõtte vastava näite koha pealt alamfailis on järgmine: <naide>endis/aeg, **praegus**/ = nüüdis/aeg.</naide>

Kuna programm kontrollib, kas lühendatud liitsõna sisaldab märksõna alamstringina, siis antud olukorras sisaldub märksõna *aeg* täiendsõnas. Seega käsitleb programm sõna *praegus*/täispika liitsõnana ning sõnalt eemaldatakse ainult liitsõnamärk.

Programmi väljund liitsõnaga *aeg* on järgmine:

339 marksona = aeg; endisaeg, **praegus**, nüüdisaeg, ajapikku, päevaaeg, ööaeg, suveaeg, talveaeg, öitse(mis)aeg, saabumisaeg, väljumisaeg, esiaeg, muinasaeg, ürgaeg, kiviaeg, pronksiaeg, rauaaeg, vanaaeg, keskaeg, uusaeg, antiikaeg, feodaalaeg, orduaeg, tsaariaeg, jääaeg, teoorjusaeg, orjaaeg, paganusaeg, kriisiaeg, ärkamisaeg, stagnaaeg, stagnatsioonaaeg, seisakuaeg, marjaaeg, seeneaeg, õieaeg, jooksaeg, innaaeg, pesitsusaeg, korjeaeg, karistusaeg, sundeaeg, kooliaeg, õppeaeg, õpiaeg, tööaeg, puhkeaeg, lennuaeg, garantiiaeg, tagatiseaeg, poolaeg, kolmandikaeg, lisaeg, sünniaeg, surmaaeg, aegilmutus, ajakadu, ajakaotus, ajakulu, ajavaru, ajaühik, ajamärk, ajanõuded, ajaluule, ajaühitus;

- *mine*-lõpuliste ja märksõnade puhul, kus lühenenud liitsõna ei sisaldu loetelus vaid asetseb eraldi näite tagide vahel.

Näiteks märksõna *arenemine*, mille sõnaartikkel on järgmine:

```
<art nr="1959"><hdr><mrks>arenemine</mrks> <grm><tyyp_grp><tyyp>12</tyyp>
<vorm>-mise</vorm></tyyp_grp></grm>.</hdr> <naited><naide>Isendi
arenemine</naide> <sel>isendiareng, ontogenees,</sel> <naide>liigi
arenemine</naide> <sel>liigiareng, fülogenees.</sel> <naide>Ala/, taand/,
üle/arenemine</naide> <sel>ala-, taand-, üleareng.</sel>
<naide>Arenemis/</naide> <sel>arengu-:</sel> <naide>/aste <hom>-
.astme</hom>, /järk, /käik, /lugu, /võime, /võimeline</naide></naited>
</art>
```

Kuna programm vaatab eelkõige liitsõnade loetelusid ning lähtub loetelu alguses või lõpus olevates täispikkadest liitsõnadest, siis antud juhul ei suudeta genereerida õiget liitsõna. Kui märksõna sisaldub täiendsõnas, siis peaks olema loetelu alguses liitsõna täiskuju, mitte lühendatud sõna. Seetõttu programm ei suuda genereerida täiskujusid antud stiilis lühendatud liitsõnadega. Taolised sõnad on veel *avaldamine, kaevamine, kartulipanemine, kartulivõtmise, kõnelemine, puurimine, õppimine, üleandmine*. Samuti leiduvad üksikud näited tegusõnadest: *katsetama, kirjastama, levitama, nühkima, paarituma, pihustama, pikeerima, sorteerima ja ümber kehastuma*.

Liitsõnade genereerimisel peab arvestama kõigi liitsõnade lühenemiskujude võimalustega, kuid kindlasti peab arvestama ka erandolukordadega, mis võivad liitsõnaartiklis esineda. Käesoleva bakalaureuse töö raames ei suudetud kõiki keerulisi olukordi lahendada ja korrekture teha, kuid tööst võiks olla kasu arvutisõnastike edasiarendamisel.

Kokkuvõte

Arvutisõnastike kasutamine on muutunud erinevatel elualadel üha sagedasemaks. Eriti suurt kasutuspiinda on leidnud sõnastike elektroonilised variandid üliõpilaste ja kooliõpilaste seas. Põhjuseid on mitmeid, kuid tähtsamateks neist on mugavus ja kiirus. Ühe näitena võiks mainida viitamist sõnaartiklite vahel. Arvutisõnastike eelis trükitud sõnastike ees on mahukus ning informatsioonirohkus. Paberkujul olevad sõnaraamatud peavad olema kompaktsed, mistõttu võib mõne sõna kohta informatsiooni otsimine osutuda suisa võimatuks.

Elektrooniline sõnastik pole küll trükitud raamatu koopia, kuid see peab säilitama näiliselt paberandjaga sarnase kuju. Seetõttu on säilinud arvutisõnastikes liitsõnade lühendamine, mis on aga eelkõige mõeldud kompaktsel paberkujul oleva sõnaraamatu jaoks. Käesolevas töös püütakse leida lahendus, kuidas arvutisõnastikust kõikide liitsõnade täiskuju üles leida.

Töö teoreetilises pooles antakse üldine ülevaade sõnaraamatus sisalduvast informatsioonist ning sõnaartikli sisust. Sõnaartikkel võib sisaldada märksõnu, grammatilist ja pragmaatilist infot, viiteid ning näiteid. Töös kirjeldatud sõnaartikli osi kirjeldatakse kolme erineva sõnastiku (ÕS, Seletussõnaraamat ja Õpilase ÕS) baasil. Valitud kolm sõnastikku on oma sõnaartiklis sisalduva info poolest väga erinevad. Seletussõnaraamatu põhieesmärk on anda märksõnale võimalikult palju ja erinevaid seletusi. ÕS aga on orienteeritud grammatilisele informatsioonile ning liitsõnadele. Kõige primitiivsemaks osutus Õpilase ÕS, kus on märksõna järel küllalt suur grammatilise info hulk ning lühike seletus. Liitsõnade väljatoomist pole aga Õpilase ÕS-is vajalikuks peetud.

Alates teisest peatükist antakse ülevaade töö praktilisest poolest, kus püütakse täita käesoleva töö põhieesmärk-lühendatud liitsõnade täiskujude genereerimine. Töö käigus vaadatakse läbi erinevad sõnade lühendamiste stiilinäidised. Kuna Seletussõnaraamatu lühendatud liitsõnade kuju erineb ainult algfailis liitsõnaloendi ees asuvast sõnaühendist *Ls*, siis antakse ülevaade programmi tööst ÕS-i baasil. Analüüsi käigus tuuakse näiteid nii sõnastiku algfailist kui ka elektroonilisel kujul olevast variandist. Liitsõnade genereerimiseks kirjutatakse vastav programm ning antakse ülevaade selle tööpõhimõttest. Programmilõikude abil kirjeldatakse, millistel tingimustel milliste tulemusteni jõuti ning tuuakse ka näide lahenduse väljundfailist.

Töö käigus tuuakse välja olukorrad, mida programm suudab kas korrektselt või ebakorrektselt lahendada. Käesoleva töö autori arvates suudab programm suhteliselt hästi lahendada ka sellised olukorrad, kus on tegemist eesti keele ebareeglipäraste sõnadega (nt. ader, põder). Töö autori arvates väärib kõige suuremat tähelepanu liitsõnade moodustamine lüheneva tüvega nimisõnade puhul, kuid käesoleva töö raames selle probleemi juures pikemalt ei peatunud. Rasketeks osutusid genereerimisel omastavas käändes oleva täiendsõnaga genereeritavad liitsõnad, mis asusid mõne teistsuguse loeteluga samas näites, kus märksõna esines põhisõnana. Üksikud liitsõnad, mida genereerida ei suudetud, esinesid sõnaartiklites, milles märksõnaks oli kas tegusõna algvorm või *mine*-lõpuline nimisõna.

Kokkuvõtteks võib öelda, et üldiselt saavutati käesoleva töö praktilises pooles algselt püstitatud eesmärk. Töö käigus ilmnisid siiski mõned erandlikud olukorrad, mille lahendamine vajab edaspidi pikemat analüüsi.

Generating The Full Form of The Compound words

Bachelor thesis

Maarja-Liisa Natus

Abstract

Computer dictionaries have become very useful in different fields. The electronic versions of dictionaries are mostly used among young students. The most important reasons why students prefer electronic versions are ease of use and speed of navigation. There are also many other advantages of the computer dictionaries compared to printed ones. Most relevant are capaciousness and abundance of information. As the printed dictionaries must be compact, searching for specific information can sometimes be quite complicated. The electronic dictionary is not an electronic copy of the printed book, however, the visual form of the printed version must be maintained. That is why the abbreviation of the compound words is used in the computer dictionary. In this bachelor thesis, the author is looking for solutions to find the full form of the compound words.

In the theoretical part of the thesis, an overview of the computer dictionary and detailed description about a dictionary entry is presented. Each dictionary entry may include the head word, grammatical and pragmatic information, also references and examples. The theoretical part is based on the master's thesis of Margit Langemets, the bachelor's theses of Kristel Uiboaed and Külli Kivi. Elements of the dictionary entry are described on the basis of three dictionaries (ÕS, Seletussõnaraamat and Õpilase ÕS). The three chosen dictionaries are very different from each other in terms of the information contained in the entries. The main purpose of Seletussõnaraamat is to give as many explanations as possible for the head word. ÕS, in the other hand, is oriented to grammatical information and compound words. The most trivial of the three is Õpilase ÕS, where after the head word there is some grammar information and also a little explanation. The information about the compound words has not been provided because it is not necessary for this kind of dictionary.

In chapter 2 an overview of the practical work of the thesis is provided. This chapter deals with the main objective of the work- to generate the full form of the compound words. The author has analysed all styles of the appearances of the shortened compound words. As Seletussõnaraamat is similar to ÕS in terms of structure, the review of the program has been given based on the examples of ÕS. In chapter 3 the author has described the work of the algorithm. In chapter 4 the author has provided the analysis on the efficiency of the program that generates the full forms of the compound words.

Finally, the author analyses the cases that the program can solve and the cases it is unable to solve. The examples are enhanced with the opinions of the author. The main conclusion is that when the unsolved cases of the dictionaries have been fixed, the solution helps improve the computer dictionaries in the future.

Viited

- [Kiv93] Kivi, Külli. Sõnastikud struktureeritud tekstiandmebaasina. Diplomitöö Tartu, 1993 .
- [Lan00] Langemets, Margit. Sõnaraamatu arvutilingvistiline analüüs: väitekiri magistrikaadi taotlemiseks. Tallinn, 2000.
- [Tal05] Talve, Birgit. Tekstide kaudu tuvastatud eesti keele tesaurusest puuduvad sõnatähendused. Bakalaureusetöö. Tartu, 2005.
- [Uib05] Uiboaed, Kristel. Sõnaseletuste genereerimine Tesauruse põhjal. Bakalaureusetöö. Tartu, 2005.
- [ÕS99] Eesti Keele Instituut. Eesti Keele Sõnaraamat. ÕS 1999. Tallinn 1999.

URL-id

- [url:AM] Autor: Kadri Muischnek, Heili Orav. Eesti keele korpused ja arvutileksikonid — mis on olemas ja mida veel vaja on. (25.09.2002)
<http://vana.am.ee/3859>
- [url:EKI] <http://www.eki.ee/keeletehnoloogia/projektid/estmorf/estmorf.html>
- [url:IBS] <http://dict.ibs.ee/syn/about.html>
- [url:VES29_05] http://www.vabaeestisona.com/arhiiv/VES29_05/index.html
- [url:WIK] http://et.wikipedia.org/wiki/Eesti_keeles_grammatika
- [url:QS1999] <http://kotkas.ebc.ee/login/?d=qs1999>
- [url:SELETAV] <http://kotkas.ebc.ee/login/?d=seletav>
- [url:QPILASE_QS] http://kotkas.ebc.ee/login/?d=qpilase_qs

Lisa 1

Sõnaraamatu algfail, HTML kuju ja programmi väljund

ÕS-i näited:

1.

- Sõnaraamatu algfail:

```
<art nr="6"><hdr><mrks>aabits</mrks> <grm><tyyp_grp><tyyp>8</tyyp> <vorm>-  
a</vorm></tyyp_grp></grm>.</hdr> <naited><naide><halb_v>Liikuv aabits</halb_v>  
<etem_p>ladumis/aabits</etem_p></naide> <sel>tähetasku ja lipikud.</sel> <naide>Kuke/aabits. Käitumise,  
kunsti aabits <lyh>v</lyh> aabitsa/kursus</naide> <stiil>ülek</stiil> <sel>(algteadmised).</sel>  
<naide>Aabitsa/kirjandus = /raamatud = aabitsad, /tükk, /tarkus</naide></naited> </art>
```

- HTML kuju

aabits < 8 -a>. ◇ {Liikuv aabits} -> ladumis/aabits *tähetasku ja lipikud*. Kuke/aabits. Käitumise, kunsti aabits v aabitsa/kursus *ülek (algteadmised)*. Aabitsa/kirjandus = /raamatud = aabitsad, /tükk, /tarkus

- Programmi väljund

6 marksona=aabits;ladumisaabits,kukeaabits, aabitsakursus,aabitsakirjandus,aabitsaraamatud, aabitsatükk,aabitsatarkus;

2.

- Sõnaraamatu algfail:

```
<art nr="13"><hdr><mrks>aade</mrks> <grm><tyyp_grp><tyyp>31</tyyp> <vorm>.aate,  
aadet</vorm></tyyp_grp></grm></hdr> <tahendus> <sel>ülev mõte.</sel> </tahendus>  
<naited><naide>Vabaduse, karskuse aade = vabadus/, karskus/aade. Aate/kaaslane, /mees</naide>  
<sel>aatleja</sel></naited> </art>
```

- HTML kuju

aade < 31 .aate, aadet> *ülev mõte*. ◇ Vabaduse, karskuse aade = vabadus/, karskus/aade. Aate/kaaslane, /mees *aatleja*

- Programmi väljund

13 marksona = aade; vabadusaade, karskusaade, aatekaaslane, aatemees;

3.

- Sõnaraamatu algfail:

```
<art nr="254"><hdr><mrks>ader</mrks> <grm><tyyp_grp><tyyp>18</tyyp> <vorm>adra, .atra; .atru  
<lyh>ja</lyh> .atrasid</vorm></tyyp_grp></grm>.</hdr> <naited><naide>Hoiab atra vaol. Ei tema kõlba atra  
taha</naide> <stiil>ülek</stiil> <sel>(maatööle).</sel> <naide>Võttis suure tüki maad atra alla</naide>  
<stiil>ülek</stiil> <sel>(haris üles).</sel> <naide>Anna aega atra seada</naide> <stiil>ülek</stiil>  
<sel>(valmistuda).</sel> <naide>Käsi/, hobu/, traktori/, aia/, kraavi/, uudismaa/, hark/, vannas/, hõlm/, konks/,  
mutt/ader. Adra tera, hõlm, nuga <lyh>v</lyh> adra/tera, /hõlm, /nuga. Adra/põllundus</naide>  
<term>etn</term> <sel>künnipõllundus.</sel> <i class="lyh">vt ka</i> <viit kuhu="">adra</em>-  
liitsõnad</em></viit></naited> </art>
```

- HTML kuju:

ader < 18 adra, .atra; .atru ja .atrasid>. ◇ Hoiab atra vaol. Ei tema kõlba adra taha *üle* (*maatööle*). Võttis suure tüki maad adra alla *üle* (*haris üles*). Anna aega atra seada *üle* (*valmistuda*). Käsi/, hobu/, traktori/; aia/, kraavi/, uudismaa/; hark/, vannas/, hõlm/, konks/, mutt/ader. Adra tera, hõlm, nuga v adra/tera, /hõlm, /nuga. Adra/põllundus *etn künnipõllundus. vt adra/ -liitsõnad*

- Programmi väljund:

254 marksona = ader; käsiader, hobuader, traktoriader, aiaader, kraaviader, uudismaaader, harkader, vannasader, hõlmader, konksader, muttader, adra/tera, adrahõlm, adranuga, adrapõllundus;

4.

- Sõnaraamatu algfail:

```
<art nr="322"><hdr><mrks>aed</mrks> <grm><tyyp_grp><tyyp>22</tyyp> <vorm>aia, .aeda, .aeda
<lyh>ja</lyh> aiasse; .aedu <lyh>ja</lyh> .aedasid, .aedadesse <lyh>ja</lyh>
aiusse</vorm></tyyp_grp></grm>.</hdr> <naited><naide>Karjub nagu siga aia vahel. Töö läks aia
taha</naide> <stiil>üle</stiil> <sel>(ebaõnnestus).</sel> <naide>Viljapuu/ = puuvilja/aed, marja/, kapsa/,
tapu/, roosi/aed</naide> <sel>rosaarium,</sel> <naide>roht/</naide> <stiil>vmo</stiil> = <naide>rohu/aed.
Kooli/, kodu/aed. Suve/ ja talve/aed. Lust/</naide> <stiil>vmo</stiil> = <naide>ehis/ = ilu/aed. Laste/aed.
Taime/aed</naide> <term>aiand, mets.</term> <naide>Looma/aed. Karja/, lamba/, kana/, jooksu/aed.
Surnu/aed</naide> <sel>kalmistu.</sel> <naide>Paradiisi/aed.</naide> <term>kal:</term> <naide>vabe/,
vesi/aed</naide> <sel>püünise tiib.</sel> <naide>Pleek/aed</naide> <stiil>vmo</stiil> <sel>(kangaste
pleegitamiseks).</sel> <naide>Uss/aed</naide> <stiil>murdes.</stiil> <sel>Tarad:</sel> <naide>vitstest
(punatud), paekividest (laotud), võrgust aed = vits/, paekivi/, võrk/aed; teivas/ = hirs/ = pistand/aed; piirde/aed.
Aia/lill, /mari <hom>-marja</hom>, /tee, /tool.</naide> <k>Aiatööriistu:</k> <naide>aia/käärid</naide>
<bb>mitm,</bb> <naide>/kärü, /nuga, /prits, /traktor. Aia/roigas, /teivas, /latt</naide> <sel>taralatt,</sel>
<naide>/post</naide> <sel>tarapost.</sel> <naide>Aia/alune, /poolne, /tagune, /äärne.</naide>
<term>bot:</term> <naide>aed/hortensia</naide> <sel>(ilupõõsas),</sel> <naide>/-liivatee</naide>
<sel>(ravimtaim),</sel> <naide>/oblikas</naide> <sel>(kõögiviljataim)</sel></naited> </art>
```

- HTML kuju:

aed < 22 aia, .aeda, .aeda ja aiasse; .aedu ja .aedasid, .aedadesse ja aiusse>. ◇ Karjub nagu siga aia vahel. Töö läks aia taha *üle* (*ebaõnnestus*). Viljapuu/ = puuvilja/aed, marja/, kapsa/, tapu/, roosi/aed *rosaarium*, roht/ *vmo* = rohu/aed. Kooli/, kodu/aed. Suve/ ja talve/aed. Lust/ *vmo* = ehis/ = ilu/aed. Laste/aed. Taime/aed *aiand*, *mets*. Looma/aed. Karja/, lamba/, kana/, jooksu/aed. Surnu/aed *kalmistu*. Paradiisi/aed. *kal:* vabe/, vesi/aed *püünise tiib*. Pleek/aed *vmo* (*kangaste pleegitamiseks*). Uss/aed *murdes*. *Tarad:* vitstest (punatud), paekividest (laotud), võrgust aed = vits/, paekivi/, võrk/aed; teivas/ = hirs/ = pistand/aed; piirde/aed. Aia/lill, /mari □-marja□, /tee, /tool. *Aiatööriistu:* aia/käärid *mitm*, /kärü, /nuga, /prits, /traktor. Aia/roigas, /teivas, /latt *taralatt*, /post *tarapost*. Aia/alune, /poolne, /tagune, /äärne. *bot:* aed/hortensia (*ilupõõsas*), /-liivatee (*ravimtaim*), /oblikas (*kõögiviljataim*)

- Programmi väljund:

322 marksona = aed; viljapuu/aed, puuvilja/aed, marja/aed, kapsa/aed, tapu/aed, roosiaed, rohtaed, rohu/aed, kooliaed, kodu/aed, suve/aed, talve/aed, lusta/aed, ehisaed, iluaed, lasteaed, taime/aed, looma/aed, karja/aed, lamba/aed, kana/aed, jooksu/aed, surnu/aed, paradiisiaed, vabe/aed, vesiaed, pleekaed, ussaed, vitsa/aed, paekiviaed, võrkaed, teivasaed, hirs/aed, pistandaed, piirdeaed, aialill, aiamari, aiatee, aiatool, aiakäärid, aiakärü, aianuga, aiaprits, aiatraktor, aiaroigas, aiateivas, aialatt, aiapost, aiaalune, aiapoolne, aiataagune, aiaäärne, aedhortensia, aed-liivatee, aedoblikas;

5.

- Sõnaraamatu algfail:

```
<art nr="39356"><hdr><mrks>sõda</mrks> <grm><tyyp_grp><tyyp>16</tyyp> <vorm>sõja, sõda, .sõtta
<lyh>ja</lyh> sõjasse; sõdu <lyh>ja</lyh> sõdasid, sõjusse <lyh>ja</lyh>
```

sõdadesse</vorm></tyyp_grp></grm>.</hdr> <naited><naide>Inglismaa astus sõtta. Mehed läksid sõtta. Risti/
rööv/, pärilus/, vallutus/, vabadus/, kodu/, maailma/sõda. Mere/, õhu/, tuuma/sõda. Tolli/, kaubandus/sõda.
Lume/, padja/, vee/, sule/<naide> <stiil>üle</stiil> <naide>sõna/sõda. Sõja/hirm/, /oht/, /propaganda/, /õhutaja,
/ähvardus/, /kolle;/ /kavalus/, /plaan/, /saladus/, /õnn/, /jumal/, /saak;/ /haavad/, /kohus <hom>-kohtu/<hom>, /kool
<hom>-kooli/<hom>, /laager/, /purustus/, /sõnum = /teade <hom>-teate/<hom>, /sündmus/, /tee/, /tehas;/ /film/, /jutt
<hom>-jutu/<hom>, /mäng/, /tants. Sõja/hüüd/, /keeris/, /kisa/, /kära/, /möll,</naide> <stiil>üle</stiil>
<naide>sõja/leek/, /tuli. Sõja/kaaslane, /invaliid/, /ohver/, /veteran. Sõtta/astumine, /mineja,
/minek</naide></naited> </art>

- HTML kuju:

sõda < 16 sõja, sõda, .sõtta ja sõjasse; sõdu ja sõdasid, sõjusse ja sõdadesse>. ◇ Inglismaa astus sõtta. Mehed
läksid sõtta. Risti/, rööv/, pärilus/, vallutus/, vabadus/, kodu/, maailma/sõda. Mere/, õhu/, tuuma/sõda. Tolli/
kaubandus/sõda. Lume/, padja/, vee/, sule/ üle< /stiil> sõna/sõda. Sõja/hirm/, /oht/, /propaganda/, /õhutaja,
/ähvardus/, /kolle;/ /kavalus/, /plaan/, /saladus/, /õnn/, /jumal/, /saak;/ /haavad/, /kohus -kohtu/, /kool -kooli/, /laager/, /purustus/
/sõnum = /teade -teate/, /sündmus/, /tee/, /tehas;/ /film/, /jutt -jutu/, /mäng/, /tants. Sõja/hüüd/, /keeris/, /kisa/, /kära,
/möll, üle< : sõja/leek/, /tuli. Sõja/kaaslane, /invaliid/, /ohver/, /veteran. Sõtta/astumine, /mineja, /minek

- Programmi väljund:

39356 marksona = sõda; ristsõda, röövsõda, pärilussõda, vallutussõda, vabadussõda, kodusõda, maailmasõda,
meresõda, õhusõda, tuumasõda, tollisõda, kaubandussõda, lumesõda, padjasõda, veesõda, sulesõda, sõnasõda,
sõjahirm, sõjaoht, sõjapropaganda, sõjaõhutaja, sõjaähvardus, sõjakolle, sõjakavalus, sõjaplaan, sõjasaladus,
sõjaõnn, sõjajumal, sõjasaak, sõjahaavad, sõjakohus, sõjakool, sõjalaager, sõjapurustus, sõjasõnum, sõjateade,
sõjasündmus, sõjatee, sõjatehas, sõjafilm, sõjajutt, sõjamäng, sõjatants, sõjahüüd, sõjakeeris, sõjakisa, sõjakära,
sõjamöll, sõjaleek, sõjatuli, sõjakaaslane, sõjainvaliid, sõjaohver, sõjaveteran, sõttaastumine, sõttamineja,
sõttaminek;

6.

- Sõnaraamatu algfail:

```
<art nr="26411"><hdr><mrks>nuga</mrks> <grm><tyyp_grp><tyyp>16</tyyp> <vorm>noa, nuga; nuge
<lyh>ja</lyh> nugasid</vorm></tyyp_grp></grm>.</hdr> <naited><naide>Noa tera, selg ja pea. On ämmaga
nugade peal</naide> <stiil>üle</stiil> <sel>(tulis).</sel> <naide>Läheb noa alla, käis noa all</naide>
<stiil>argi</stiil> <sel>(operatsioonil).</sel> <naide>Mees sai nuga</naide> <stiil>argi</stiil> <sel>meest
löödi, haavati noaga.</sel> <naide>Aedniku/, klaasi/, köögi/, liigend/, tasku/nuga; ketas/nuga. Noa/hoop, /teriti,
/tupp, /töö</naide> <sel>nt vestmine,</sel> <naide>/vineer</naide> <term>eh</term> <sel>höövlivineer,</sel>
<naide>/võll</naide> <term>tehn</term> <sel>(höövelmasinal).</sel> <naide>Nuga/terav. Nuga/kala</naide>
<term>zool</term></naited> </art>
```

- HTML kuju:

nuga < 16 noa, nuga; nuge ja nugasid>. ◇ Noa tera, selg ja pea. On ämmaga nugade peal üle< (tulis). Läheb noa
alla, käis noa all argi (operatsioonil). Mees sai nuga argi meest löödi, haavati noaga. Aedniku/, klaasi/, köögi/,
liigend/, tasku/nuga; ketas/nuga. Noa/hoop, /teriti, /tupp, /töö nt vestmine, /vineer eh höövlivineer, /võll tehn
(höövelmasinal). Nuga/terav. Nuga/kala zool

- Programmi väljund:

26411 marksona = nuga; aednikunuga, klaasinuga, kööginuga, liigendnuga, taskunuga, ketasnuga, noahoop,
noateriti, noatupp, noatöö, noavineer, noavõll, nugaterav, nugakala;

7.

- Sõnaraamatu algfail:

<art nr="19122"><hdr><mrks>ladu</mrks> <grm><tyyp_grp><tyyp>16</tyyp> <vorm>lao, ladu, .lattu </lyh>ja</lyh> .laosse</vorm></tyyp_grp></grm></hdr> <tahendus> <sel>kauba jms hoidla;</sel> <term>trük:</term> <sel>ladumine, ladumisviis; ladum, laotu.</sel> </tahendus> <naited><naide>Kauba/, relva/, vilja/, kütte/ = kütuse/, kesk/ladu. Foto/, kirjuti/, käsi/, masin/ladu. Lao/ruum, /tööline, /tõstuk, /käive</naide> <term>maj</term></naited> </art>

- HTML kuju:

ladu < 16 lao, ladu, .lattu ja .laosse> *kauba jms hoidla; trük: ladumine, ladumisviis; ladum, laotu.* ◊ Kauba/, relva/, vilja/, kütte/ = kütuse/, kesk/ladu. Foto/, kirjuti/, käsi/, masin/ladu. Lao/ruum, /tööline, /tõstuk, /käive *maj*

- Programmi väljund:

19122 marksona = ladu; kaubaladu, relvaladu, viljaladu, kütteladu, kütuseladu, keskladu, fotoladu, kirjutiladu, käsiladu, masinladu, laoruum, laotööline, laotõstuk, laokäive;

8.

- Sõnaraamatu algfail:

<art nr="23970"><hdr><mrks>mitu</mrks> <grm><tyyp_grp><tyyp>32</tyyp> <vorm>.mitme, mitut</vorm></tyyp_grp></grm>.</hdr> <naited><naide>Mitut moodi, viisi = mitmel moel, viisil. Mitmel pool, mitmele poole, mitmelt poolt. Mitme otsaga asi</naide> <sel>(keerukas).</sel> <naide><niru_v>Mitmete aastate tagant; kukkus mitmed meetrid; nägi mitmeid ilusaid kirikuid</niru_v> <etem_p>mitme aasta tagant; kukkus mitu meetrit; nägi mitut ilusat kirikut</etem_p>. Mitme/astmeline, /haruline, /kandiline, /kihiline, /kohaline, /kõiteline, /meheline, /osaline, /suunaline, /tahuline, /toaline, /tuhandeline, /värviline, /õieline. Mitme/keelne, /nädalane, /tonnine. Mitme/laadiline = /laadne, /tunniline = /tunnine. Mitme/hobuse/rakend, /põlve/elamu, /päeva/sõit</naide> <term>sport,</term> <naide>/välja/süsteem</naide> <term>põllum,</term> <naide>/värvi/trükk</naide> <term>trük</term></naited> </art>

- HTML kuju:

mitu < 32 .mitme, mitut>. ◊ Mitut moodi, viisi = mitmel moel, viisil. Mitmel pool, mitmele poole, mitmelt poolt. Mitme otsaga asi (*keerukas*). Mitmete aastate tagant; kukkus mitmed meetrid; nägi mitmeid ilusaid kirikuid, parem: mitme aasta tagant; kukkus mitu meetrit; nägi mitut ilusat kirikut. Mitme/astmeline, /haruline, /kandiline, /kihiline, /kohaline, /kõiteline, /meheline, /osaline, /suunaline, /tahuline, /toaline, /tuhandeline, /värviline, /õieline. Mitme/keelne, /nädalane, /tonnine. Mitme/laadiline = /laadne, /tunniline = /tunnine. Mitme/hobuse/rakend, /põlve/elamu, /päeva/sõit *sport*, /välja/süsteem *põllum*, /värvi/trükk *trük*

- Programmi väljund:

23970 marksona = mitu; mitmeastmeline, mitmeharuline, mitmekandiline, mitmekihiline, mitmekohaline, mitmekõiteline, mitmemeheline, mitmeosaline, mitmesuunaline, mitmetahuline, mitmetoaline, mitmetuhandeline, mitmevärviline, mitmeõieline, mitmekeelne, mitmenädalane, mitmetonnine, mitmelaadiline, mitmelaadne, mitmetunniline, mitmetunnine, mitmehobuserakend, mitmepõlveelamu, mitmepäevasõit, mitmeväljasüsteem, mitmevärvitruk;

9.

- Sõnaraamatu algfail:

<art nr="13428"><hdr><mrks>kelp</mrks> <grm><tyyp_grp><tyyp>21</tyyp> <vorm>kelba, .kelpa</vorm></tyyp_grp></grm></hdr> <tahendus> <sel>maja kolmnurkne otskatus;</sel> <stiil>harv</stiil> <sel>(voodi, ree) otslaud.</sel> </tahendus> <naited><naide>Katuse/, maja/kelp. Kelp/katus</naide></naited> </art>

- HTML kuju:

kelp < 21 kelba, .kelpa> *maja kolmnurkne otskatus; harv (voodi, ree) otslaud.* ◇ Katuse/, maja/kelp. Kelp/katus

- Programmi väljund:

13428 marksona = kelp; katusekelp, majakelp, kelpkatus;

10.

- Sõnaraamatu algfail:

```
<art nr="32279"><hdr><mrks>põder</mrks> <grm><tyyp_grp><tyyp>18</tyyp> <vorm>põdra, .põtra; .põtru
<lyh>ja</lyh> .põtrasid</vorm></tyyp_grp></grm>.</hdr> <naited><naide>Ema/põder = põdra/lehm,
isa/põder = põdra/pull. Põhja/põder. Põdra/jaht, /kari <hom>-karja</hom>, /kasvataja, /kasvatus, /liha, /nahk,
/rakend, /vasikas.</naide> <term>bot:</term> <naide>põdra/juur</naide> <sel>(rohttaim),</sel>
<naide>/kanep, /mokk</naide> <sel>(seen).</sel> <term>zool:</term> <naide>põdra/kärbes,
põder/põrnikas</naide></naited> </art>
```

- HTML kujuga:

põder < 18 põdra, .põtra; .põtru ja .põtrasid>. ◇ Ema/põder = põdra/lehm, isa/põder = põdra/pull. Põhja/põder. Põdra/jaht, /kari -karja, /kasvataja, /kasvatus, /liha, /nahk, /rakend, /vasikas. *bot*: põdra/juur (*rohttaim*), /kanep, /mokk (*seen*). *zool*: põdra/kärbes, põder/põrnikas

- Programmi väljund:

32279 marksona = põder; emapõder, põdralehm, isapõder, põdrapull, põhjapõder, põdrajaht, põdrakari, põdrakasvataja, põdrakasvatus, põdraliha, põdrannahk, põdrarakend, põdravasikas, põdrajuur, põdrakanep, põdramokk, põdrakärbes, põderpõrnikas;

Seletussõnaraamatu näited:

1.

- Sõnaraamatu algfail:

```
<art> <mrks>aade</mrks><grm>aate <tyyp>19</tyyp> </grm><sl>s. </sl><sel>ülev, paleuslik juhtmõte v.
idee. </sel><naide>Vabaduse, võrdsuse, vendluse aade. Isamaa kasu oli tema ülim aade. Eluvõõrad aated. Vana
professor elas oma aadete maailmas. *Siis noores hinges löid leekima kõrged aated. Ma nägin uut ilma, lahtist
maad ja laia väljavaadet, ning nägin vaba rahvast tõusvat. </naide><o>F. Tuglas. </o><l>Ls. karskus-, rahuaade.
</l> </art>
```

- HTML kujuga:

aade <aate 19> *s. ülev, paleuslik juhtmõte v. idee.* Vabaduse, võrdsuse, vendluse aade. Isamaa kasu oli tema ülim aade. Eluvõõrad aated. Vana professor elas oma aadete maailmas. Siis noores hinges löid leekima kõrged aated. Ma nägin uut ilma, lahtist maad ja laia väljavaadet, ning nägin vaba rahvast tõusvat. (*F. Tuglas*)

Liitsõnad: karskus-, rahuaade.

- Programmi väljund:

marksona=aade;karskusaade;

2.

- Sõnaraamatu algfail:

<art> <mrks>aine </mrks><grm><tyyp>19</tyyp> </grm><sl>s. </sl><nr>1.</nr> <term>ka fütis. keem. </term><sel>see, millest koosneb füüsiline keha; mateeria vorm, mida iseloomustab stabiilsus ja seisumassi olemasolu. </sel><naide>Tahke, vedel, gaasiline, kristalne, amorfne, pulbriline, vahajas, vees lahustuv, kergesti lenduv, pindaktiivne, plahvatusohtlik, kõrgmolekulaarne, madalmolekulaarne aine. Keemiline aine. Orgaanilised, anorgaanilised ained. Radioaktiivsed, antiseptilised, kantserogeensed ained. Puidust, metallist vm. ainest ese. Ainete ringkäik (looduses) ehk aineringe. Kõik peotoitudeks vajalikud ained on hangitud, ostetud. </naide><l>Ls. aher-, alg-, anti-, ase-, ballast-, ekstraktiiv-, elus-, immutus-, jääk-, kasvu-, kiud-, kleep-, kontrast-, koostis-, kummi-, kütte-, lagu-, liht-, liim-, liit-, lisa-, lubi-, lõhke-, lõhna-, lähte-, mineraal-, mõju-, mõru-, määarde-, mürk-, park-, puit-, põhi-, põletus-, pärilikkus-, rasv-, ravi-, ründe-, sarv-, side-, tarbe-, tehis-, toime-, toit-, toor-, täite-, uimastus-, valk-, varu-, võõr-, väetus-, värv-, väärisaine; kuiv-, maitse-, tang-, toidu-, võrtsaine. </l><nr>2.</nr> <sel>objekt, allikas, lähtematerjal. </sel><nr>a.</nr> <sel>kunstiteose teema aluseks olev lähtematerjal, kujutamiseks valitav tegelikkusloik. </sel><naide>Romaani, novelli, näidendi aine on kirjanik saanud oma kodukohast. Tänuväärne aine filmi, ooperi jaoks. Suur Isamaasõda on pakkunud rohkesti ainet kõikidele kunstidele. "Kalevipoja", antiikmütoloogia ainetel loodud maalid, muusikateosed. </naide><nr>b.</nr> <sel>kõneaine, kõneluse (ka mõtiskluse, arutluse, vaidluse jne.) teema. </sel><naide>Arutluse aineks oli sotsialistlik võistlus. Jutt kaldus argipäevastele ainetele. Sobiv aine mõtisklusteks, vaidlusteks. Sündmus andis ainet kuulujuttudeks, keelepeksuks. *Muidugi oli pajatuseks nüüd niisugune aine valitud, et koolipreili sellest osa võtta ei võinud. </naide><o>E. Vilde. </o><nr>c.</nr> <sel>uurimise ainevald v. objekt. </sel><naide>Väitekirja aine. Rahvaluuleteaduse aineks on traditsiooniline suuline looming. </naide><l>Ls. jutu-, kõne-, uurimis-, vestlusaine. </l><nr>3.</nr> <sel>õppeaine. </sel><naide>Kohustuslikud, fakultatiivsed ained. Paljudele on matemaatika üks raskemaid aineid. Ta sai kõigis ainetes head hinded. </naide><txh_varjund>●</txh_varjund> <sel>(ainelõigu kohta). </sel><naide>Vastajad olid [keemias] oma hinded kätte saanud ja uus aine oli edasi võetud. </naide><o>A. Pervik. </o><l>Ls. eksami-, eri-, eriala-, fakultatiiv-, humanitaar-, kõrval-, lemmik-, meelis-, pea-, põhi-, reaali-, õppeaine. </l> </art>

- HTML kuju:

aine <19> s.

1. ka fütis. keem. see, millest koosneb füüsiline keha; mateeria vorm, mida iseloomustab stabiilsus ja seisumassi olemasolu. Tahke, vedel, gaasiline, kristalne, amorfne, pulbriline, vahajas, vees lahustuv, kergesti lenduv, pindaktiivne, plahvatusohtlik, kõrgmolekulaarne, madalmolekulaarne aine. Keemiline aine. Orgaanilised, anorgaanilised ained. Radioaktiivsed, antiseptilised, kantserogeensed ained. Puidust, metallist vm. ainest ese. Ainete ringkäik (looduses) ehk aineringe. Kõik peotoitudeks vajalikud ained on hangitud, ostetud.

Liitsõnad: aher-, alg-, anti-, ase-, ballast-, ekstraktiiv-, elus-, immutus-, jääk-, kasvu-, kiud-, kleep-, kontrast-, koostis-, kummi-, kütte-, lagu-, liht-, liim-, liit-, lisa-, lubi-, lõhke-, lõhna-, lähte-, mineraal-, mõju-, mõru-, määarde-, mürk-, park-, puit-, põhi-, põletus-, pärilikkus-, rasv-, ravi-, ründe-, sarv-, side-, tarbe-, tehis-, toime-, toit-, toor-, täite-, uimastus-, valk-, varu-, võõr-, väetus-, värv-, väärisaine; kuiv-, maitse-, tang-, toidu-, võrtsaine.

2. objekt, allikas, lähtematerjal.

a. kunstiteose teema aluseks olev lähtematerjal, kujutamiseks valitav tegelikkusloik. Romaani, novelli, näidendi aine on kirjanik saanud oma kodukohast. Tänuväärne aine filmi, ooperi jaoks. Suur Isamaasõda on pakkunud rohkesti ainet kõikidele kunstidele. "Kalevipoja", antiikmütoloogia ainetel loodud maalid, muusikateosed.

b. kõneaine, kõneluse (ka mõtiskluse, arutluse, vaidluse jne.) teema. Arutluse aineks oli sotsialistlik võistlus. Jutt kaldus argipäevastele ainetele. Sobiv aine mõtisklusteks, vaidlusteks. Sündmus andis ainet kuulujuttudeks, keelepeksuks. Muidugi oli pajatuseks nüüd niisugune aine valitud, et koolipreili sellest osa võtta ei võinud. (E. Vilde)

c. uurimise ainevald v. objekt. Väitekirja aine. Rahvaluuleteaduse aineks on traditsiooniline suuline looming.

Liitsõnad: jutu-, kõne-, uurimis-, vestlusaine.

3. õppeaine. Kohustuslikud, fakultatiivsed ained. Paljudele on matemaatika üks raskemaid aineid. Ta sai kõigis ainetes head hinded.

● (ainelõigu kohta). Vastajad olid [keemias] oma hinded kätte saanud ja uus aine oli edasi võetud. (A. Pervik)

Liitsõnad: eksami-, eri-, eriala-, fakultatiiv-, humanitaar-, kõrval-, lemmik-, meelis-, pea-, põhi-, reaali-, õppeaine

- Programmi väljund:

marksona=aine ,aheraine ,algaine ,antiaaine ,aseaine ,ballastaine ,ekstraktiivaine ,elusaine ,immutusaine ,jäakaine ,kasvuaine ,kiudaine ,kleepaine ,kontrastaine ,koostisaine ,kummiaine ,küteteaine ,laguaine ,lihtaine ,liimaine ,liitaine ,lisaine ,lubiaaine ,lõhkeaine ,lõhnaaine ,lõheteaine ,lõheteaine ,mineraalaine ,mõjuaine ,mõruaine ,määrdeaine ,mürkaine ,parkaine ,puitaine ,põhiaaine ,põletusaine ,pärilikkusaine ,rasvaine ,raviaaine ,ründeaine ,sarvaine ,sideaine ,tarbeaine ,tehisaine ,toimeaine ,toitaine ,tooraine ,täiteaine ,uimastusaine ,valkaine ,varuaine ,võõraaine ,väetusaine ,värvaaine ,kuivaine ,maitseaine ,tangaine ,toiduaine ,jutuaine ,kõneaine ,uurimisaine ,eksamiaine ,eriaaine ,erialaaine ,fakultatiivaine ,humanitaaraine ,kõrvalaine ,lemmikaine ,meelisaine ,peaaine ,põhiaaine ,reaalaine ;

3.

- Sõnaraamatu algfail:

```
<art> <mrks>aasta </mrks><grm><tyyp>1</tyyp> </grm><sl>s. </sl><nr>1.</nr> <sel>ajavahemik, mille
vältel Maa teeb tiiru ümber Päikese; kalendriaasta (365, lisapäeva-aastal 366 ööpäeva), lähtepunktiks 1. jaanuar.
</sel><naide>1980. aasta, aasta 1980. Mõõdunud, eelmine, käesolev, tulev, eelolev aasta. Aasta algus, lõpp.
Kongress peetakse kas vana aasta lõpus või uue aasta alguses. August on aasta 8. kuu. Mis aastal sa sündinud
oled? Milline näeb Tallinn välja 2100. aastal, aastal 2100? Aastail 1941--1944. Uut aastat vastu võtma. Öne ja
edu algavaks aastaks! Tulevast aastast alates. Sideeriline aasta </naide><term>astr. </term><sel>'Päikese näiva
liikumise vältus tähtede suhtes, täheaasta (365 päeva 6 t. 9 min. 9 sek.)'. </sel><naide>Troopiline aasta
</naide><term>astr. </term><sel>Päikese näiva liikumise vältus kevadpunktist kevadpunktini (365 päeva 5 t.
48 min. 46 sek.)'. </sel><naide>Aastad ei ole vennaksed. Vs. </naide><txh_varjund>●</txh_varjund>
<vormikood>pl. </vormikood><sel>(mõnest v. paljudest aastatest koosneva perioodi kohta).
</sel><naide>Mõõdunud sajandi 60-ndad aastad. Suure Isamaasõja karmidel aastatel. </naide><nr>2.</nr>
<sel>12 kuu pikkune ajavahemik, lähtepunktiks ükskõik missugune moment. </sel><naide>22. veebruaril
möödus 10 aastat onu surmast. See juhtus kümmekond aastat tagasi, mõned head aastad tagasi. Aasta paari
pärad. Päevast päeva, aastast aastasse ikka üks ja sama töö. Ma pole enam aasta otsa teatrisse saanud. Siit jätkub
tööd terveks, kogu aastaks, aastaks otsaks. Suusatajad saavad nüüd trennida aasta ringi. Ta sai varguse eest viis
aastat. Fuksia õitseb peaaegu aasta läbi. Teise aasta ristikupõld. Ta elas aastate viisi välismaal. Aastate jooksul
on kogunenud ülearust koli. Aastate pikku kerkis uus linnaosa Õismäele. Akadeemiline aasta
</naide><sel>õppeaasta kõrgemas õppeasutuses'. </sel><naide>Parem aasta oodata kui kaks kahetseda. Vs.
</naide><txh_varjund>●</txh_varjund> <sel>(vanuse, ea, eluaastate kohta). </sel><naide>Poiss on juba seitse
aastat vana. Ligi 1000 aastat vana toomkirik. 3. mail saab laps aasta vanaks. Ta võib olla nii aasta
nelja(teistkümmene)-viieteistkümmene. Aastaid võib tal olla juba üle neljakümne. Niipalju aastaid ei oskaks talle
küll anda. Ta on oma aastate kohta väga nooruslik. Aastate poolest alles noor mees, aga pea hall otsas. Tal juba
70 aastat turjal. Poiss võis olla aastat viisteist vana. Parimad aastad jäävad juba seljataha. Kes daamide aastaid
oskab või julgeb arvata! Juustes on tal juba aastate hõbedat. Küürus aastate koorma all. Surm ei küsi aastatest.
Armastus ei päri aastaid taga. Vs. </naide><naide>*Andrese ja Mari esimene poeg Indrek käis juba mõnda kuud
neljandat aastat, kui sündis järgmine poeg.. </naide><o>A. H. Tammsaare. </o><f>aastates ~ aastais
</f><sel>(täiendiga:) eas; (täiendita:) eakas, elatanud; mitte enam noor; </sel><stiil>hrv. </stiil><sel>(esemete
vms. kohta:) vana. </sel><naide>Parimates, keskmistes aastates mees. Ta on mees parimates aastates.
Keskealine mees ja umbes samades aastates naine. Tema aastates veel abielluda! Ta juba aastates mees. Tüse,
juba aastates, kuid mitte veel vana naisterahvas. Hallipäine aastais direktor. *Päike soojendas aastates maja
sammaldunud katust.. </naide><o>E. Maasik. </o><l>Ls. abielu-, alg-, algkooli-, algus-, ameti-, aruande-,
asutamis-, eelarve-, elu-, finants-, heina-, ikaldus-, ilmumis-, juubeli-, kalendri-, kasvu-, keskkooli-, kiriku-,
kooli-, kriisi-, kuu-, lapsepõlve-, leina-, lese-, liig-, lisapäeva-, maapao-, majandus-, marja-, mõõna-,
neiupõlve-, noorus-, nälja-, okupatsiooni-, olümpia-, orjus-, pagendus-, pensioni-, pool-, proovi-,
põllumajandus-, põua-, päikese-, raamatu-, rahandus-, rahu-, revolutsiooni-, rännu-, saagi-, surma-, sõja-,
sünni-, teenistus-, tegevus-, tõusu-, tähe-, töö-, valgus-, valitsemis-, vanadus-, vangla-, veerand-, vilja-, õne-,
õpi-, õpingu-, õpipoisi-, õppe-, õuna-, ülikooliaasta; uus-, vana-aasta. </l> </art>
```

- HTML kuju:

aasta <1> s.

1. ajavahemik, mille vältel Maa teeb tiiru ümber Päikese; kalendriaasta (365, lisapäeva-aastal 366 ööpäeva), lähtepunktiks 1. jaanuar. 1980. aasta, aasta 1980. Mõõdunud, eelmine, käesolev, tulev, eelolev aasta. Aasta algus, lõpp. Kongress peetakse kas vana aasta lõpus või uue aasta alguses. August on aasta 8. kuu. Mis aastal sa sündinud oled? Milline näeb Tallinn välja 2100. aastal, aastal 2100? Aastail 1941–1944. Uut aastat vastu võtma. Öne ja edu algavaks aastaks! Tulevast aastast alates. Sideeriline aasta *astr.* *'Päikese näiva liikumise vältus*

tähtede suhtes, täheaasta (365 päeva 6 t. 9 min. 9 sek.)'. Troopiline aasta astr. 'Päikese näiva liikumise vältus kevadpunktist kevadpunktini (365 päeva 5 t. 48 min. 46 sek.)'. Aastad ei ole vennaksed. (Vanasõna)

• <pl. > (mõnest v. paljudest aastatest koosneva perioodi kohta). Mõõdunud sajandi 60-ndad aastad. Suure Isamaasõja karmidel aastatel.

2. 12 kuu pikkune ajavahemik, lähtepunktiks ükskõik missugune moment. 22. veebruaril möödus 10 aastat onu surmast. See juhtus kümmekond aastat tagasi, mõned head aastad tagasi. Aasta paari pärast. Päevast päeva, aastast aastasse ikka üks ja sama töö. Ma pole enam aasta otsa teatrisse saanud. Siit jätkub tööd terveks, kogu aastaks, aastaks otsaks. Suusatajad saavad nüüd treenida aasta ringi. Ta sai varguse eest viis aastat. Fuksia õitseb peaaegu aasta läbi. Teise aasta ristikupõld. Ta elas aastate viisi välismaal. Aastate jooksul on kogunenud üleaurust koli. Aastate pikku kerkis uus linnaosa Õismäele. Akadeemiline aasta 'õppeaasta kõrgemas õppeasutuses'. Parem aasta oodata kui kaks kahetseda. (Vanasõna)

• (vanuse, ea, eluaastate kohta). Poiss on juba seitse aastat vana. Ligi 1000 aastat vana toomkirik. 3. mail saab laps aasta vanaks. Ta võib olla nii aasta nelja(teistkümmene)-viieteistkümmene. Aastaid võib tal olla juba üle neljakümne. Niipalju aastaid ei oskaks talle küll anda. Ta on oma aastate kohta väga nooruslik. Aastate poolest alles noor mees, aga pea hall otsas. Tal juba 70 aastat turjal. Poiss võis olla aastat viisteist vana. Parimad aastad jäävad juba seljataha. Kes daamide aastaid oskab või julgeb arvata! Juustes on tal juba aastate hõbedat. Kүүrus aastate koorma all. Surm ei küsi aastatest. Armastus ei päri aastaid taga. (Vanasõna)

Andrese ja Mari esimene poeg Indrek käis juba mõnda kuud neljandat aastat, kui sündis järgmine poeg.. (A. H. Tammsaare)

aastates ~ aastais (täiendiga:) eas; (täiendita:) eakas, elatanud; mitte enam noor; hrv. (esemete vms. kohta:) vana. Parimates, keskmistes aastates mees. Ta on mees parimates aastates. Keskealine mees ja umbes samades aastates naine. Tema aastates veel abielluda! Ta juba aastates mees. Tüse, juba aastates, kuid mitte veel vana naisterahvas. Hallipäine aastais direktor. Päike soojendas aastates maja sammaldunud katust.. (E. Maasik)

Liitsõnad: abielu-, alg-, algkooli-, algus-, ameti-, aruande-, asutamise-, eelarve-, elu-, finants-, heina-, ikaldus-, ilmumis-, juubeli-, kalendri-, kasvu-, keskkooli-, kiriku-, kooli-, kriisi-, kuu-, lapseõlve-, leina-, lese-, liig-, lisapäeva-, maapao-, majandus-, marja-, mõõna-, neiupõlve-, noorus-, nälja-, okupatsiooni-, olümpia-, orjus-, pagendus-, pensioni-, pool-, proovi-, põllumajandus-, põua-, päikese-, raamatu-, rahandus-, rahu-, revolutsiooni-, rännu-, saagi-, surma-, sõja-, sünni-, teenistus-, tegevus-, tõusu-, tähe-, töö-, valgus-, valitsemise-, vanadus-, vangla-, veerand-, vilja-, õnne-, õpi-, õpingu-, õpipoisi-, õppe-, õuna-, ülikooliaasta; uus-, vana-aasta.

• Programmi väljund:

marksona=aasta ;abieluaasta ,algaasta ,algkooliaasta ,algusaasta ,ametiaasta ,aruandeaasta ,asutamisaasta ,eelarveaasta ,eluaasta ,finantsaasta ,heina-,ikaldusaasta ,ilmumisaasta ,juubeliaasta ,kalendriaasta ,kasvuaasta ,keskkooliaasta ,kirikuaasta ,kooliaasta ,kriisiaasta ,kuuaasta ,lapseõlveaasta ,leina-,leseasta ,liigaasta ,lisapäeva-,maapaoaasta ,majandusaasta ,marja-,mõõna-,neiupõlveaasta ,noorusaasta ,nälja-,okupatsiooniaasta ,olümpia-,orjusaasta ,pagendusaasta ,pensioniaasta ,poolaasta ,prooviaasta ,põllumajandusaasta ,põua-,päikeseaasta ,raamatuasta ,rahandusaasta ,rahuasta ,revolutsioonista ,rännuaasta ,saagiaasta ,surma-,sõja-,sünniaasta ,teenistusaasta ,tegevusaasta ,tõusuaasta ,täheaasta ,tööaasta ,valgusaasta ,valitsemisaasta ,vanadusaasta ,vangla-,veerandaasta ,vilja-,õnneaasta ,õpiaasta ,õpinguaasta ,õpipoisiasta ,õppeaasta ,õuna-,uusaasta ,vanaaasta;

marksona=aasta-aastalt ,aastaaeg ,aastaruanne ,aastarv ,aastabibliograafia ,aastaeelarve ,aastahinne ,aastainventuur ;

4.

• Sõnaraamatu algfail:

<art> <mrks>abi </mrks><grm><tyyp>11</tyyp> </grm><sl>s. </sl><nr>1.</nr> <sel>aitamine, abistamine, abistus; see, mis aitab. <lyh>vt ka</lyh><ptr>appi</ptr>. </sel><naide>Vastastikune abi. Sõbralik, seltsimehelik, omakasupüüdmatu, ootamatu, kiire abi. Materiaalne, rahaline, majanduslik, sõjaline abi. Vältimatu meditsiiniline abi. Tehnilise abi auto. Abi arengumaadele, loodusõnnetuste piirkondadele. Kelleltki abi paluma, otsima, saama. Kellelegi abi osutama, andma. Hädas vajatakse abi. Õpilane vajab abi matemaatikas. Sõprade abi kulus marjaks ära. Hea nõu on kõige parem abi. Mitte kusagilt polnud abi loota. Haavatud oigasid ja palusid abi. Kas teil abi ka tarvis läheb? Abi jäi hiljaks. Naabrite abist ta keeldus. Poiss võttis küüned abiks. Ei maksa teiste abile lootma jääda. Ta tuleb toime ilma teie abita. Naabrite abiga said kartulid võetud.

</naide><txh_varjund>•</txh_varjund> <mrks>abil </mrks><sel>(postpositsioonile lähedasena:) millegi, kellegi abiga, midagi kasutades, abiks, appi võttes, millegi vahendusel, millegagi. </sel><naide>Juurte abil

hangib taim mullast toitaineid. Väentaimesid kinnituvad väentaimes abil. See on katsete abil kindlaks tehtud. Sugestiooni abil saab inimest mõjustada. Inglisekeelset kirjandust loeb ta sõnaraamatu abil. Vinnasin end käte abil purdele. </naide><l>Ls. arsti-, esma-, finants-, häda-, kaas-, kiir-, oma-, sünnitus-, välis-, õigusabi. </l><nr>2.</nr> <sel>[millestki, kellestki] kergendus, kasu. </sel><naide>Ravimist, rohust, süstidest oli varsti abi märgata. Kaitsekiivrist, turvavööst on kindlasti abi. Ettevaatusest oli abi. Sellest on vähe abi. Lapsed alles väikesed, paljude neist abi on. Minust pole neile abi, ma ei tunne ümbrust ega inimesi. Poisist on isale juba suur abi. Arvasin, et kompress võiks haigele jalale abi tuua. Kui ei ole surmatõbi, siis saab ikka viinast abi. Vs. </naide><nr>3.</nr> <sel>abiline (tihti elukutse, ameti poolest); abistaja(d), aitaja(d); vrd. asetäitja. </sel><naide>Ministri, direktori, metsaülevaataja, kapteni, vedurijuhi, laojuhataja abi. Vallavanema, kirjatöötaja abi. USA riigisekretäri abi Lähis-Ida küsimustes. Remondiga ei saa ma üksi hakkama, tuleb abi võtta. Ta ei jõua seda üksi ära teha, keegi peaks talle abiks minema. Tütar on emale köögis abiks. Ta püüdis kogu aeg nõu ja jõuga abiks olla. Juhtival töötajal on sageli asetäitja(d) või abi(d). *Sauna Mari käis peres ühtepuhku abiks, sest perenaise tervis oli selle lapsega vilets. </naide><o>A. H. Tammsaare. </o><l>Ls. eri-, meistri-, noorem-, vanemabi. Ka: direktori-, kapteni- jne. abi. </l> </art>

- HTML kuju:

abi <l1> s.

1. *aitamine, abistamine, abistus; see, mis aitab. vt ka appi.* Vastastikune abi. Sõbralik, seltsimehelik, omakasupüüdmatu, ootamatu, kiire abi. Materiaalne, rahaline, majanduslik, sõjaline abi. Vältimatu meditsiiniline abi. Tehnilise abi auto. Abi arengumaadele, loodusõnnetuste piirkondadele. Kelleltki abi paluma, otsima, saada. Kellelegi abi osutama, andma. Hädas vajatakse abi. Õpilane vajab abi matemaatikas. Sõprade abi kulus marjaks ära. Hea nõu on kõige parem abi. Mitte kusagilt polnud abi loota. Haavatud oigasid ja palusid abi. Kas teil abi ka tarvis läheb? Abi jäi hiljaks. Naabrite abist ta keeldus. Poiss võttis küüned abiks. Ei maksa teiste abile lootma jääda. Ta tuleb toime ilma teie abita. Naabrite abiga said kartulid võetud.

- **abil** (*postpositioonile lähedasena:*) millegi, kellegi abiga, midagi kasutades, abiks, appi võttes, millegi vahendusel, millegagi. Juurte abil hangib taim mullast toitaineid. Väentaimesid kinnituvad väentaimes abil. See on katsete abil kindlaks tehtud. Sugestiooni abil saab inimest mõjustada. Inglisekeelset kirjandust loeb ta sõnaraamatu abil. Vinnasin end käte abil purdele.

Liitsõnad: arsti-, esma-, finants-, häda-, kaas-, kiir-, oma-, sünnitus-, välis-, õigusabi.

2. *[millestki, kellestki] kergendus, kasu.* Ravimist, rohust, süstidest oli varsti abi märgata. Kaitsekiivrist, turvavööst on kindlasti abi. Ettevaatusest oli abi. Sellest on vähe abi. Lapsed alles väikesed, paljude neist abi on. Minust pole neile abi, ma ei tunne ümbrust ega inimesi. Poisist on isale juba suur abi. Arvasin, et kompress võiks haigele jalale abi tuua. Kui ei ole surmatõbi, siis saab ikka viinast abi. (*Vanasõna*)

3. *abiline (tihti elukutse, ameti poolest); abistaja(d), aitaja(d); vrd. asetäitja.* Ministri, direktori, metsaülevaataja, kapteni, vedurijuhi, laojuhataja abi. Vallavanema, kirjatöötaja abi. USA riigisekretäri abi Lähis-Ida küsimustes. Remondiga ei saa ma üksi hakkama, tuleb abi võtta. Ta ei jõua seda üksi ära teha, keegi peaks talle abiks minema. Tütar on emale köögis abiks. Ta püüdis kogu aeg nõu ja jõuga abiks olla. Juhtival töötajal on sageli asetäitja(d) või abi(d). Sauna Mari käis peres ühtepuhku abiks, sest perenaise tervis oli selle lapsega vilets. (*A. H. Tammsaare*)

Liitsõnad: eri-, meistri-, noorem-, vanemabi. Ka: direktori-, kapteni- jne. abi.

- Programmi väljund:

marksona=abi ,abil ;arstiabi ,esmaabi ,finantsabi ,hädaabi ,kaasabi ,kiirabi ,omaabi ,sünnitusabi ,välisabi ,eriabi ,meistriabi ,nooremabi ,direktoriabi ,kapteniabi ;

5.

- Sõnaraamatu algfail:

<art> <mrks>tsikaad</mrks><grm>-i <tyyp>22</tyyp> </grm><sl>s. </sl><term>zool. </term><sel>siristiga putukas (eriti troopikas ja subtropikas). </sel><naide>Tsikaadide vali sirin. Õlipuudel siristasid tsikaadid. </naide><l>Ls. liht-, mägitsikaad. </l> </art>

- HTML kuju:

tsikaad <-i 22> s. zool. siristiga putukas (eriti troopikas ja subtroopikas). Tsikaadide vali sirin. Õlipuudel siristasid tsikaadid.

Liitsõnad: liht-, mägitsikaad.

- Programmi väljund:

marksona=tsikaad;lihttsikaad;

6.

- Sõnaraamatu algfail:

```
<art> <mrks>aer</mrks><grm>-u <tyyp>22</tyyp> </grm><sl>s. </sl><sel>paadi-, varem ka
laevasõudevahend, eeskätt tulli külge kinnitav paariline sõudevahend. </sel><naide>Pikad, rasked aerud.
Kahe, nelja aeruga paat. Süstal on kahelabaline, kanuul ühelabaline aer. Kalur sõudis nii, et aerud raksusid. Üks
meestest asus aerudele, teine tüüris. Suruge aerudele! Tugevad käed vedasid, kiskusid aere. Rand on lähedal,
tõmba aerud paati. *Juhkam tõmbab aerudega nii ägedasti, et neiule vett näkku läigib. </naide><o>J. Lintrop.
</o><naide>Kakskümmend meest oli aerudel, kümme kummalgi pool. </naide><o>F. Tuglas. </o><l>Ls.
kanuu-, paaris-, pära-, rooli-, süsta-, üksikaer. </l> </art>
```

- HTML kuju:

aer <-u 22> s. paadi-, varem ka laevasõudevahend, eeskätt tulli külge kinnitav paariline sõudevahend. Pikad, rasked aerud. Kahe, nelja aeruga paat. Süstal on kahelabaline, kanuul ühelabaline aer. Kalur sõudis nii, et aerud raksusid. Üks meestest asus aerudele, teine tüüris. Suruge aerudele! Tugevad käed vedasid, kiskusid aere. Rand on lähedal, tõmba aerud paati. Juhkam tõmbab aerudega nii ägedasti, et neiule vett näkku läigib. (J. Lintrop) Kakskümmend meest oli aerudel, kümme kummalgi pool. (F. Tuglas)

Liitsõnad: kanuu-, paaris-, pära-, rooli-, süsta-, üksikaer.

- Programmi väljund:

marksona=aer;kanuu-aer,paaris-aer,pära-aer,rooli-aer,süsta-aer;

7.

- Sõnaraamatu algfail:

```
<art> <mrks>agentuur</mrks><grm>-i <tyyp>22</tyyp> </grm><sl>s. </sl><nr>1.</nr> <sel>(luure)agentide
võrk, nende organisatsioon. </sel><naide>Imperialistlikud riigid loovad oma agentuure teistes riikides.
Menševikud tegutsesid töölisliikumises kodanluse agentuurina. </naide><kasutusjuht>◦</kasutusjuht>
<stiil>piltl. </stiil><naide>Agentuur töötab hästi: ülemusel on varsti teada, mis alluvad temast räägivad.
</naide><l>Ls. luureagentuur. </l><nr>2.</nr> <sel>mõne asutuse kohalik esindus, osakond vms.
</sel><naide>Kindlustusseltsil oli agentuure paljudes linnades. </naide><l>Ls. posti-, sideagentuur.
</l><nr>3.</nr> <sel>ajakirjandusele uudiseid, informatsiooni hankiv organisatsioon. </sel><naide>TASS,
Reuter, UPI jt. suured agentuurid. </naide><l>Ls. informatsiooni-, teadete-, telegraafiagentuur. </l> </art>
```

- HTML kuju:

agentuur <-i 22> s.

1. (luure)agentide võrk, nende organisatsioon. Imperialistlikud riigid loovad oma agentuure teistes riikides. Menševikud tegutsesid töölisliikumises kodanluse agentuurina.

◦ piltl. Agentuur töötab hästi: ülemusel on varsti teada, mis alluvad temast räägivad.

Liitsõnad: luureagentuur.

2. mõne asutuse kohalik esindus, osakond vms. Kindlustusseltsil oli agentuure paljudes linnades.

Liitsõnad: posti-, sideagentuur.

3. *ajakirjandusele uudiseid, informatsiooni hankiv organisatsioon.* TASS, Reuter, UPI jt. suured agentuurid.
Liitsõnad: informatsiooni-, teadete-, telegraafiagentuur.

- Programmi väljund:

marksona=agentuur;postiaagentuur,informatsiooniagentuur,teadeteagentuur;

8.

- Sõnaraamatu algfail:

```
<art> <mrks>ahel</mrks><grm>-a <tyyp>2</tyyp> </grm><sl>s. </sl><nr>1.</nr> <sel>hrv.
</sel><vormikood>pl. </vormikood><sel>kett (näit. endisajal vangide) köitmiseks, millegi kinnitamiseks.
</sel><naide>Ahelate kõlin. Ahelaisse panema, neetima. Ahelais orjad. Vangid vabastati ahelaist. Tulekolde
kohale kookude või ahelate otsa riputati keedupajad. Ta on nagu ahelatega kodu külge needitud. *Kõlistas [koer]
korra ahelaga ja vigises jälle. </naide><o>A. Maripuu. </o><txh_varjund>•</txh_varjund> <stiil>piltl.
</stiil><sel>(millegi kammitseva, köitva, takistava v. ikestava, rõhuva kohta). </sel><naide>Kapitalismi,
pärisorjuse ahelad. Kolonialismi ahelaist vabanenud arengumaad. Revolutsioon vabastas ahelaist rõhutatud rahvad.
*Kui kesktalvel külm õige pakaseks läheb, raputab vangistatud Peipsi oma ahelaid. </naide><o>Juh. Liiv.
</o><l>Ls. jala-, kaela-, käe-, paja-, raud-, vangiahel(ad); abielu-, jää-, orjaahel(ad). </l><nr>2.</nr>
<sel>üksteisele järgnevate esemete, nähtuste, faktide, sündmuste, ka inimeste jne. pidev rida.
</sel><naide>Jõekallast palistav hõbepajude ahel. Rahvatantsijate värvikad ahelad staadionimurul. Terve ahel
kauneid järvi. Paljudes kõrgmolekulaarsetes ainetes paiknevad süsinikuaatomid ahelana. Assotsiatsioonide,
järelduste ahel. Sündmuste ahel. Helgemad hetked elupäevade hallis ahelas. See oli vaid üks lüli mu tööde ja
ettevõtmiste pikas ahelas. </naide><l>Ls. arengu-, elektri-, koralli-, kõrval-, pea-, süsiniku-, toitumisahel. </l>
</art>
```

- HTML kujut:

ahel <-a 2> s.

1. hrv. <pl. > *kett (näit. endisajal vangide) köitmiseks, millegi kinnitamiseks.* Ahelate kõlin. Ahelaisse panema, neetima. Ahelais orjad. Vangid vabastati ahelaist. Tulekolde kohale kookude või ahelate otsa riputati keedupajad. Ta on nagu ahelatega kodu külge needitud. Kõlistas [koer] korra ahelaga ja vigises jälle. (A. Maripuu)

• *piltl. (millegi kammitseva, köitva, takistava v. ikestava, rõhuva kohta).* Kapitalismi, pärisorjuse ahelad. Kolonialismi ahelaist vabanenud arengumaad. Revolutsioon vabastas ahelaist rõhutatud rahvad. Kui kesktalvel külm õige pakaseks läheb, raputab vangistatud Peipsi oma ahelaid. (Juh. Liiv)

Liitsõnad: jala-, kaela-, käe-, paja-, raud-, vangiahel(ad); abielu-, jää-, orjaahel(ad).

2. *üksteisele järgnevate esemete, nähtuste, faktide, sündmuste, ka inimeste jne. pidev rida.* Jõekallast palistav hõbepajude ahel. Rahvatantsijate värvikad ahelad staadionimurul. Terve ahel kauneid järvi. Paljudes kõrgmolekulaarsetes ainetes paiknevad süsinikuaatomid ahelana. Assotsiatsioonide, järelduste ahel. Sündmuste ahel. Helgemad hetked elupäevade hallis ahelas. See oli vaid üks lüli mu tööde ja ettevõtmiste pikas ahelas.

Liitsõnad: arengu-, elektri-, koralli-, kõrval-, pea-, süsiniku-, toitumisahel.

- Programmi väljund:

marksona=ahel;jalaahel,kaelaahel,käeahel,pajaahel,raudahel,abieluahel,jääahel,arenguahel,elektriahel,koralliahel ,kõrvalahel,peaahel,süsinikuahel;

9.

- Sõnaraamatu algfail:

```
<art> <mrks>ahnus</mrks><grm>-e <tyyp>5</tyyp> </grm><sl>a. s. </sl><sel>&lt; ahne.
</sel><naide>Täitmatu ahnus. Liigkasuvõtja ahnusel pole piiri. Nälginu ahnusega asus ta toidu kallale.
</naide><l>Ls. au-, kasu-, raha-, saagi-, võimuahnus. </l> </art>
```


- HTML kuju:

ahnus <-e 5> a. s. < *ahne*. Täitmatu ahnus. Liigkasuvõtja ahnusel pole piiri. Nälginu ahnusega asus ta toidu kallale.

Liitsõnad: au-, kasu-, raha-, saagi-, võimuahnus.

- Programmi väljund:

marksona=ahnus;auahnus,kasuahnus,rahaahnus,saagiahnus;

10.

- Sõnaraamatu algfail:

<art> <mrks>tentriline</mrks><grm>-se <tyyp>5</tyyp> </grm><sl>adj. </sl><sel>-keskene; keskmesse puutuv v. rakendatud. </sel><l>Ls. ego-, geo-, helio-, selenotsentriline. </l> </art>

- HTML kuju:

tentriline <-se 5> adj. *-keskene; keskmesse puutuv v. rakendatud.*

Liitsõnad: ego-, geo-, helio-, selenotsentriline.

- Programmi väljund:

marksona=tentriline;egotsentriline,geotsentriline,heliotsentriline;

Lisa 2

Programm liitsõnade täiskuju leidmiseks

```
#Programm ÕS-i liitsõnade täiskuju leidmiseks
#Autor: Maarja-Liisa Natus

use Encode;
use charnames ':full';
open (SISENDFAIL, "<:utf8","qs9.orig");
Srida = <SISENDFAIL>;
open (SONAFAIL, ">:utf8","valjund.orig");
while (<SISENDFAIL>) {
    $marksonad="";
    $rea_alguse="";
    @marks={};
    $ridal=$rida;
    $ridal=~m/nr\="(.*)"><hd/gi;
    $number=$1;
    print $number;
    $marksonad=lc &PR_leia_marksonad($ridal);
    @marks=split(/;/,$marksonad);
    $marksonad = $number.'.'. "marksona = ";
    print $marksona;
    for $i (@marks) {$i=&PR_tootle_marksona($i);$i=&PR_tootle_naide($i); $abi=$i;$marksonad=$marksonad.$i.".";};
    $marksonad = substr($marksonad,0,(length $marksonad)-1).";"; # viimase ";" võtan ära
    while ($rida =~</naide>/)
    {
        $pos_alguse = index($rida,<naide>); #mitmendal positsioonil <
        $pos_lopu = index($rida,</naide>); #mitmendal positsioonil loputagi <
        $rea_alguse=lc substr($rida, $pos_alguse+7, $pos_lopu-($pos_alguse+7));#pärast <naide> tagi, mis tuleb
        $rea_alguse=&PR_tootle_naide($rea_alguse); # töötlen stringi
        @naited=split(/[,; ]/,$rea_alguse);
        for $i (@marks){
            for $j (@naited) {
                while ($j=~m/^(i){$j=substr($j,1);} # võtan tühikud eest ära
                    $skaldu=index($j,'/');
                if ($skaldu >=0 and length $j>0){
                    if (index($j,$i)<0) { # Märksõna pole näite liikmes
                        $pos2=index($j,substr($i,0,3)); # vaatan, kas on sarnast stringi !NB Vea annab suurte, väikeste tähtede korral
                        if ($pos2>=0 and index($j, '(')=-1){
                            if ($skaldu>$pos2) # vaatan, kas see string asub enne või pärast / kriipsu
                            {
                                $j=substr($j,0,$skaldu).substr($j,$skaldu+1) and $i=substr($j,0,$skaldu);
                            }
                        }
                        else
                        {
                            $i=substr($j,$skaldu+1,length $j);
                            $j=substr($j,0,$skaldu).substr($j,$skaldu+1);}
                    }
                }
            }
        }
        else {
            $pos1=index($j,'"');
            if ($pos1>=0){
                if ($pos1==0) # kas / asub alguses?
                {
                    $j=$i.substr($j,1);}
            }
            else
            {
                if ((length $j)-1==$skaldu) # kas / asub lõpus?
                {
                    $i=$abi;
                    $j=substr($j,0,$skaldu).$i;
                }
            }
            else
            {
                $nelopp=index($j, 'ne' );
                $kusne=substr($j,$nelopp, (length $j-$nelopp+2));
                if($kusne=='ne' and $skaldu>$pos2){
                    $i=substr($j,0,$skaldu);}
                $j=substr($j,0,$skaldu).substr($j,$skaldu+1);
            }
        }
    }
}
else
{
```

```

        $j=substr($j,0,$skaldu).$i; # võtan / välja
    }
}

else
{
if ((($skaldu>0) and (substr($j,0,$skaldu) ne $i or substr($j,$skaldu+1) ne $i))
{
if (index(substr($j,$skaldu+1),$i)<0)
{
if (substr($j,0,$skaldu) ne $i and index($j,'(')<0 and index(substr($rea_alguse,index($rea_alguse,$j)+$skaldu+1),')')>=0)
{
$si=substr($j,0,$skaldu);
}
$si=substr($j,0,$skaldu);
$j=substr($j,0,$skaldu).substr($j,$skaldu+1); # võtan / välja
}
}
else
{
if(index($j, '(')==-1){
$si=substr($j,$skaldu+1,length $j);
$j=substr($j,0,$skaldu).$i;
}}
}
}
else
{
$j=substr($j,0,$skaldu).substr($j,$skaldu+1); # võtan / välja
}
}
}
$marksonad = $marksonad.$j.' ';
}
}
}
}

$rida=substr($rida, $pos_lopu+8);
}

$marksonad =substr($marksonad,0,(length $marksonad)-1)."; ";
print SONAFAIL $marksonad,"\n";# print SONAFAIL $sona;
$rida = <SISENDFAIL>;
}
close(SISENDFAIL);close(SONAFAIL);
sub PR_leia_marksonad {
if($_[0]=~m/(<mrks>[\.a-zžöäõü\W\-\_\.]*</mrks>)/gi) { #mall
$sona1=$1;}
if($_[0]=~m/(<mrks_z>[\.a-zžöäõü\W\-\_\.]*</mrks_z>)/gi) { #mall
$sona1=$1;
$sona1=~s/<mrks_z>([\.a-zžöäõü\W\-\_\.]*</mrks_z>/<mrks>$1</mrks>)/gi;
}
while ($sona1 =~/<mrks>/)
{
$pos_lopp = index($sona1,'</mrks>');
$rea_lopp=substr($sona1,$pos_lopp+7);
$rida3=substr($rea_lopp,$pos_lopp+9);#läheme 2 sümbolit edasi
$sona1=~m/(<mrks>[\.a-zžöäõü\W\-\_\.]*</mrks>)/i;
$rea_algus=$1; #kõik märksõna tagidest kuni marksona tagide lõpuni
$pos_alg = index($rea_algus,<mrks>);
#$alg_indeks=index($sona1, $rea_algus);
$pos_lopp = index($rea_algus,</mrks>');
if (length $rea_algus > 0){
#marksona tagide vahelt sõna <mrks> "sona" </mrks>
$marksonad = $marksonad.substr($rea_algus,$pos_alg+6,$pos_lopp-($pos_alg+6)).';';
}
$sona1 = $rea_lopp ;
}
} return $marksonad;
}
sub PR_tootle_naide { # 1. korjan liigsed tagid
$mr_rida="";
$mr_rida=$_[0];
while (index($mr_rida,'<')>=0)
{# 1. lühendid välja
if (index($mr_rida,'<lyh>')>=0)
{ $mr_rida=substr($mr_rida,0,index($mr_rida,'<lyh>')).substr($mr_rida,index($mr_rida,'</lyh>')+6); }
else { $mr_rida=substr($mr_rida,0,index($mr_rida,'<')).substr($mr_rida,index($mr_rida,'>')+1);}
}
return $mr_rida;
}
sub PR_tootle_marksona {

```

```

$m_riida=""
$m_riida=$_[0];
while ($m_riida =~ m/^\./) {$m_riida=substr($m_riida,1);} # võtan / eest ära
while ($m_riida =~ m/^\(\.\)/) {$m_riida=substr($m_riida,3);} # võtan (.) eest ära
while ($m_riida =~ m/^\./) {$m_riida=substr($m_riida,1);} # võtan . eest ära
while ($m_riida =~ m/\$/) {$m_riida=substr($m_riida,0,(length $m_riida) -1);} # võtan _ tagant ära
if (index($m_riida, '/') >= 0) {$m_riida=substr($m_riida,0,index($m_riida, '/')).substr($m_riida,index($m_riida, '/')+1);} # võtan / keskelt ära NB!
if (index($m_riida, '.') >= 0) {$m_riida=substr($m_riida,0,index($m_riida, '.')).substr($m_riida,index($m_riida, '.')+1);} # võtan . keskelt ära NB!
if (index($m_riida, '"') >= 0) {$m_riida=substr($m_riida,0,index($m_riida, '"')).substr($m_riida,index($m_riida, '"')+1);} # võtan " keskelt ära
if (index($m_riida, "□") >= 0) {$m_riida=substr($m_riida,0,index($m_riida, "□")).substr($m_riida,index($m_riida, "□")+1);} # võtan . keskelt ära
return $m_riida; }

```