

## KATSELINE SEMANTIKA: PLANEERIMINE JA TEOSTUS

Anni Jürine, Jane Klavan, Ann Veismann

**Ülevaade.** Artiklis leiab käsitlemist katseliste meetodite kasutamine keeleteaduses, täpsemalt semantikas. Pikemalt peatutakse kahel etapil – katse kavandamise ja läbiviimise etapil. Tuues näiteid katselise semantika valdkonnast, arutletakse muuhulgas järgmiste teemade üle: hüpoteeside püstitamine ja operatsionaliseerimine, uurimuse valiidsus ja reliaablus, mõõdetavad tunnused, valimi esinduslikkus ja juhuslikkus, valimi maht, katseplaani valik, katsematerjali koostamine ja esitus, katsealuste värbamine ja instrueerimine. Artikli eesmärk ei ole pakkuda vastuseid kõigile küsimustele, mis keeleteadlastel katset planeerides tekkida võiks. Paljuski oleneb katse planeerimine ja teostus konkreetsest uurimisküsimusest ja iga katsega kaasneb rida spetsiifilisi küsimusi. Siinses artiklis on peetud oluliseks välja tuua vaid mõningad põhiaspektid, millega iga uurija, kes eksperimentaalset meetodit tähenduse uurimisel rakendab, peaks arvestama.\*

**Võtmesõnad:** katselised meetodid, semantika, kognitiivne keeleteadus

### 1. Sissejuhatus

Käesolev artikkel jätkab teemat, mida alustas varem ilmunud artikkel (Klavan jt 2013). Seal toodi välja põhjuseid, miks võiks katselisi meetodeid kasutada kõrvuti introspeksiooni ja korpusanalüüsiga, ning kirjeldati eri tüüpi katseid (nt vaba moodustuse katse, sunnitud valiku katse, väljamõeldud stiimuliga moodustuskatse, klassifitseerimiskatse, sobituskatse, vastuvõetavuse hindamise katse, ajaliselt mõõdetud verifitseerimine). Kui nimetatud artikli eesmärgiks oli anda ülevaade mõningatest võimalikest katsetest semantikas, täpsemalt peamiselt ruumisemantikas, siis käesoleva artikli eesmärk on arutleda selle üle, kuidas praktiliselt tähendust puudutavaid keelelisi katseid kavandada ja läbi viia. Ruuminappusest sõltuvalt keskendume ka siinses artiklis vaid semantika uurimiseks kasutatavatele katsetele ning seega jääb

\* Artikkel on valminud sihtfinantseeritava teema SF0180056s08 toel.

siinse kirjutise fookusest välja näiteks loetelukatse kui välimeetod (suuline intervjuu), mille eesmärgiks on välja selgitada teatud valdkonna (põhi)sõnavara ja/või teadmiste struktuuri (nt Sutrop 2001, Vainik 2004, Uusküla 2008 jt).

Keeleteadlastel nagu teistelgi uurijatel on võimalus andmeid koguda kas vaatluse või katse teel. Teaduslikult huvi pakkuvat nähtust võib vaadelda kas selle loomulikus keskkonnas ilma sekkumata või luua katseolukord ja uurida, kuidas mingite tingimuste muutmine mõjutab uuritavat nähtust. Võimalus tingimusi muuta ongi olemuslik erinevus nende kahe meetodi vahel. Vaatlus annab uurijale kahtlemata loomulikumat andmed, katse muudab loomulikku olukorda ja seda võib ühelt poolt pidada katseliste meetodite puuduseks, teisalt on just see põhjus, miks üldse katseid teha – isoleerida erinevad tegurid ja kontrollida nende mõju.

See omakorda on põhjus, miks katse kavandamisel tuleb olla eriti hoolikas. Iga konkreetse katse ülesehitus oleneb palju uurimisküsimusest ja katse tüübist ning väga üldisi must-valgeid reegleid pole võimalik leida. Siiski on olemas mõningad põhiaspektid, millega iga uurija, kes eksperimentaalset meetodit tähenduse uurimisel rakendab, peaks arvestama. Nende aspektide üle siinses käsitluses pikemalt arutletaksegi keeleteadusliku semantika vaatenurgast.

Kuna katselised meetodid on nii Eesti kui ka rahvusvahelises keeleteaduses suhteliselt uus ja vähelevinud valdkond (kui pidada silmas eelkõige semantikat ja süntaksit, mitte foneetikat, sotsiolingvistikat ja psühholingvistikat), on raske leida selleteemalisi üldkäsitusi. Paljud keeleuurijad, kes tegelevad süntaksi või semantikaga ja tahaksid kasutada katselisel teel saadud andmeid, tunnevad, nagu neil tuleks endal jalgratast leiutama hakata. Rahvusvahelistes erialastes ajakirjades leidub küll hulgaliselt erinevate (psühho)lingvistiliste katsete tulemusi kirjeldavaid artikleid, kuid põhirõhk on seal arusaadavalt asetatud tulemuste analüüsile ja sisulistele küsimustele ning konkreetsete katsete kavandamise ja läbiviimise etappi on kirjeldatud vaid paari lausega. Autorid leiavad, et selline tühik väärrib täitmist ning käesolev artikkel oleks esimeseks sammuks sellel teel. Kõige enam võtabki sinne käsitlus aluseks psühholoogide vastavasisulisi kirjutisi (nt Cobb 1998, Field, Hole 2006, MacLin, Solso 2007, Kantowitz jt 2009). Kuna nii eesti- kui ka ingliskeelsed õpikud katselise semantika alal puuduvad, siis on kasutatud ka erinevaid autorite poolt läbitud loengukursuste ning tööpadade konspekte ja materjale (nt E. Dąbrowska (2010), E. Käärik (2011) ja A. Luuk (2012)).

Mööda ei saa vaadata tõsiasjast, et süntaksi uurimisel on võrdlemisi laialt kasutusel vastuvõetavuse hinnangu laadsed katsed<sup>1</sup>, mille kohta leidub nii monograafiaid (nt Schütze 1996, Cowart 1997) kui ka artikleid (nt Sprouse, Almeida (ilmumas)). Sellele vaatamata võib öelda, et puudub üks terviklik eestikeelne ülevaade võimalustest ja valikutest, kuidas semantikauurimisel oma hüpoteese katseliselt kontrollida.

Katset planeerides tuleb läbi mõelda väga paljud erinevad punktid (katsealuste valik, katseühikute hulk, sõltuvad ja sõltumatud muutujad, saadud andmete tüüp, andmeanalüüsi võimalused ja piirangud, kui nimetada vaid mõned). Osa katse ülesehitust puudutavatest küsimustest on väga üldised ja puudutavad igat eksperimenti, teised on spetsiifilisemad ja sõltuvad uurimisküsimusest. Alljärgnevalt kirjeldatakse pikemalt kaht etappi – katse kavandamist ja läbiviimist – ning arutletakse mõningate praktiliste küsimuste üle, mis nende kahe etapiga kaasnevad.

<sup>1</sup> Autorid on tänulikud ühele retsensentidest, kes juhtis tähelepanu, et vastuvõetavuse hindamise katse on olemuselt pigem generatiivse kui kognitiivse keeleteadusega sobiv. Autorid leiavad, et katse tüüp pole siiski seotud kindla koolkonnaga. Generatiivse koolkonnaga sobivalt saab vastuvõetavuse hindamise katset kasutada grammatiliselt korrektsete/vastuvõetavate keelendite kindlaksmääramiseks, samas kui kognitiivse koolkonna esindajaid huvitavad selle katsetüübi puhul kõnelejate eelistused ja kasutuse varieeruvuse ulatus.

## 2. Kavandamise etapp

### 2.1. Probleemi sõnastus ja ülesande püstitus

Enne katse tegemist tuleks hoolega kaaluda, milleks katset teha. Katse tegemine ei saa olla eesmärk iseeneses. Keeleandmete kogumiseks on ka muid võimalusi ja katse tegemist võib pidada üheks vaevarikkamaks ja riskantsemaks neist. Seepärast on vaja põhjalikult läbi mõelda, millist hüpoteesi katse on mõeldud kontrollima. Nagu maailmas üldiselt võib ka keeles märgata mitmesuguseid korrelatsioone. Ent korrelatiivne seos ei tähenda tingimata põhjuslikku seost. Seepärast luuaksegi katseolukord, kus saab kontrollida arvatavat põhjust ja välja selgitada, kas sellel on ka tegelik mõju kaasnevale nähtusele. Seejuures ei tohi siiski unustada, et tegelik elu on enamasti palju keerulisem kui katseolukord.

Hästi läbimõeldud eesmärgid ja konkreetselt sõnastatud uurimusküsimused tagavad uurimuse parema kvaliteedi ehk selle suurema usaldusväärsuse. Konkreetselt sõnastatud uurimisküsimustest kasvavad omakorda välja **hüpoteesid** ehk oletuslikud põhjendused mingi nähtuse olemuse ja toimumise seletamiseks. Väga oluline samm on **hüpoteeside operatsionaliseerimine** ehk hüpoteeside viimine sellisele kujule, et neid on võimalik mõne konkreetse katsega kontrollida (s.o kinnitada või ümber lükata). See tähendab, et iga teoreetilise mõistega tuleb seada vastavusse mingi konkreetne mõõdetav parameeter (operatsionaliseerimisest semantikas vt ka Geeraerts 2006, Geeraerts 2010, Stefanowitsch 2010). Hüpoteeside püstitamine ja operatsionaliseerimine nõuavad ühelt poolt suurt loomingulisust ja teiselt poolt ranget loogilisust, sest iga uurimisküsimuse kohta on võimalik püstitada erinevaid hüpoteese ja hüpoteesi kontrollimisel on võimalikud mitmed operatsionaliseeringud, kuid keeleteadur peab suutma selgelt põhjendada, miks just mingid kindlad valikud on tehtud.

Geeraerts (2010: 72–73) toob välja empiirilise (semantika)uurimuse kuus tunnust: 1) põhineb kogutud andmetel; 2) andmed on saadud kas vaatluse või katse teel; 3) kasutab kvantitatiivset analüüsi; 4) sisaldab korrektselt formuleeritud hüpoteesi; 5) hüpotees on operatsionaliseeritud; 6) empiirilise uurimusega kaasneb empiiriline tsükkel, mis tekib korduvast andmete kogumisest, hüpoteesi testimisest ja tulemuste tõlgendamisest. Nii võib näiteks tuua, et kui tuntud polüseemia võrgustikanalüüsile sõnast *over* (Brugman 1988, Lakoff 1987) on ette heidetud intuitsioonil põhinemist (nt Sandra, Rice 1995), siis võrgustikanalüüsi saaks pidada esimeseks sammuks hüpoteesi püstitamise ja selle operatsionaliseerimise teel semantikas (vt Glynn 2010: 241–242). Näiteks võiks eesti keele sõna *üle* kohta testida, kas keelekasutajad eristavad/tajuvad selle sõna tähendustes uurimuses (Veismann 2009) välja toodud kolme kujundskeemi. Semantika alase katse teeb keeruliseks just objektiivselt mõõdetavate parameetrite ja seoste leidmine. Sorteermiskatse ja (tähtsuse) sarnasuse hindamise katse on üks võimalus kontrollida, milliseid eristusi polüseemse üksuse tähendustes kõnelejad tajuvad (vt Sandra, Rice 1995), kuid pole päris selge, kuidas seada need andmed vastavusse kujundskeemidega (kujundskeemide psühholoogilisest reaalsusest katsetes vt Gibbs, Colston 1995).

Igasuguse teadusliku uurimuse puhul tuleb silmas pidada kahte olulist nüanssi – uurimuse valiidsust ja reliaablust. **Valiidsust** võib defineerida kui kasutatava meetodika paikapidavust, kehtivust või adekvaatsust. Valiidsus näitab,

missugusel määral mõõdab katse seda, mida ta on plaanitud mõõtma. Toome näiteks mittevaliidsest katsest sellise vastuvõetavuse hindamise katse, kus katseisikuil tuleb hinnata sõnade *kõrval*, *juures*, *lähedal*, *küljel* sobivust pildi kohta, kus pall on puu kõrval. Selle katse tulemusena saame me küll pildi sobivate keelendite hulgast ning võimalikest kasutuseelistustest, kuid me ei saa järeldada, et kõneleja tegelikult kaaluks kõikide nende kasutamist vastavas olukorras. Kuna kõik need variandid on pildi kohta võimalikud, siis võib oletada, et katsealuste vastustest tuleb esile hoopis suurem varieeruvus väljendites kui see tegelikus keelekasutuses reaalne võiks olla. Seega võime tulemuste põhjal järeldada vaid seda, mis oli juba enne teada – kõik need sõnad on selle situatsiooni kirjeldamiseks võimalikud. Siinse katse mittevaliidsus tuleneb sellest, et puudub selgesõnaliselt esitatud testitav hüpotees, mis muudab võimatuks ka operatsionaliseerimise. Näiteks võiks siin püstitada hüpoteesi, et objektide kaugus üksteisest ja/või nende suuruse suhe määrab, millist kaassõna kasutada. Sel juhul saame manipuleerida objektide suuruse või omavahelise kaugusega ja kontrollida, kas hüpotees peab paika. Hea näide valiidselt katsest on selline, millel on selgesõnaline eesmärk, kontrollitavad hüpoteesid ja nende operatsionaliseeringud. Näiteks tahtes kontrollida, mille alusel valib kõneleja sünonüümsete sõnade *hulgas* ja *seas* vahel, saame me püstitada hüpoteesi (nt et valik sõltub objektide arvust või ühtsusest) ja vastavat tunnus manipuleerides selgitada kas vastuvõetavuse hindamise või produktsioonikatsega välja, millisel juhul millist keelendit eelistatakse (vt Klavan 2008).

**Reliaablus** on kasutatava meetoodika stabiilsus ja järjekindlus. See näitab, kui suure tõenäosusega saadaks katse kordamisel sama tulemus. Nii nagu ei ole olemas ideaalselt objektiivset uurimust, pole olemas ka ideaalse valiidsuse ja reliaablusega uurimust. Uurimuse reliaablus ehk usaldusväärsus on suuresti seotud valimi varieeruvusega (vt ptk 2.3). Kuna keeleliste katsete puhul on katsealusteks inimesed ja inimesed on teatavasti väga erinevad, siis ei saa me ilmselt kunagi ühtegi katset korrates täpselt sama tulemust. Rakendades aga adekvaatseid statistilisi andmeanalüüsi meetodeid on võimalik iga katse tulemusi analüüsides juhuslik viga sisse arvestada ja leida, kas valimi põhjal saadud tulemused on üldistatavad populatsioonile ehk suuremale kõnelejaskonnale või mitte. Madala valiidsusega uuring, olenemata sellest, kas reliaablus on kõrge või madal, on ühtviisi halb.

## 2.2. Mõõdetavad tunnused

Keeleuurijaid huvitab tihtilugu, kuidas keeles midagi muutub, mis selle muutuse põhjustab ja kuidas seda muutust mõõta. Muutumine ei pea olema mõistetud aja(loo)liselt, nii võib näiteks semantikas uurida, kuidas muutub ümbritsev kontekst vastavalt sellele, kas kasutatakse sünonüümipaarist üht või teist liiget (näiteks millal kasutatakse kaassõna *peal* ja millal alalütlevat käänat (vt Klavan 2012)); või seda, millistel tingimustel kalduvad katsealused kirjutama kaassõna kokku sellele eelneva käändsõnaga (vt Jürine 2011); või kuidas määrata polüsemse sõna põhitähendust, millega saaks siduda kõik selle sõna ülejäänud tähendused (vt Veismann 2008, 2009). Kõike seda, mida uurijad mõõdavad, nimetatakse **tunnuseks** ehk **muutujaks** (ingl *variable*)<sup>2</sup>. Tunnus on mõõtmise, küsitluse,

<sup>2</sup> Termineid *muutuja* ja *tunnus* kasutatakse artiklis paralleelselt, esimene on levinud rohkem psühholoogia ja teine statistikaalastes kirjutistes.

katse või vaatluse tulemusena saadud arvuline või mittearvuline näitaja. Tunnuse väärtus sõltub sellest, mida mõõdetakse või keda küsitletakse, ning mõõtmisel omistatakse igale mõõdetavale objektile mõõdetava tunnuse väärtus (vrd hüpoteeside operatsionaliseerimine). Keeleteadlastele, kelle huviks on tähendus – midagi, mida ei saa käega katsuda – on mõõdetavate tunnuste leidmine kriitilise tähtsusega. Erinevad meetodid, millega keelenähtusi mõõdetakse, võivad erineda selles osas, millist liiki andmeid nende abil saab koguda. Katse koostamisel on väga oluline teadvustada, mis liiki tunnustega on tegemist ja millisel skaalal tunnuste väärtusi katsega mõõdetakse.

Kõigepealt tehakse vahet sõltumatutel ja sõltuvatel muutujatel. **Sõltumatud muutujad** on muutujad, mis on uurija poolt manipuleeritavad; s.t muutujad, mis uurija ise on välja valinud vastavalt püstitatud hüpoteesile ja mille väärtusi on tal endal võimalik määrata. **Sõltuvad muutujad** on need muutujad, mille väärtusi uurija mõõdab, eeldades, et nende mõõtmise tulemustes ilmnev varieeruvus on uurija poolt välja valitud sõltumatu muutuja erinevate väärtuste manipuleerimise tagajärg. Teisisõnu, uurija eeldab, et sõltumatu muutuja väärtuste tahtlik varieerimine uurimuses põhjustab sõltuva muutuja väärtuse varieerumise. Näitena sõltuvast ja sõltumatutest muutujatest võib tuua vastuvõetavuse hindamise katse, mis kontrollib, kas kaassõna *peal* ja alalütleva käände kasutus kombineerituna fraasi pikkuse, sõnajärje ja objekti semantikaga mõjutab statistiliselt oluliselt vastuvõetavust (Klavan 2012). Sellises uurimuses on sõltuvaks muutujaks konkreetne vastuvõetavuse hinnang ja sõltumatuteks muutujateks, mis eeldatavasti vastuvõetavuse hinnangut kuidagi mõjutavad, on konstruktsioon (adessiiv või *peal*), fraasi pikkus, sõnajärg ja objekti semantika.

Muutujaid on võimalik mõõta erinevatel mõõteskaaladel. Üldiselt eristatakse **nelja mõõteskaalat**: nimetuste skaala ehk nominaalskaala, järjestikyskaala ehk ordinaalskaala, vahemikkskaala ehk intervallskaala ja suhteskaala (skaalade kohta vt Tooding 2007: 32–36). Erinevate mõõtmiskaaladega on seotud ka erinevad **tunnusetüübid**. Tehakse vahet **arvulistel** ehk kvantitatiivsetel ja **mittarvulistel** ehk kvalitatiivsetel tunnustel. Arvulistel tunnustel on konkreetne arvuline väärtus ja neid võib omakorda jagada **diskreetseteks** ja **pidevateks** arvtunnusteks. Diskreetset arvtunnused saadakse loendamisel ja tegemist on lõpliku arvu täisarvulise väärtusega, nt mitu lausete gruppi on katsealune lausete sorteerimise katsel moodustanud. Pidevad arvtunnused saadakse otsesel mõõtmisel ja tegemist on lõpmatu arvu võimalike väärtustega, nt reaktsiooni kiirus mõistmiskatses. Mittearvulistel tunnustel puudub arvuline väärtus ja neid võib jagada **järjestustunnusteks** ja **nominaaltunnusteks**. Järjestustunnuste väärtuste vahel on võimalik objektiivne järjestus, nt hinnangud etteantud skaalal (hästi sobiv – sobiv – pigem ei sobi – üldse ei sobi) või fraasi pikkus (lühike, keskmine, pikk). Nominaaltunnuste väärtuste vahel aga ei ole sisulist järjestust, nt sõnaliik (nt *üle* võib olla kas kaassõna, määrsõna või abimäärsõna), elus või eluta objekt, semantiline roll.

Igal skaalal ja tunnuseliigil on oma eripära ning mõõtmiskaala ja tunnuse tüüp määravad andmeanalüüsi võimalused – eri tüüpi tunnustele on rakendatavad erinevad meetodid. Arvulistele tunnustele on rakendatav suurem hulk andmetöötlusmeetodeid, mittearvuliste tunnuste puhul on meetodite hulk piiratum. Näiteks nominaaltunnuste puhul on ainuke mõeldav matemaatiline tehe sageduste lugemine, s.t loetakse kokku, kui mitu liiget igas nominaalses kategoorias esineb ja

võrreldakse erinevaid kategooriaid vastavalt nende suurusele. Järjestustunnuste puhul saab lisaks sagedusele järeltõlge teha ka tunnuste järjekorra järgi. Samas puudub täpne info selle kohta, kui suur on erinevate väärtuste vahe. Viimast liiki infot näitavad arvulised tunnused. Oluline on meeles pidada, et mitteamvuliste tunnuste väärtuste ümber kodeerimine arvudeks (sobib hästi = 5, üldse ei sobi = 0) ei muuda neid arvulisteks tunnusteks.

Keeleteaduse andmed erinevad paljuski standardsetest, kvantitatiivsete meetodite sissejuhatavates õpikutes esitatud näidisandmestikest. Suur hulk keeleandmeid on mitteamvulised ja suure variatiivsusega. See omapära seab piirid andmeanalüüsile ja tulemuste tõlgendamisele. Näiteks on korpusanalüüsi meetodeid kasutades saadud andmed enamasti seotud sagedusega. **Sagedus** on aga oma olemuselt mitteamvuline, kvalitatiivne tunnus – loetletakse, mitu korda mingi keelenähtus mingis kategoorias esines, kuid sellisele andmestikule ei ole võimalik rakendada andmeanalüüsi meetodeid, mis põhinevad arvulistel väärtustel. Ka mitmete katseteliste meetodite teel kogutud andmestikes esineb sageli just mitteamvulisi tunnuseid, nt sunnitud valiku katse (mitu korda esines ühte vastust võrreldes teiste vastustega), moodustuskatse (mitu korda moodustati ühte vastust võrreldes teiste vastustega, mitu korda anti õige või vale vastus), klassifitseerimiskatse (millistesse kategooriatesse andmed jagunesid ja/või kui palju üksuseid kuhugi klassifitseeriti). Arvulisi tunnuseid leidub keeleteaduse andmestikes suhteliselt vähe. Selliste tunnuste näideteks on keeleüksuste pikkus sõnades või silpides ja reaktsioonikiirus mingi vastuse andmisel.

## 2.3. Valim

### 2.3.1. Valimi esinduslikkuse ja juhuslikkuse printsiibid

Empiirilise uurimuse puhul on oluline teha vahet kahel kesksel mõistel – üldkogum ehk populatsioon ja valim. **Üldkogum** on “uurimisülesande sisuga piiritletud kõigi uurimisobjektide kogum”, **valim** aga “kindlal viisil eraldatud osa üldkogumis, mida analüüsis kasutatakse üldkogumi asemel” (Tooding 2007: 123). Valimi moodustamisel, s.t nii katsematerjali koostamisel kui ka katses osalevate keelekõnelejate leidmisel, tuleb silmas pidada, et valim oleks uuritava üldkogumi suhtes **esinduslik**. Oluline on selgeks teha, mis on see uuritav nähtus, mille kohta katse koostatakse ja kas katses osalevad isikud esindavad kõnelejaskonda, mida uuritakse. Üldjuhul on soovitatav kasutada juhuslikkuse printsiibil põhinevaid meetodeid valimi moodustamiseks. Rakendades **lihtsa juhuvalimi** (ingl *simple random sample*) meetodit on populatsiooni igal liikmel võrdne tõenäosus saada valimisse kaasatud. Valimi moodustamise meetodid, mis eiravad juhuslikkuse printsiipi, põhjustavad sageli nihkeid, kuna selline valim võib osutuda ebatüüpiliseks või mitteesinduslikuks. Seega ei ole selliste valimite puhul võimalik teha statistiliselt korrektseid järeltõlge populatsiooni kohta. **Mugavusvalim** (ingl *convenience sampling*) ja **lumepallivalim** (ingl *snowball sampling*) on näited mitte-juhuslikest valimitest. Esimese meetodi puhul moodustatakse valim “mugavalt”, näiteks hõlpsa kättesaadavuse, leitavuse, käsitletavuse, koostöövalmiduse alusel (nt üliõpilased). Teise meetodi puhul leitakse vaid mõned inimesed, kes vastavad tingimustele ning igauks neist

toob järgmised esindajad, kes omakorda peavad kaasama veel mõned. Sellist liiki valimi moodustavad näiteks sotsiaalvõrgustike kaudu leitud katsealused ja sel teel levitatud katsed. Muutujad nagu sugu, vanus, haridustase, sotsiaalne klass jne on soovitatav muuta kontrollitud muutujateks.

### 2.3.2. Valimi maht

Ka valimi mahu määramisel on oluline silmas pidades, milliseid andmeanalüüsi meetodeid on soov hiljem kogutud andmete kirjeldamiseks ja analüüsimiseks kasutada. Valimi mahu määramisel peab lähtuma püstitatud statistilisest hüpoteesist ja selle kontrollimiseks kasutatava teststatistiku kujust<sup>3</sup>. Tabelis 1 on välja toodud mõned sagedamini kasutatavad andmeanalüüsi meetodid ja soovitatav valimi maht (tähistatud  $n$ -ga). Üheks rusikareegliks peetakse, et katses osalevaid isikuid peaks olema vähemalt 15 iga katsetingimuse kohta. Sealjuures soovitatakse alati koguda andmeid rohkematelt isikutelt, kuna uurija peab arvestama ka võimaliku andmete kaoga (Gries 2009: 54). Soovitatud on ka varasemate uurimuste kasutamist eeskujuna. Hilisemat statistilist andmeanalüüsi silmas pidades ei ole ka liiga suur katsealuste hulk soovitatav, sest väga suur andmehulk võib muuta väiksemagi erinevuse statistiliselt oluliseks.

**Tabel 1.** Rusikareeglid valimi mahu määramisel (Wilson Van Voorhis, Morgan 2007: 48)

Meetod	Soovitatav valimi maht
Rühmakeskmiste erinevus	$n = 30$ (rühmas)
Seoste hindamine (korrelatsioon, regressioon)	$n = 50$
$\chi^2$ -test	$n = 20$

### 2.4. Katseplaani

Keeleuurijal on katse koostamisel valida mitme erineva katseplaani vahel. Katseplaani valimisel on taaskord oluline läbi mõelda, mida uuritakse ja milleks. Oluline on ka selgeks teha erinevate plaanide eelised ja kitsaskohad. Kaks peamist katseplaani, mille vahel tuleb otsus langetada, on **osalejatevaheline planeering** (ingl *between-subjects design*) ja **osalejatesisene planeering** (ingl *within-subjects design*). Kõige lihtsama katse malli puhul on uurijal kaks katsealuste rühma – katserühm ja kontrollrühm. Neid kahte rühma koheldakse katse käigus mingil viisil erinevalt, nt katserühma mõjutatakse kuidagi, samas kui kontrollrühma ei mõjutata. Seejärel uuritakse, kuidas ja kas nende kahe rühma tulemused erinevad. Kui erinevus on olemas, võib järeldada, et selle tingis(id) katserühmale avaldatud mõju ehk katses rakendatud sõltumatu(d) muutuja(d). Sellise malli puhul on tegemist osalejatevahelise planeeringuga. Näiteks kui testida hüpoteesi, et pikemas fraasis eelistavad eesti keele kõnelejad kasutada pigem adessiivi kui kaassõna, siis ühele katsealuste rühmale antakse pikad fraasid hinnata või moodustada, teisele lühikesed.

<sup>3</sup> Vabavaralised kalkulaatorid valimi mahu määramiseks: <http://www.danielsoper.com/statcalc> ja <http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize> (12.02.2013).

Levinud on ka osalejatesisene katseplaani, kus sama katsealune osaleb katse ühes osas katsetingimustes (katserühma liikmena) ja teises osas kontrolltingimustes (kontrollrühma liikmena). Peamiseks eeliseks sellise katseplaani juures on see, et uurijal on suurem tõenäosus leida tulemustes erinevusi, kuna juhusliku vea suurus on vähendatud (nt individuaalsed erinevused). Teiseks eeliseks on see, et võrreldes osalejatevahelise katseplaani nõuab osalejatesisene katseplaani katse väiksemat arvu osalejaid. Peamiseks probleemiks on siin aga nn (stiimuli) järjekorra mõju, millest tuleb täpsemalt juttu allpool (vt ptk 2.5). Üldiselt on võimalik järjekorra mõju vähendada, kasutades juhuslikus järjestuses stiimulit, s.t katse- ja kontrolltingimused vahelduvad samas katseseerias osaleja jaoks juhuslikus järjestuses.

Kuna osalejatesisene katseplaani on efektiivsem (osalejatevaheline indiviidi tasandi varieeruvus on vähendatud), siis üldiselt soovitatakse rakendada just sellist katsemalli. Osalejatesisene katseplaani ei sobi, kui katsel on katsealustele püsiv mõju (nt katsealuste treenimine vastavalt kahele erinevale katsetingimusele) või kui katse on väga pikk. Rakendatav on ka nn segakatseplaani (ingl *mixed design*), kus korraga on kasutusel mitu erinevat katseplaani, nt nii osalejatevaheline kui -sisene planeering. Näiteks võib osalejatevaheliseks sõltumatuks muutujaks olla sõnajärg (üks rühm hindab ühe sõnajärgiga ja teine teistsuguse sõnajärgiga stiimulit) ja osalejatesiseseks sõltumatuks muutujaks fraasipikkus (kõik osalejad hindavad nii pikki kui lühikesi fraase).

Lisaks kahele peamisele eelpool nimetatud katseplaanile on uurija käsutuses veel suur hulk erinevaid katseplaine, nt mitme rühmaga katseplaani (ingl *multiple group design*), faktorplaanid (ingl *factorial design*), kvaasieksperimentaalne katseplaani (ingl *quasi-experiments*), plokk-plaanid (ingl *block design*), ladina ruut (ingl *Latin square*)<sup>4</sup>. Milline katseplaani valida sõltub uurimuse iseloomust, uurimisküsimusest, hüpoteesidest, sõltumatutest ja sõltuvatest muutujatest jne.

## 2.5. Katsematerjali koostamine ja esituse järjekord

Keeleteadusliku katsematerjali ehk **stiimuli** koostamisega seotud praktilisi nõuandeid ja kasulikke näpunäiteid pakuvad nt Schütze (1996: 184–186) ja Cowart (1997: 89–102)<sup>5</sup>. Siinkohal tahame välja tuua mõned põhieeglid: stiimul (nt katse kasutatavad laused, fraasid, sõnad jms) tuleks katsealustele esitada juhuslikus järjekorras; uurija peaks stiimulit kontrollima selle leksikaalse sisu, tähenduse, kompleksuse, sageduse, pikkuse ja struktuuri osas (Schütze 1996: 150–168); uurijal tuleks iga katsetingimuse jaoks välja mõelda mitu erinevat stiimulit (Schütze, Sprouse (ilmumas): 13). Kui uurija ei suuda ülalmainitud aspektides stiimulit kontrollida, võib tulemuste tõlgendus osutada keeruliseks, kuna ei saa olla kindel, millised stiimuli aspektid katsealustele mõju avaldasid. Siinjuures on paslik meelde tuletada uurimuse valiidsuse küsimust (vt ptk 2.1). Nii näiteks tekib tulemuste tõlgendamisel probleeme, kui katses kasutatavad keelelised üksused erinevad leksikaalse sisu poolest, aga soovitakse hoopis uurida keelelise üksuse pikkuse mõju konstruktsiooni/lause vastuvõetavusele. Kui näiteks tulemustes tuleb esile oluline erinevus lühikeste ja pikkade üksuste vahel, siis ei ole uurijal võimalik kindlaks teha, kas erinevused tulemustes on tingitud just stiimuli erinevast pikkusest (uurimuse

<sup>4</sup> Mainitud katseplaanide kohta võib lugeda mõnest sotsiaalteadlastele mõeldud katsete planeerimise käsiraamatust (nt Cobb 1998, Field, Hole 2006, MacLin, Solso 2007, Kantowitz jt 2009).

<sup>5</sup> Nii Schütze (1996) kui Cowart (1997) piirduvad ainult vastuvõetavushinnangu katsetega, kuid paljuski kehtivad autorite poolt välja toodud soovitused ja hoiatused ka teist liiki keeleteaduslike katsete puhul.

seisukohast oluline tulemus) või hoopis stiimuli leksikaalsest sisust (uurimuse seisukohast ebaoluline tulemus, nn kõrvalmuutuja). Üks võimalus selle probleemi vältimiseks on, et katsematerjali tuuakse sisse kontroll-laused, kus katsealustel lastakse hinnata sama leksikaalse sisuga lauseid mingis kolmandas lausemallis. Kui erinevate leksikaalse sisuga katsestiimulite vahel ei tule selles kolmandas kontrolltingimuses oluliselt erinevaid tulemusi, võib järeldada, et katsetingimustes esinenud erinevus on tõepoolest tingitud sõltumatust muutujast “pikkus” ja mitte sellest, et laused on erineva leksikaalse sisuga.

Katse pikkus, s.t stiimuli **maht**, oleneb suuresti sellest, kuidas sõltuvat tunnusust mõõdetakse (kas nominaalsel, järjestus- või intervallskaalal), millist andmeanalüüsi meetodit plaanitakse hiljem tulemuste analüüsimisel rakendada. Samuti sõltub stiimuli maht valimi suurusest (vt ptk 2.3). Üldiselt on välja toodud, et nominaalse skaala puhul on optimaalne kasutada vähemalt nelja katseüksust ühe katsetingimuse kohta. Igal juhul tuleb arvestada, et katse ei saaks nii pikk, et katsealune tüdineb vastamast.

Katse ettevalmistamisel on oluline ka küsimus, millisesse **järjekorda** stiimulid (laused, visuaalne info vm) seada. Stiimulid ise peavad olema hoolega kontrollitud, kuid nende esitamise järjekord peab olema juhuslikustatud. Interneti teel läbiviidavas katses on võimalik juhuslikustada iga katsealuse katse eraldi. Paberil läbiviidava katse korral võiks kasutusel olla erinevad versioonid katsest (Coward 1997: 96). Seda seepärast, et mitme stiimuli esitamisel järjest võib eelneva stiimuli vastus mõjutada järgneva stiimuli vastust. Erinevate mõjudena on välja toodud (Schütze 1996: 134–140): harjumine – vastus igale järgnevale stiimulile tuleb kergemini; väsimine – igale järgnevale stiimulile vastamine väsitab katsealust ja ta võib vastamisest tüdineda; ülekanne – eelnev stiimul mõjutab samamoodi vastama järgnevale; tundlikkus – iga järgneva stiimuliga aimab katsealune üha enam, millist vastust talt oodatakse. Nende mõjude vältimiseks on soovitatud (Coward 1997: 91–92, Schütze 1996: 134–140) lasta katsealustel enne katset harjutada; teha katse nii lühike ja põnev kui võimalik; pikendada aega katsetingimuse vahel (vältides nii ülekanne mõju); koostada katse nii, et katsealune ei märkaks stiimulite varieeruvust (vt ka ptk 3.1); juhuslikustada stiimulite järjekord. Kõige selle juures on oluline jälgida, et katse läbiviimise tingimused ei muutuks.

### 3. Läbiviimise etapp

Peale katse hoolikat ettevalmistamist saab asuda katse läbiviimise juurde. Ka selles etapis on olulisi nüansse, mida silmas pidada. Siinses artiklis on eristatud katse ettevalmistamise ja läbiviimise etappi. Tegelikult ei ole aga need teemad üksteisest rangelt eristatavad ja nii mõnelegi seigale, millele siinses peatükis viitame, tuleb mõelda juba katse planeerimise ajal. Esiteks tahame toonitada, et olenemata sellest, mis tüüpi katsega on tegu, tuleb igal juhul esmalt koostada pilootkatse. Pilootkatse viiakse läbi väikse hulga katsealustega ning selle eesmärgiks on kontrollida, kas koostatud stiimul töötab. Kui millegagi on probleeme (nt ei ole juhis piisavalt selge või on stiimulis vajakajäämisi), tuleb probleemne koht parandada ning viia läbi uus pilootkatse. Pilootkatsest saab teada ka seda, kui kaua katse keskmiselt aega võtab.

### 3.1. Katsealuste värbamine ja instrueerimine

Kõige enam kasutatud sihtrühm katsealuste värbamisel on üliõpilased, samas märgivad Gonzalez-Marquez jt, et katsealuste leidmiseks on ka teisi võimalusi ja kohti (katsealuste värbamise viiside kohta vt nt 2007: 74). Katsealuste värbamisel tuleb aga kindlasti silmas pidada valimi koostamise põhimõtteid, mida käsitleti ptk-s 2.3. Inimeste motiveeritus katses osaleda sõltub sellest, kui võrd nad peavad selleks pingutama. Kui katse ja selle läbiviimiseks vajalik tehnika võimaldab uurijal endal katsealuste juurde minna, on osalejaid lihtsam värvata. Samuti võib katsealuste leidmist hõlbustada preemiade olemasolu. (Gonzalez-Marquez jt 2007: 74) Oluline on katsealuste värbamisel ausalt öelda, kui kaua katses osalemine aega võtab. Internetis täidetavate katsete puhul peaks katse tutvustuses olema kirjas aus prognoos katse kestuse kohta. Vastasel juhul võib katsealune vastamisest tüdida, katse pooleli jätta või järgmisel korral osalemast keelduda. Eetilistel põhjustel on oluline rõhutada, et katses osalemine on alati vabatahtlik.

Katsealuste kohta tuleb koguda ka taustaandmeid. Keeleteaduslike katsete puhul on eeskätt tähtsad demograafilised andmed nagu emakeel, teised keeled, päritolu, vanus, sugu, haridus. Need andmed annavad ülevaate sellest, kui homo- või heterogeenne valim on. Katsealuse emakeele olulisus on semantikat puudutavates uurimustes ilmselge, kuid ka teiste keelte oskus võib katsealuse keeletunnetust mõjutada, kui indiviid on teatud keelt õppinud väga pikka aega. Ka sotsiaalsed muutujad nagu vanus, sugu ja haridus, samuti regionaalne päritolu võivad tulemusi teatud suunas mõjutada. (Nende muutujate kontrollitavuse kohta vt ptk 2.3). Üldiselt ei eelda enamik semantika uurimiseks kasutatavatest katsetest isikuandmete (nt nimi) kogumist, vaid piisab demograafilistest andmetest. Kui identifitseerimist võimaldavaid andmeid on mingil põhjusel tarvis koguda, tuleb isikuandmeid sisaldavate failidega hoolikalt ümber käia, sest uurija peab tagama katseisikute anonüümsuse (Field, Hole 2006: 101, Gonzalez-Marquez jt 2007: 74).

Enne katsega alustamist tasub lausuda mõni sissejuhatav lause, et osalejaid motiveerida. Samas kaasneb igasuguse info jagamisega uurimuse kohta alati oht katsealuste vastuseid mingil viisil mõjutada. (Cowart 1997: 87) Nimelt on täheldatud, et kui katsealused taipavad, milliseid vastuseid neilt teatud küsimustele oodatakse, vastavadki nad nii, et nende vastused oleksid sotsiaalselt vastuvõetavamad, s.t nad püüavad uurijat aidata (Gries 2009: 49). Vainik (2004: 130) sedastab oma kogemusest, et katsealused on väga vastutulelikud, täites neile antud viimase ülesande innuga ka juhul, kui nad on eelnevast katsest väsinud ja neile pakutakse võimalust lõpetada. Katse läbiviija võib tulemusi mõjutada ka näiteks kehakeele abil katsealustele vihjeid andes. Seetõttu oleks hea kasutada topelt-pimedat meetodit, mil katset viib läbi keegi, kes ei tea, mis on uurija(te) hüpoteesid. Üks võimalus tulemusi mõjutada ehk katseisikuid praimida on teatud keelendi varasem mainimine samas olukorras.

Tulemuste mõjutamise välistamiseks kasutatakse sotsiaal- ja humanitaarteadustes sageli varjamist. See tähendab, et katsealustele ei avaldata olulisi detaile katse eesmärkide kohta või öeldakse, et katse eesmärk on miski muu, kui see tegelikult on (Kirk 2013: 23). Samas tuleb meeles pidada, et katsealuste petmine ja nende eest informatsiooni varjamine on taunitav eetilistel põhjustel. Katses osalejad peavad uurijale andma informeeritud nõusoleku, mis tähendab, et nad mõistavad,

millega nad ennast seovad ja nõustuvad osalema. On aga küsitav, kas informeeritud nõusolekut saab anda ilma (korrektse) infota (Field, Hole 2006: 99). Field ja Hole osutavad, et psühholoogias pole teatud uurimisküsimuste puhul sellest midagi halba, kui katsealune teab, mis on katse eesmärk, ning see ei mõjuta eksperimendi tulemusi soovimatus suunas (2006: 99). Keeleteadusest võib näiteks tuua, kui katsega uuritakse kahe konstruktsiooni varieerumist keeles (nt kaassõna *peal* ja adessiivi varieerumist ruumisuhete väljendamisel (Klavan 2012)). Sellisel juhul, kui ei ole alust arvata, et keelenorming saadavaid tulemusi mingil moel mõjutaks, ei pea eksperimendi eesmärki katsealuste eest varjama ning katsealuseid võib suunata teadlikult kahe konstruktsiooni vahel valima. Kuigi äärmustena on teada näiteid sotsiaalpsühholoogia vallast, kus osalejaid tulebki uurimuse eesmärkide osas eksiteele juhtida, siis üldiselt tuleks otsest katsealuste petmist võimalusel vältida (Field, Hole 2006: 99–100). Kui tahetakse varjata katse otsest eesmärki, võiks katse eesmärgi sõnastada võimalikult üldiselt (nt katses uuritakse keelekasutust või ruumiväljendeid vms). Oluline on, et katsealustele selgitatakse katse läbiviimise protseduuri piisaval määral, et nad teaks, mis juhtuma hakkab (Gonzalez-Marquez jt 2007: 74), samuti tuleb pärast katse sooritamist avaldada katse tegelik eesmärk.<sup>6</sup>

Katse eesmärgi varjamise huvides on kasutusel ka kas täitelaused (ingl *filler*) või sõltuvalt stiimulist muud täiteüksused (nimetatakse ka distraktoriteks). Need on katseüksused, mis pole seotud uurimisküsimusega (Schütze, Sprouse (ilmumas): 14). Täiteüksuste eesmärgiks ongi katsealustes katse eesmärgi osas segadust tekitada, et katsealune ei aimaks katse konkreetset eesmärki ja uurija hüpoteese ega püüaks vastata uurijale meelepäraselt. Cowarti (1997: 92) järgi peab täiteüksuseid olema vähemalt sama palju kui neid üksusi, mis katse eesmärki täidavad, ideaaljuhul on neid aga 2–3 korda rohkem. Täiteüksusi kasutades tuleb aga silmas pidada, et katse liiga pikaks ei läheks, samuti on tähtis täiteüksuste järjekord (vt ka ptk 1.5). Täiteüksused peaksid olema paigutatud teiste katseüksuste vahele pooljuhuslikus järjekorras selliselt, et katse esimene stiimul on täiteüksus ning et täiteüksus eraldaks kõiki eesmärki täitvaid katseüksusi.

### 3.2. Juhis ja stiimul

Uurija ei tohiks kunagi alahinnata mõju, mida tema instrueerimine katsealustele avaldab (Cowart 1997: 56). Seetõttu tuleb hoolikalt läbi mõelda nii see, mida katsealustele enne katset öelda, kui ka kirjalik tööjuhise, kus on kirjas, mida neilt oodatakse. Praktikas kättesaamatu, kuid eesmärk, mille poole püüelda, on olukord, kus kõik katsealused saavad ühtmoodi õigesti aru, mida nad tegema peavad. Iga väiksemgi erinevus selles, kuidas iga indiviid ülesandele läheneb, suurendab varieeruvust andmetes (Cowart 1997: 89).

Cowarti (1997: 56–57) järgi on üks olulisemaid küsimusi, mis tekib naiivsete keelekasutajate kasutamisel katsealustena, see, missugune juhise toob esile vajalikud otsustused, s.t sellised, mis ei ole mõjutatud keelevälistest faktoritest. Teatud uurimisküsimuste, näiteks lause vastuvõetavuse hindamise puhul võib keelekasutajate vastuseid mõjutada teadmine koolis õpitud grammatikast, kuid uurija

<sup>6</sup> Enne katse läbiviimist soovitatakse ka uurida (Field, Hole 2006: 98), kas institutsioonis on eetikakomitee, kellele tuleb esitada hindamiseks kavandatava uurimistöö plaan, ning kes peab andma nõusoleku uurimistöö läbiviimiseks. Eestis tegutseb Tartu Ülikooli inimuuringute eetika komitee, kes menetleb nii ravimi kliiniliste uuringute ja biomeditsiiniliste, psühholoogiliste käitumisteaduste alase uurimistöö projekte kui ka üliõpilastööde projekte (<http://www.ut.ee/et/teadus/eetikakomitee>, 12.02.2013).

on huvitatud hoopis tegelikust keelekasutusest. Sellisel juhul võib juhised suunata katsealust ignoreerima õpitud reegleid ja tegema otsustusi ükskõik missugustele muudele kriteeriumidele toetudes. Näiteks uuris Jürine (2011) sunnitud valiku katse abil, kuidas on kaassõnaühendite norminguvastane kokkukirjutamine seotud tähendusega. Kuna kaassõnaühendite kokkukirjutamine on praeguse normingu kohaselt reeglitevastane, suunas juhised katsealuseid valima kirjalikult selle alusel, mis lausesse kõige paremini *sobib*, mitte õpitud reeglitele toetudes. Sellist juhiseid nimetatakse intuitsiooni-juhiseks, alternatiiv on kasutada preskriptiivset juhiseid, mille puhul suunatakse katsealust toetuma keelenormile. Samas nendib Cowart (1997: 57–58), et pole piisavalt tõendeid väitmaks, et katsealused suudavad lauseid hinnates rakendada erinevaid standardeid, s.t hinnagud võivad olla samad nii intuitsiooni- kui ka preskriptiivse juhise puhul. Siiski on oluline, et juhised, mis katsealustele antakse, on võimalikult selge ja ühene ning suunab nende tähelepanu sellele ülesandele, mille täitmist neilt oodatakse (Cowart 1997: 91).

Materjali esitamiseks katsealustele on erinevaid võimalusi. Cowarti (1997: 63) järgi võib materjali esitamise viis katse tulemusi teatud aspektides mõjutada. Stiimuli võib esitada nii kirjalikult, piltidena, reaalsete objektidena kui ka auditiiivsel kujul. Tavalisim vorm visuaalse stiimuli esitamiseks on prinditud küsitluslehed, kus stiimul on esitatud kirjalikult ja/või piltidena. Näiteks võib vastuvõetavuse hindamise katse, milles uuritakse ruumisemantikat, kasutada ruumi kirjeldamiseks pilti ning esitada katsealustele iga pildi kohta lause (Carlson, Hill 2007: 253). Kirjaliku katse plussina võib nimetada asjaolu, et selle ettevalmistus ja läbiviimine on lihtne ja see võimaldab koguda andmeid suurelt hulgal katsealustelt korraga. Ka võimaldab kirjalik vorm sama küsitluslehe erinevaid versioone korraga esitada. Sellise katse puhul on miinuseks see, et tulemused sõltuvad inimeste lugemisoskusest, mille tase on üldkogumis varieeruv. Samuti see, et inimeste ootused kirjutatud keelele võivad olla üsna erinevad suulise keelega võrreldes. Näiteks vastuvõetavuse hindamise katse puhul võivad teatud laused suuliselt esitatuna olla vastuvõetavamad kui kirjalikult. (Cowart 1997: 64) Ainult kujundliku stiimuli puhul kirjalikust tulenevaid probleeme ei esine. Küll aga on ka piltide kasutamisel omad miinused, näiteks võivad pildid olla raskesti tõlgendatavad ning anda suurema varieeruvusega andmeid. Vahel kasutatakse visuaalse stiimuli esitamiseks ka reaalseid objekte, sellise katse miinuseks on see, et katse läbiviijal on suurem võimalus katsealuseid kogemata mõjutada, kui katset korraldatakse mitmes järgus. Auditiiivne stiimul võib osutada kasulikuks, kui on oluline juhtida katsealuse tähelepanu võimalikult kaugele kirjakeele normeringutest.

Katse stiimuli võib esitada ka arvuti abil. Näiteks kasutatakse sellist esitlusviisi interneti teel levitatavates küsitlustes. Üldiselt on arvuti abil täidetavatel küsitluste tüüpi katsetel samad plussid ja miinused, mis ülalpool esitatud, kuid siin käsitletavat erinevust tulenevad just interneti kui teistsuguse meediumi rollist. Gonzalez-Marquez jt toovad välja, et ideaalis tuleb tagada, et kõik katsealused on võrdses situatsioonis (2007: 74). Internetis levivate katsete puhul puudub aga kontroll katseks kulutatud aja, katsealuste fookuseerituse ja motivatsiooni üle. Olenvalt võimalustest, mis internetikatse jaoks kasutatavad tehnilised abivahendid pakuvad, on uurijal siiski teatud juhtudel võimalik tuvastada katse ajaline kestus. Internetikatsetel on ka plussid. Katsetes osalemine interneti teel on katsealusele palju mugavam mitmel põhjusel. Kuna katse tegemise koht ja aeg ei ole määratud, saab

katsealune sooritada katse talle sobival ajal ja sobivas kohas. See tähendab, et katse situatsioon on loomulikum ja vabam. Peale selle võimaldab interneti teel levitav katse leida hõlpsasti katsealuseid.

## 4. Kokkuvõte

Artikli eesmärgiks oli anda ülevaade katselise semantika võimalustest ja probleemidest, tuues konkreetseid näiteid seikadest, millega semantikas katseid tehes peaks arvestama. Katse kavandamise juures toodi välja olulised etapid, mis, nagu iga muugi teadusliku katse puhul, tuleb läbida ka tähendust uurides. Probleemi sõnastuse ja ülesande püstituse juures on oluline hüpoteeside operatsionaliseerimine, mis semantika omapära arvestades on üks keerulisemaid ülesandeid tähenduste uurimisel empiirilisel, kuid just sellest sõltub uurimuse valiidsus ja reliaablus. Lähemalt peatuti ka muutujate/tunnuste valiku küsimusel ja erinevatel tunnuste tüüpidel. Mitteamarvuliste tunnuste ülekaal ei iseloomusta mitte ainult semantikat, vaid kogu keeleteaduslikku andmestikku (v.a foneetika). Katset ette valmistades tuleb läbi mõelda ka valimi koostamise põhimõtted ja valimi optimaalne suurus. Artiklis püüti näidata, et keeleuurijal on võimalik valida mitme erineva katseplaani vahel ja et konkreetne valik sõltub uurimuse iseloomust. Olulist rolli mängib keelise katse kavandamisel loomulikult ka see, milline on katsematerjal, mille uurija katsesse valib. Sellega seoses toodi välja, mida arvestada katsematerjali koostamisel ja esitamisel. Viimaks kirjeldati probleeme, millega tuleb arvestada katset läbi viies – katsealuste värbamine ja instrueerimine ning sellega seonduvad eetilised küsimused, juhis ja katse stiimuli esitamise viis.

Selge on, et iga uurimuse puhul kerkivad katset tehes üles konkreetselt just vastava uurimusküsimusega seotud probleemid ja sinne ülevaade suutis anda vaid valiku olulisematest ja tavalisematest aspektidest, mida kaaluda. Igasuguse keelise katse puhul jääb aga kehtima vanasõna “üheksa korda mõõda, ükskord lõika”. Selleks, et katselisel teel tähenduse uurimiseks adekvaatseid andmeid keelekõnlejatelt kätte saada, tasub planeerimise peale kulutada proportsionaalselt rohkem ressursi kui teostusele ja andmeanalüüsile. Kui katse ülesehituses on tehtud viga, ei päästa uurimust ei see, kui katse viiakse läbi suure hulga katsealustega, ega see, kui andmeid analüüsitakse mitmekesiste statistiliste andmeanalüüsi meetoditega. Potentsiaalseid puudusi, mida tuleb katse ülesehitamisel ette ka kogenenumatel katsekoostajatel, aitab tuvastada pilootkatse.

### Viidatud kirjandus

- Brugman, Claudia 1988 [1981]. *The Story of over: Polysemy, Semantics, and the Structure of the Lexicon*. New York: Garland Press.
- Carlson, Laura A.; Hill, Patrick L. 2007. Experimental methods for studying language and space. – Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson, Michael Spivey (Eds.). *Methods in Cognitive Linguistics. Human Cognitive Processing* 18. Amsterdam, Philadelphia: John Benjamins, 250–276.
- Cobb, George W. 1998. *Introduction to Design and Analysis of Experiments*. New York: Springer.

- Cowart, Wayne 1997. *Experimental Syntax: Applying Objective Methods to Sentence Judgments*. London, New Delhi: Sage.
- Dąbrowska, Ewa 2010. *Masterclass in Experimental Methods in Linguistics*. Loengumaterjalid.
- Field, Andy; Hole, Graham 2006. *How to Design and Report Experiments*. London: Sage.
- Geeraerts, Dirk 2006. *Methodology in Cognitive Linguistics*. – Gitte Kristiansen, Michel Achard, René Dirven, Francisco J. Ruiz de Mendoza Ibáñez (Eds.). *Cognitive Linguistics: Current Applications and Future Perspectives. Applications of Cognitive Linguistics 1*. Berlin and New York: Mouton de Gruyter, 21–49.
- Geeraerts, Dirk 2010. *The Doctor and the Semanticist*. – Dylan Glynn, Kerstin Fischer (Eds.). *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches. Cognitive Linguistics Research 46*. Berlin, New York: Mouton de Gruyter, 63–78.
- Gibbs, Raymond W. Jr.; Colston, Herbert L. 1995. *The cognitive psychological reality of image schemas and their transformations*. – *Cognitive Linguistics*, 6 (4), 347–378. <http://dx.doi.org/10.1515/cogl.1995.6.4.347>
- Glynn, Dylan 2010. *Testing the hypothesis. Objectivity and verification in usage-based Cognitive Semantics*. – Dylan Glynn, Kerstin Fischer (Eds.). *Quantitative Cognitive Semantics: Corpus-Driven Approaches*. Berlin: Mouton de Gruyter, 239–270.
- Gonzalez-Marquez, Monica; Becker, Raymond B.; Cutting, James E. 2007. *An introduction to experimental methods for language researchers*. – Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson, Michael Spivey (Eds.). *Methods in Cognitive Linguistics. Human Cognitive Processing 18*. Amsterdam, Philadelphia: John Benjamins, 53–86.
- Gries, Stefan Th. 2009. *Statistics for Linguistics with R: A Practical Introduction*. Berlin, New York: Mouton de Gruyter. <http://dx.doi.org/10.1515/9783110216042>
- Jürine, Anni 2011. *Vormierinevused tähenduserinevuse ilminguna. Kaassõnähendite grammatiseerumine konstruktsioonides*. – *Keel ja Kirjandus*, 12, 899–916.
- Kantowitz, Barry H.; Roediger, Henry L., III; Elmes, David G. 2009. *Experimental Psychology*. 9th edition. Belmont: Wadsworth, Cengage Learning.
- Kirk, Roger E. 2013. *Experimental Design: Procedures for the Behavioral Sciences*. 4th ed. Thousand Oaks: Sage.
- Klavan, Jane 2008. *Medial Region Adpositions in English and Estonian: A Cognitive Grammar Perspective*. Magistritöö. Tartu: Tartu Ülikooli Kirjastus.
- Klavan, Jane 2012. *Evidence in Linguistics: Corpus-Linguistic and Experimental Methods for Studying Grammatical Synonymy*. *Dissertationes linguisticae Universitatis Tartuensis 15*. Tartu: Tartu Ülikooli Kirjastus.
- Klavan, Jane; Veismann, Ann; Jürine, Anni 2013. *Katselised meetodid tähenduse uurimisel*. – *ESUKA / JEFUL*, 4 (1), 17–34.
- Käärik, Ene 2011. *Empiiriline andmeanalüüs*. Loengumaterjalid.
- Lakoff, George 1987. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Luuk, Aavo 2012. *Eksperimentaalne psühholoogia*. Loengumaterjalid.
- MacLin, M. Kimberly; Solso, Robert L. 2007. *Experimental Psychology*. 8th ed. Boston: Pearson.
- Sandra, Dominiek; Rice, Sally 1995. *Network analyses of prepositional meaning: Mirroring whose mind – the linguist's or the language user's?* – *Cognitive Linguistics*, 6 (1), 89–130. <http://dx.doi.org/10.1515/cogl.1995.6.1.89>
- Schütze, Carson T. 1996. *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Chicago: University of Chicago Press.
- Schütze, Carson T.; Sprouse, Jon (ilmumas). *Judgment Data*. – Devyani Sharma, Rob Podesva (Eds.). *Research Methods in Linguistics*. Cambridge: Cambridge University Press.
- Sprouse, Jon; Almeida, Diogo (ilmumas). *Power in Acceptability Judgment Experiments and the Reliability of Data in Syntax*.

- Stefanowitsch, Anatol 2010. Empirical cognitive semantics: Some thoughts. – Dylan Glynn, Kerstin Fischer (Eds.). *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*. *Cognitive Linguistics Research* 46. Berlin: Mouton de Gruyter, 355–381.
- Sutrop, Urmas 2001. The list task and a cognitive salience index. – *Field Methods*, 13 (3), 263–276. <http://dx.doi.org/10.1177/1525822X0101300303>
- Tooding, Liina-Mai 2007. *Andmete analüüs ja tõlgendamine sotsiaalteadustes*. Tartu: Tartu Ülikooli Kirjastus.
- Uusküla, Mari 2008. *Basic Colour Terms in Finno-Ugric and Slavonic Languages: Myths and Facts*. *Dissertationes linguisticae Universitatis Tartuensis* 9. Tartu: Tartu Ülikooli Kirjastus.
- Vainik, Ene 2004. *Lexical Knowledge of Emotions: The Structure, Variability and Semantics of the Estonian Emotion Vocabulary*. *Dissertationes linguisticae Universitatis Tartuensis* 5. Tartu: Tartu Ülikooli Kirjastus.
- Veismann, Ann 2008. *Eesti kaassõnade tähenduste kirjeldamise probleemidest*. – *Keel ja Kirjandus*, 5, 335–352.
- Veismann, Ann 2009. *Eesti keele kaas- ja määrsõnade semantika võimalusi*. *Dissertationes linguisticae Universitatis Tartuensis* 11. Tartu: Tartu Ülikooli Kirjastus.
- Wilson Van Voorhis, Carmen; Morgan, Betsy 2007. Understanding power and rules of thumb for determining sample sizes. – *Tutorials in Quantitative Methods for Psychology*, 3 (2), 43–50.

## Võrgumaterjalid

- Tartu Ülikooli Inimuuringute Eetikakomitee. <http://www.ut.ee/et/teadus/eetikakomitee> (12.02.2013).
- Daniel S. Soperi statistika kalkulaator. <http://www.danielsoper.com/statcalc> (12.02.2013).
- William D. Duponti ja Walton D. Plummer Jr-i statistika kalkulaator. <http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/PowerSampleSize> (12.02.2013).

**Anni Jürine** (Tartu Ülikool) on uurinud morfosüntaktilisi protsesse keeles, erialased huvid seotud eelkõige tänapäeva eesti keele ja keele muutumisega.  
[anni.jyrine@ut.ee](mailto:anni.jyrine@ut.ee)

**Jane Klavan** (Tartu Ülikool) on uurinud eesti keele kohakaassõnade ja kohakäänete paralleelset kasutust. Peamiseks keeleteaduslikuks huviks on muuhulgas korpusandmete ja katselisel teel saadud keeleandmete võrdlus.  
[jane.klavan@ut.ee](mailto:jane.klavan@ut.ee)

**Ann Veismann** (Tartu Ülikool) tegeleb kognitiivse (ruumi)semantikaga, polüseemiaga, grammatiliste üksuste tähendustega. On uurinud eesti keele kaas- ja määrsõnade semantikat.  
[ann.veismann@ut.ee](mailto:ann.veismann@ut.ee)

# EXPERIMENTAL SEMANTICS: DESIGN AND IMPLEMENTATION

Anni Jürine, Jane Klavan, Ann Veismann

University of Tartu

In semantics, conducting corpus linguistic analyses is much more common than conducting experiments. There are various reasons for this bias in empirical methods, including, among other things, the conviction that conducting an experiment is very difficult and risky. The aim of the present article is to discuss some of the practical aspects involved in experimental semantics. The focus is on two crucial stages; planning and conducting linguistic experiments. The article is concerned, *inter alia*, with the following topics: postulating and operationalising hypotheses, the validity and reliability of an empirical quantitative study, different types of variables, the representativeness of the sample, simple random sample, sample size, choosing an experimental design, constructing and presenting experimental materials, recruiting and instructing participants, and ethical considerations in experimental research. The aim here is not to give a complete list of all the possible practicalities one may encounter when conducting linguistic experiments. The specific design and implementation of an experiment depends largely on the specific research question. The focus is instead on some of the basic aspects that every researcher needs to consider when doing experimental semantics.

One of the crucial points in designing any empirical study is the postulation and operationalisation of hypotheses and variables. Taking into account the complex nature of semantic studies, this aspect requires both creativeness as well as vast experience. At the same time, the quality of an experiment (its validity and representativeness) depends largely on the decisions the researcher makes in these early stages of research planning. Although it may seem tedious to spend too much energy and resources on detailed planning, it is worth paying meticulous attention to the crucial step of experimental design. Neither large quantities of data nor advanced statistical methods will save a linguistic experiment that suffers from dubious validity.

**Keywords:** experimental methods, semantics, cognitive linguistics