

KLIENDI SOOVIDE AUTOMAATNE TUVASTAMINE EESTIKEELSETES INFODIALOGIDES

**Olga Gerassimenko, Tiit Hennoste, Riina Kasterpalu,
Mare Koit, Andriela Rääbis, Krista Strandson,
Maret Valdisoo, Evely Vutt**

Ülevaade. Käsiloleva uurimuse eesmärk on leida keelelisi märguandeid, mida arvuti saaks kasutada dialoogiaktide tuvastamisel. Artiklis piirdume kliendi soovidega kui dialoogi modelleerimisel kesksete aktidega. Analüüsimiseks valisime 144 institutsionaalset telefonikõnet Tartu Ülikooli Eesti dialoogikorpusest. Need jagunevad nelja situatsioonirühma: infotelefon, reisibüroo, polikliiniku registratuur, takso tellimine. Analüüsitud dialoogides esitavad kliendid kaht liiki soove: infosoovid (helistaja vajab kindlat infot, nt telefoninumbrit) ja ametniku tegevust taotlevad soovid (reserveerida aeg arsti vastuvõtule, saata takso). Analüüs näitab, et kliendi soovides esineb piiratud hulk verbivorme ja kindlad verbid on süstemaatiliselt seotud kindlate kontekstitingimustega. Dialoogide struktuuri kui dialoogiaktide järgnevuse esitame Markovi ahelana, mis aitab arvutil kasutaja aktidele adekvaatselt reageerida.*

Võtmesõnad: dialoogikorpus, dialoogiaktide tuvastamine, vestlusanalüüs, eesti keel

1. Sissejuhatus

Suhtlemisel loovad ja tõlgendavad osalejad dialoogiakte. Selleks et modelleerida suhtlust arvutil, tuleb analüüsida, kuidas dialoogiakte keeleliselt väljendatakse ja milliseid keelelisi tunnuseid saaks kasutada arvuti kui vestluspartner lausungites esitatud dialoogiaktide tõlgendamiseks.

Tuntakse kahte dialoogiaktide automaatse interpreteerimise arvutimudelite klassi (Jurafsky, Martin 2000).

* Tööd on osaliselt toetanud Eesti Teadusfond (grant nr 5685).

Esimene klass on märguandepõhine ehk tõenäosuslik. Siin eeldatakse, et kuulaja kasutab erinevaid leksikaalseid, kollokatsioonilisi, süntaktilisi, prosoodilisi või vestluse struktuuriga antud märguandeid, mis aitavad tal otsustada, kuidas kõneleja akti tõlgendada. Selle lähenemisviisi aluseks on nn mikrogrammatikate idee (Goodwin 1996), mille kohaselt saab iga dialoogiaktiga siduda omaette väikese grammatika. Selliseid märguandeid on sagedasemate dialoogiaktide puhul laialdaselt kirjeldatud (Jurafsky jt 1998). Märguandepõhised mudelid lähenevad akti tõlgendamisele kui klassifitseerimisülesandele ja lahendavad seda statistiliste klassifitseerijate treenimisega korpustel, kus on märgendatud dialoogiaktid.

Teine mudelite klass kasutab järeldamist. Kui analüüsida nt lauset *Kas te oskate mulle öelda X?*, siis selle otsene tõlgendus oleks küsimus: *Kas te olete võimeline mulle ütleva X?* Kõneleja tegeliku soovi *Ütelge mulle X* tuletab kuulaja otsese küsimuse töötlemise tulemusel. Järeldamismudelid põhinevad arvamusloogikal ja rakendavad loogilist tuletust, et arutleda kõneleja kavatsuste üle.

Selles artiklis keskendume märguandepõhisele mudelile. Analüüsime eestikeelsetes ametikõnedes esinevad soove, kus kõneleja eesmärk on kas saada kuulajalt teatavat infot või saavutada, et kuulaja teeks teatava tegevuse. Seega piirdume siin võimalike soovide ühe alamhulgaga.

Teine piirang on see, et me vaatleme ainult kliendi soove. See on arusaadav, sest meie kaugem eesmärk on luua selline dialoogsüsteem, mis suudaks täita infoametniku rolli ja seega interpreteerida kasutaja akte ning nendele adekvaatselt reageerida.

Meie otseseks eesmärgiks on leida kliendi soovide tuvastamiseks sobivad leksikaalsed ja süntaktilised märguanded ning nende kasutamise tingimused. Käesolevas artiklis keskendume kliendi soovides predikaadi funktsioonis esinevatele verbidetele.

2. Mis on soov?

Maailmas on juba mitmeid aastaid läbi viidud erinevaid projekte korpuste märgendamiseks ja ühiste standardite loomiseks. On välja töötatud mitmeid dialoogiaktide märgenduskeeme, millest tuntuim on DAMSL (Dialog Act Markup in Several Layers, vt Allen jt 1997).

Eesti dialoogikorpuse märgendamiseks oleme Tartu Ülikoolis välja töötanud oma dialoogiaktide tüpologia, mille aluseks on vestlusanalüüsi põhimõtted (Hennoste, Rääbis 2004). See erineb mõnevõrra märgenduskeemist DAMSL, sest ta lähtub eestikeelsete suuliste dialoogide analüüsist ja vastab täpselt meie eesmärgile – uurida eestikeelset suulist suhtlust. Kui skeemis DAMSL esineb akt *request*, mis võib tähendada nii küsimust kui ka infosoovi (s.t direktiivi), siis meie tüpologia teeb vahet küsimuste ja direktiivide vahel. Küsimustel on eesti keeles selged formaalsed tunnused (küsisõnad, intonatsioon, sõnajärg), mida saab kasutada märguannetena nende automaatsel tuvastamisel. Direktiividel aga selliseid tunnuseid ei ole. Seega meie tüpologia kohaselt on *Kas te oskate mulle öelda X?* küsimus, aga *Ütelge mulle X* või *Tahan teada X* direktiiv (soov).

Meie tüpoloogias on üldse kolme liiki direktiive: soovid, pakkumised ja ettepanekud. Soov väljendab kõneleja kavatsust saada teatavat infot või kutsuda esile adressaadi teatav tegevus (H – helistaja, klient, V – vastaja, ametnik)¹:

¹ Näidetes kasutame vestlusanalüüsi transkriptsiooni. Vt ka Hennoste, Rääbis 2004, Strandson (käesolevas kogumikus), Rääbis (käesolevas kogumikus), <http://math.ut.ee/~koit/Dialoog/TranscriptionMarks.rtf> (26.03.2007).

H: → tere, hh' ütle mulle palun Parkeri kabinet number.

V: mt üheksakend neli neli null seitse.

Ettepanekud ja pakkumised (mingi tegevuse tegemiseks) on teineteisega sarnased, mõlemal juhul on sobivad reaktsioonid nõustumine või keeldumine. Seejuures on ettepaneku puhul tegevuse tegijaks kuulaja (*helistage hiljem*), pakkumise puhul aga kõneleja ise (*ma saadan teile programmi*) (Hennoste, Rääbis 2004).

Soovide esitamiseks saab kasutada erinevaid verbe, sõltuvalt sellest, missugust reaktsiooni oodatakse kuulajalt (*tõuse püsti, võta leiba* jne). Käesolevas artiklis, nagu eespool öeldud, piirdume kliendi soovidega, kus

- 1) klient vajab infot (nt telefoninumbrit) või
- 2) klient ootab, et ametnik teeks teatava tegevuse (nt saadaks talle takso).

Seejuures kaasneb tegevuse sooritamisega alati ka info andmine: ametnik informeerib klienti, et ta on tegevuse lõpetanud (*jaa, takso tuleb teile*) või et ta ei ole suuteline seda tegema.

3. Empiiriline materjal ja meetod

Meie analüüs põhineb Eesti dialoogikorpusel EDiK.² Korpus sisaldab umbes 800 suulist dialoogi, sh üle 700 telefonikõne. Korpuses on märgendatud dialoogiaktid. Iga dialoogi on teineteisest sõltumatult märgendanud kaks isikut (kokku on märgendajaid olnud 10) ja seejärel on märgendused ühtlustatud.

Käesoleva artikli jaoks valisime korpusest juhuslikult 144 telefonikõnet (kokku 19 938 tekstisõna). Kõnedes on esindatud neli suhtlussituatsiooni: 1) infotelefon (küsitakse telefoninumbreid, aadresse jms), 2) takso tellimine, 3) kõned reisibüroosse ja 4) kõned polikliinikusse. Kokku leidus dialoogides 130 kliendi soovi (tabel 1), millest üks jäi edasisest analüüsist välja, sest ta oli esitatud vene keeles.

Tabel 1. Kasutatud korpus

Dialoogi tüüp	Dialooge	Teksti-sõnu	Kliendi soove
Infotelefon	60	4384	56
Reisibüroo	36	12104	33
Polikliinik	26	2422	22
Takso	22	1028	19
Kokku	144	19938	130

Analüüsimiseks kasutame kahte meetodit. Esmalt püüame kliendi soovides leida tüüpilisi märguandeid lihtsa statistika abil. Eesti dialoogikorpuse analüüsimiseks on välja töötatud tarkvara, mis muuhulgas võimaldab hõlpsasti loendada mitmesuguseid dialoogides esinevaid nähtusi: sõnu või sõnajärgjendeid, transkriptsiooni-elemente, dialoogiakte (Treumuth 2005).

Teiseks, eesti keeles kasutatakse soovide väljendamiseks erinevaid verbivorme (tabel 2). Meie eesmärk on leida, missugustel tingimustel erinevaid variante kasutatakse. Selleks viime läbi sõnavormide kasutuse mikroanalüüsi vestlusanalüüsi

² <http://math.ut.ee/~koit/Dialoog/EDiK.html> (12.09.2006)

meetodil (vestlusanalüüsi kohta vt nt Hutchby, Wooffitt 1998). Infotelefoni- ja reisibürookõnedes ootavad kliendi soovid vastuseks info andmist (nt *ma paluks filo 'soofiateaduskonna`dekanaadi`numbrit.*) ning polikliiniku- ja taksokõnedes ametniku tegevust (*ma palun`taksot`Ringtee`kuuskend kaheksa`bee.*). Kui klient polikliinikusse helistades või taksot tellides vajab infot, siis vormistab ta oma kavatsuse alati küsimusena, mitte aga soovina (*kas teil`nihukest taksot`ka on kuhu`viis inimest peale mahuks.*).

4. Verbivormide analüüs

Kliendi soovid võib jaotada kahte klassi: sellised, mis sisaldavad verbi (predikaadi funktsioonis), ja sellised, mis verbi ei sisalda (tabel 2). Kliendi soovides esineb üksnes piiratud kogus verbe. Kõige sagedasemad on *soovima* (s.h laiendiga variandid *teada soovima, küsida soovima*) – 35 juhtu ning *paluma* – 24 juhtu (kokku 46% soovidest). Lisaks nendele esinevad veel verbid *tahtma* (ja selle variandid) – 15 korda ning *ütleva* – 8 korda, kokku 23 korda (18%). Nimetatud neli verbi moodustavad 64% kõigist juhtudest. Veel neli verbi esinevad 4-6 korda (kokku 19 korda ehk 15%). Ülejäänud verbe on soovides kasutatud 1-2 korda.

Eesti keeles on võimalik soovi väljendada nii käskiva kõneviisiga (*istu/istuge*), kindla kõneviisiga (*ma tahan/me tahame*) kui ka tingiva kõneviisiga (*ma tahaks/me tahaksime*) (vt EKG II: 174, 178). Vormide arv aga, milles eri verbid kliendi soovides predikaadi funktsioonis esinevad, on piiratud. Vaatame neid lähemalt.

Verbi **soovima** tähendus sisaldab soovi, palvet. Lisaks sellele sisaldab ta ka viisakust (EKSS). Meie korpuses esinevates soovides kasutatakse seda verbi ainsuse esimeses isikus ning kas tingivas või kindlas kõneviisis (*ma sooviks/soovin*). Küsimustes võib seda verbi kasutada teises isikus. Käskiva kõneviisi kasutamine on võimalik käsu korral – et partner ise sooviks midagi (*soovi X*), mida infodialogides mõistagi ei esine.

Tabel 2. Kliendi soovide arv korpuses (erinevate verbivormidega ja ilma verbita)

Verb	Kõneviis			Kokku
	kindel	tingiv	käskiv	
soovima	2	24		26
teada soovima		8		8
küsida soovima		1		1
tahtma	3	7		10
teada tahtma		4		4
küsida tahtma		1		1
paluma	12	12		24
ütleva			8	8
võtma		6		6
vaja olema		5		5
huvitama	1	3		4
huvitatud olema	2			2

Verb	Kõneviis			Kokku
	kindel	tingiv	käskiv	
panema	1	1	2	4
andma			2	2
küsima	1	1		2
oskama	2			2
saama	1			1
olema pakkuda	1			1
vaatama			1	1
vabandama			1	1
Verbiga kokku	26	73	14	113
Verbita				16
Kokku				129

Ühendit *teada soovima* oleme automaatse tuvastuse vajadustest lähtudes käsitlenud verbi *soovima* variandina ning samamoodi toimunud ka teiste infiniitsete verbilaienditega konstruktsioonide puhul. Siin on lisatud verb *teada*, mis osutab soovile saada teadmisi. *Teada soovima* esineb soovides niisamuti nagu *soovima* (meie korpuses ainult tingivas kõneviisis): *hh e sooviksin teada: natuke `Austria suusareisi kohta*.

Analoogiliselt on verbi *soovima* variant *küsida soovima*. Meie korpuses on selle kohta üksainus näide ja tingivas kõneviisis (*sooviksin küsida Petseri `kloostri kohta*).

Verb ***tahtma*** on verbi *soovima* sünonüüm, kuid tal on tugevam tähendus: pingsalt, vältimatult vajama (EKSS). Soovides esineb *tahtma* ainult tingivas kõneviisis. Kindel kõneviis väljendaks käsku, aga see on institutsionaalsetes dialoogides ebatavaline. Tingiv kõneviis pehmendab verbiga *tahtma* väljendatud soovi (EKG II: 35, Quirk jt 1972).

Teada tahtma on *tahtma* variant ja teda kasutatakse samal viisil. Meie korpuse näited on ainult tingivas kõneviisis: *ma=taks] teada `Ee Kaubamaja ilusa `longi telefoninumbrit*.

Küsida tahtma on samuti *tahtma* variant ja tedagi kasutatakse samamoodi. Meie korpuse ainus näide on tingivas kõneviisis: *tahaks=küsida seda: Alek `sandri auto=ja=kaatrikeskuse `telefoninumbrit*.

Verb ***paluma*** on eesti keeles tähelepanuväärne. Kõigepealt kasutatakse teda ainsuse esimeses isikus viisakussõnana *palun*, teiseks võib teda kasutada tavalise verbina (EKSS). Tema tähendus on sarnane verbi *soovima* tähendusega, kuid *paluma* on pehmem.

Soovides kasutatakse verbi *paluma* ainult kindla või tingiva kõneviisi esimeses isikus (*palun/paluksin*), teise isiku korral (*sa palud/paluksid*) ei oleks tegu sooviga. Käskivas kõneviisis saab seda verbi kasutada ainult kuulajale käsu edastamiseks, et too endale midagi paluks (*palu nüüd...*).

Verbi *paluma* kasutuses on selged erinevused olenevalt sellest, kas tegu on telefoninumbri või muu info soovimisega. Klient küsib telefoninumbreid enamasti nii, et *paluma* on lausungis ainus verb (*palun haigla*), harva lisab ta sõna *number* (*palun haigla number*). Selliseid info tüüpi osutusega lausungeid ei kasutata aadresside soovimisel või takso tellimisel (*palun taksot*). Infotelefonile helistades

tahavad kliendid sageli teada just telefoninumbreid, seetõttu on ametniku jaoks loomulik, et klient oma soovi vormistamisel info tüüpi ei väljenda.

Verbide *soovima* ja *paluma* vahel on erinevus: kui telefoninumbrit pärides kasutatakse verbi *soovima*, siis lisatakse alati nimisõna *number*. Seega võib teha järelduse, et verbi *palun* kasutatakse sellistes situatsioonides enamasti vormelina.

Verbi **ütleva** semantika käskivas kõneviisis (*ütlev/ütelge*) kirjutab ette tema tõlgendamise infosoovina, mitte aga tegevussoovina. Tingiva või kindla kõneviisi ainsuse esimeses isikus ei saa seda verbi soovi väljendamiseks kasutada (*ma ütleks/ma ütlen*).

Ranget käsku võib väljendada kindla kõneviisi teise isiku abil (*sa ütled, kuhu sa asjad peitsid*), kuid meie korpuses sellised näited puuduvad.

Tingiva kõneviisi teist isikut võib soovides kasutada üksnes koos küsiva partikliga *kas* (*kas te ütlete...*), kuid selliseid lausungeid tõlgendame me küsimustena (mis aga pole käesoleva artikli teema).

Võtma on verb, mis semantiliselt pole seotud sooviga. *Ma võtaks/ma võtan* näitab, et kõneleja soovib midagi saada. Samal ajal näitab verbi tähendus seda, et soovitud saab võtta teatud kohast ja et partner on kohustatud selle kõneleajale andma. Seetõttu esineb *võtma* infosoovides ainult tingivas kõneviisis, et soovi pehmenendada. Seda verbi ei saa infosoovides kasutada käskiva või tingiva kõneviisi teises isikus (*võta/võtke, sa võtad/võtaks*). Käskiva kõneviisiga soov ei oleks infosoov, vaid väljendaks käsku adressaadile teha midagi (nt *võta raha*). Kindlas kõneviisis saab seda verbi kasutada ainult ühendverbi koosseisus, väljendamaks ranget käsku, mitte aga (info)soovi (*sa võtad selle üles!*).

Adverbi ja verbi kombinatsioon **vaja olema** osutab vajadusele millegi või kellegi järele (EKSS). Soovides saab seda verbi kasutada ainult kindla või tingiva kõneviisi esimeses isikus (*hh mul oleks vaja sõita aprillis Budapesti*). Meie korpuses on kõik näited tingivas kõneviisis. Siin ei kasutata kindlat kõneviisi, sest see mõjuks kui nõudmine, kohustamine (mul on vaja, seepärast sa pead andma).

Huvitama on verb, mis ei näita kõneleja konkreetset soovi, vaid ainult huvi millegi vastu (EKSS). Selles funktsioonis saab teda kasutada kindla või tingiva kõneviisi esimeses isikus (*min:d uvitavad turismireisid Lõuna=Eestis*).

Olema huvitatud on verbi *huvitama* variant ja teda kasutatakse samal viisil.

Kumbagi vormi ei kasutata konkreetsetes infosoovides, vaid ainult sissejuhatavates lausungites, mis alustavad teemat, näitamaks, millest jutt tuleb (*hh olen uvitatud reisidest Skandi'naaviamaadesse=h.; ma olen huvitatud=õ selle: ses informatsioonist=et > ku palju võiks < maksta sõit Inglismaale*).

Verbi **panema** tähendus on eesti keeles lai ja hägune (EKSS). Kõigepealt võib ta tähendada füüsilist tegevust (*pane maha*). Infosoovides on *panema* ühendverbi *kinni panema* (siin tähenduses 'reserveerima') osa. Teda kasutatakse sellistes dialoogides, kus kliendil on vaja reserveerida arsti vastuvõtuaega (*aga paneme :: 'edasi ka 'ravi juba 'kinni*). Seda ühendverbi kasutatakse mõnikord (sõltuvalt kontekstist) ilma adverbina. Meie infosoovides kasutatakse *panema* käskiva kõneviisi teises isikus (*pane/pange*), tingivas kõneviisis *paneks* ja kindla kõneviisi mitmuse esimeses isikus *paneme*. Kindla kõneviisi ainsuse teise isiku kasutamine oleks tõlgendatav kui range käsk (*sa paned mulle selle numbri!*). Meie andmetes selliseid näiteid ei esine.

Ülejäänud verbid esinevad soovides igauks ainult 1-2 korda.

Verbil **andma** on eesti keeles mitmesuguseid üldisi tähendusi (EKSS). Soovides saab teda kasutada ainult käskivas kõneviisis (*andke `laua oma `ka*). Küsimustes võib seda verbi kasutada tingiva kõneviisi teises isikus (*kas sa annaks*).

Oskama ja **saama** väljendavad mõlemad muu hulgas subjekti võimelisust tegevust sooritada (EKSS). Infosoovides kasutatakse mõlemat modaalses funktsioonis konstruktsioonides *oskama öelda, saama öelda, oskama soovitada*. Kõik sellised konstruktsioonid sisaldavad lisaks ka sõna *äkki* või *võib-olla*, mis pehmen-davad soovi ja annavad ametnikule võimaluse infot mitte anda või mitte omada (*et võibolla oskate midagi `soovitada. // äkki te oskate `aadressi öelda*).

Vaatama ja **andma** esinevad käskivas kõneviisis. Kindla kõneviisi esimeses isikus väljendavad nad kõneleja enda tegevust (*ma vaatan/annan*). Infotelefo-nile helistades väljendab *vaadake* käsku ametnikule otsida midagi oma andme-baasist.

Olema pakkuda eeldab, et partner pakuks midagi. Sellist konstruktsiooni ei saa kasutada konkreetse info soovimiseks, vaid ainult üldises soovis, kus kõne-leja ei tea kindlalt, mida vastuseks oodata (*aga võibolla teil on: `teisi jaanuaris variante ka `pakkuda*).

Küsima ja **vabandama** ei ole meie infosoovides predikaadi funktsioonis kasutatust leidnud (vt ptk 5).

5. Eri verbivormide kasutamise tingimused

Eri verbe kasutatakse dialoogides erinevatel tingimustel. Tüüpilisel infosoovil on meie korpuses järgmised tunnused.

- 1) Suhtlus leiab aset võõraste vahel. Meie korpuses on vähe tuttavatevahelisi kõnesid (nt arst räägib polikliiniku registratuuri töötajaga). See mõjutab verbivormide esinemust.
- 2) Soovi esitab klient suhtluse algul. Enamasti järgneb soov vahetult rituaalsele sissejuhatusele (kutsungi vastuvõtmine, tutvustus, tervitus):

V: info´ telefon=Kersti=tere

H: → tere ma palun: ´juuksuriäri Mag´ noolia.

Vahel kasutatakse sissejuhatava osa lõpus ja soovi esitamise eel veel ühte lausungit, mis näitab, et järgneb soov. Mõnikord esitab klient küsimuse, et otsustada, kas planeeritud soovi on võimalik väljendada (kas ametnik suudab anda oodatavat infot, nt eraisikute telefoninumbreid jms):

H: .hh tere. .hh ee (.) ´kas te tegelete ´lennupiletitega ´ka.(.)

V: ikka.

H: → .hh ee mul oleks vaja sõita aprillis ´Budapesti.

Lisaks esmastele soovidele esineb ka soove, mis tekivad suhtluse käigus.

- 3) Tüüpilise soovi kolmas tingimus on, et klient on selle valmis mõelnud enne helistamist. Seega on soov helistamise põhjuseks.

Tüüpilised soovid saab jaotada kahte rühma. Esiteks, konkreetset soovid, kus kõneleja eesmärk on kas saada teatavat täpset infot (telefoninumber) või algatada

ametniku teatav tegevus (takso saatmine). Teiseks, üldised soovid, mis üksnes juhatavad sisse teema. Sellisele soovile järgneb küsimus-vastus-alamdialoog, kus kogutakse infot.

Vaatame erinevate verbide kasutamise tingimusi kliendi soovides (vt ka tabelid 3–6).

Verbivormi **sooviks** kasutatakse erinevates dialoogitüüpides nii konkreetsete kui ka üldiste tüüpiliste soovide väljendamiseks.

Kindel kõneviis *ma soovin* esineb meie korpuses kahel korral ja mõlemal juhul alustab pooleli jäävat lausungit. Seetõttu võib seda tõlgendada kui viga, mille kõneleja ise parandab (*ma soovin=e mingi `Tar `Furgo on siin kuskil `Ringteel mis laenutab fur `goone.*)

sooviks teada ja *sooviks küsida* kasutatakse ainult üldiste soovide väljendamiseks (*.hh=em sooviksin: teada natuke sellest `Soome reisist aasta `vahetusel.*). Sellised soovid esinevad reisibüroodialoogides.

Tabel 3. Verbide esinemine taksodialoogides

Verb	Kõneviis			Kokku
	Kindel	Tingiv	Käskiv	
soovima		8		8
tahtma	1			1
paluma	6	2		8
Verbita				3
Kokku	7	10	0	20

Tahaks kasutatakse kõigis dialoogitüüpides ja enamasti konkreetsete soovide väljendamiseks, mis on helistamise tegelikuks põhjuseks (meie andmetes on sellest ainult kaks erandit).

Vormi *ma tahan* kasutatakse ainult konkreetsete soovide väljendamiseks, mis on helistamise tegelikud põhjused, kuid üksnes tuttavate suhtluses (*tead mina tahan ka ühte kaarti ja ühte arvet tänase kuupäevaga.*)

Verbiühendeid *tahaks teada* ja *tahaks küsida* kasutatakse erineval viisil. Kõik kasutused esinevad võõrastevahelises suhtluses: *ma=taks teada `Ee Kaubamaja ilusa `longi telefoninumbrit.*

Nii **paluks** kui ka **palun** kasutatakse peaaegu eranditult takso tellimiseks või infotelefonile helistamisel (reisibüroo- ja polikliinikudialoogides on kummaski üksainus näide). Kõik kasutused väljendavad konkreetseid soovide.

Verbivormi **öelge** kasutatakse peaaegu eranditult tüüpilistes soovides ning ta väljendab infotelefonidialoogides konkreetseid soovide. Ainult ühel korral on seda verbi kasutatud hilisemas soovis eelmise soovi täpsustamiseks (*vot `öelge=mulle=säl: mingit `info `numbrit või .hh.*)

Võtaks esineb ebatüüpilistes konkreetsetes soovides. Meie andmetes kasutatakse seda vormi kahes rollis. Esiteks, kokkuvõtva otsuse tegemiseks, kui kõneleja teab, et soovitud info leitakse (*.hh ee no ma igaks juhaks võtaks selle `interneti aadressi.*). Teiseks võib *võtaks* osutada, et kõneleja vajab rohkem infot (*võtaks `veel mõne.*)

Oleks vaja esineb erinevates dialoogitüüpides ainult tüüpilistes soovides ja väljendab kas konkreetset või üldist soovi (*oleks vaja* `Pärnust `Arko´rok nõude `laenus telefoninumbrit; *hh mul oleks vaja* `sõita aprillis `Budapesti).

Tabel 4. Verbide esinemine registratuuridialoogides

Verb	Kõneviis			Kokku
	Kindel	Tingiv	Käskiv	
soovima		6		6
tahtma	2	4		6
paluma		1		1
vaja olema		1		1
panema	1	1	2	4
Verbita				3
Kokku	3	13	2	21

Tabel 5. Verbide esinemine infotelefonidialoogides

Verb	Kõneviis			Kokku
	Kindel	Tingiv	Käskiv	
soovima	2	5		7
tahtma (sh teada -, küsida -)		5		5
paluma	6	8		14
ütleva			8	8
võtma		5		5
vaja olema		1		1
andma			2	2
küsima	1			1
oskama	1			1
saama	1			1
vaatama			1	1
vabandama			1	1
Verbita				7
Kokku	11	24	12	54

Mind huvitavad / mind huvitaks / ma olen huvitatud esinevad ainult reisibüroodialoogides ja väljendavad peaaegu alati suhtluse alguses üldisi soove (*min:d uvitavad turismireisid* `Lõuna=Eestis). Suhtluse käigus tekkinud soovi kohta on üksainus näide, mis samuti väljendab üldist soovi (*hh vot neid m-m* `vaatamisväärsusi ja `neid nagu `vitakski).

Panema (kinni) kasutatakse ainult registratuurikõnedes ja ainult konkreetsete ebatüüpiliste soovide väljendamiseks, mis on tekkinud suhtluse käigus. Meie andmetes esineb kolm vormi.

Vorm *paneme* osutab koostööle ja on tüüpiline meditsiinidialoogides (*just, just. aga paneme::* `edasi ka `ravi juba `kinni).

Vormi *paneks* kasutatakse samuti võõrastevahelises suhtluses (*hh`paneks ühe`raviaja ka`kinni=et niigu`nii ma pean`tulema*). Ilma isikule viitava lõputa tingiva kõneviisi kasutamine võib osutada kõneleja ebakindlusele. Samuti on see viisakusvahend, mis vähendab survet soovi täitjale.

Vormi *pane* kasutatakse tuttavate suhtluses (*e no pane üksteist=e`neliküm-mend vast siis=e,`Kutsar`Maire tuleb siia`tagasi*).

Aja reserveerimise soove väljendatakse tüüpiliselt verbivormi *ma sooviks* abil.

Tabel 6. Verbide esinemine reisibüroodialoogides

Verb	Kõneviis			Kokku
	Kindel	Tingiv	Käskiv	
soovima (sh teada -, küsida -)		14		14
tahtma (sh teada -)		3		3
paluma		1		1
võtma		1		1
vaja olema		3		3
huvitama	1	3		4
huvitatud olema	2			2
küsima		1		1
oskama	1			1
olema pakkuda	1			1
Verbita				3
Kokku	5	26	0	34

Andke leiab kasutust infotelefonikõnedes dialoogi kestel tekkivate konkreetsete soovide väljendamiseks (*andke`laua oma`ka*).

Oskate soovitada /oskate öelda /saate öelda esinevad dialoogi käigus tekkivates ebatüüpilistes soovides, kus klient ei oska soovi ise esitada või ei ole kindel, et infoandja saab vastata (*et võibolla oskate midagi soovitada*).

Vaadake kasutatakse ainult dialoogi käigus tekkivate konkreetsete soovide väljendamiseks (*hh aga`vaadake sellist=nagu`Pee`Ef`Tee*).

On pakkuda esineb üldises soovis (*aga võibolla teil on: `teisi jaanuaris variante ka`pakkuda*).

Küsin/küsiksin kasutatakse ainult lausungi selles osas, mis annab eelinfot, näitamaks, et järgneb soov. Lausungi sooviosa ei sisalda predikaati (*sellist`asja küsin`Pärnust`lapsehoidjat*).

Vabandage võib kasutada predikaadina käskiva kõneviisi korral või vormelina, mis eesti keeles väljendab viisakust. Meie andmetes kasutatakse teda vabandava vormelina. Lausungis on soovi sisaldav osa esitatud kui deklaratiivne lause kindlas kõneviisis (*vabandage`Tartus on mingi`söögi`koht nimega`Maailm*).

Need harvad verbid esinevad infotelefonidialoogides (*vaadake, tahaks küsida, saate, andke, vabandage*), reisibüroodialoogides (*sooviks teada, olen huvitatud, on pakkuda, sooviks küsida, huvitab/huvitaks*) või mõlemas nimetatud dialoogitüübis (*võtaks, tahaks teada, küsin/küsiks, oskate*).

6. Verbita soovid

Kuusteist soovi ei sisalda verbi. Sellised soovid saab jaotada kolme rühma. Esimese rühma moodustavad 9 noomenifraasi (`teisipäeva `päraslõuna pärast `kolme), teise rühma 6 dialoogipartiklit *jah/jaa/mhmh* ja kolmandasse rühma kuulub fraas *no nii, nii*.

Rolli alusel võib needsamad soovid jaotada kaheks. Esimese rühma moodustavad 13 soovi, mis on reaktsioonid ametniku eelmisele voorule. Ametnik on midagi pakkunud ja kliendi nõusolek on samaaegselt ka soov.

V: ahah, ma vaatan ´äkki mõni ´sidepunkt. (2.5) näiteks siin on ´Kuressaare sidepunkt ja (0.8) ja: ´liikuv ´postkontor. (.)

H: → no nii, [nii?]

V: [.h e] neli viis kaks?

Ülejäänud kolm soovi, mis on vormistatud noomenifraasidena, esitavad tüüpilisi konkreetseid soove infotelefoni- ja taksodialoogides:

V: `Maria=Takso tere.

H: → õõ `Lossi: `kolmteist.

Niisiis on kuueteistkümnest verbita soovist kolmteist (81%) dialoogi käigus tekkinud ebatüüpilised soovid.

7. Testimine

Viisime läbi kaks eksperimenti, moodustades selleks Tartu Ülikooli Eesti dialoogikorpusel baasil testkorpusel, kuhu võtsime 505 dialoogi (kokku 57 585 tekstisõna). Testkorpusel leidis 466 kliendi soovi.

Esimese eksperimendi eesmärgiks oli kontrollida, kui paljud meie analüüsitud korpusel esinevad tingivas kõneviisis verbid on kliendi infosoovides predikaatideks. Püstitasime hüpoteesi, et kui arvuti leiab lausungis sellise tingivas kõneviisis verbivormi, siis see lausung on tõenäoliselt soov. Tulemused on esitatud tabelis 7. Nagu näha, eristavad mõned märguanded soove väga hästi (*sooviks(in)*, *paluks(in)*). Keskmine soovide tuvastusprotsent oli analüüsitud korpusel 81 ja testkorpusel 78. Muidugi on soovide tuvastamiseks vaja leida lisaks tingivas kõneviisis verbidele ka muid märguandeid.

Teises eksperimendis testisime nelja verbivormi, mis esinesid tüüpilistes konkreetsetes soovides: *öelge*, *palun*, *paluks*, *sooviks*. Eesmärk oli teada saada, kui paljud neid verbivorme sisaldavad kliendi lausungid on soovid. Vaatlesime testkorpusel ainult neid kliendi soove, mis vahetult järgnesid rituaalsele algusele. Selliseid soove oli 270. Verb *öelge* esines 70 korda (neist 63 korda soovides), *palun* 110 korda (105 soovides), *paluks* 50 korda (48 soovides), *sooviks* 61 korda (54 soovides). Seega võib järeldada, et verbivormid on päris head märguanded kliendi esimeste konkreetsete soovide tuvastamiseks.

Muidugi võivad needsamad verbid esineda ka küsimustes, kuid me saame eeldada, et arvuti tuvastab küsimused enne soove, kasutades selleks muid, küsimustele spetsiifilisi keelelisi märguandeid.

Tabel 7. Tingivas kõneviisid verbide esinemiste arv analüüsitud korpuses ja testkorpuses

	Analüüsitud korpuses	Soovides	Testkorpuses	Soovides
sooviksin	11	7	27	25
sooviks	18	17	29	27
tahaks	11	7	6	2
paluksin	2	2	18	8
paluks	11	11	31	29
võtaks	9	6	24	24
oleks vaja	5	5	8	2
huvitaks/ uvitaks	6	3	13	7
paneks	1	1	9	7
küsisin	1	1	1	0
tahaks teada	5	4	2	1
tahaks küsida	1	1	2	1
sooviks teada	1	1	3	0
sooviksin teada	8	7	1	1
sooviksin küsida	1	1	6	6
Kokku	91	74/81%	180	140/78%

8. Arutelu

Nagu korpuse analüüs näitab, esinevad infosoovides enamasti kindlad verbid kindlates vormides.

Esiteks võib sellised verbivormid jaotada kahte rühma.

Esimese rühma verbide puhul kasutatakse käskivat kõneviisi (*üttele, pane, anna, vaata*; 13 soovi). Neidsamu verbe saab kasutada kavatsuse väljendamiseks tingivas kõneviisis ainult *kas*-küsimustes (*kas te (ei) vaataks/annaks/paneks/ütteleks*).

Teise rühma kuuluvad verbivormid tingiva või kindla kõneviisi esimeses isikus. Sellisel viisil on formuleeritud 99 soovi. 73 neist sisaldavad verbi tingivas kõneviisis ning ülejäänud 26 kindlas kõneviisis.

Tingiv kõneviisi ongi üldiselt seotud sooviga, lisades viisakust. Eesti keeles on tingival kõneviisil teatavasti täpne morfoloogiline tunnus *-ks-* (*-ksi-*), mida saab automaatsel tuvastamisel kasutada (EKG I: 235).

Kindel kõneviis on deklaratiivse akti universaalne vorm (Quirk jt 1972). Kõigil analüüsitud juhtudel on ta paralleelvariandiks tingiva kõneviisi kõrval. Ainsaks erandiks on ühend *huvitatud olema*, mis esineb meie andmetes kaks korda.

Mõningaid kindla kõneviisi vorme saab küll kasutada infosoovides, kuid mitte institutsionaalsete dialoogide puhul, sest nad väljendavad ranget ja ebaviisakat käsku (*sa ütled*).

Teiseks, soovid võib jaotada kahte rühma: tüüpilised ja ebatüüpilised. Tüüpiline soov rahuldab järgmisi tingimusi:

- suhtlus leiab aset võõraste vahel;
- soov esitatakse suhtluse alguses;
- autor on selle valmis mõelnud enne suhtluse algust ja sellisena on see helistamise tegelik põhjus.

Tüüpilised soovid saab jaotada kahte liiki: 1) soovid, kus kõneleja taotleb konkreetset infot või tegevust, 2) üldised soovid, mis juhatavad sisse teema.

Meie analüüs näitab, et soovides kasutatakse kindlaid verbe kindlatel tingimustel:

sooviks, paluks, palun, öelge tüüpilistes konkreetsetes soovides (üheainsa erandiga),

sooviks teada, sooviks küsida; mind huvitab /huvitaks /ma olen huvitatud ainult tüüpilistes üldistes soovides (meie andmetes esinesid sellised soovid üksnes reisibüroodialoogides).

Mõlemat liiki soove saab formuleerida vormi *oleks vaja* abil.

Vorme *võtaks, paneks /pane/paneme, andke, vaadake* kasutatakse ainult konkreetsetes soovides, mis tekivad dialoogi käigus.

Seevastu vorme *oskate soovitada, oskate öelda, saate öelda, on pakkuda* kasutatakse üldistes soovides, mis tekivad dialoogi käigus.

Tüüpiliste ja ebatüüpiliste soovide erinevus tuleb selgelt esile registratuuridialoogides. Tüüpiliste soovide esitamiseks kasutatakse enamasti vormi *ma sooviks(in)*. Klient kasutab vormi *tahaks* sellisel juhul, kui pakutud aeg on ebasobiv. Kui suhtlus leiab aset tuttavate vahel (arst ja registratuuritöötaja), siis kasutatakse soovide väljendamiseks vorme *ma tahan, tahaks*.

Verbi *panema* kõik kasutused on spetsiifilised ja ebatüüpilised (klient helistab tagasi või formuleerib soovi, mis on tekkinud dialoogi käigus).

Kaks kasutust (lisaks veel *palun*) ei ole soovlause predikaadid: *küsiksin/küsin* on eelteade, *vabandage* on viisakusvormel.

Kolmandaks, nagu nägime, vormistatakse tüüpilised soovid enamasti tingivas või käskivas kõneviisis. Ainsad erandid on *palun*, mille tähendus sisaldab viisakust, ja *mind huvitab/ma olen huvitatud*, mis rõhutab kõneleja huvi.

Verbi mittedisaldavad lausungid väljendavad enamasti ebatüüpilisi konkreetseid soove.

9. Dialoogi modelleerimine arvuti

Kui arvuti on tuvastanud predikaadi funktsioonis esineva verbi alusel kliendi soovi, siis verbi laiendid aitavad tal mõista selle soovi sisu (*mul oleks vaja sõita aprillis Budapesti; sooviks taksot Puurmanni viisteist.; ma palun: juuksuriäri Mag'noolia* jne), et soovi täita. Muidugi ei ole otstarbekas, et arvuti täidaks korraga erinevate ametnike rolle (infotelefonil vastaja, taksodispetšer jne), ühe ainevaldkonnaga piirdumine võimaldab oluliselt lihtsustada kliendi soovide prognoosimist ja tõlgendamist.

Enamasti ei saa ametnik kliendi soovi kohe täita, vaid peab esitama täpsustavaid küsimusi, s.t algatama alamdialoogi. Selliseid alamdialooge võib algatada ka klient³:

H: sooviks taksot `Puurmanni `viisteist. | DIE: SOOV |
(0.5)

V: → ja `kelle `nimele. | KYE: AVATUD | | VTE: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

³ Dialoogiaktide märgendid vt nt <http://math.ut.ee/~koit/Dialoog/EDiC.html> (12.09.2006).

H: Ülle? | KYJ: INFO ANDMINE | | VTJ: VASTUSE TINGIMUSTE TÄP-
SUSTAMINE |
(.)
V: `Ülle `nimele. | KYE: VASTUST PAKKUV | | PPE: ÜLEKÜSIMINE |
H: jah. | KYJ: JAH | | PPJ: LÄBIVIIMINE |
V: ja `tuleb teile `auto. | DIJ: MUU |

Arvuti peab suhtlejana kasutama teadmisi dialoogi struktuuri kohta – dialoogiaktide võimalikke järgnevusi –, mis aitab arvutil nii kasutaja akte tuvastada kui ka nendele adekvaatselt reageerida.

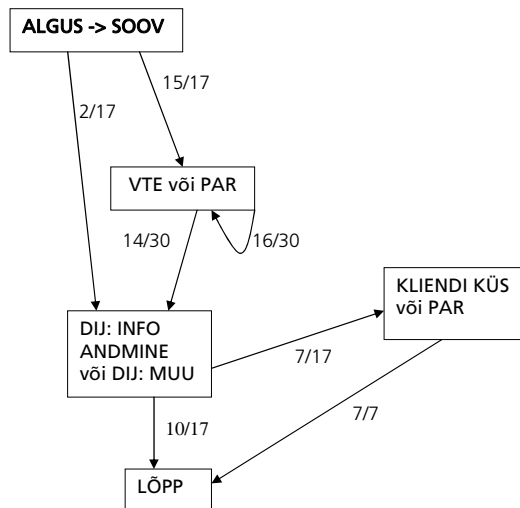
Üks võimalus dialoogi struktuuri esitada on nn Markovi ahel – graaf, mille sõlmedele vastavad dialoogiaktid ja servadele võimalikud üleminekud ühelt aktilt teisele koos üleminekutõenäosustega.

Uurisime oma analüüsitava korpuses kliendi sooviga algavate dialoogide struktuuri ja moodustasime vastavad Markovi ahelad (joonised 1–4). Sooviga algavaid taksodialooge on 17, registratuuridialooge 15, infotelefonidialooge 48 ja reisibüroodialooge 23. Ülejäänud 41 dialoogis esitas klient pärast rituaalset sissejuhatust küsimuse. Lihtsuse huvides on joonistel ära jäetud rituaalne algus ja lõpp ning alamdialoogide struktuuri ei ole eraldi välja toodud.

Joonistel on kasutatud Eesti dialoogiaktide tüpologia aktinimesid ja järgmisi lühendeid:

- VTE – vastuse tingimuste täpsustamine (küsimus – vastus),
- PAR – parandussekvents (paranduse algatus – läbiviimine – [hindamine]),
- KÜS – (saadud infot) täpsustav küsimus ja vastus,
- DIJ – direktiivi (antud juhul soovi) järelliige.

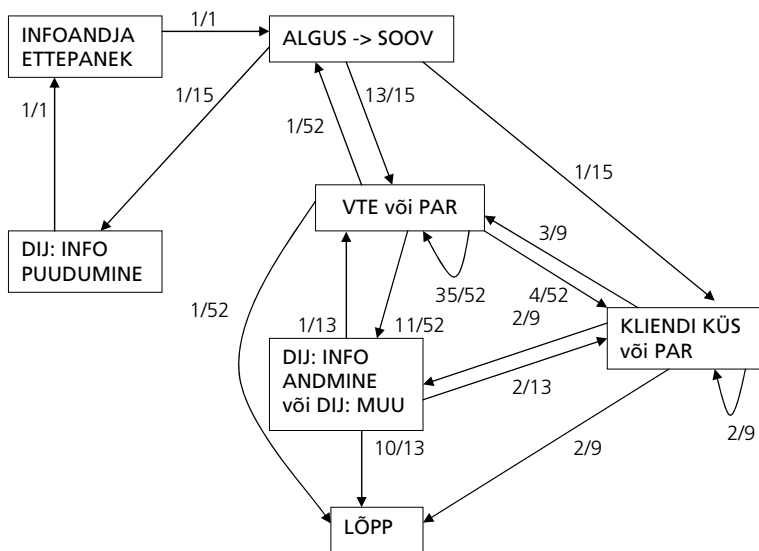
Nagu joonistelt näha, on taksokõned suhteliselt lihtsa ülesehitusega. Ülejäänud dialoogitüüpides on hargnemisvõimalusi ja erinevaid üleminekuid tunduvalt rohkem. Eriti keeruline on reisibüroodialoogide struktuur, need on ka tunduvalt pikemad kui muud dialoogid (vrd tabel 1).



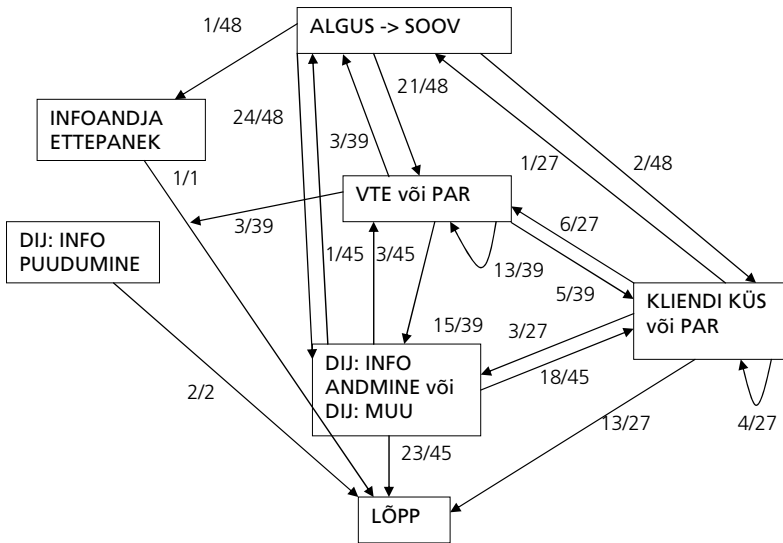
Joonis 1. Taksodialoogide mudel

Järgmises näites on paksus kirjas toodud see osa taksodialoogist, mida kajastab joonis 1: kliendi soovile järgneb ametniku vastuse tingimuste täpsustamine ja kliendi vastus (joonisel tähistab seda alamdialogi VTE), misjärel ametnik teatab soovi täitmisest (DIJ: MUU):

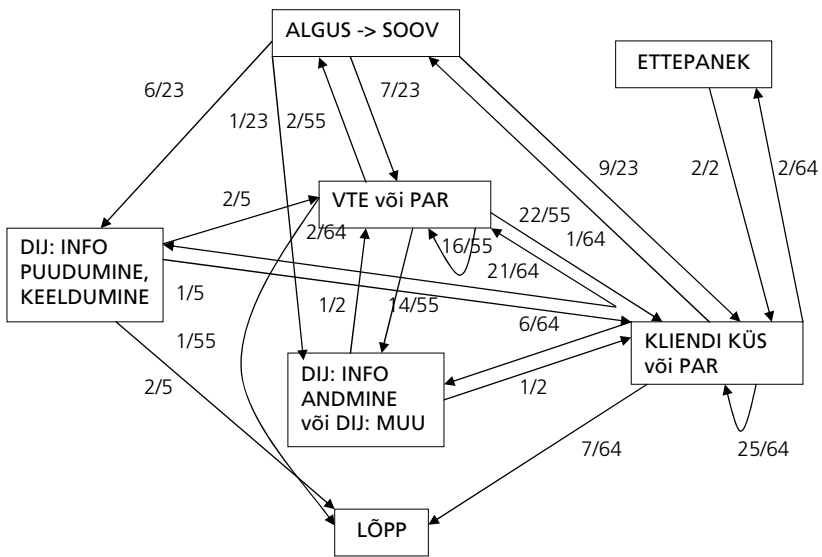
((kutsung)) | RIE: KUTSUNG |
 V: `Maria takso | RIJ: KUTSUNGI VASTUVÕTMINE | | RY: TUTVUS-
 TUS |
 tere? | RIE: TERVITUS |
 H: tervist, | RIJ: VASTUTERVITUS |
sooviks taksot `Puurmanni `viisteist. | DIE: SOOV |
(0.5)
V: ja `kelle `nimele. | KYE: AVATUD | | VTE: VASTUSE TINGI-
MUSTE TÄPSUSTAMINE |
H: Ülle? | KYJ: INFO ANDMINE | | VTJ: VASTUSE TINGIMUSTE
TÄPSUSTAMINE |
(.)
V: `Ülle `nimele. | KYE: VASTUST PAKKUV | | PPE: ÜLEKÜ-
SIMINE |
H: jah. | KYJ: JAH | | PPJ: LÄBIVIIMINE |
 V: ja `tuleb teile `auto. | DIJ: MUU |
 (.)
 H: ai[täh.] | RIE: TÄNAN |
 V: [`kõi]ke=ead teile. | RIE: SOOVIMINE |
 (.)
 H: `näge`mist | RIE: HÜVASTIJÄTT |



Joonis 2. Registratuurialoogide mudel



Joonis 3. Infotelefonidialogide mudel



Joonis 4. Reisibüroodialoogi mudel

Analüüsitud dialoogide tüüpilise struktuuri võib esitada regulaaravaldisega:

kliendi soov (ametniku algatatud täpsustus või parandus)*
ametniku vastus (kliendi algatatud täpsustus või parandus)*.

Tärn (*) tähistab siin sulgudes toodud alamdialoogi esinemist null või enam korda, s.t kliendi soovile võib järgneda (aga ka mitte järgneda) ametniku algatatud alamdialooge, seejärel annab ametnik soovitud info ja pärast vastuse saamist võib klient omakorda algatada täpsustavaid alamdialooge.

Sellisele muustrile vastavad kõik sooviga algavad taksodialoogid, 85% infotelefon- ja 80% registratuurikõnedest, kuid ainult 60% reisibüroodialoogidest. Põhiline muustrist kõrvalekalle seisneb selles, et ühe osaleja algatatud alamdialoogi järel (või ka sees) võib partner omakorda algatada alamdialooge, s.t initsiatiiv liigub ühelt osalejalt teisele, küsija ja vastaja rollid vahelduvad. Näiteks esitab ametnik enne infoandmist täpsustava küsimuse (algatades sellega alamdialoogi) ja klient jätkab kohe pärast vastamist omakorda täpsustava küsimusega:

V: mis `kuupäev teid `uvitab. | KYE: AVATUD | | VTE: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

H: ee praegu `kuupäeva võibola täpselt ei `oskagi `öelda, | KYJ: INFO PUUDUMINE | | VTJ: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

H: .hh aga sooviks teada=et mitu `inimest on ültse `rühmas. | KYE: AVATUD |

H: kui suured grupid teil `on. | KYE: AVATUD |

V: `Varmõnsheik on `lennuki `täis sada= `seitsekend `inimest. | KYJ: INFO ANDMINE |

Või esitab klient pärast info saamist uue küsimuse ja ametnik täpsustab omakorda, enne kui vastab sellele küsimusele:

H: hh aga kas sellist varianti on `ka=et=ee näiteks kahekümne= `öeksandal minna. | KYE: JUTUSTAV KAS |

V: nii | VR: NEUTRAALNE VASTUVÕTUTEADE |

V: ja siis mitu päeva `olla. | KYE: AVATUD | | VTE: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

H: noh ma=i=tea. | KYJ: EDASILÜKKAMINE | | VTJ: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

H: viis= `kuus | KYJ: INFO ANDMINE | | | VTJ: VASTUSE TINGIMUSTE TÄPSUSTAMINE |

V: äm=ma usun `küll et on selline variant ka `võimalik. | KYJ: JAH |

Koostatud Markovi ahelad vajavad muidugi täpsustamist, enne kui neid dialoogsüsteemis saab kasutada.

10. Kokkuvõte ja edasine töö

Valisime analüüsimiseks 144 institutsionaalset telefonikõnet Tartu Ülikooli Eesti dialoogikorpuselt. Need jagunevad nelja situatsioonirühma: infotelefon, reisibüroo, polikliiniku registratuur, takso tellimine.

Meie analüüsitud dialoogides esitavad kliendid kaht liiki soove. Ühed on infosoovid (helistaja vajab kindlat infot, nt telefoninumbrit), teised on ametniku tegevust taotlevad soovid (reserveerida aeg arsti vastuvõtule, saata takso). Sellest seisukohast jagunevad dialoogid kahte rühma. 1) Infotelefoni- ja reisibüroodialoogides on helistaja soovid eranditult infosoovid. 2) Takso- ja registratuurialoogides esinevad soovid on eranditult tegevust ootavad soovid.

Sealjuures on oluline, et ka tegevuse soovile vastatakse ikkagi keeleliselt, infoga, mis ütleb, et soov täidetakse või mis tingimustel saab seda soovi täita.

Analüüsi tulemusel saab teha järgmised järeldused.

- Soovid vormistatakse väheste verbide abil.
- Keskne on tingiv kõneviis ning selle sees teatud verbid kindlates vormides (tabel 2).
- Kindlates situatsioonides on omad vormelistunud konstruktsioonid: info-telefonis *palun/paluks, öelge, sooviks*; takso tellimisel *palun, sooviks*, registratuuris *sooviks*.
- Teatud situatsioonides juhatab soov sisse pikema jutu, nagu reisibürookõnedes.
- Teatud sõnavormid on kasutusel vestluse keskel, kokkuvõttena või uue teemaosa algatajana.

Seega näitas meie analüüs, et kliendi soovid vormistatakse eesti institutsionaalsetes dialoogides piiratud arvu verbivormide abil, mis esinevad lausungis kui predikaadid. Verbi võimaliku kasutuse määrab tema semantika.

Lisaks analüüsisime kliendi sooviga algavate dialoogide struktuuri, mille esitasime Markovi ahelatenä.

Küsimusi ja nende tuvastamise märguandeid on analüüsitud artiklis (Hennoste jt 2005) ning rituaale on käsitletud töös (Hennoste, Rääbis 2004).

Edaspidi kavandame analüüsida rohkem dialooge, et täpsustada verbivormide kasutustingimusi. Samuti tuleb moodustada detailsed Markovi ahelad, mille alusel saaks dialoogsüsteem kasutaja dialoogiakte tõlgendada ja neile adekvaatselt reageerida.

Kirjandus

- Allen, James; Core Mark 1997. Draft of DAMSL: Dialog Act Markup in Several Layers. <http://www.cs.rochester.edu/research/cisd/resources/damsl/RevisedManual/RevisedManual.html> (4.07.2006).
- EKG I = Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1995. Eesti keele grammatika I. Morfoloogia. Sõnamoodustus. Eesti Teaduste Akadeemia Eesti Keele Instituut. Tallinn.
- EKG II = Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1993. Eesti keele grammatika II. Süntaks, lisa: kiri. Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut. Tallinn.
- EKSS = Eesti kirjakeele seletussõnaraamat. 1988 –. Eesti Keele Instituut. Tallinn.
- Goodwin, Charles 1996. Transparent vision. – Elinor Ochs, Emanuel A. Schegloff, Sandra A. Thompson (Eds.). Interaction and Grammar. Cambridge: Cambridge University Press, 370–404
- Hennoste, Tiit; Rääbis, Andriela 2004. Dialoogiaktid eesti infodialoogides: tüpoloogia ja analüüs. Tartu Ülikooli arvutiteaduse instituut. Tartu Ülikooli Kirjastus.

- Hennoste, Tiit; Gerassimenko, Olga; Kasterpalu, Riina; Koit, Mare; Rääbis, Andriela; Strandson, Krista; Valdisoo, Maret 2005. Questions in Estonian Information Dialogues: Form and Functions. – Václav Matousek, Pavel Mautner (Eds.). Text, Speech and Dialogue. 6th International Conference TSD 2005. Springer, 420–427.
- Hutchby, Ian; Wooffitt, Robin 1998. Conversation Analysis. Principles, Practices and Applications. Cambridge: Polity Press.
- Jurafsky, Daniel; Martin, James H. 2000. Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Upper Saddle River (NY): Prentice Hall.
- Jurafsky, Daniel; Shriberg, Elizabeth E.; Fox, Barbara; Curl, Traci 1998. Lexical, prosodic, and syntactic cues for dialog acts. – Proceedings of ACL/COLING–98 Workshop on Discourse Relations and Discourse Markers. ACL, 114–120.
- Quirk, Randolph; Greenbaum, Sidney; Leech, Geoffrey; Svartvik Jan 1972. A Grammar of Contemporary English. London, New York: Longman.
- Treumuth, Margus 2005. A software tool for the Estonian Dialogue Corpus. – Proceedings of the Second Baltic Conference on Human Language Technologies. Tallinn, 341–346.

Olga Gerassimenko (Tartu Ülikool) on uurinud tagasisidevahendeid eesti ja vene suulises suhtluses. olga.gerassimenko@ut.ee

Tiit Hennoste (Helsingi Ülikool, Tartu Ülikool) uurimisvaldkondadeks on kommunikatsiooniteooriad, suuline keel jm. tiit.hennoste@ut.ee

Riina Kasterpalu (Tartu Ülikool) uurimisvaldkonnad on suuline suhtlus, dialoogipartiklid. riina.kasterpalu@ut.ee

Mare Koidu (Tartu Ülikool) erialaks on keeletehnoloogia; on uurinud dialoogi modelleerimist arvutil. mare.koit@ut.ee

Andriela Rääbise (Tartu Ülikool) uurimisvaldkonnad on suuline kõne, telefonisuhtlus, infodialoogide struktuur. andriela.raabis@ut.ee

Krista Strandson (Tartu Ülikool) on uurinud parandusi suulises eesti keeles. krista.strandson@ut.ee

Maret Valdisoo (Tartu Ülikool) on uurinud dialoogi modelleerimist arvutil. maret@ut.ee

Evelyn Vutt (Tartu Ülikool) on uurinud dialoogi modelleerimist arvutil. nurm@ut.ee

AUTOMATIC RECOGNITION OF CUSTOMER REQUESTS IN ESTONIAN INFORMATION DIALOGUES

**Olga Gerassimenko, Tiit Hennoste, Riina Kasterpalu,
Mare Koit, Andriela Rääbis, Krista Strandson, Maret Valdisoo,
Evely Vutt**

The goal of the paper is to find out linguistic cues which can be used for automatic recognition of dialogue acts by the computer. We limit us with customer requests here which are the most important acts for modelling the dialogue on the computer. We have chosen 114 dialogues from the Estonian Dialogue Corpus of the University of Tartu. Four situational groups are represented in the dialogues: directory inquiries (phone numbers, addresses, etc.), calls for taxi, calls to travel agencies, calls to outpatients' offices. There were 130 customer requests in our analysed corpus. Customers ask either for information or for an action of the operator. As the analysis demonstrates, only certain modes and persons of certain verbs can be used in order to express requests. Certain verbs are systematically related to certain contextual conditions. We represent the structure of a dialogue as a sequence of dialogue acts using Markov chain. It helps the computer to recognise the user acts and adequately react to them.

Keywords: dialogue corpus, dialogue act recognition, conversation analysis, Estonian