

EESTI KEELE SÕNAVARA AFEKTIKALDUVUS JA MIS SELLEGA PEALE HAKATA

ENE VAINIK

Arvatavasti ei kerkiks sõnavara afektikalduvus eraldi jututeemana kunagi üles, kui me suhtleksime ainult oma emakeeles ja ainult lähedaste inimestega, sest taoline suletud minikommuun toimiks jagatud emotsioonide ja tähenduste tõttu niigi sujuvalt. On iseenesest mõistetav, et inimesena me suudame aru saada nii meile öeldud lause sisust kui ka sellest, millises meeleolus on lause autor. Asjade loomuliku kulu korral on keeleliste sõnumitega kaasnevate tundetoonide tabamine osa keelilisest ja suhtluskompetentsist, mille emakeele valdaja omandab juba lapsena vaikimisi koos keeleoskuse ja üldise sotsialiseerumisega. Eri kultuure uurides on leitud, et afekti verbaalne väljendamine hakkab peale üksiksõna tasandist ja kõik grammatilised vormid positiivse ja negatiivse afekti väljendamiseks omandatakse juba enne viiendat eluaastat (Ochs 1988).

Me ei ela aga ammu enam hõimuühiskonnas ning suhtlemine üle keele- ja kultuuripiiride toob kaasa ka vajaduse arvestada sõnavara emotsioone mõjutava mõõtmega. On valdkondi, kus sõnavara afektikalduvus on tavalisest aktuaalsem ja selle teadvustamisele on vaja kaasa aidata. Lisaks spetsiifilistele ja eriotstarbelistele emotsioonide üle reflekteerimise kontekstidele, nagu seda on näiteks psühhoteraapia seansid, mispuhul rikka ja nüansseeritud emotsioonisõnavara valdamine tuleb kasuks nii terapeutidele kui patsiendile, puutume igapäevaselt kokku ka argise sõnavara võimega mõjuda meie emotsioonidele. Võtkem kas või arutelud selle üle, kas sõnad *neeger*, *mustlane* ja *okupant* on eesti keeles solvava tähendusega või mitte.¹ Vajadusele osutada sõnadega kaasaskäivatele tundetoonidele on reageerinud juba ka eesti leksikograafid, kes koostavad eesti keelt teise keelena õppijale suunatud nn põhisõnavara sõnastikku (vt Kallas, Tuulik 2011; Vainik 2012).

Tänapäeva maailmas aga on meil vaja suhelda veelgi laiemalt – mitte ainult inimestega võõrastest kultuuridest ja võõrkeeltes, vaid ka arvuti ja muude tehnoloogiliste seadmetega, mispuhul meil on samuti kasutusmugavusest tingitud ootus, et need seadmed kõneleksid meiega võimalikult ühte moodi, s.t meie emakeeles, ja et nad teeksid seda võimalikult loomuliku kõne sarnaselt. Kuna inimese loomulik kõne on enamasti ilmestatud kõneleja emotsioonide, suhtumiste ja hoiakutega, siis ootame ka arvutilt, et ta ei annaks edasi mitte ainult sõnumite sisu, vaid teeks seda emotsionaalselt usutaval viisil. Sellest ootusest johtuvalt on tarvis pöörata ka tavalisest suuremat tähelepanu eesti keele sõnavara potentsiaalsele afektikallakule – selle taga seisab praktiline vajadus muuta eestikeelne sünteeskõne loomulikumaks ja seega paremini jälgitavamaks (Altrov, Pajupuu 2008).

¹ <http://arvamus.postimees.ee/833486/kari-kasper-neeger-ja-mustlane-meedias/> (8. V 2012) ja <http://www.delfi.ee/news/paevauudised/eesti/kolvart-sona-okupatsioon-kasutamise-solvab-siinset-vene-kogukonda.d?id=64363615> (9. V 2012).

Emotsioonide väljendumise uurimine eestikeelses kõnes ja uurimistulemustest lähtuv modelleerimine on Eesti Keele Instituudis praegu mitmete inimeste tööks,² sest probleem on mitmetahuline ja ülesanne koosneb mitmest etapist. Kui näiteks oletada, et meil juba on olemas süntesaator, mis oskab lisaks neutraalsele hääletoonile produtseerida rõõmsat, vihast ja kurba hääletooni, siis sellest üksi ei piisa. Kui etteloletav tekst ei ole emotsioonide suhtes kuidagi eelmärgendatud, siis läheb kindlasti vaja mingit sensorit,³ mis kirjalikus vormis esitatud teksti põhjal esmalt tuvastaks, millist registrit neist oleks igal üksikul juhul tarvis rakendada. Selline sensor peaks sisendiks oleva kirjaliku teksti läbi analüüsima ning etteantud ühiku kaupa otsustama, kas tegu on neutraalse või emotsionaalse varjundiga järjendiga, ning emotsionaalse järjendi puhul otsustama, kas see liigitub pigem rõõmu, kurbuse, viha või mingi muu emotsiooni alla. Eestikeelse kõne süntees toimub lause kaupa (Mihkla 2007) ja seetõttu on ka sensori tööühikuks vaikimisi lause.

Laiema taustana tuleks arvesse võtta, et emotsioonidel ei ole kõnes mingit neile ainuomast kindlat väljendusviisi – jälgi emotsioonidest võib leida pea kõigil nn traditsioonilistel keeletasanditel: foneetikas, leksikas, süntaksis, tuletuses, morfoloogias jne (vt Foolen 1997; Wilce 2009). Inimene lähtub seega kõnes väljenduva emotsiooni äratundmisel väga erinevat laadi infokildudest ja paneb neist kokku oletusliku tervikpildi, mis edasises suhtluses saab kas kinnitust või tühistatakse uute, teist laadi infokildude poolt. Kõne mõistmine ja kirjaliku teksti lugemine selle mõttest ja tundetoonist aru saades on dünaamiline ja eri tegureid arvestav aktiivne protsess, mingil juhul ei kujuta see endast vaid kõneleja poolt teele saadetud „postipaki” passiivset avamist.⁴ Inimene tugineb nii oma maailmatundmisele kui ka spetsiifilisematele teadmistele suhtlusprotsessist, selle osalistest ja nende emotsioonidest ja vahekordadest ning suhtluse dünaamikast, lisaks läheb tarvis empaatiat, sotsiaalset kompetentsi ning järeldusvõimet, sest mitte alati ei ole emotsioon sõnaselgelt väljendatud. Need inimlikud teadmised ja võimed, omandatud keele kasutamise ja üldise sotsialiseerumise protsessis, pole loomult mitte deklaratiivset, vaid protseduurilist laadi. Arvutile vastuvõetavad juhised kirjutatakse aga programmidenä, kus kõik vajalik täht-tähelt kirjas. Järelikult tuleb esmalt muuta meie endi protseduurilised teadmised deklaratiivseks, tuua nad keelekäitumusliku alateadvuse aladelt tähelepanu teadlikku fookusesse.

Järgnevalt keskendutakse leksikaalsele tasandile ning arutletakse selle üle, millised sõnaklassid on seotud emotsioonide (ehk üldisemalt afekti) väljendamisega ja mille suhtes peaks seega tundlikud olema tekstist afektiivsuse märke otsiva sensori „retseptorid”. Tervikuna põhineb lähenemine optimistlikul eeldusel, et ka spetsiaalselt eelmärgendamata tavalises tekstis leidub siiski piisavalt informatsiooni teksti autori võimaliku emotsiooni kohta.

² Sellega tegeldakse SF projekti nr SF0050023s09 „Eesti keele alusuuringud keeletehnoloogiliste rakenduste teenistuses”, samuti keeletehnoloogia projekti „Kõne ja teksti emotsionaalsuse statistilised mudelid” raames. Hille Pajupuu kaasjuhendamisel on käsil kaks doktoritööd emotsioonidest: Rene Altrovi „Eesti emotsionaalse kõne korpuse loomine ja emotsioonide tajumine” ning Kairi Tamuri „Eesti põhiemotsioonide akustiline analüüs ja modelleerimine”.

³ Mujal (Vainik 2010) nimetatud ka automaatseks emotsioonituvastajaks (AET).

⁴ Lähemalt lingvistilise kommunikatsiooni olemuse mõistmist eksitavast „postipaki” metafoorist vt Lakoffi ja Johnsoni raamatust „Metafoorid, mille järgi me elame” (2011).

Lisaks eeldatakse, et emotsioonile viitav informatsioon on tekstis ka leksikaalselt kuidagi representeeritud ja et arvutit saab vastava programmi abil panna leksikaalseid afektimarkereid registreerima, analüüsima ning neist järeldusi tegema.

Fookus on sõnavaral

Tekstides sisalduva afektikalduvuse (hinnangu, ekspressiivsuse, emotsioonide) märkamise ja mõõtmise vastu tunnevad huvi erinevate valdkondade teadlased. Keeleteaduses on tekstide hinnangulisuse ja afektiga tegeldud peamiselt tekstilingvistika vallas (nt Biber, Finegan 1989; Hunston, Thompson 2000). Rakenduslikku suunda, mille raames tegeldakse meetodite väljatöötamisega, selleks et oleks võimalik automaatselt tuvastada tekstides sisalduvat hoiakut, hinnangut või emotsiooni, nimetatakse tundmusanalüüsiks (*sentiment analysis*). Tehisintellekti loomise paradigma raames tegeldakse sarnaste asjadega ka afektirehkenduse nime all (*affective computing*) (vt nt Picard 1997). Tekstide afektiivsuse mõõtmisega tegeldakse aga ka psühholoogia vallas, mispuhul näiteks analüüsitakse ravile tulnute kirjutisi ravi alguses ja lõpus, ning tuvastada on võimalik muutusi teatud sõnaklasside (pronoomenid, funktsioonisõnad) kasutuses. Järelikult ei ole nende sõnaklasside kasutamise muster iseloomulik mitte üksnes konkreetsele inimesele, vaid ka psühholoogilisele seisundile (Pennebaker 2011). Psühhodiagnostilise otstarbega teksti analüüsi programme on loodud mitmeid, näiteks Linguistic Inquiry and Word Count (LIWC)⁵, EMOTAIX, VAAL⁶.

Katsetes määrata teksti, lõigu, lause ja isegi osalause emotsiooni on üksmeelselt leitud, et taolise süsteemi osaks peab olema sõnade tüüpilisi emotsionaalseid lisatähendusi esitav sõnastik (ülevaadet vt Vainik 2010). Kuigi teoreetilised lähenemised pole üksmeelel, kas sõnade emotsionaalseid omadusi tuleks pidada pigem osaks keele (juba realiseerunud) struktuurist või osaks inimese mentaalsest struktuurist, kus see esineb alles potentsiaalina, ei ole see praktikuid seganud sõnadega kaasnevaid emotsionaalseid laenguid määramast ja kokku rehkendamast. Tekstist emotsiooni tuvastamise lahendustes kasutatakse peamiselt kas nn emotsionaalseid võtmesõnu (ingl *emotional keywords*), s.t emotsioonide otseseid nimetusi (nt *viha, rõõm, kadedus*) või nn leksikaalset külgetõmbavust (*lexical affinity*), mis teatud sõnadel, nagu *vähk* või *majanduskriis*, emotsioonide suhtes on.

Võtmesõnadele keskenduvad lähenemised lähtuvad eeldusest, et hoolimata emotsionaalsete tähenduste varieeruvusest ei kanna vähemalt teatud osa keele sõnavarast mitte üksnes konventsionaalseid tähendusi ja/või funktsioone, vaid edastab ka teatud konsensust informatsiooni emotsionaalse laengu osas. Psühholoogias ei kahelda, et ka üksiksõna nagu iga muugi stiimul võib esile kutsuda emotsionaalseid reaktsioone ja et neil reaktsioonidel võib olla mingil määral stabiilne ja kinnistunud iseloom. Selle, lähtekohalt biheivioristliku mõtteviisi klassikaks on C. E. Osgoodi, G. J. Suci ja P. H. Tannenbaumi ([1957] 1975) uurimus, milles esmakordselt kirjeldati nn semantilise diferent-

⁵ vt <http://www.liwc.net/> (24. V 2012).

⁶ EMOTAIX (Piolat, Bannour 2009) ja VAAL, vt <http://www.vaal.ru/proekt/vaal2000.php> (14. V 2012).

siaali meetodit ja selgitati välja semantilise ruumi kolm universaalset dimensiooni: *hinnang*, *tugevus* ja *aktiivsus*, mille suhtes paigutuvad väidetavalt niisugasti keele kõik sõnad kui ka mittesõnalised mõisted – nimed, pildid, narratiivid – ja üldse kõik, mida küsimustiku ja polaarsest adjektiividest moodustatud skaalade abil mõõdetakse. Kirjanduses arutletakse pigem selle üle, kuidas üksiksõnade iseloomulike valentside põhjal määrata lause ja kogu teksti valentsi (Bestgen 1994).

Mis puudutab eesti keelt, siis juba on loodud katselisi süsteeme, mis ka tekstilõigu emotsionaalse valentsi kokku rehkendavad (Pajupuu jt 2012). H. Pajupuu ja kolleegide loodud emotsioonidetektor⁷ põhineb eesti keele põhisõnavara sõnastiku (vt Kallas, Tuulik 2011) märksõnastikul, mille jaoks on katsete teel tehtud kindlaks nende tundetoon ehk emotsionaalne valents (Vainik 2012). Pajupuu jt väidavad, et tekstilõigu emotsiooni määramiseks ca 75 % kindlusega piisab väikesest, umbes 600 sagedasemast võtmesõnast koosnevast sõnastikust. Autorid nendivad, et tulemused on piisavalt head ning muude tasandite arvessevõtt peale sõnavara tundetoonide oleks ebaotstarbekalt vaevarikas (Pajupuu jt 2012: 184). Seega lõigu tasandil on põhjust juba edusammude üle rõõmu tunda. Kuid kuna süntesaator loeb teksti ette lausete kaupa, siis jätkuvalt on aktuaalne teema, kuidas teha kindlaks konkreetse lause ettelugemiseks sobiv emotsionaalne register.

Emotsiooni kindlakstegemisel üksiklauses tuleb kindlasti arvestada, et sõnade arv võib varieeruda väga suures ulatuses ning lühemates lausetes ei pruugi nn võtmesõnu ega ka emotsionaalset valentsi kandvaid sõnu üldse sisalduda. Lisaks sõltub lausete emotsionaalsuse määr tekstižanrist. Inglisekeelse tekstikorpuse (500 teksti) läbisõelumise tulemusel leiti, et ainult 3 % uuritud tekstidest sisaldasid märkimisväärselt afektile viitavat sõnavara (adverbid, adjektiivid ja verbid). Kõige enam viidati seejuures emotsioonidele isiklikes kirjades (Biber, Finegan 1989). Käesoleva artikli autori meelest ei saa seetõttu eeldada, et mis tahes eestikeelset tekstilauset selle žanrilisest taustast olenemata iseloomustaks leksikaalselt just võtmesõnadega markeeritud emotsionaalsus.

Problemaatiliseks võib seega osutuda asjaolu, et jooksva tekstilause sõnadest vaid osa on nn sisusõnad, mis üldse põhimõtteliselt saaksid olla kas võtmesõnadeks või oma valentsi poolest „emotsioone külgetõmbavateks” sõnadeks. Tuginedes Kerge ja Pajupuu (2010) statistikale eestikeelsete tekstižanride sõnaliigilisest koosseisust, võib väita, et suhteliselt neutraalse (aga siiski mitte kantseliitliku) teksti puhul, nagu seda on üleskirjutatud TV-monoloog (päevalehtede uudiste tutvustus), on 10 sõnast seitse „sisusõnad” (nimisõnad, verbid, adjektiivid) ja kolm funktsioonisõnad (deiktikud, adpositsioonid, hüüdsõnad jmt). Pisut ekspressiivsema tekstiliigi puhul, nagu seda on spontaanse raadiodialoogi üleskirjutus, on 4,5 sisusõna kohta 5,5 funktsioonisõna. Ameerika sotsiaalpsühholoog ja ekspressiivse kirjutamise pikaajaline uurija James Pennebaker (2011) viitab korduvatele tulemustele inglise keele kohta, kus see proportsioon on sama, s.t üle poole igapäevases jooksvas suhtluses kasutatavatest sõnadest on mitte sisu-, vaid funktsioonisõnad. Sealjuures inglise keeles on kohustuslik grammatiliste artiklite kasutamine, mis funktsioonisõnade suhtelist arvukust automaatselt tõstab. Teine oluline osa Pennebakeri ja tema töögrupi uurimistulemustest on aga avastus, et just

⁷ Igaüks võib seda katsetada: <http://peeter.eki.ee:5000/valence> (22. V 2012).

erinevate funktsioonisõnade (pronoomenite, prepositsioonide, kognitiivsete sidesõnade ja artiklite) kasutusmustrid reedavad teksti autori emotsionaalset seisundit kõige rohkem.

See avastus tundub paljutootavana, sest eesti keele puhul, mis on tüpoloogiliselt aglutineeriv-flekteeriv ja kus puuduvad grammatilised artiklid, on funktsioonisõnadel suhteliselt väiksem koormus süntaktiliste seoste osutajana (selleks on käänded, pöörded ja ühildumine) ja oletatavasti selle võrra suurem roll sõnumi interpreteerimise asjaoludele viitajatena. Viimast väidet ei tule muidugi võtta absoluutsena, sest on erandeid, näiteks postpositsioonid ja afiksaaladverbid, mis väljendavad mõistetevahelisi seoseid sõnumi sisu tasandil. Kuid oletada võib, et valdava tendentsina on nn mittetäistähenduslikud ehk funktsioonisõnad kõnesituatsiooni asjaolude (sh emotsionaalse suhtumise) väljendamise teenistuses.

Veel üheks problemaatiliseks asjaoluks on küsimus, kelle emotsiooni siis ikkagi tahetakse tekstilause põhjal määrata: kas lause autori emotsiooni/hinnangut (nt Biber, Finegan 1989; Hunston, Thompson 2000) või rahuldab ka nn mitteautori (ehk lause keskse tegelase) emotsiooni tuvastamine (nt Bednarek 2008)? Kuigi nii mõnedki emotsioonituvastuse süsteemid ennast selle eristusega ei vaeva, oleks eesti keele süntesaatori otstarvet silmas pidades ilmselt oluline aru saada, millal lause otseselt väljendab selle autori emotsiooni (ja järelikult tuleb rakendada mõnd spetsiifilist akustilist registrit) ja millal lause täiesti emotsioonivabalt kirjeldab või mainib kellegi kolmanda isiku – protagonist – emotsiooni (mispuhul eri registrit rakendada poleks tarvis). Näiteks selle artikli kirjutamise ajal vahendas uudisteportaal Delfi uudiseid massimõrvar Breiviki kohtuprotsessist, mispuhul kirjeldustes mainiti ära kurjategija iga naeratus, nutuvärelus ja muu emotsionaalne reaktsioon: „Prokuröri jutt ajas Breivikile naeru peale.” „Omatehtud propagandavideo pani Breiviki pisaraid valama.” „Breivik hõiskas rõõmust, kui Utøyal inimesi tappis.”⁸ Oleks äärmiselt kohatu, kui süntesaator vastavaid lauseid kas nuuksudes, väriseva või rõõmsa hääletooniga ette kannaks. Mõnikord on lause keskse tegelase emotsioonidest distantseerumine lausa hädatarvilik.

Seepärast arvab siinkirjutaja, et kui tahetakse tuvastada lausega kaasneva võivat emotsiooni, siis tuleb esmalt otsustada, kelle emotsiooni soovitakse tuvastada, ning vastusest lähtuvalt vaadata lauses sisalduvat sõnavara märksa täpsemalt. Tundmusanalüüsi tuleks kaasata nii emotsioonisõnavara (vt Vainik 2002a) ning positiivset või negatiivset valentsi kandev tundetooni-line üldsõnavara (vt Vainik 2012) kui ka erinevaid abistavaid rolle täitvad funktsioonisõnad.

Sõnavara afektikalduvus kui laiem mõiste

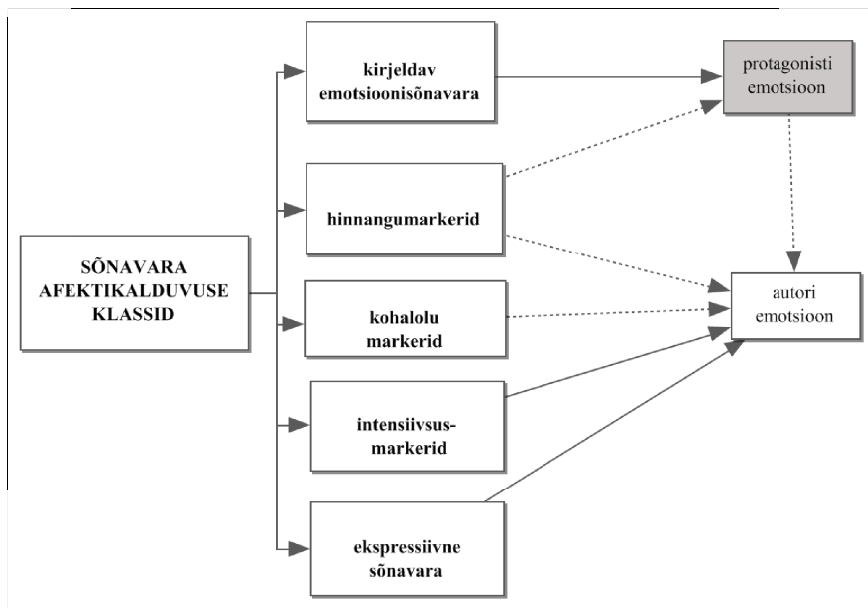
Emotsioonide leksikaalset väljendumist käsitlevas kirjanduses ollakse üksmeelel, et põhimõtteliselt jaguneb see kahte suurde kategooriasse vastavalt sellele, kas emotsioone esmajoones väljendatakse või kirjeldatakse (ülevaadet

⁸ <http://www.delfi.ee/news/paevauudised/valismaa/tunnistaja-breivik-hoiskas-roomust-kui-utoyal-inimesi-tappis.d?id=64374581> (14. V 2012), <http://www.delfi.ee/news/paevauudised/valismaa/omatehtud-propagandavideo-pani-breiviki-pisaraid-valama.d?id=64259757> (14. V 2012), <http://www.delfi.ee/news/paevauudised/valismaa/prokurori-jutt-ajas-breivikile-naeru-peale.d?id=64259641> (14. V 2012).

vt Bednarek 2008, eesti keeles vt Vainik 2002b). Näiteks hüüatused ja hüüdlused, samuti piltlik keelekasutus seostuvad kahtlemata keele ekspressiivse funktsiooniga, kujutades endast emotsiooni verbaalset „väljaelamist”, mispuhul semantiline sisu on tihti isegi teisejärguline. Seevastu otseste emotsiooni- ja tundenimetuste (nagu *rõõm*, *viha*, *kurbus*, *nõrdimus*, *kadedus* jne) kasutamine (eriti kellegi kolmanda emotsioonide kirjeldamiseks) kujutab endast põhimõtteliselt teistsugust tegevust ja seda seostatakse keele deskriptiivse funktsiooniga. Nagu eelnevast arutelust nähtus, ei tarvitse kitsalt emotsioonisõnavarale lootma jäämine teksti tundmusanalüüsis ennast õigustada, mistõttu tuleks luua sensor, mis on tundlik eri laadi sõnavaraüksuste suhtes, mis ühel või teisel viisil kas siis teksti autori või keske tegelase emotsioone võivad reeta.

Hõlmamaks võimalikult mitmekesisest sõnavara on siinkirjutaja esialgu tööterminina võtnud kasutusele leksikaalse afektikalduvuse (*lexical bias of affect*) mõiste (Vainik 2012). Tegemist on puhtalt funktsionaalse ja rakenduslikust otstarbest lähtuva kategooriaga, mis on koosseisult heterogeenne. Selle alla mahuvad eri viisid, kuidas sõna või väljend saab aidata kaasa emotsioonide (ja üldisemalt afektiivsuse, hinnangu, ekspressiivsuse) kirjeldamisele, väljendamisele, mõjutamisele ja tekstis tuvastamisele.

Esialgu tundub, et relevantne sõnavara on mõttekas jagada viide laiemasse afektikalduvuse klassi (vt joonist 1). Need afektikalduvuse klassid moodustavad kontinuumi, mille otsteks on eespool välja toodud kaks põhilist emotsioonide leksikaalse väljendamise viisi – deskriptiivne ning ekspressiivne sõnavara. Nende otspunktide vahele jäävad aga muud, keele üldsõnavarasse kuuluvad rühmad, mis samuti, iga rühm ise viisil, töötavad kaasa afekti keelelisele väljendamisele. Milline sõnavara täpsemalt kuulub nendesse leksi-



Joonis 1. Sõnavara afektikalduvuse klassid ja nende kasutatavus emotsioonide tuvastamisel.

kaalsetesse klassidesse, sellest alljärgnevalt lähemalt, kusjuures tutvustamist alustatakse siin kontinuumi keskosast, mis on võrdselt kaugel nii otsest deskriptiivsest kui ka ekspressiivsest äärmusest.

Kohalolumarkerid

Emotsiooni äratundmisest kirjalikus tekstilauses saab rääkida alles pärast seda, kui oleme ära tundnud, et emotsiooni kandja – inimene – on selles lauses tõepoolest „kohal” ning on osavõtlik lause sisu suhtes. Autori kohalolu markeerimine ei tähenda ilmtingimata, et lause oleks väga emotsionaalne, see on vaid eeltingimus. Kohalolumarkeriteks on deiktilised aja- ja kohamäärused (nt *nüüd, siin, praegu, täna, eile, homme*), 1. isiku personaalpronoomenid (*mina, ma, meie, me*), kognitiivsed ja modaalverbi konstruksioonid 1. isikuga seondult (*ma ei tea, minu arvates, pean, mul tuleb*), ebamäärase sisuga asesõnad (*niimoodi, selline, igasugune*). Need on väljendid, mille referendid on täpsemalt teada vaid asjaosalistele ja seega määratud ainult kõnesituatsiooni siseselt. Nende esinemine suurendab lausete kontekstisidusust ja kõneleja–kuulaja seose tihedust.

Omamoodi kohalolumarkeriks on ka nimisõnu laiendavate täiendite (omadussõnade) ning verbe ja omadussõnu laiendavate määruste (adverbide) kasutamine, sest igasugune kõnealuse situatsiooni osalisi või nende tegevusi iseloomustav sõnavara reedab lausungi autori ja tema teadmiste, arvamuste ning väärtushinnangute sekkumist, mille all on tihti hoomatav vaikimisi omaksvõetud „normatiivne” nähtuste esinemise laad või määr. Tähelepanu juhtimine vaikenormist ühes või teises suunas hälbimisele väljendab autori positsiooni. Näiteks kas või süütuna näivad omadussõnad *oranž* ja *ruuduline* lisavad situatsiooni vaatleja positsiooni, kes need omadused oma teadmistele, varasematele kogemustele ning ootustele tuginedes määratleb.

Hinnangumarkerid

Lausungi autori veelgi kindlamat kohalolu ja lisaks tema suhtumist kõnealustesse nähtustesse saab järeldada sellest, kui vastava situatsiooni iseloomustamiseks kasutatakse hinnangulist leksikat. Siinses käsitluses hinnangumarkeriteks nimetatav sõnavara klass langeb kokku kirjanduses kasutatava terminiga *leksikaalne külgetõmbavus* ja suure osaga sellest, mida eesti keeles on iseloomustatud terminiga *tundetoon* (vt Vainik 2012). Lisaks hinnangulistele omadussõnadele, mis moodustavad polaarseid paare (nt *hea–halb, tugev–nõrk, ilus–kole*), leidub hulganisti ka nimisõnu ja verbe, milles väärtushinnang on kas tähenduse osa (nt *sõber, vabadus, preemia* vs. *vaenlane, vangla, karistus*) või on sellega seotud konnotatiivselt (nt *aitama, koostöö, puhkus* vs. *segama, sõda, haigestuma*).

Eesti keele põhisõnavara märksõnastiku tundetoonide uurimine andis tulemuseks, et umbes kolmandik täistähenduslikest märksõnadest (nimisõnad, adjektiivid, verbid) omas mingisugust tundetooni. Katsete abil määrati positiivne, negatiivne ja vastuoluline tundetoon ning neid esines enam-vähem võrdses proportsioonis (Vainik 2012). Kuivõrd põhisõnavara sõnastiku on koondatud just sagedasemad ja harilikumad eesti keele sõnad, siis on tõenäoline, et mõni neist ka tavalisse tekstilausesse satub.

Kui lause sisaldab taolisi hinnangulisi nimisõnu, verbe ja omadussõnu, siis järeldub sellest, et lause autoril on olemas väärtushinnangutel põhinev

suhtumine. Hinnang ise ei ole veel emotsioon, küll aga kujutab endast emotsiooni komponenti (Lazarus 1991). Interpreteerija võib oma maailmatundmisele ja samastumisvõimele tuginedes järeldada hinnangust ka oletatava emotsiooni, mis on eeldatavasti samamärgiline hinnanguga (s.t positiivne või negatiivne). Enda emotsioonide väljendamine mingil teemal hinnanguliste sõnadega rääkides on tegevus, mida suhtlustreenerid jm psühholoogia praktikud nimetavad „siltide kleepimiseks”.

Intensiivsusemarkerid

Teist laadi afektikalduvust väljendab osavõtt kõnealusest situatsioonist sellisel viisil, et lisatakse kirjeldusse selle keskse tegevuse intensiivsust rõhutada või vähendavaid vahendeid. Seda, et autor on kohal ja rõhutab mingi omaduse määra, reedab näiteks adjektiivide võrdevormide (*suurim, värskem, kõige töökam*) ja suurendussõnade ehk nn augmentiivsete määrsõnade (nt *tohutult, kohutavalt, meeletult, hullult, metsikult*) kasutamine. Intensiivsust lisavad väidetele ka kategoorilisust väljendavad täiendid ja määruised (nt *kõik, alati, mitte kunagi, mitte ükski, ainus, mitte keegi, iga kord*). Leidub terve omaette sõnaklass, rõhumäärsõnad, mille liikmete ülesandeks on väljendada kõneleja suhtumist ja ilmestada kas kogu lauset tervikuna või tõsta esile selles mingit üksiksõna (Veski 1982: 31). Eesti keel on rõhumäärsõnade poolest rikas: *vaevalt, arvatavasti, iganes, ka, koguni, küll, muide, muidu, nimelt, nähtavasti, ometi, tõepoolest, õige, õigupoolest, aina, ainult, isegi, ju, just, lausa, suisa, tõepoolest, üksnes, hoopis*.

Intensiivsuse dimensioon võib põimuda ka hinnanguga, kuivõrd näiteks rõhumäärsõnad võivad olla otseselt jaatavad ning sõnumi sisu tõepärasust kinnitavad (*igatahes, jaa, jah, kindlasti, muidugi*) või eitavad (*ei, sugugi, mitte*). Teises, intensiivsust vähendavas suunas modifitseerivad sõnumit aga kahtlust või ebakindlust väljendavad rõhumäärsõnad ehk, *tõenäoliselt, vaevalt, vahest, vist, vististi, võib-olla*. Emotsionaalne hoiak võib olla eksplitsiitselt väljendatud ka tõenäosust või ihaldatavust edasi andvate kommunikatiivmodaalsete kiiludena: *kahjuks, õnneks, tõsijutt, ausõna, kas pole, eks ole, arvatavasti, tegelikult, samuti, nagu jne*.

Lisaks lausungi intensiivsust ühes või teises suunas modifitseerivatele elementidele on loomu poolest afektiivsed ka terve hulk keelelisi väljendusvahendeid, milles implitsiitselt põimuvad keele kirjeldav ja ekspressiivne funktsioon. Sellisel puhul on kõnealuse situatsiooni kirjeldamiseks käibel vahendid, mis juba iseenesest sisaldavad autori emotsionaalset laetust ja viiteid tema subjektiivsetele tajudele ja kujutlustele. Siia kuuluvad onomatopoeetiline sõnavara (nt *ludinal, plärtsti, mütakei, põmdi, hopsti, kõmm*), piltlikud väljendid (nt *pani käpa peale, kees vihast, suu kõrvuni, kadedusest roheline, sent surmale võlgu jne*), deminutiivid ehk vähendussõnad (*lapsukene, poisu, rumalukene, tibatilluke, pisitasa, väheldane, pikaldane jne*) ja augmentiivid ehk suurendussõnad (nt *jurakas, pirakas, kantsakas, ilmatu suur, superkallis, hirmkallis jne*). Need väljenditüübid on siin paigutatud intensiivsusemarkerite alla, sest ekspressiivsele kirjeldusele on iseloomulik, et sõnumi sisutasandit kas võimendatakse (hüperbool, absurd) või lisatakse sellele paralleelse teistest kogemusvaldkondadest, kasutades selleks jäljendamist, võrdlust ja metafoori. Kujundlikkus lisab sõnumile tõhusust (Schnall 2005).

Ekspressiivne sõnavara

Autori emotsionaalses kaasahaaratuses ei ole mingit kahtlust siis, kui kasutama hakatakse otseselt ekspressiivset sõnavara, millel muud funktsiooni peale „väljaelamise” ei olegi. Ekspressiivses sõnavaras on võimalik eristada mitmeid eri rühmi, olgu siin nimetatud mõningad neist. Näiteks kujutavad endast autori emotsiooni spontaanseid, kuid siiski konventsionaalseid väljendusi hüüdsõnad (nt *oh, ahaa, oi, ai, vau, noo*), sõnaklass, millel üldise arusaama kohaselt täpsemini määratletav semantiline sisu puudub. Pigem suulisele kõnele on iseloomulikud vande- ja kirumissõnad (nt *pagan, raisk, kurat*) ning sõnad, mis on otseselt suunatud kuulaja emotsioonide mõjutamisele. Sellised sõnad on söimusõnad (nt *värdjas, lits*) ja hellitusnimed (*maimuke, päikesekiir, musirull*), mida ei pruugi kirjalikus tekstis kuigi tihti esile tulla. Siiski, tekstides, milles vahendatakse otsest kõnet, võib neid kohata, samuti kui naljatlevat ja halvustavat leksikat, mis on samuti suunatud oma adressaadi emotsioonide mõjutamisele kas siis positiivses või negatiivses suunas. Halvustava, laitva või põlgliku tähendusega on pejoratiivsed sõnad,⁹ näiteks *möks, pläga, mögin, loru, luhva, lakard, juhmard, logard, õgard, kãpard, uimerdis, sahmerdis, kekutis, nãrukael, tattnina, kobakãpp, paksmagu*, naljatlev on näiteks *kurikael (kurjategija asemel)*.

Loomult samuti ekspressiivsed, kuigi intensiivsusest nõrgemad on rühm kinnistunud vormeleid, nagu *vabandage, aitãh, palun, palun vãga, võta heaks, jumal tãnatud, taevale tãnu* jne. Nendegi väljendite kasutamise motiiviks on kommunikatiivmodaalne eneseväljendus ja mitte semantiline sisu. Ent otseste kõne vahendamise puhul võib neid ka kirjalikus tekstis ette tulla.

Kirjeldav emotsioonisõnavara

Kõige selle kõrval, kuidas afekt tekstis implitsiitselt avalduda saab, on keeles olemas lõpuks ka spetsiaalsed sõnad afektiivsete seisundite (tunnete, emotsioonide, meeleolude jne) otseseks reflekteerimiseks. Need on ühest küljest keeles leiduvad arvukad emotsiooninimetused, nagu *viha, armastus, rõõm, kurbus, kadedus, raev, ahastus, ängistus, ärevus, iha, kirg, hirm, tusk, vaimustus, eufooria* jne, ja teisest küljest emotsioonide tavaliste väljendusliigutuste kirjeldused, nagu *nutt, naer, naeratus, suudlus, kallistus, värin, pisarad, aplaus, ovatsioonid, peksma, möllama, kãratama, söimama* jne. Emotsioonid ja nende väljendustegevused ei pruugi jutuaineks saada üksnes nimisõnaliselt väljendatuna (nt „Mul on *hirm*”), vaid ka verbide kujul (nt „Ma *kardan*”) või omadussõnaga viidatuna (nt „Ma olen *arg*”). Selles osas, mis puudutab emotsioonide väljendustegevuste kirjeldusi, läheb see sõnavararühm sujuvalt üle ekspressiivseks sõnavaraks, sest paljud emotsionaalsete väljendusliigutuste kirjeldused on ise kujundlikud ja kõneleja-kirjeldaja afektiivset laengut kandvad, hüperboolsed, lisaks liitub tihti veel hinnanguline konnotatsioon. Näiteks võime kirjeldada kellegi imestusega seonduvat näoilmet väljenditega *suu ammuli* (neutraalne), *mokk tõllakil* (halvustav) või *karp lahti* (hümoorikas). Viimased kaks kuuluvad juba selgesti ekspressiivsete kirjelduste rubriiki. Ekspressiivse emotsioonikirjelduse puhul esineb afekt seega nii kõnealuse sündmuse tasandil (millest rããgitakse) kui ka kõnesituatsiooni tasandil (kuidas rããgitakse, antud juhul kaasaelamisega, kujundlikult ja mõjusalt).

⁹ Eesti keele käsiraamat (Erelt jt 1997), veebiversioon aadressil <http://www.eki.ee/books/ekk09/> (18. V 2012).

Mida leksikaalse afektikalduvusega peale hakata?

Eelnevast tutvustusest nähtus, et konkreetse sõna afektikalduvuse laad võib olla funktsionaalselt erinev: sõna võib emotsioonile kas referentsiaalselt otse osutada, seda kirjeldada, vihjata selle positiivsele või negatiivsele valentsile või anda lihtsalt märku, et autori kohalolu ning sündmuste kirjeldamise viisi tõttu pole lause emotsionaalselt neutraalne. Lause emotsionaalne toon selgub mitmete markerite või nende puudumise koosmõjust. Üksiku sõna afektikalduvus võib seega olla kas vältimatu (nt söimusõnadel) või toimida üksnes koosmõjus teiste sõnade ja muude kontekstist tulenevate teguritega (nt hinnangumarkerid toimivad üksnes koos kohalolumarkeritega). Sissejuhatavas osas tutvustatud „võtmesõnad” ja „külgetõmbega” sõnad moodustavad seega vaid osa sõnavarast, mis emotsioonide keelelisele vahendamisele kaasa aitavad.

Leksikaalse afektikalduvuse klasside arvessevõtt laseks tekstisensoril kasutada teadmist, et eri tüüpi sõnad annavad lause üldisesse afektiivsusesse erineva panuse olenevalt sellest, kas tegu on nn sisusõnadega või nn funktsioonisõnadega, mille rolliks on vahendada autori kohalolu ja sellega kaasnevat lause üldist ekspressiivsust. Veel võimaldab afektikalduvuse klasside eristamine sensoril teha järeldusi selle kohta, kas lauses sisaldub asitõendeid pigem protagonististi või lause autori emotsiooni kohta (vt joonist 1), sest erinevad afektikalduvuse klassid töötavad eri moodi, vahendades afektiga seotust kas siis üksnes protagonistiga toimuva sündmuse tasandil (emotsioonikirjeldused, kui need ei seostu 1. isikuga) või üksnes lause autori tasandil (ekspressiivne sõnavara ja intensiivsusemarkerid, emotsiooninimetused seoses 1. isikuga); kolmanda võimalusena tuleb kõne alla autori kaasaelamine kõnealusele situatsioonile (hinnangumarkerid ja/või kohalolumarkerid). Joonisel 1 protagonististi emotsioonide juurest autori emotsiooni juurde viiv katkendnool osutab autori sekundaarsele emotsioonile, mis tuleneb keskele tegelasele kaasaelamisest.

Tuleb arvestada, et arvuti on põhimõtteliselt inimesest erinev. Ta ei ole meiega koos üles kasvanud ja inimlikus suhtluses keelt omandanud, ta ei jaga meiega ühist subjektiivselt tajutavat reaalsust ja ilma vastava tarkvarata puudub tal võime keelise sõnumeid mis tahes tasandil interpreteerida. Kui vaadata suvalist tekstilauset n-ö arvuti silmaga, siis koosneb see üksnes digitaalsetest kahendsüsteemsetest (0 ja 1) märkidest. Millised instruksioonid tuleks siis arvutile sõnavara afektikalduvuse klasside kohta anda, arvestades, et tema „loeb” teksti hoopis teistmoodi kui inimene? Mida ta põhimõtteliselt oleks suuteline „puudu” vs. „kohal” (0 vs. 1) tasandil tekstilausest registreeerima ja mille põhjal järeldusi tegema?

Selleks et kindlaks teha, millise hääletooniga lauset ette lugeda, peab süsteem esmalt vastavalt struktureeritud sõnastikule tuginedes kontrollima, kas ja millistesse afektikalduvuse klassidesse kuuluvat sõnavara lause sisaldab. Pidades silmas eesmärki, et sensor õpiks ära tundma esmajoones lause autori emotsiooni – mida on vaja sünteeskõnes matkida – ja mitte reageerima protagonististi emotsiooni neutraalsetele kirjeldustele – mida ei ole tarvis matkida –, tuleks alustada afektikalduvuse kõige algelisemast tasandist, lausungi autori kohalolu kindlakstegemisest. Kui autor on „kohal”, siis võib ta: a) väljendada autentset autoriemotsiooni; b) elada kaasa lause protagonististi

emotsioonidele; c) olla ka täiesti neutraalne. Kui autorit „kohal” pole, siis jääb üle ainult neutraalsus. Kui autor on kohal, siis peab sensor jõudma järeldusele, kas lausung on pigem emotsionaalne või neutraalne, milleks jällegi pakuvad abi afektikalduvuse klassid. Kui lauses leidub hinnangu- või intensiivsusmarkereid, ekspressiivset sõnavara ja/või kirjeldavat sõnavara (viimane loeb ainult siis, kui seostub 1. isikuga), on suurem tõenäosus, et tegu on ka emotsionaalselt laetud lausega. Kui ühtegi neist markeritest ei leidu, tuleks järeldada, et tegu on pigem neutraalse lausega. Kui lausung osutub emotsionaalseks, siis tuleb kindlaks teha, kas tegu on pigem eheda autori-emotsiooniga (kirjeldav emotsioonisõnavara seoses 1. isikuga, ekspressiivne sõnavara, intensiivsusmarkerid) või kaasaelamisega lause keskele tegelasele (hinnangumarkerid, intensiivsusmarkerid).

Laused, milles puuduvad kohalolumarkerid, hinnangumarkerid, intensiivsusmarkerid ja ekspressiivne sõnavara, võib ette lugeda neutraalse, osavõtmatu hääletooniga. Kui selliseid lauseid on järjestikku mitu, on üsna tõenäoline, et tekst kuulub formaalsesse registrisse ja selle autori emotsioonidest on soovitatav distantseeruda. Sellises lauses sisalduvatel võtmesõnadel ei ole üksikuna võttes erilist kaalu. Näiteks oleks arvatavasti väär hakata lauset *Päike loojub pool kuus õhtul* sünteesima rõõmsatoonilisena üksnes seetõttu, et sõna *päike* seostatakse positiivse tundetooniga (Vainik 2002a). Lauses puuduvad markerid, mis osutaksid kuidagi selle autorile, saati siis tema emotsionaalsele seisundile. Võrreldagu seda lausega *Päike loojub ju pool kuus õhtul*, mis puhul väike rõhumäärsõna *ju* lisab kõnealusele situatsioonile seose kõne-situatsiooniga, esindades lausungi autori kohalolu ja vaikimisi tema teistsuguseid eelistusi päikese loojumise kellaaja suhtes. Arvuti neist vaike-eelistustest aru ei saa, kuid teda saab õpetada registreerima, et autor on lauses kohal, seega emotsiooni tõenäosus suurem. Kui meil on lause *Naaber oli eile vihane ning karjus oma naise peale*, siis kohalolumarkeri *eile* põhjal võib sensor järeldada autori kohalolu, mis on eelduseks, et lause ei pruugi olla täiesti neutraalne. Kuid kuna puuduvad viited 1. isikule, siis emotsiooni kirjeldavatest sõnadest *vihane* ja *karjus* ei saa teha järeldust, et lause autor elab kaasa sellisel määral, et lause tuleks lugeda ette vihase hääletooniga. Arvatavasti on sünteeskõne jaoks vaja sünteesida eraldi toonid autentsete autori-emotsioonide (viha, rõõm, kurbus, hirm, üllatus) ning vähem väljendusrikkaste, kuid protagonistiga emotsiooniga resoneeruvate kaasaelamisemotsioonide tarbeks, nt lihtsalt reibas, terav ja tuhm hääletoon.

Lõpetuseks

Käesoleva artikli autor on veendunud, et emotsioon saab olla ainult inimesel, emotsiooni ei saa olla lausel ega tekstilõigul. Seega, kui tahame, et lauseid loetaks ette emotsionaalses mõttes sobiva intonatsiooniga, ei tule modelleerida mitte üksnes konkreetse emotsiooni vokaalseid parameetreid, vaid mingil määral ka selle kandjat – inimest lause taga.¹⁰ Käesoleva artikli põhiline sõnum on tõdemus, et leksikaalne afektikalduvus algab kohalolust. Kirjalike tekstilausete taga peituva inimliku emotsiooni rekonstrueerimisel on seetõttu oluline esmalt teha selgeks, kas autor on lauses „kohal” või on tegu auto-

¹⁰ Mingil määral seda ka tehakse, nt hingamispauside lisamisega (Kerge jt 2008).

rist distantseeritud lausungiga. Tekstižanrid erinevad selle poolest, kui palju on neis tuntav autori kohalolu. Näiteks isiklikud kirjad on žanr, milles autori kohalolu moodustab olemusliku osa, samuti on see kindlasti esindatud spontaanse raadiovestluse salvestistes. Teleuudiste diktoritekstis või koosoleku protokollis me autori kohalolu markeerivaid keelendeid ei eelda. Teadustekstides on need taunitavad, juriidilis-formaalsetes dokumentides (v.a volitus) lausa välistatud. Autori kohalolu seostub kindlasti negatiivselt formaalsusnäitajatega, mille mõõtmiseks vähemasti pikemates tekstilõikudes on meetodid olemas (Kerge jt 2007). Üksiklause tasandil autori kohalolu ja selle määra mõõtmiseks valideeritud vahendeid teadaolevalt veel ei ole, mõningal määral võiksid abiks olla siin nn kohalolumarkeritena esiletoodud leksikaalsed tunnused; lisaks lasevad autori kohalolu ja isegi sündmusest osavõttu järeldada intensiivsusemarkerid, hinnangumarkerid ja ekspressiivne sõnavara. Emotsioone kirjeldav sõnavara (emotsiooninimetused, väljendusliigutuste kirjeldused), eriti kui see on kasutatud kolmanda isiku kohta, ilmingimata lause autori kohalolu ega emotsionaalset osavõttu järeldama ei sunni.

Siinne lühikirjeldus esitas mõistagi vaid afekti leksikaalse tuvastamise põhimõtted. Kontrollimist ja täpsustamist vajavad veel nii afektikalduvuse klasside koosseis kui ka loogiline järjestus, mille alusel saaks sensor markerite kohalolu ja puudumise alusel kõige tõenäosemad järeldused genereerida. Tuleviku teema on lahendada ka eitusest ning paradoksaalsete osalausestega liitlausestest tulenevad komplikatsioonid. Ning loomulikult on omaette lahendamist vajav probleem lausesisene tsiteerimine (otsene ja kaudne kõne), sest siis sisaldub lausung lausungis, autoreid ja ka emotsioone võib korraga olla mitu. Tööpõld on lai ning täis kive ja kände, siinkirjutajal ei ole selles suhtes mingeid illusioone.

Artikkel on valminud sihtfinantseeritava teadusteema nr SF0050023s09 „Eesti keele alusuuringud keeletehnoloogiliste rakenduste teenistuses” rahalisel toel.

Kirjandus

- Altrov, Rene, Pajupuu, Hille 2008. The Estonian emotional speech corpus: release 1. – The Third Baltic Conference on Human Language Technologies. Vilnius: Vytauto Didžiojo Universitetas; Lietuvių kalbos institutas, lk 9–15.
- Bednarek, Monika 2008. Emotion Talk across Corpora. Houndmills–New York: Palgrave Macmillan.
- Bestgen, Yves 1994. Can emotional valence in stories be determined from words? – Cognition and Emotion, kd 7 (1), lk 21–36.
- Biber, Douglas, Finegan, Edward 1989. Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. – Text, kd 9, nr 1, lk 93–124.
- Erelt, Mati, Erelt, Tiiu, Ross, Kristiina 1997. Eesti keele käsiraamat. Tallinn: Eesti Keele Instituut.
- Foolen, Ad 1997. The expressive function of language: Towards a cognitive semantic approach. – S. Niemeier, R. Dirven (toim), The Language of Emotions. Conceptualization, Expression, and Theoretical Foundation. Amsterdam–Philadelphia: John Benjamins Publishing Company, lk 15–32.

- Hunston, Susan, Thompson, Geoff (toim) 2000. *Evaluation in Text. Authorical Stance and the Construction of Discourse*. Oxford: Oxford University Press.
- Kallas, Jelena, Tuulik, Maria 2011. Eesti keele põhisõnavara sõnastik: ajalooline kontekst ja koostamispõhimõtted. – Eesti Rakenduslingvistika Ühingu aastaraamat, nr 7. Tallinn: Eesti Keele Sihtasutus, lk 59–75.
- Kerge, Krista, Pajupuu, Hille, Altrov, Rene 2007. Tekst, kontekstuaalsus ja kultuur. – Keel ja Kirjandus, nr 8, lk 624–637.
- Kerge, Krista, Pajupuu, Hille, Tamuri, Kairi 2008. Where should TTS-synthesizer pause and breath? – The Third Baltic Conference on Human Language Technologies. Vilnius: Vytauto Didžiojo Universitetas; Lietuvių kalbos institutas, lk 143–149.
- Kerge, Krista, Pajupuu, Hille 2010. Text-types in speech technology and language teaching. – B. Alonso, J. L., Dolores González Álvarez, Ú. K. Torrado, A. E. Martínez Insua jt (toim), *Analizar datos > Describir variación / Analysing data > Describing variation*. Vigo: Universidade de Vigo, lk 380–390.
- Lakoff, George, Johnson, Mark 2011. *Metafoorid, mille järgi me elame*. Tallinn: Tallinna Ülikooli Kirjastus.
- Lazarus, Richard S. 1991. *Emotion and Adaption*. – M. Jenkins, K. Oatley, N. L. Stein (toim), *Human Emotions: A Reader*. Oxford: Blackwell Publishers, lk 73–78.
- Mihkla, Meelis 2007. *Kõne ajalise struktuuri modelleerimine eestikeelsele tekst-kõne sünteesile – Modelling the temporal structure of speech for the Estonian text-to-speech synthesis*. (Dissertationes linguisticae Universitatis Tartuensis 8.) Tartu: Tartu Ülikooli Kirjastus.
- Ochs, Elinor 1988. *Culture and Language Development: Language Acquisition and Language Socialisation in a Samoan Village*. Cambridge: Cambridge University Press.
- Osgood, Charles, E., Suci, George J., Tannenbaum, Percy H. [1957] 1975. *The Measurement of Meaning*. Urbana, Chicago: University of Illinois Press.
- Pajupuu, Hille, Kerge, Krista, Altrov, Rene 2012. Lexicon-based detection of emotion in different types of text: Preliminary remarks. – Eesti Rakenduslingvistika Ühingu aastaraamat, nr 8. Tallinn: Eesti Keele Sihtasutus, lk 171–184.
- Pennebaker, James W. 2011. *The Secret Life of Pronouns: What Our Words Say About Us*. New York–Bloomsbury Press.
- Picard, Rosalind W. 1997. *Affective Computing*. Mass.: The MIT Press.
- Piolat, Annie, Bannour, Rachid 2009. An example of text analysis software (EMOTAIX-Tropes) use: the influence of anxiety on expressive writing. – *Current Psychology Letters*, kd 25, nr 2, lk 2–19.
- Schnall, Simone 2005. The pragmatics of emotion language. – *Psychological Inquiry*, nr 16, lk 28–31.
- Vainik, Ene 2002a. Kas eestlased on „kuumaverelised”? Eestlaste rahvalikust emotsioonikategooriast. – *Emakeele Seltsi aastaraamat*. Tallinn: Emakeele Selts, lk 63–86.
- Vainik, Ene 2002b. Millest on tehtud eestlaste emotsionisõnavara? – *Keel ja Kirjandus*, nr 8, lk 537–553.
- Vainik, Ene 2010. Kuidas õpetada kõnesüntesaatorile empaatiat? Emotsiooni automaatse tuvastuse võimalustest eestikeelses kirjalikus lauses sisalduva info põhjal. – Eesti Rakenduslingvistika Ühingu aastaraamat, nr 6. Tallinn: Eesti Keele Sihtasutus, lk 327–347.

- Vainik, Ene 2012. Kuidas määrata eesti keele sõnavara tundetoone? – Eesti Rakenduslingvistika Ühingu aastaraamat, nr 8. Tallinn: Eesti Keele Sihtasutus, lk 257–274.
- Veski, Asta 1982. Eesti keele grammatika II. Tartu: Tartu Riiklik Ülikool.
- Wilce, James M. 2009. *Language and Emotion*. Cambridge–New York: Cambridge University Press.

Bias of Affect in Estonian Lexis and What to do with It

Keywords: Estonian, emotion, speech synthesis, affective computing, lexical bias of affect

A practical need to make Estonian synthetic speech more natural and easier to follow has directed researcher attention to human emotions as reflected in natural speech. The article contemplates how the emotion or, more generally, affectivity of the author of a written sentence could be recognized from lexical cues. It appears that it is not only content words (nouns, verbs, adjectives) that deserve attention in this respect, but also various function words (pronouns, deictics, intensifiers etc.) Five classes of the lexical bias of affect are distinguished: descriptive emotion vocabulary, expressive emotion vocabulary, evaluative markers, presence markers and intensity markers; also, their contribution to the expression either of the author of the utterance or of the protagonist of the sentence is demonstrated. A conclusion is made that different tones of voice should probably be synthesized for the author's authentic emotions (anger, joy, sadness, fear, surprise) and for the less expressive emotions resonating to those of the protagonist, including, e.g., a brisk, sharp or dull tone of voice. The main message of the article is that recognition of the lexical bias of affect should begin from the function words referring to the presence of the author of the utterance.

Ene Vainik (b. 1964), PhD, Institute of the Estonian Language, Department of Language Technology, senior research fellow, ene@eki.ee