

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond
Tarkvarateaduse instituut

Greta Kalvi 153522IABM

**EESTI KEELEKESKKONNA
AGENDIPÕHINE SIMULATSIOON
PROGRAMMEERIMISVAHENDIGA
NETLOGO**

Magistritöö

Juhendaja: prof Kuldar Taveter

Konsultant: prof Martin Ehala,
Tartu Ülikool

Tallinn 2017

Autorideklaratsioon

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autor: Greta Kalvi

08.05.2017

Annotatsioon

Magistritöö eesmärk on lihtsustada eesti keele püsijäämise uurimist Eesti keelekeskkonnas, kavandades ja koostades intuitiivse agendipõhise simulatsiooni. Eelnevalt loodud simulatsioonid ei ole agendipõhised ja protsesside tulemus ei ole prognoositav. Lisaks ei võimalda eksisteeriv simulatsiooniprogramm keeleteadlastel ja teiste valdkondade ekspertidel eksperimenteerides analüüsida erinevate Eestis kõneldavate keelte elujõulisust ja keelevalikute muutumist globaliseerumise kontekstis.

Eesti keele püsijäämise uurimiseks tutvub töö autor sarnaste süsteemide koostamise meetodikatega ning mudelite abil saab analüüsida probleemi ja otsida võimalikke lahendusi.

Töö tulemusena selgub, millise meetodika põhjal ja kuidas kavandatakse mudelite abil probleemvaldkonna analüüs. Autor valideerib koostatud simulatsioonisüsteemi ning teeb esimesed järeldused. Magistritöö tulemusel valmiv simulatsioon on Tartu Ülikooli institutsionaalse uurimisteema "Eesti keele kestlikkus avatud maailmas" üheks osaks.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 52 leheküljel, 5 peatükki, 16 joonist, 18 tabelit ja koodi magistritööga kaasas oleval CD-l.

Abstract

Agent-based simulation of the language environment in Estonia with the programming tool NetLogo

The aim of this master thesis is to find a solution for the survival of the Estonian language in multi-lingual language environment, while designing and compiling an intuitive agent-based simulation. Previously created simulations are not agent-based and the process results are not predictable. In addition, the existing simulation program does not have the possibility for language researchers and other interested parties to analyze by experimenting with different Estonian speakers' vitality and change in language preferences in the context of globalization.

To study Estonian language survival the author will research the methodologies of similar systems and analyzes existing models to find a solution to the problem.

As a result of this thesis the methodology and the proposed model is designed using the analysis of the problem area and model design will be analyzed and described for use in the problem area study. The outcome of the master's thesis will be part of University of Tartu institutional research topic "Sustainability of Estonian in the Era of Globalization".

The thesis is in Estonian and contains 52 pages of text, 5 chapters, 16 figures, 18 tables and code on the CD which is added to the thesis.

Lühendite ja mõistete sõnastik

| | |
|---------------------------|--|
| Agendipõhine simulatsioon | Vastavalt kavandatud ja teostatud simulatsioon, mille abil saab uurida keerukaid süsteeme |
| Agent | Üksus, mis on keskkonnas, kus ta tegutseb ja mida ta mõjutab [1, p. 36] |
| AOM | <i>Agent-Oriented Modeling</i> , agentorienteeritud modelleerimine |
| BPMN | <i>Business Process Model and Notation</i> , notatsioon protsesside kirjeldamiseks |
| NetLogo | Agendipõhine programmeeritav modelleerimiskeskond |
| ODD | <i>Overview, Design Concepts, and Details protocol</i> , agendipõhiste mudelite kirjeldamise protokoll [2, p. 115] |

Sisukord

| | |
|--|----|
| 1 Sissejuhatus | 10 |
| 1.1 Taust ja probleem | 11 |
| 1.2 Ülesande püstitus | 11 |
| 1.3 Metoodika..... | 12 |
| 1.4 Ülevaade tööst | 13 |
| 2 Agentorienteeritud modelleerimine ja agendipõhine simulatsioon | 15 |
| 2.1 Agentorienteeritud modelleerimise ja simulatsiooni põhimõisted | 15 |
| 2.2 Agentorienteeritud modelleerimine ja simulatsioonisüsteemi kavandamine | 16 |
| 2.3 Agentorienteeritud modelleerimine ja simulatsioon keeleökoloogias..... | 18 |
| 2.4 Agendipõhise simulatsiooni platvormid..... | 20 |
| 3 Probleemvaldkonna analüüs ja simulatsioonisüsteemi disain | 22 |
| 3.1 Analüüsimudelid..... | 22 |
| 3.1.1 Eesmärgimudel | 23 |
| 3.1.2 Rollimudelid | 24 |
| 3.1.3 Organisatsioonimudel..... | 25 |
| 3.1.4 Domeenimudel | 25 |
| 3.2 Disainimudelid..... | 26 |
| 3.2.1 Agendimudel | 26 |
| 3.2.2 Agentide tutvuste mudel..... | 26 |
| 3.2.3 Suhtlusmudel | 27 |
| 3.2.4 Agentide teadmiste mudel | 28 |
| 3.2.5 Stsenaariumid | 29 |
| 3.3 Simulatsiooni tase..... | 33 |
| 3.3.1 Simulatsioonisüsteemi eesmärk..... | 34 |
| 3.3.2 Üksused, olekumuutujad ja kaalud..... | 34 |
| 3.3.3 Protsessi ülevaade ja ajakava..... | 35 |
| 3.3.4 Lähtestamine..... | 41 |
| 3.3.5 Sisendandmed..... | 43 |
| 4 Simulatsioonisüsteemi valideerimine | 44 |

| | |
|---|----|
| 4.1 Simulatsiooni parameetrite häälestamine ja määramine | 44 |
| 4.2 Valideerimise protseduur ja simulatsioonisüsteemi valideerimine | 46 |
| 4.3 Võrgustiku struktuuri mõjude analüüs | 56 |
| 4.4 Simulatsiooni rakendamise ja edasiarendamise võimalused | 58 |
| 5 Kokkuvõte | 60 |
| Kasutatud kirjandus | 62 |
| Lisa 1 – Simulatsioonisüsteemi seadistatud parameetrid | 64 |
| Lisa 2 – Simulatsioonisüsteemi kood | 65 |

Jooniste loetelu

| | |
|---|----|
| Joonis 1. Üldine eesmärgimudel..... | 23 |
| Joonis 2. Organisatsioonimudel..... | 25 |
| Joonis 3. Domeenimudel. | 25 |
| Joonis 4. Agentide tutvuste mudel..... | 26 |
| Joonis 5. Suhtlusmudel..... | 27 |
| Joonis 6. Agentide teadmiste mudel. | 28 |
| Joonis 7. Algandmete faili avamise protsess. | 36 |
| Joonis 8. Agentide lähtestamise protsess..... | 36 |
| Joonis 9. Partneri valimise protsess..... | 36 |
| Joonis 10. Suhtlemise protsess. | 38 |
| Joonis 11. Suhtluse kontrollimise protsess..... | 39 |
| Joonis 12. Suhtluse iteratsiooni protsess. | 40 |
| Joonis 13. Faili eksportimise protsess. | 41 |
| Joonis 14. Simulatsiooni algne vaade..... | 44 |
| Joonis 15. Katsete 1–4 võrgustiku struktuurid. | 51 |
| Joonis 16. Katsete 5–8 võrgustiku struktuurid. | 55 |

Tabelite loetelu

| | |
|--|----|
| Tabel 1. Vaatepunktide raamistik agendipõhise simulatsiooni kavandamiseks ja teostamiseks probleemvaldkonna jaoks..... | 17 |
| Tabel 2. Vaatepunktide raamistiku mudelid..... | 22 |
| Tabel 3. Starteri roll..... | 24 |
| Tabel 4. Partneri roll..... | 24 |
| Tabel 5. Agendi tüüp – Suhtleja..... | 26 |
| Tabel 6. Edukalt suhtlemise stsenaarium..... | 29 |
| Tabel 7. Suhtluse kontrollimise stsenaarium..... | 31 |
| Tabel 8. Tõenäosuste muutmise stsenaarium..... | 32 |
| Tabel 9. Suhtleja keeleoskuse muutmise stsenaarium..... | 33 |
| Tabel 10. Esialgsed keeleoskuste keskmised väärtused..... | 48 |
| Tabel 11. Katse 1 keskmised keeleoskuste väärtused..... | 48 |
| Tabel 12. Katse 2 keskmised keeleoskuste väärtused..... | 49 |
| Tabel 13. Katse 3 keskmised keeleoskuste väärtused..... | 50 |
| Tabel 14. Katse 4 keskmised keeleoskuste väärtused..... | 50 |
| Tabel 15. Katse 5 keskmised keeleoskuste väärtused..... | 53 |
| Tabel 16. Katse 6 keskmised keeleoskuste väärtused..... | 53 |
| Tabel 17. Katse 7 keskmised keeleoskuste väärtused..... | 54 |
| Tabel 18. Katse 8 keskmised keeleoskuste väärtused..... | 54 |

1 Sissejuhatus

Käesoleva magistr töö teemaks on eesti keele püsijäämise uurimiseks Eesti keelekeskkonna arvutipõhise simulatsiooni kavandamine ja koostamine selliselt, et selle mudeli abil oleks võimalik teha erinevaid sotsiolingvistilisi simulatsioonikatseid. Töö käigus analüüsitakse probleemvaldkonda kirjeldavate mudelite abil ning kavandatakse seeläbi simulatsioonisüsteem, mis töö käigus valideeritakse, kasutades sisendandmeid varasematest uuringutest. Hetkel on Eesti keelekeskkonnas uurimisel mitmeid probleeme, mida käsitletakse käesolevas magistr töö. Töö käigus otsitakse lahendusi Eesti keelekeskkonna praeguseks teostatud simulatsiooniprogrammi realisatsioonis olevatele puudustele, millest üks on puuduv agendipõhisus, mistõttu ei ole võimalik kirjeldada autonoomseid agente ning ei saa simulatsioonide abil ennustada nende vahelist ilmnevat käitumist. Samuti leitakse lahendus probleemile, millest tulenevalt keeleteadlastel ja teistel huvigruppidel hetkel puudub eksperimenteerimisvõimalus, kasutades simulatsiooni erisuguste parameetrite erinevaid väärtusi.

Käesolevas magistr töö käsitletavasse Eesti keelekeskkonda kuuluvad inimesed, kelle alaliseks elukohaks on Eesti ning kelle puhul uuritakse nende keeleoskusi kolme keele osas ning nende keelte omavahelist koostoimimist. Uuritavateks keelteks on riigikeel eesti keel, teine riigis räägitum keel vene keel ja globaliseerumise keel inglise keel.

Agendipõhine simulatsioon hõlmab modelleeritud iseseisvaid agente ning nende interaktsioonidest tulenevaid seoseid, mida saab tänu koostatud arvutisimulatsiooni muudetavatele parameetritele suunata ja jälgida.

Eesti keelekeskkonna püsijäämist globaliseerumise kontekstis uurib Tartu Ülikooli professor ja keeleteadlane Martin Ehala, kes juhib vastavat uurimisrühma Tartu Ülikoolis. Antud magistr töö juhendaja on Tallinna Tehnikaülikooli agentorienteeritud modelleerimise ja tarkvaratehnika professor Kuldar Taveter, kes osaleb nimetatud Tartu Ülikooli projekti uurimisrühmas.

1.1 Taust ja probleem

Eesti keele püsijäämiseks on vaja uurida erinevate keelte elujõulisust ja keelevalikute muutumist Eesti keelekeskkonnas globaliseerumise kontekstis.

Käesolev magistritöö on aktuaalne, sest üha globaliseeruv maailmas, milles on domineerivaks keeleks inglise keel, peavad väikesed riigid oma keele püsijäämiseks uurima teiste keelte mõju oma keelekeskkonnale.

Antud lõputöö on vajalik keeleteadlastele ja informaatikutele, kuna see pakub põhjalikke mudeleid ja seadistatavat simulatsiooni Eesti keelekeskkonna uurimiseks. Programmeerimiskeskonna NetLogo intuitiivne kasutajaliides võimaldab keeleteadlastel sooritada simulatsiooniekspimente ning simulatsioonisüsteemi kasutajaliidese abil muuta simulatsiooni parameetreid.

Magistritöö tulemusena valmib Eesti keelekeskkonna arvutipõhine mudel, mille abil saab analüüsida eesti keele dünaamikat ja püsijäämist ning erinevate keelte koosmõju Eestis. Samuti saab Eesti keelekeskkonna mudeli abil läbi simuleerida mitmesuguseid võimalikke tulevikustsenaariumeid.

Magistritöö valmib Tartu Ülikooli institutsionaalse uurimisteema “Eesti keele kestlikkus avatud maailmas“ raames.

1.2 Ülesande püstitus

Magistritöö eesmärgiks on kavandada ja koostada agendipõhine arvutisimulatsioon, mis võimaldaks keeleteadlastel ja teistel huvigruppidel analüüsida erinevate Eestis kõneldavate keelte elujõulisust ja keelevalikute muutumist globaliseerumise kontekstis. Magistritöö raames teostatava agendipõhise simulatsiooni abil on võimalik leida vastuseid järgmistele uurimisküsimustele:

1. Kuidas mõjutavad keelekasutajate omadused keelekasutajate omavahelist suhtlemist? Keelekasutajate omaduste all mõeldakse keeleoskust erinevate keelte osas, keelevalikute eelistust ja agendi konformismi (ehk suhtluses suhtluspartneri keelega kohanemise) taset.

2. Kuidas mõjutavad keelekeskkonna omadused keelekasutajate omadusi? Keelekeskkonna omadused on erinevates keeltes suhtlemist eelistavate keelekasutajate geograafiline jaotus Eestis.
3. Kuidas mõjutavad eesti ja vene keelt mitte oskavad keelekasutajad (näiteks pagulased ja välisüliõpilased) Eesti elanike keelevalikuid?
4. Millised on vene ja inglise keele mõjud peamiselt eesti keelt kõnelevale keelekeskkonnale?
5. Millised on eesti ja inglise keele mõjud peamiselt vene keelt kõnelevale keelekeskkonnale Eestis? Millised tegurid mõjutavad keelekeskkondi enim?
6. Millised on keelekasutajate võrgustike kujunemise protsessid?

Käesoleva magistr töö eesmärgi saavutamise erinevaid aspekte väljendavad järgmised alameesmärgid:

- Kavandada Eesti keelekeskkonna simulatsioonisüsteem, järgides probleemvaldkonna analüüsiks agentorienteeritud modelleerimise ehk AOM-metoodikat (*Agent-Oriented Modeling*) [1] ja agendipõhise simulatsiooni kavandamiseks ODD-metoodikat (*Overview, Design Concepts, and Details protocol*) [2], [3].
- Koostada Eesti keelekeskkonna simulatsioonisüsteem agendipõhiste simulatsioonide programmeerimise keskkonnas NetLogo.
- Kavandada ja seadistada loodud simulatsioonisüsteem, et selle abil oleks võimalik leida vastuseid uurimisküsimustele 1–2.
- Koostada ja seadistada loodud simulatsioonisüsteem, et selle abil oleks võimalik leida vastuseid uurimisküsimustele 3–6.

Kõige olulisemaks kvaliteedieesmärgiks magistr töö eesmärkide saavutamise juures on võimalikult adekvaatselt kujutada keelekasutajate omavahelist suhtlemist simulatsioonisüsteemi agentide omavahelise suhtlemise kaudu.

1.3 Metoodika

Kasutatavaks uurimismeetodiks on disainiteaduse meetod [4], sest töö tulemusena disainitakse keeruline tehis – agendipõhine simulatsioonisüsteem. Vastavalt disainiteaduse meetodile [4] rakendatakse olemasolevaid teadmisi (mis väljenduvad

teooriates, raamistikes ja metoodikates) simulatsiooni kavandamiseks ja loomiseks. Samas kajastab nimetatud mudeli loomine sotsiaalse keskkonna vajadusi, olles rakendatud selles keskkonnas. Magistriöös kasutatava simulatsiooni sotsiaalseks keskkonnaks on Eesti keelekeskkond, mis on loodud selgelt defineeritud parameetrite väärtustega isikutest, kes moodustavad suhtlusgruppe. Uurimaks Eesti keelekeskkonna evolutsiooni, rakendatakse käesolevas magistritöös ühelt poolt keeleökoloogia teooriaid, raamistikke ja meetodeid ning teiselt poolt kahte agentorienteeritud modelleerimise ja simulatsiooni metoodikat selleks, et kavandada ja töötada välja Eesti keelekeskkonna agendipõhine simulatsioon.

Mainitud Eesti keelekeskkonna agendipõhises simulatsioonis modelleeritakse agentidena keelekasutajaid. Simulatsiooni käigus pöörduvad agendid üksteise poole suhtlust algatava agendi poolt eelistatud keeles ning muudavad vajaduse korral oma keele valikut lähtudes suhtluspartneri eelistustest. Käesoleva magistritöö raames loodud Eesti keelekeskkonna agendipõhise simulatsiooni valideerimiseks kasutatakse keeleteadlaste poolt koostatud simulatsiooniekspimente, mis annavad uusi teadmisi sotsiolingvistiliste protsesside kohta, nagu näiteks keeltevaheline konkurents ja keelevahetus.

1.4 Ülevaade tööst

Magistritöö teises peatükis kirjeldab autor, mis on agentorienteeritud modelleerimine ja agendipõhine simulatsioon ning kuidas kasutatakse agentorienteeritud modelleerimist ja agendipõhist simulatsiooni keeleökoloogias. Teises peatükis selgitatakse, millised on agendipõhise simulatsiooni keskkonnad ja põhjendatakse, milline nendest on sobivaim antud töös ette seatud ülesannete lahendamiseks.

Kolmandas peatükis autor esitab ja kirjeldab agentorienteeritud probleemvaldkonda Eesti keelekeskkonna jaoks AOM-metoodika vaatepunktide raamistikus kirjeldatud tüüpi mudelite abil. Samas peatükis esitab autor probleemvaldkonna analüüsi põhjal Eesti keelekeskkonna agendipõhise simulatsiooni disaini ning kirjeldab simulatsiooni taset, millele annavad aluse eelnevalt kirjeldatud analüüsi- ja disainimudelid.

Neljandas peatükis määrab autor simulatsioonisüsteemi olulised parameetrid ning kirjeldab nende mõju simulatsiooni väljundile. Täiendavalt analüüsib autor Eesti

keelekeskkonna agendipõhise simulatsiooni käitumist erinevate simulatsiooni-
parameetrite väärtuste korral, valideerides valminud simulatsiooni algandmete ja
oodatava väljundi abil. Vastavalt eelnenud analüüsile toob autor välja mõned valminud
simulatsioonisüsteemi edasiarendamise võimalused ja selle süsteemi võimaliku kasu
keelekeskkonna dünaamika analüüsis.

2 Agentorienteeritud modelleerimine ja agendipõhine simulatsioon

Käesolevas peatükis antakse ülevaade, mis on agentorienteeritus, agentorienteeritud modelleerimine ja agendipõhine simulatsioon ning selgitatakse agendi mõistet. Tuuakse välja simulatsioonisüsteemi kavandamise meetodikad. Kirjeldatakse, kuidas kasutatakse agendipõhist modelleerimist keeleökoloogias. Tuuakse välja erinevad agendipõhise simulatsiooni platvormid, loetletakse olulisemad nõuded nendele ja selgitatakse, milline nendest vastab kõige paremini esitatud nõuetele.

2.1 Agentorienteeritud modelleerimise ja simulatsiooni põhimõisted

Agentorienteeritus on agentide ja objektide eristamine, mille kohaselt agendid on iseseisvad aktiivsed üksused, mis võivad tajuda sündmusi, täita tegevusi ja suhelda [1, p. 7], [5, p. 13]. Tavalised objektid on passiivsed üksused, millel ei ole agentidele omast võimekust [1, p. 7], [5, p. 13].

Agentorienteeritud modelleerimine on mingi probleemvaldkonna analüüs ja modelleerimine, mille keskseks üksuseks on agent ehk aktiivne olem [1]. Antud meetodika alusel modelleerimise eesmärgiks on sotsiaalsete ja tehniliste süsteemide, mis koosnevad inimestest ja tehnilistest elementidest, analüüsimine ja kavandamine [1], [6].

Agendipõhine simulatsioon on arvutipõhise simulatsiooni liik, mis hõlmab autonoomseid omavahel suhtlevaid agente ning nende suhtlusest tulenevaid dünaamilisi protsesse. Multiagentsüsteemid koosnevad agentidest, kes on iseseisvad ja autonoomsed ning suhtlevad üksteisega mingis keskkonnas [7, p. 1842].

Agendipõhine simulatsioon võimaldab simuleerida üksikute agentide käitumisest ja interaktsioonidest väljakujunevat niinimetatud ilmnevat käitumist. Agendipõhist simulatsiooni kasutatakse selleks, et simuleerida üksikute agentide otsustusprotsesse ning nendel põhinevat agentide sotsiaalset ja organisatoorset käitumist [8, p. 6].

Agendid on suletud funktsionaalsusega, mis tähendab, et agendi vorm ja struktuur on teiste agentide eest peidetud [9, p. 6], millest tulenevalt agent ei tea, millist infot teine agent valdab ja kuidas ta käitub. Agendipõhist simulatsiooni võib kasutada näiteks selleks, et simuleerida keerukaid süsteeme, milleks on ka antud lõputöös käsitletud sotsiaalsed võrgustikud.

Tüüpilisel agendipõhisel mudelil on kolm aspekti:

- Agendid, nende omadused ja käitumine;
- Agentide suhted ja suhtluse meetodid (kuidas ja kellega suheldakse);
- Agentide keskkond [8, p. 7].

Agent on üksus, mis on asetatud mingisse keskkonda, mida ta tajub sensorite abil (mille alla käib ka informatsiooni vastuvõtmine teistelt agentidelt), teostab tajutu põhjal loogilist järeldamist ning vastavalt selle tulemustele mõjutab keskkonda täiturite (*actuator*) kaudu [1, p. 36].

Iseseisev agent tegutseb vastavalt oma käitumisreeglitele, teadmistele ja suhtlemisele teiste agentidega. Agentide parameetrite väärtuste muutumine tähistab nende arengut simulatsiooni käigus ning nende õppimisvõimet [8, p. 9]. Agentide kohandumine sõltub nende parameetritest, mis ei ole otseselt vajalikud suhtluses, kuid mis mõjutavad agendi pikaajalist käitumist ehk arengut (näiteks keelekasutajat esitav agent kohandub suhtluses ja vastab suhtluspartneri keeles) [8, p. 9]. Individuaalne õppimine ja kohanemine (ehk akommodatsioon) on simulatsioonisüsteemis esitatud agendi parameetrina, mis muutub simulatsiooni käigus vastavalt agendi arengule [8, p. 9].

2.2 Agentorienteeritud modelleerimine ja simulatsioonisüsteemi kavandamine

Oma töös kasutab autor Eesti keelekeskkonna probleemvaldkonna kontseptuaalseks analüüsiks agentorienteeritud modelleerimise metoodikat [1] ning kontseptuaalse analüüsi põhjal Eesti keelekeskkonna agendipõhise simulatsiooni kavandamiseks ODD-metoodikat [2], [3]. Nimetatud metoodikate kohaselt teostatakse agentorienteeritud probleemvaldkonna analüüs ja koostatakse agendipõhine simulatsioon lähtudes kolmest abstraktsioonitasemest: analüüs, disain ja simulatsioon ning kolmest vaatest: informatsioon, interaktsioonid ja käitumine. Probleemvaldkonna analüüsi ning

agendipõhise simulatsiooni kavandamise ja teostamise kolm loogilist abstraktsioonitaset ühelt poolt ning informatsiooni, interaktsioonide ja käitumise vaated teiselt poolt võtab kokku tabelis 1 kujutatud vaatepunktide raamistik [1, p. 185], [10].

Tabel 1. Vaatepunktide raamistik agendipõhise simulatsiooni kavandamiseks ja teostamiseks probleemvaldkonna jaoks.

| Abstraktsiooni-tase | Vaatepunkti aspekt | | |
|---------------------|---|---|---|
| | Interaktsioonid | Informatsioon | Käitumine |
| Analüüs | Milliseid rolle peab simulatsioonimudel hõlmama? Mis on nende rollide kohustused ja piirangud? Millised on suhted rollide vahel? | Millist staatilist informatsiooni peab simulatsioonimudel hõlmama? Milline on selle informatsiooni struktuur? | Mis on simulatsioonimudeli eesmärk? Millised on simulatsioonimudeli eesmärgi saavutamise erinevaid aspekte esitavad alameesmärgid? Mis on simulatsioonimudeli kvaliteedieesmärgid? |
| Disain | Millised agenditüübid vastavad millistele rollidele? Millist tüüpi agendid suhtlevad millist tüüpi agentidega? Mis on agentide vahelise suhtluse protokollid? | Millised on agentide niinimetatud privaatsed olekumuutujad või atribuudid? Millised on erinevat tüüpi agentide poolt jagatud globaalsed muutujad? | Kuidas vastavad agentide tegevused eesmärkidele? Milliseid tegevusi sooritavad millist tüüpi agendid? Millises järjekorras sooritavad agendid tegevusi? Kuidas agentide tegevused muudavad agentide olekumuutujate ja globaalsete muutujate väärtusi? |
| Simulatsioon | Millised simulatsiooni-platvormi konstruktsioonid vastavad agenditüüpidele? Kuidas tekitada agenditüüpide eksemplare? Kuidas teostada simulatsiooni-platvormil agentidevahelisi interaktsioone? | Millised simulatsiooni-platvormi muutujad vastavad agentide olekumuutujatele ja globaalsetele muutujatele? | Kuidas teostada simulatsiooniplatvormil agentide tegevusi? Kuidas teostada simulatsiooniplatvormil agentide olekumuutujate ja globaalsete muutujate väärtuste muutmist tegevuste poolt? |

Vaatepunktide raamistiku (tabel 1) ridade ja veergude ristumiskohtadel esitatakse vastaval abstraktsioonitasemel ning vaates küsimusi, millele tuleb vastata simulatsioonisüsteemi kavandamisel ja teostamisel mingi antud probleemvaldkonna jaoks.

Tabelis 1 toodud informatsiooni vaade hõlmab agentide lokaalsete ja globaalsete muutujate kirjeldust. Agendi lokaalseteks muutujaks on näiteks keeleoskus erinevate keelte osas, keelevalikute eelistused ja agendi konformismi tase. Globaalsete muutujate kaudu kirjeldatakse muuhulgas kommunikatsiooni katkemise lävendi väärtus ning erinevate keelevalikute eelistustega agentide geograafiline jaotus, milleks on Eesti kontekstis vastavalt eesti- ja venekeelse suhtluse eelistajate ülekaal eri piirkondades. Interaktsioonide vaade esitab eri keeli eelistavate agentide omavahelisi interaktsioone, mis leiavad aset nende otsese suhtluse teel. Käitumise vaates on toodud käitumisreeglid, mis määravad agendi keelevaliku nende omavahelistes interaktsioonides ning ka võimalikud muutused keelevaliku eelistustes akumulunud interaktsioonide tulemusena.

2.3 Agentorienteeritud modelleerimine ja simulatsioon keeleökoloogias

Keele püsijäämise uurimisega tegeleb keeleökoloogia, mis käsitleb kõnelejat, keelt ja keelekeskkonda. Kõnelejas koondatakse tunnused, mida saab objekti küsitledes teada (näiteks vanus, sugu, rahvus, keelteoskus, keelekasutus ja keelevaliku eelistus) [11, p. 8]. Keele tasand kirjeldab keelt iseloomustavad tunnused nagu antud keele kõnelejate arv ja nende demograafiline käitumine [11, p. 8]. Keskkonnatasand keskendub konkreetse piirkonna keelekasutajate suhtluse muutumisele ning hõlmab näiteks keelte mainet ja erinevate keelekasutajate jaotust [11, p. 8]. Keeleökoloogia eesmärgiks on tuvastada tegurid, mis põhjustavad muudatusi kõneleja keelevalikutes. Kõneleja keelevalikuid mõjutavad nii keelekeskkond (näiteks teiste keeleoskus) kui ka tema enda seisukohad, mis kujunevad suhtlemisel (näiteks keele kasutamine erinevate suhtlejatega ja tema keeleoskused erinevate keelte osas). Välise keelekeskkonna mõjutuste alusel muudavad agendid oma esimest keelevalikut, mis pikas perspektiivis põhjustab keelevahetust [11, p. 9].

Keelte konkurents on laialt levinud nähtus, mis on pälvinud paljude keeleteadlaste ja antropoloogide tähelepanu. Keelte konkurentsi uurimiseks tuleb keskenduda üksikosade

vahelisele suhtlusele, mitte keelekeskkonna osadele [12, p. 22]. Sotsiaalse suhtluse moodustamiseks kasutatakse võrgustikke, milles vastavate ühenduste abil agendid saavad üksteisega suhelda. Keelevalik sõltub sotsiaalse võrgustiku suhetest ja võrgustikus kasutatavast keelest. Keelevaliku üks aspekt on keele lojaalsus ehk positiivne suhtumine keele päritolusse [12, p. 27]. Keeleökoloogia mudelid võtavad muutlikkuse parameetri alusel arvesse keelelojaalsust, mis kirjeldab, kuidas keeled konkureerivad ja kuidas neid suhtluses mõjutatakse [12, p. 28]. Madal keele dünaamilisus võib olla põhjustatud tugevast keelelojaalsusest ja üksikute agentide isikuomadustest, samas kui kõrge muutlikkus on tingitud vestluskaaslaste keelekeskkondade etnilisest erinevusest [12, p. 28]. Muutlikkus põhjustab agendi keelevahetust suhtluspartneri keelele ehk akommodatsiooni, aga kõik suhtlejad ei ole samal määral kohanemisvõimelised [12, p. 28]. Akommodatsioon on sotsiolingvistiliselt oluline mõiste, mis tähendab, et agent kohaneb suhtluses oma keelest erineva keelega, et vestluskaaslasega edukalt suhelda [12, p. 28]. Oluliseks suhtleja omaduseks on ka etnosuletus (*ethnic closure*) [13], mis väljendab sotsiaalset tendentsi, et inimesed eelistavad suhelda nendega, kes jagavad sama rassi ning sama kultuurilist ja etnilist tausta ehk inimesed kipuvad olema etniliselt suletud [14], [15]. Etnilist suletust mõjutab tugevalt ka erinevate agendigruppide ja nendega seotud teemade meediakajastus [12, p. 27].

Keelevalikut on eelnevalt modelleerinud statistikaarvutuste keeles J koostatud Eesti keelekeskkonna simulatsiooniprogrammi abil Andres Karjus [16]. Peamine erinevus selles teadustöös on agendipõhisuse puudumine, sest simulatsioonides ei kasutata autonoomseid üksusi – agente, mis simuleerivad keelekeskkonna üksusi, vaid seal on agendid ühe suure tabeli read ning tegemist on monoliitse programmiga. Sellest tulenevalt ei ole Andres Karjuse simulatsioonis võimalik kirjeldada autonoomseid agente, mistõttu keeleteadlastel ei ole võimalik jälgida agentidevahelistest interaktsioonidest tingitud käitumist, mida ei saa ette ennustada. Lisaks on mainitud simulatsiooniprogrammis modelleeritud vaid teatud piirkonna elanike suhtlemist omavahel, mitte nende interaktsioone ka väljaspool oma keelekeskkonda. Realistlikkuse huvides aga on interaktsioonide modelleerimine keelekeskkondade vahel vajalik.

Agendipõhine simulatsioon on vajalik vahend keerukate olukordade uurimiseks nagu keelevahetus ja keeltevaheline konkurents. Mainitud uurimismeetod modelleerib keelekasutaja keelevalikuid erinevate keelekeskkondade vahel kui ka keelekeskkonna sees. Agendipõhine simulatsioon võimaldab jäljendada erinevaid stsenaariume, sest

sellega saab muuta simulatsiooni käigus erinevate parameetrite väärtusi ning seeläbi analüüsida vastavate muutuste võimalikke mõjusid [11, p. 14].

Seevastu ülalpool kirjeldatud Andres Karjuse teadustöö realisatsioonis ei ole võimalik luua erinevate nuppude, liugmenüüde ja graafikutega intuiitivset ja vajadusel hõlpsalt muudetavat kasutajaliidest, mis on vajalik keeleteadlastele sisendite muutmiseks ja simulatsiooniekspriimentide sooritamiseks erinevatel algingimustel.

Selleks, et modelleerida keeleökoloogilisi protsesse nagu keelte säilimine ja kõrgeima keeleoskusega keele vahetus, tuleb esmalt modelleerida keelekasutajate käitumist ja eriti suhtluskeele valiku protsesse [17]. Keelevaliku protsesside modelleerimiseks tuleb analüüsida kõnelejat ja keskkonnategureid (nagu keelelävi suhtluses), mis mõjutavad kõneleja keele valikut erinevate keelte osas konkreetsetes suhtlussituatsioonis [17].

Varasemad uuringud on näidanud, et agendid, kes kuuluvad enamusrühma ja on rohkem mõjutatud eelarvamustest, pigem väldivad agente, kellel juba on vähemusrühmast sõpru ja seega neid ka ei tutvustata, mis tingib nende sõbrunemise vaid enamusrühma sõprade sõpradega [14, p. 127]. Sõbrunemine ei sõltu niivõrd isiklikest omadustest (rahvus või eelarvamus), kui tutvustavast kolmandast osapoolast [18], [19]. Sõprussuhted moodustuvad, sest agendid hindavad vastastikust sõprust ja suurema tõenäosusega sõbrunetakse oma sõbra sõbraga (*triadic closure*) [20], kes on nii antud agendile endale kui ka tema sõpradele sarnasemad oma omaduste ja eelistuste poolest [14, p. 128], [21, p. 278].

2.4 Agendipõhise simulatsiooni platvormid

Agendipõhise simulatsiooni platvormid, mis on laiemalt kasutusel, on Repast, Swarm, NetLogo ja Mason [22, p. 2]. Nimetatud platvormidest on Swarm esimene omalaadne, mis nõuab kõrgemal abstraktsioonitasemel koodi kirjutamist [8, p. 7]. Swarmi lihtsam versioon on Repast, mis põhineb programmeerimiskeelel Java ja on varustatud simulatsioonide näidetega [22, p. 5]. Repasti kiiremaks alternatiiviks on Mason, mis põhineb täielikult programmeerimiskeelel Java [22, p. 5]. NetLogo on rohkete näidetega platvorm, millel on sisseehitatud graafiline kasutajaliides [22, p. 14].

Simulatsioonikeskkonna valimisel lähtutakse järgmistest nõuetest: lihtne paigaldada, agentide käitumise vaadet on mugav visualiseerida, keskkonna keelt on lihtne õppida ja

kasutada ning platvorm sisaldab vahendeid agentidevaheliseks kommunikatsiooniks, teadete vahetuseks ja suhtlemiseks [23, p. 264]. NetLogo vastab mitmele nõudele – seda on lihtne alla laadida, keskkonna keel on funktsionaalne ja õppimine on suhteliselt kiire. Lisaks sisaldub NetLogo programmeerimiskeskonnas teek (*library*) paljude näidetega. Samuti võimaldab NetLogo visualiseerida arendatud süsteemi, mis on väga oluline multiagentsüsteemide puhul, sest soodustab paremat arusaamist simulatsioonist ning annab simulatsiooni kohta kohest tagasisidet [23, p. 266].

NetLogo peamised üksused on paigad (*places*), agendid ehk kilpkonnad (*turtles*), vaatleja (*observer*) ja ühendused (*links*). Vaatleja kontrollib eksperimenti, milles agendid ja paigad aktiivselt osalevad. Paigad on statsionaarsed agendid ehk võrgustiku komponendid, mille peal on agendid, kes suudavad liikuda – kilpkonnad (*turtles*). Nii paigad kui ka liikuvad agendid suudavad mõjutada neid ümbritsevat keskkonda, näiteks tuvastada teiste agentide olemasolu, vaadelda ümbritsevaid agente ja muuta keskkonda. Agente saab organiseerida kasutaja poolt määratud nimedega gruppidesse. Üks oluline tunnus, mis suurendab NetLogo simulatsioonivõimekust, on asjaolu, et iga paiga ja agendi jaoks võib kasutaja defineerida muutujaid. NetLogo võimaldab määrata ja kontrollida iga paiga ja agendi käitumist. NetLogo pakub palju erinevaid programmeerimise primitiive, tehteid ujukomaga ja juhuslike arvude genereerimise ning joonestamise võimekusi. Põhilised andmestruktuurid on loendid (*lists*) ning maatriks (*matrix*), mis suudavad lihtsasti talletada suuri andmemassiive. NetLogo toetab nii funktsioone, mida kutsutakse raporteerijateks, kui ka protseduure. Agente hallatakse üksuse poolt, mida nimetatakse vaatlejaks, mis käsib igal agendil täita konkreetset arvutuslikku ülesannet [23, p. 266]. NetLogo sobib agendipõhiste simulatsioonide loomiseks, sest iga agendi autonoomia avaldub ainult tema kohalike muutujatena ja kõik muu on globaalne, mis tähendab, et rakendus on lihtsalt hallatav ja kasutajaliideses on agendid sõltumatud vastavalt klassikalisele agentide määratlusele [1, p. 7], [8, p. 6]. NetLogo on vahend kiireks prototüüpimiseks ja esmaseks katsetamiseks multiagentsüsteemidega ning modelleeritud süsteemide animatsioonideks [23, p. 267].

Ülaltoodud põhjustel valis käesoleva töö autor NetLogo ka Eesti keelekeskkonna agendipõhise simulatsiooni loomise platvormiks.

3 Probleemvaldkonna analüüs ja simulatsioonisüsteemi disain

Käesolevas peatükis kirjeldatakse probleemvaldkonna analüüsi ja simulatsioonisüsteemi disaini agentorienteeritud modelleerimise metoodika alusel.

Agentorienteeritud modelleerimise keskmes on tabelis 2 esitatud vaatepunktide raamistik. Vaatepunktide raamistiku analüüsimudelid kirjeldavad antud probleemvaldkonnale püstitatud funktsionaalseid ja kvaliteedieesmärke, nende eesmärkide saavutamiseks vajalikke rolle ning süsteemis esitatavat teavet. Agentorienteeritud disainimudelid kirjeldavad, millist tüüpi agendid vastavad millistele rollidele, millist privaatsset ja jagatud teavet need agendid omavad ja töötlevad ning kuidas nad suhtlevad ja käituvad. Platvormist sõltuvad disainimudelid kirjeldavad simulatsioonisüsteemi disaini ja realisatsiooni konkreetsel platvormil [6, p. 210].

Tabel 2. Vaatepunktide raamistiku mudelid.

| | Vaatepunkti aspekt | | |
|--------------------|--|--------------------------|------------------|
| Abstraktsioonitase | Interaktsioonid | Informatsioon | Käitumine |
| Analüüs | Rollimudelid, organisatsioonimudel | Domeenimudel | Eesmärgimudel |
| Disain | Agendimudel, agentide tutvuste mudel, suhtlusmudel | Agentide teadmiste mudel | Stsenaariumid |
| Simulatsioon | Interaktsioonide spetsifikatsioonid | Muutujate kirjeldused | Protsessimudelid |

Antud peatüki järgmistes alapeatükkides on üksahaaval modelleeritud tabelis 2 esitatud mudelitüübid.

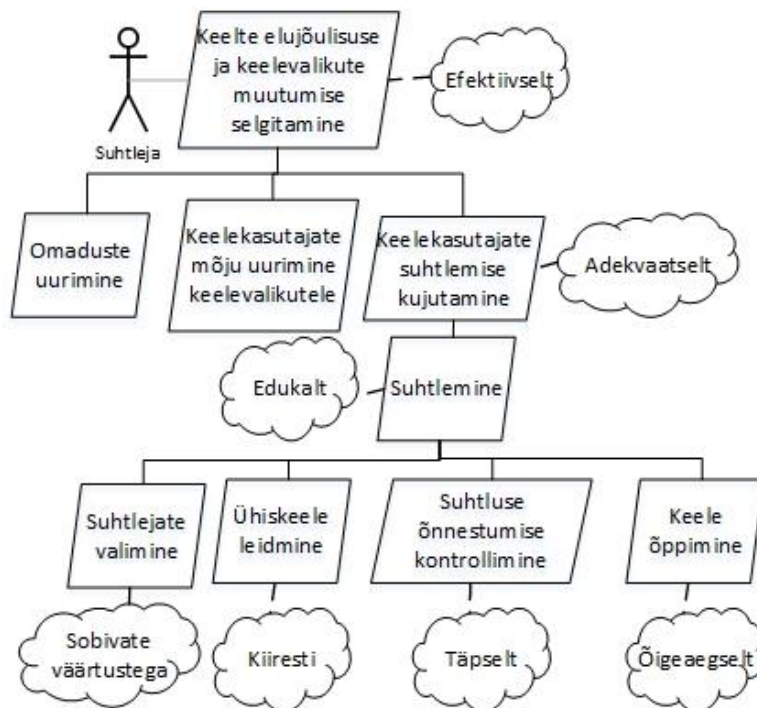
3.1 Analüüsimudelid

Analüüsimudelid vastavad tabelis 2 esitatud vaatepunktide raamistiku probleemvaldkonna analüüsi reale. Analüüsimudeliteks on eesmärgimudel, rollimudelid, organisatsioonimudel ja domeenimudel. Järgnevates alapeatükkides esitatakse üksahaaval

erinevat tüüpi probleemvaldkonna analüüsi mudeleid, kasutades Microsoft Visio tarkvara, mis võimaldab mudeleid esitada vaatepunktide raamistik [1, p. 185], [6, p. 210] kohaselt.

3.1.1 Eesmärgimudel

Eesmärgimudeli ülesanne on kirjeldada, milliseid funktsionaalseid eesmärke loodav süsteem tervikuna peab täitma, milliseid kvaliteediesmärke tuleb ühe või teise funktsionaalse eesmärgi täitmise juures silmas pidada ning millised on funktsionaalsete eesmärkide saavutamiseks vajalikud rollid. Nii eesmärke kui ka kvaliteediesmärke võib edasi jaotada väiksemateks alameesmärkideks ja -kvaliteediesmärkideks, kus madalama taseme eesmärk väljendab temaga vahetult seotud kõrgema taseme eesmärgi saavutamise mingit aspekti või tahku [1, p. 66]. Joonisel 1 on kujutatud magistritöö raames loodud simulatsioonisüsteemi eesmärgimudel. Sellelt mudelilt on näha, et olulisim eesmärk on Eestis kõneldavate keelte elujõulisuse ja keelevelikute muutumise selgitamine. Selle jaoks tuleb uurida keelekasutajate omadusi ja nende (eesti ja vene keelt mitte oskavate keelekasutajate) mõju keelevelikutele ning kujutada adekvaatselt keelekasutajate suhtlemist. Adekvaatne keelekasutajate suhtlemise kujutamine tähendab alati lihtsustust, mis toob esile keelte elujõulisuse ja keelekasutajate muutumise selgitamise seisukohalt olulise, milleks antud juhul on keelekasutajate eduka suhtlemise esitamine.



Joonis 1. Üldine eesmärgimudel.

Eduka suhtluse alameesmärkideks on suhtlejate valimine, ühise suhtluskeelega leidmine, suhtluse õnnestumise kontrollimine ning keele õppimine.

3.1.2 Rollimudelid

Rolli kirjeldab rollimudel, mis koosneb järgmistest elementidest:

- Rolli nimi, mis identifitseerib rolli;
- Rolli sõnaline kirjeldus;
- Kohustused, mida rolli täitev agent peab täitma, et saavutada eesmärged ning nendega seotud kvaliteedieesmärged;
- Piirangud, mida rolli täitev agent peab arvestama oma kohustusi täites [1, p. 73].

Peamised rollid suhtlemisel on suhtluse Alustaja ehk Starter ja Partner. Käesolevas magistritöös kasutatakse rollide nimetusi Alustaja ja Starter sünonüümidega. Tabelites 3 ja 4 kirjeldatakse vastavaid rolle detailsemalt. Tabel 3 esitab Starteri rolli.

Tabel 3. Starteri roll.

| | |
|-------------------|--|
| Rolli nimi | Starter |
| Kirjeldus | Suhtluse algataja roll |
| Kohustused | Valib suhtluseks partneri. Alustab suhtlust oma kõrgeima keeleoskusega keeles. Teise võimalusena pakub suhtluseks oma keeleoskuse tasemelt järgmise keele. |
| Piirangud | Edukaks suhtluseks peab keeleoskus ületama keelelääve. Mida kõrgem keeleoskus on suhtlejal, seda tõenäolisemalt suhtlus õnnestub. |

Suhtluse algataja valib endale suhtluspartneri ning saab alustada suhtlust oma kõrgeima keeleoskusega keeles ehk reeglina emakeeles. Kui suhtlus esimesel katsel ei õnnestu, pakub Starter oma keeleoskuse tasemelt järgmise keele. Selleks, et suhtlus õnnestu, peab suhtlejate keeleoskus ületama keelelääve. Tabel 4 kirjeldab Partneri rolli.

Tabel 4. Partneri roll.

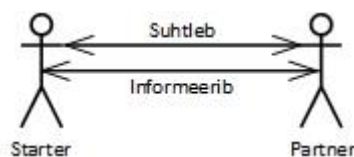
| | |
|-------------------|-----------------------------|
| Rolli nimi | Partner |
| Kirjeldus | Suhtluskatsele vastaja roll |
| Kohustused | Vastab suhtluse algatajale. |

| | |
|------------------|--|
| | Kui suhtluse algataja poolt pakutud keel ei sobi, siis pakub oma kõrgeima keeleoskusega keele. |
| Piirangud | Edukaks kiireks suhtluseks esimesel katsel peab suhtluse algataja kõrgeima keeleoskusega keele keelelävi olema piisavalt kõrgel tasemel. Mida sarnasem on partner suhtluse algatajaga ja mida kõrgem on partneri keeleoskus, seda tõenäolisemalt suhtlus õnnestub. |

Suhtlusele vastaja saab suhtluskatse algatajale vastata samas keeles või pakkuda oma kõrgeima keeleoskusega keelt. Edukaks suhtluseks peab mõlema suhtleja keeleoskus olema piisavalt kõrgel tasemel ehk ületama keeleläve.

3.1.3 Organisatsioonimudel

Organisatsioonimudel iseloomustab seoseid erinevate rollide vahel [1, p. 76]. Joonisel 2 on kujutatud süsteemi osapooled eri rollides. Jooniselt on näha seoste tüübid rollide vahel.

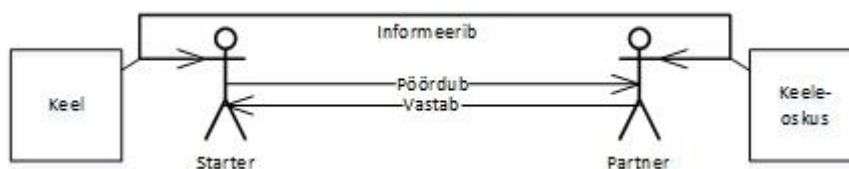


Joonis 2. Organisatsioonimudel.

Vastavalt joonisel 2 kujutatud organisatsioonimudelile suhtlejad Starter ja Partner suhtlevad omavahel, mille käigus vahetatakse informatsiooni oma keeleoskuste ja suhtluskeelte osas.

3.1.4 Domeenimudel

Domeenimudeli (joonis 3) eesmärgiks on esitada informatsiooni, mida uuritav süsteem tervikuna peab esitama ja käsitlema. Informatsioon koosneb valdkonnaolemitest ja nendevahelistest suhetest ning suhetest valdkonnaolemite ja rollide vahel [1, p. 76].



Joonis 3. Domeenimudel.

Vastavalt joonisel 3 kujutatud domeenimudelile, on Starter suhtluse algataja, kes pöördub esimesena suhtlemiseks suhtluspartneri poole, millele Partner vastab. Suhtlejad informeerivad üksteist suhtluskeelest ning teineteise keeleoskustest, et võimalikult

edukalt suhelda. Starter ja Partner suhtlevad keelekeskkonnas, mis sisaldab kõikide suhtlejate vahelisi seoseid, mis tulenevad keelekasutajate omadustest ja nendevahelistest suhtlusseoste tugevustest.

3.2 Disainimudelid

Disainimudelid vastavad tabelis 2 esitatud vaatepunktide raamistiku probleemvaldkonna disaini reale. Disainimudeliteks on agendimudel, agentide tutvuste mudel, suhtlusmudel, agentide teadmiste mudel ning stsenaariumid. Järgnevates alapeatükkides esitatakse ükshaaval erinevat tüüpi disainimudelid, mis on loodud kasutades samuti Microsoft Visio tarkvara.

3.2.1 Agendimudel

Agendimudel teisendab probleemvaldkonna analüüsi taseme rollid disainitaseme agenditüüpideks [1, p. 79]. Tabel 5 kujutab rollide teisendamist agenditüüpideks. Tabelis 5 on toodud agendi tüüp, sõnaline kirjeldus, antud agendi tüübiga seotud rollid ja agendi tüübi seos keskkonnaga.

Tabel 5. Agendi tüüp – Suhtleja.

| | |
|-------------------------|---|
| Agendi tüüp | Suhtleja |
| Kirjeldus | Keelekasutuses suhtlejat jäljendav (<i>emulating</i>) agendi tüüp |
| Rollid | Starter, Partner |
| Seos keskkonnaga | Suhtlejate seos keskkonnaga määratakse agente kirjeldavate parameetrite sarnasuse põhjal, mille ristmõju alusel kirjeldatakse agentidevaheline keskkond. Keskkonda mõjutavad ebaõnnestunud ja õnnestunud suhtlemised. |

Suhtleja on agenditüüp simulatsioonisüsteemis, mis jäljendab keelekasutajaid suhtluses.

3.2.2 Agentide tutvuste mudel

Agentide tutvuste mudel näitab, millist tüüpi agendid suhtlevad omavahel disainitavas simulatsioonisüsteemis [1, p. 81]. Joonisel 4 on esitatud seosed agentide vahel keelekasutajate suhtlemist kujutavas simulatsioonisüsteemis.

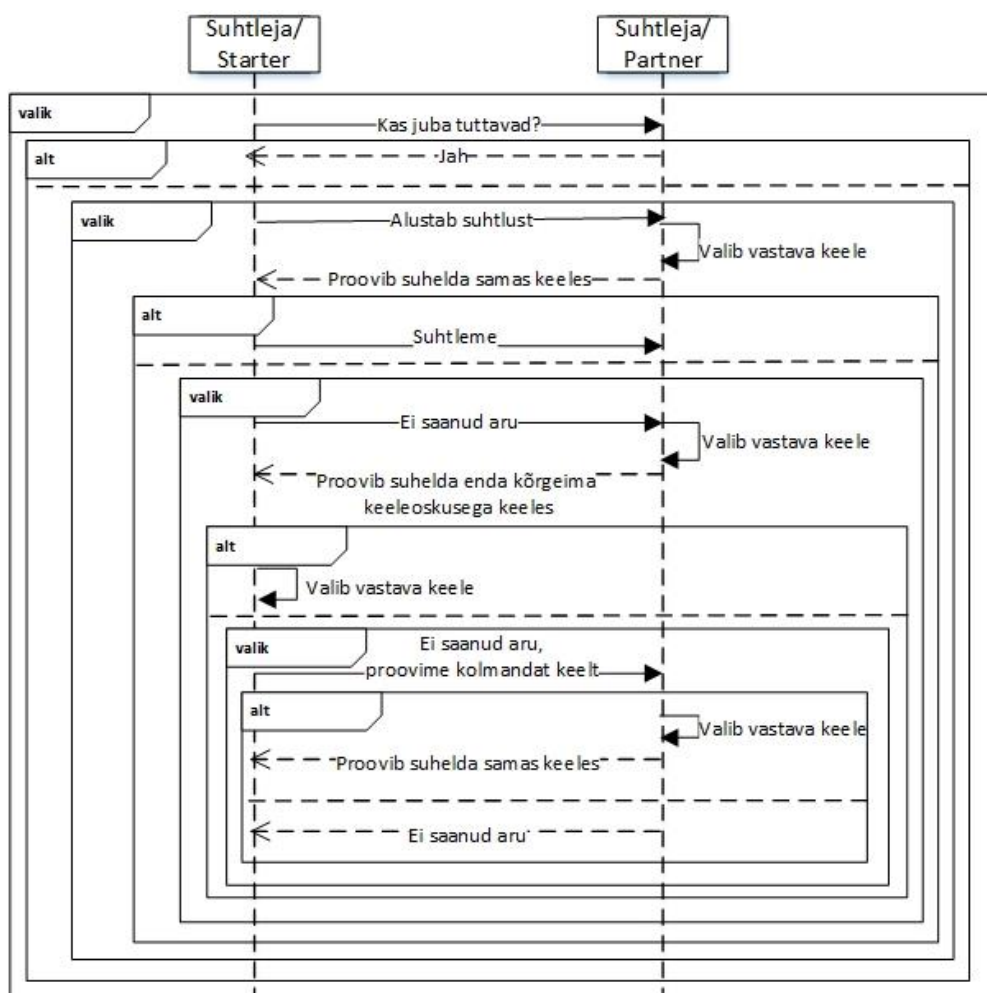


Joonis 4. Agentide tutvuste mudel.

Agentide tutvuste mudelilt on näha, kuidas suhtlemise protsessis saab üks agent korraga suhelda ainult ühe teise agentiga. Tulevastes uurimistöodes on võimalik laiendada agentidevahelist suhtlust simulatsioonisüsteemis nii, et üks agent saab korraga suhelda ka mitme teise agentiga, mis on oluline institutsionaalsete agentide (näiteks kooli ja riigi institutsioonide) puhul.

3.2.3 Suhtlusmodel

Agentide suhtlusmodel kirjeldab erinevate agentide omavahelist suhtlust kavandatavas simulatsioonisüsteemis [1, pp. 82, 87]. Joonisel 5 kujutatud suhtlusmodel kirjeldab informatsioonivahetuse toimimist erinevate agentide vahel protokollitüüpi interaktsioonimudelina.



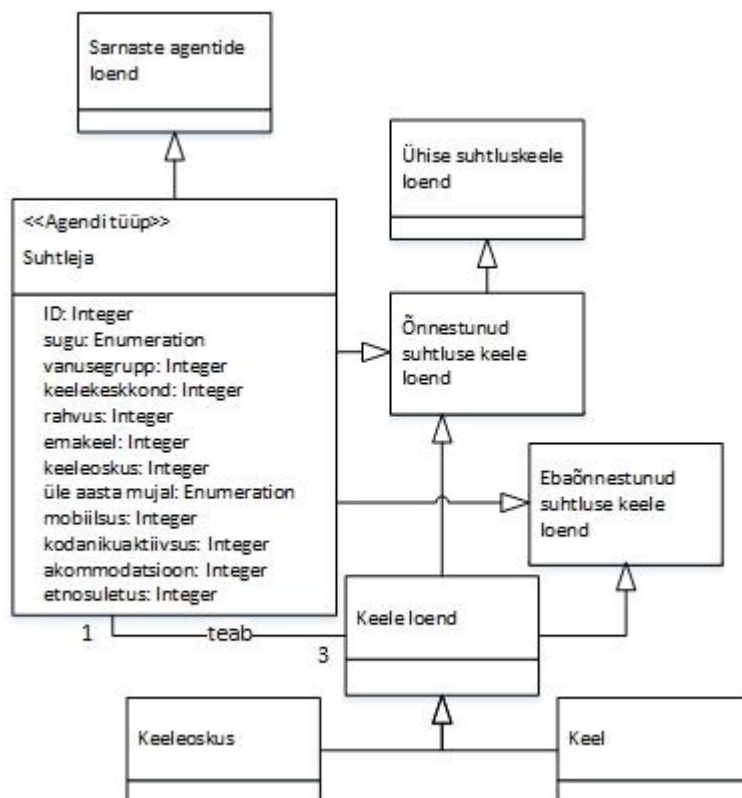
Joonis 5. Suhtlusmodel.

Vastavalt joonisel 5 kujutatud suhtlusmodelile, suhtluse alustab ja lõpetab suhtluse algataja. Kui keeleoskus ei ole suhtlemise õnnestumiseks piisaval tasemel, siis üritatakse

suhelda mõnes teises keeles kuni suhtluse õnnestumiseni või kuni suhtlemise iteratsiooni lõpuni.

3.2.4 Agentide teadmiste mudel

Agentide teadmiste mudel on domeenimudeli edasiarendus, mis kirjeldab erinevate agentide jagatud ja niinimetatud privaatsaid teadmisi [1, p. 90]. Joonisel 6 kujutatud agentide teadmiste mudel kirjeldab Suhtleja tüüpi kuuluvate agentide teadmisi.



Joonis 6. Agentide teadmiste mudel.

Joonisel 6 esitatud agentide teadmiste mudel koosneb kahte liiki teadmistest: agentide atribuutidest ja loenditest. Joonisel 6 kujutatud agentide teadmiste mudel esitab Suhtleja tüüpi kuuluvat keelekasutajat esindava agendi olulisi atribuute, mis omistatakse igale vastavale agendile. Samuti esitab agentide teadmiste mudel loendeid, mis on loodud suhtlemise jaoks või luuakse suhtlemise tulemusena. Sarnaste agentide loend moodustub agente kirjeldavate parameetrite sarnasuse põhjal ning koosneb agentide omavahelise suhtlemise tõenäosustest. Keele loendis on agentide keelte oskused ja keeled. Õnnestunud ja ebaõnnestunud suhtluse keele loendid koosnevad agentidest, kes on suhelnud ning keeltest, milles on suheldud. Ühise suhtluskeele loendis on agentide omavahelises suhtluses õnnestunud ühine keel. Suhtleja tüüpi agent ei saa teise suhtleja atribuute muuta,

kuid kaks agenti „teavad“ teineteise atribuutide väärtusi vaid siis, kui suhtluse algataja on suhtluspartneri välja valinud.

3.2.5 Stsenaariumid

Interaktsioonid agentide vahel on määratud nende tegevustega, mis sõltuvad agentide parameetrite väärtustest. Tegevuste sooritamist agentide poolt esitab agentide käitumismudel. Stsenaariumis on esitatud teatud agentide käitumismudel, mis kirjeldab, kuidas kindlaid rolle täitvad agendid saavutavad kavandatavale simulatsioonisüsteemile seatud eesmärgid. Vastavad eesmärgid on kirjeldatud eesmärkide mudeliga peatükis 3.1.1. Sarnaselt eesmärkidele võivad ka stsenaariumil olla alamstsenaariumid, mis vastavad eesmärgimudeli alameesmärkidele [1, p. 95].

Stsenaariumi saab modelleerida tabelina, mis näitab stsenaariumi eesmärki, algatajat ja päästikut ning kirjeldab stsenaariumi samme [1, p. 95]. Stsenaarium on ajendatud olukorrast, milles osaleb stsenaariumi algatanud agent. Stsenaariumis esitatavad päästikud (*trigger*) defineeritakse rollide kohustuste tabelite (peatükk 3.1.2) põhjal. Stsenaariumite modelleerimisel peab otsustama tingimused tegevuste ja alamstsenaariumite teostamiseks nagu näiteks järjestikku, paralleelselt, tsükliliselt jne. Stsenaariumid keelekasutajate suhtlust kujutava simulatsioonisüsteemi jaoks on esitatud tabelites 6–9. Iga stsenaarium esitab teatud tüüpi kuuluvate teatud rolle täitvate agentide tegevusi, kus iga tegevuse kohta esitatakse selle sooritamise tingimus (vajadusel), antud tegevuse edukaks sooritamiseks vajalikud teadmised ning antud tegevuse sooritamise jaoks olulised kvaliteedieesmärgid saadakse eesmärgimudelist [1, p. 98].

Tabel 6. Edukalt suhtlemise stsenaarium.

| Stsenaarium 1 – Edukalt suhtlemine | | | | | |
|---|--|--|-------------------------|------------------|--------------------------|
| Eesmärk | Suhelda edukalt | | | | |
| Algataja | Suhtleja | | | | |
| Päästik | Suhtluse algataja valib endale suhtluseks suhtluspartneri | | | | |
| Kirjeldus | | | | | |
| Tingimus | Samm | Tegevus | Agendi tüüp/Roll | Teadmised | Kvaliteedieesmärk |
| Vastavalt suhtlejate | 1 | Alusta suhtlust oma kõrgema keeleoskusega keeles | Suhtleja/ Starter | Keele loend | Kiiresti |

| Stsenaarium 1 – Edukalt suhtlemine | | | | | |
|---|--|---|---|--|------------------------------------|
| Eesmärk | Suhelda edukalt | | | | |
| Algataja | Suhtleja | | | | |
| Päästik | Suhtluse algataja valib endale suhtluseks suhtluspartneri | | | | |
| Kirjeldus | | | | | |
| Tingimus | Samm | Tegevus | Agendi tüüp/Roll | Teadmised | Kvaliteedi-eesmärk |
| parameetrite loendile | | | | | |
| | 2 | Vasta suhtlusele | Suhtleja/ Partner | Keele loend | Vastavalt Starteri keelele |
| Kui suhtlus ebaõnnestus | 3.1 | Proovi suhelda oma kõrgema keeleoskusega keeles | Suhtleja/ Partner | Keele loend | Edukalt |
| Kui Starteri akomodatsioon on piisav | 4.1 | Vasta suhtlusele | Suhtleja/ Starter | Keele loend | Vastavalt Partneri keelele |
| Kui suhtlus ebaõnnestus | 4.2 | Proovi suhelda oma keeleoskusest järgmise tasemega keeles | Suhtleja/ Starter | Keele loend | Edukalt |
| | 5 | Vasta suhtlusele | Suhtleja/ Partner | Keele loend | Vastavalt Starteri pakutud keelele |
| Tsükliliselt | 6 | Kontrolli suhtluse õnnestumist (stsenaarium 2) | Suhtleja/ Starter, Suhtleja/ Partner | Ühise suhtluskeele loend | Täpselt |
| | 7 | Muuda tõenäosusi järgmiseks suhtluseks (stsenaarium 3) | Suhtleja/ Starter, Suhtleja/ Partner | Ebaõnnestunud ja õnnestunud suhtluste keelte loendid | Täpselt |
| Kui keele muutmise periood on käes | 8 | Muuda keeleoskust (stsenaarium 4) | Suhtleja/ Starter, Suhtleja/ Partner | Keele loend | Õigeaegselt |

Tabel 6 kirjeldab, kuidas Starter alustab suhtlust oma kõige kõrgema keeleoskusega keeles, millele Partner vastab või pakub oma kõige kõrgema keeleoskusega keele. Sellele saab Starter omakorda vastata, kui tal on piisavalt kõrge akommodatsioon, mis tähendab vestluskaaslase keelega kohanemist, või pakkuda hoopis edukaks suhtluseks tarviliku keele. Suhtluse õnnestumist kontrollitakse (suhtlejal peavad keeleoskused ületama nõutava keeleläve) ning õnnestumisel muudetakse vastavalt tulemusele tõenäosusi järgmiseks suhtluseks ja vastava perioodi järel muudetakse keeleoskust (suurendatakse, kui keelt kasutatakse ehk õpitakse või vähendatakse, kui keelt ei kasutata või kasutatakse suhtluses ebaedukalt).

Tabel 7. Suhtluse kontrollimise stsenaarium.

| Stsenaarium 2 – Suhtluse õnnestumise kontrollimine | | | | | |
|---|---|---|---|---|---------------------------|
| Eesmärk | Kontrollida suhtluse õnnestumist vastavalt suhtluse tulemusele | | | | |
| Algataja | Suhtleja | | | | |
| Päästik | Suhtlejad on leidnud ühise suhtluskeele | | | | |
| Kirjeldus | | | | | |
| Tingimus | Samm | Tegevus | Agendi tüüp/Roll | Teadmised | Kvaliteedi-eesmärk |
| Kui keeleoskus on suurem kui keelelävi | 1.1 | Määra suhtlus õnnestunuks ja salvesta õnnestunud keel kasutatud keelena | Suhtleja/ Starter, Suhtleja/ Partner | Õnnestunud suhtluste keelte loend | Täpselt |
| Kui keeleoskus on väiksem kui keelelävi | 1.2 | Salvesta ebaõnnestunud keel kasutatud keelena | Suhtleja/ Starter, Suhtleja/ Partner | Ebaõnnestunud suhtluste keelte loend | Täpselt |
| | 2 | Salvesta keel ebaõnnestunud või õnnestunud suhtluste keelte loendisse | Suhtleja/ Starter, Suhtleja/ Partner | Ebaõnnestunud või õnnestunud suhtluste keelte loend | Täpselt |

Tabel 7 esitab suhtluse õnnestumise kontrollimist, mille jaoks on suhtlejal suhtluskeel valitud. Kui keeleoskus nii suhtluse algatajal kui ka suhtluspartneril ületab keeleläve, siis määratakse suhtlus nende vahel õnnestunuks, suhtluskeel salvestatakse kasutatud keelte loendisse ning keel salvestatakse õnnestunud keelte loendisse. Kui keeleoskus on liiga

madal ehk ei ületa määratud keeleläve, siis salvestatakse suhtlus suhtlejate vahel ebaõnnestununa ning suhtluskeel talletatakse ebaõnnestunud keelte loendisse.

Tabel 8. Tõenäosuste muutmise stsenaarium.

| Stsenaarium 3 – Tõenäosuste muutmine | | | | | |
|---|---|---|---|--------------------------------------|---------------------------|
| Eesmärk | Muuta tõenäosusi vastavalt suhtluse tulemusele | | | | |
| Algataja | Suhtleja | | | | |
| Päästik | Suhtlemine on lõppenud | | | | |
| Kirjeldus | | | | | |
| Tingimus | Samm | Tegevus | Agendi tüüp/Roll | Teadmised | Kvaliteedi-eesmärk |
| Kui suhtlus ebaõnnestus | 1.1 | Vähenda tõenäosust suhelda ebaõnnestunud suhtlejaga | Suhtleja/ Starter, Suhtleja/ Partner | Ebaõnnestunud suhtluste keelte loend | Täpselt |
| Kui suhtlus õnnestus | 1.2 | Suurenda tõenäosust suhelda õnnestunud suhtlejaga | Suhtleja/ Starter, Suhtleja/ Partner | Õnnestunud suhtluste keelte loend | Täpselt |
| Õnnestunud suhtluste keelte loend | 2 | Suurenda tõenäosust suhelda sõbra sõbraga | Suhtleja/ Starter, Suhtleja/ Partner | Õnnestunud suhtluste keelte loend | Täpselt |

Tabel 8 esitab tõenäosuste muutmist, mida tehakse peale iga suhtlust vastavalt suhtlemise tulemustele. Suhtlemise õnnestumisel järgmise suhtlemise tõenäosus suureneb, aga ebaõnnestumisel edasise suhtlemise tõenäosus langeb.

Tabel 9 esitab suhtlejate keeleoskuse muutmist, mis toimub perioodiliselt, kui teatav hulk suhtluste tulemusi on teada. Suhtluste tulemuste põhjal leitakse, kas Suhtleja peab oma keeleoskust tõstma või vähendama.

Tabel 9. Suhtleja keeleoskuse muutmise stsenaarium.

| Stsenaarium 4 – Suhtlejate keeleoskuse muutmine | | | | | |
|--|--|-----------------------------------|---|--------------------------|---------------------------|
| Eesmärk | Muuta keeleoskust vastavalt suhtluste tulemustele | | | | |
| Algataja | Suhtleja | | | | |
| Päästik | N arv suhtlemise iteratsioone | | | | |
| Kirjeldus | | | | | |
| Tingimus | Samm | Tegevus | Agendi tüüp/Roll | Teadmised | Kvaliteedi-eesmärk |
| Kui õnnestunud suhtlusi on rohkem kui ebaõnnestunud suhtlusi | 1.1 | Tõsta keeleoskust | Suhtleja/ Starter, Suhtleja/ Partner | Keele loend | Õigeaegselt |
| Kui ebaõnnestunud suhtlusi on rohkem kui õnnestunud suhtlusi | 1.2 | Vähenda keeleoskust | Suhtleja/ Starter, Suhtleja/ Partner | Keele loend | Õigeaegselt |
| | 2 | Uuenda ühise suhtluskeele loendit | Suhtleja/ Starter, Suhtleja/ Partner | Ühise suhtluskeele loend | Õigeaegselt |

Tabelis 9 esitatud keeleoskuse muutmise järel uuendab Suhtleja ühise suhtluskeele loendit, mida kasutatakse järgmisel iteratsioonil kontrollina. Järgmisel iteratsioonil on suhtluse osapooled samad, kuid nende rollid võivad olla vahetunud (suhtluse algataja võib ise osutada suhtluspartneriks).

3.3 Simulatsiooni tase

Simulatsioon tuleb disainida selliselt, et iga analüüsimudelites modelleeritud rolli võiks täita inimene või tarkvaraagent. Eksperimendid on näidanud, et agentorienteeritud analüüsi- ja disainimudelid, mis on esitatud ning selgitatud vastavates peatükkides 3.1 ja 3.2, annavad piisava aluse agentorienteeritud simulatsiooni teostamiseks [6]. Vaatepunktide raamistiku (tabel 2) simulatsioonisüsteemi taset kirjeldavad alapeatükid

3.3.1–3.3.5, kus rakendatakse simulatsioonisüsteemi kirjeldamiseks ODD-metoodikat [24]. Agentide tüüpidele vastavad peatükis 2.4 valitud simulatsiooniplatvormil NetLogo [25] agendid ehk kilpkonnad (*turtles*). Agenditüüpide eksemplare saab luua nimetatud simulatsioonisüsteemis lihtsa käsuga (*create turtles 1*), mida saab täita tsükliliselt andmete sisselugemisel. Kuna NetLogo ei võimalda otse kirjeldada agentidevahelisi interaktsioone, saab NetLogo simulatsiooniplatvormil agentidevahelisi interaktsioone teostada globaalsete muutujate kaudu (vaata alapeatükki 3.2.1). Seda saab teha tingimuslausetega, mille kohaselt suhtlejate vahel, kelle keeleoskused ületavad keeleläve ja kelle suhtlus ei ole veel õnnestunud, määratakse suhtlus õnnestunuks (*set communication-successful? true*). Seejärel salvestatakse kasutatud suhtluskeel vastavasse loendisse, kus seda kasutatakse järgnevate iteratsioonide juures. Kui erinevate suhtluste korral ei ületata keeleläve, siis määratakse suhtlus ebaõnnestunuks ning talletatakse kasutatud suhtluskeeled ebaõnnestunud suhtluse loendisse.

3.3.1 Simulatsioonisüsteemi eesmärk

Simulatsioonisüsteemi eesmärgiks on uurida (peatükk 3.1.1) Eesti keelekeskkonnas kolme keele (eesti, vene, inglise) omavahelist koostoimimist ja vastastikmõju. Simulatsioonisüsteemi abil soovitakse välja selgitada, kuidas erinevad hoiakud ja keeleoskused mõjutavad erineva emakeelega inimeste omavahelist suhtlemist. Simulatsioonisüsteem kasutab Eestis kogutud andmeid kolme konkureeriva keele kohta: riigikeel eesti keel, kohalik vähemuskeel vene keel ja globaliseerumise keel inglise keel. Algandmed on kogutud sotsioloogilise küsitlusega, millele on vastanud üle 1 000 inimese. Teostatud simulatsioonisüsteemi parameetrid ja mudeli sisend lähtuvad vastavatest küsitluse tulemustest.

3.3.2 Üksused, olekumuutujad ja kaalud

Simulatsioonisüsteem koosneb agentidest (peatükk 3.2.1), kes esindavad erinevate parameetritega keelekasutajaid. Iga agenti iseloomustavad järgmised muutujad:

- Muutujad, mis on kirjeldatud lähtestamisel (peatükk 3.3.4) kui vastavate agentide olekumuutujad (peatükk 3.2.4). Lähtestamisel kirjeldatud muutujate osakaale saab määrata simulatsioonisüsteemi kasutajaliidese abil.

- Õnnestunud suhtlust esitav loogikamuutuja kui globaalne muutuja simulatsioonisüsteemis. Nimetatud loogikamuutuja väärtust (tõene või väär) uuendatakse iga suhtluskatse lõpul.

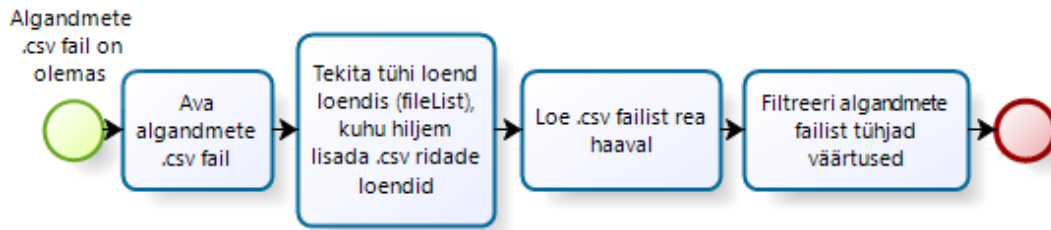
Kirjeldava simulatsioonisüsteemi ajaskaalal on üheks ajasammuks üks suhtlus ning simulatsioonis tehakse suhtluskatseid vastavalt kasutajaliidese abil valitud suhtluste arvule.

3.3.3 Protsessi ülevaade ja ajakava

Simulatsiooni käigus saavad agendid oma tegevusi teostada vastavalt käesolevas peatükis kirjeldatud protsessijoonistele, arvestades peatükis 3.2.3 kirjeldatud agentide suhtlusmudelit ja peatükis 3.2.5 kirjeldatud stsenaariume. Protsessijooniste abil esitatakse tegevuste järjekord ning olekumuutujate väärtuste muutumine. Simulatsioonisüsteemi protsessid on kirjeldatud kasutades Bizagi Modeler tarkvara, mis on laialt levinud protsesside modelleerimise tööriist ja vastab BPMN (*Business Process Model and Notation*) notatsioonile [26, p. 79]. NetLogo koodis on vajalik enne konkreetsele agendile käsu andmist pöörduda tema poole vastava NetLogo keelekonstruktsiooniga (*ask starter-self [do something]*). Kui simulatsiooniplatvormil on vaja teostada keskkonna poolt genereeritavaid sündmusi, siis ei ole vaja agendi poole pöörduda, vaid seda saab NetLogo koodis otse väljendada (näiteks *set common_language_list []*). Agentide muutujate väärtusi saab simulatsioonisüsteemis uuendada muutuja vana väärtust uue väärtusega üle kirjutades (näiteks *set success-lang 1*).

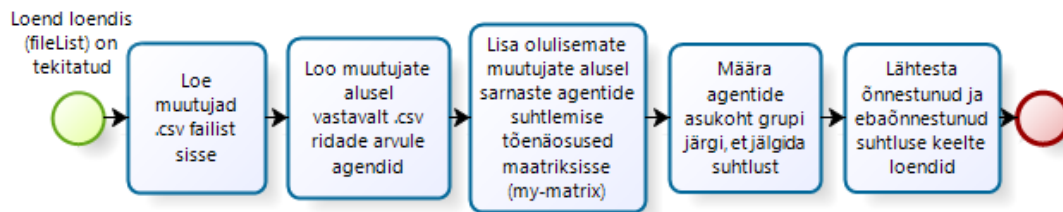
Järgnevalt on toodud simulatsioonisüsteemis olevate keelekasutajate võrgustike kujunemise protsessid algandmete väärtuste saamisest kuni suhtlusest tulenevate andmete ekspordini.

Joonis 7 illustreerib simulatsioonisüsteemi initsialiseerimist. Algtingimusena loetakse kõigepealt agentide muutujate väärtused küsitluse vastuste algandmete failist reahaaval simulatsioonisüsteemi vastavasse loendisse ning seejärel filtreeritakse algandmete failist välja tühjad väärtused.



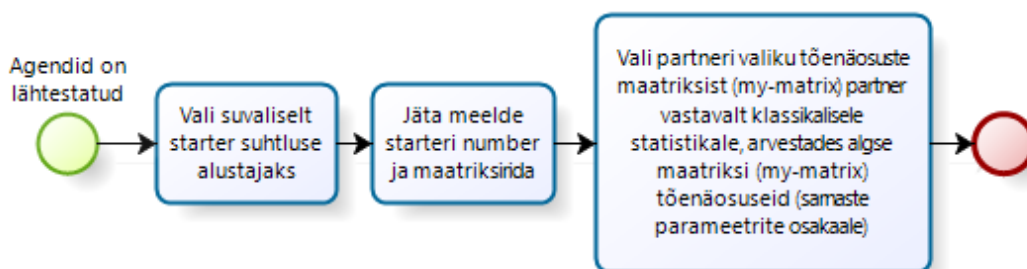
Joonis 7. Algandmete faili avamise protsess.

Joonis 8 illustreerib, kuidas järgnevalt luuakse vastavalt sisseloetud küsitluse ridadele agendid, millest igaüks esindab ühte suhtlejat. Iga agendi muutujatele omistatakse konkreetsed väärtused, vastavalt parameetrite väärtustele agendile vastavas küsitluse reas. Simulatsioonis agentide suhtluseoste loomiseks omistatakse sarnaste parameetrite väärtustega agentidele suhtlemise tõenäosused, mille ulatus sõltub parameetrite seadistatud mõju ulatusest. Agentide suhtlemise tõenäosused, mis saadakse sarnasuse alusel, lisatakse maatriksisse (*my-matrix*), mis sisaldab kõigi simulatsiooni agentide tõenäosusi. Maatriksis määratakse igale agendile asukoha number, et paremini suhtlust jälgida.



Joonis 8. Agentide lähtestamise protsess.

Kui agendid on lähtestatud, valitakse suhtluse algatajaks Starter ning talle valitakse Partner tõenäosuste maatriksist vastavalt klassikalisele statistikale, arvestades algse maatriksi (*my-matrix*) tõenäosuseid ehk sarnaste parameetrite osakaale. Partneri valiku protsessi illustreerib joonis 9. Ühe iteratsiooni jooksul sooritatavate suhtluste arv on konfigureeritav.



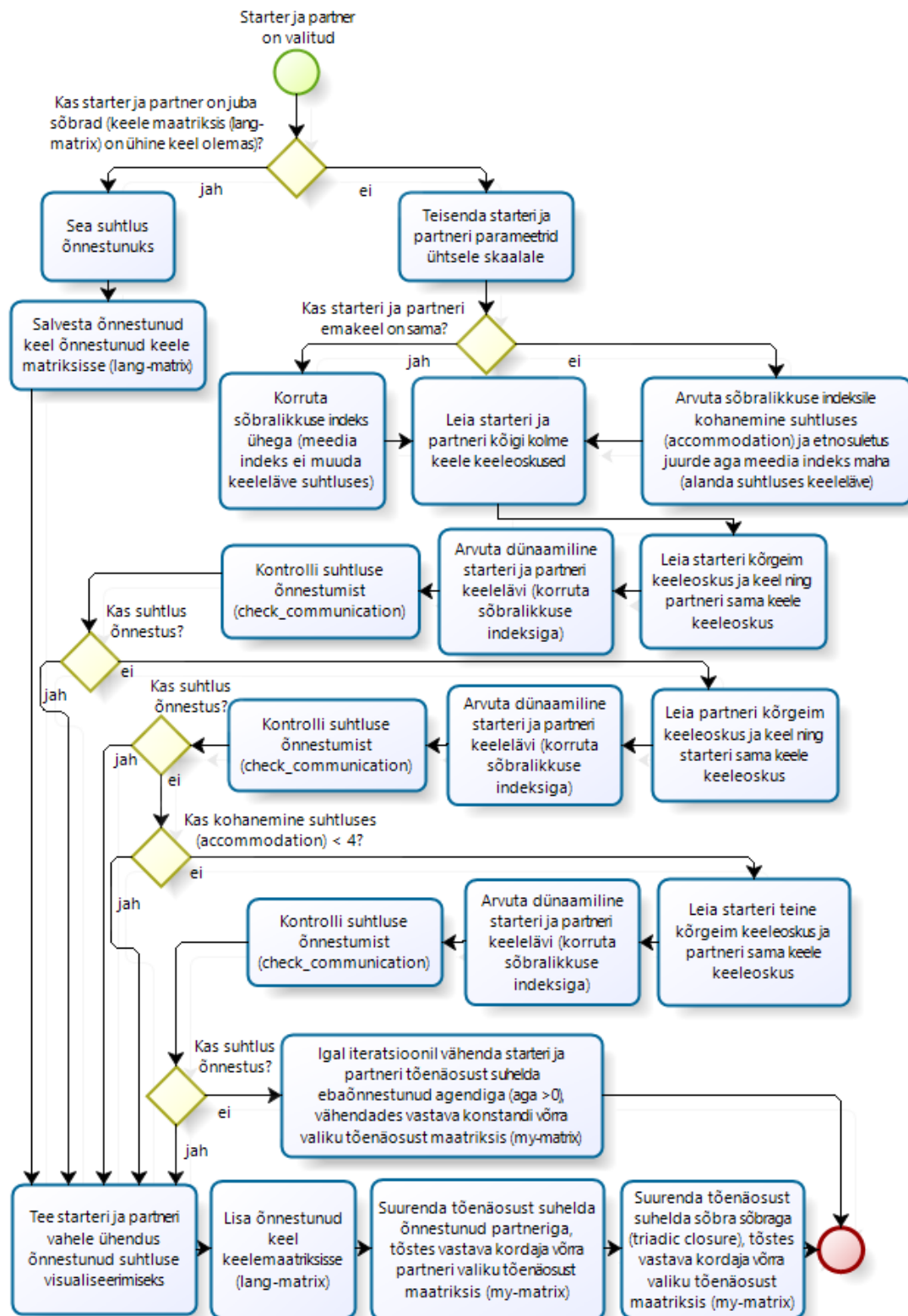
Joonis 9. Partneri valimise protsess.

Enne suhtluse algust teostatakse ennetav kontroll, kas Starter ja Partner on omavahel juba edukalt suhelnud, kasutades keelemaatriksit (*lang-matrix*). Juhul kui agendid on varem edukalt suhelnud, siis loetakse suhtlus õnnestunuks ning suurendatakse igal vastaval iteratsioonil nende valiku tõenäosust edasiseks omavaheliseks suhtlemiseks ja Partneri edukate suhtluste (sõbra sõbraga suhtlemise) tõenäosust. Juhul kui valitud Starter ja Partner ei ole omavahel varem edukalt suhelnud, siis käivitub edasine suhtlemise protsess. Järgmiseks hakkavad valitud agendid omavahel suhtlema. Nende suhtlemine toimub alljärgneva algoritmi alusel, mis sisaldab kolme keele korral kuni kolm tsüklit:

- Starter pakub oma kõige suurema keeleoskusega keele (reeglina emakeele), mille Partner võtab vastu või keeldub;
- Partner pakub oma kõige suurema keeleoskusega keele (reeglina emakeele), mille Starter võtab vastu või keeldub;
- Starter pakub oma keeleoskuse tasemelt järgmise keele, mille Partner võtab vastu või keeldub.

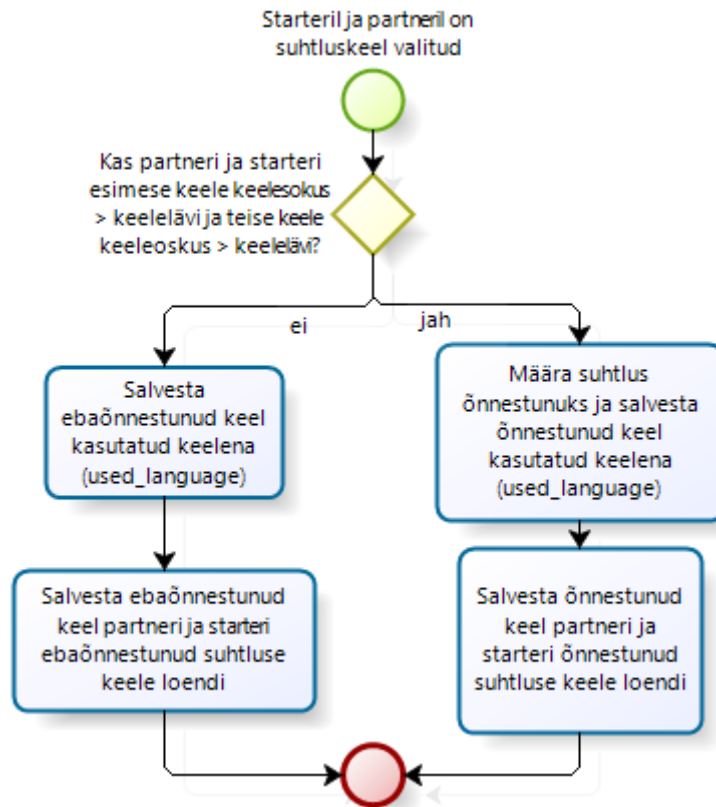
Suhtlemisel pakutud keele vastuvõtmine ja keeldumine sõltub agendi karakteristikist ehk eelkõige keeleoskusest aga ka sellest, kui avatud ta on erinevate rahvusgruppide suhtes (see tähendab agendi etnosuletuse ja akommodatsiooni väärtustest). Vastavalt Starteri ja Partneri omadustele leitakse erinevate keelekeskkondade vaheline keelelävi.

Suhtlemise protsessijoonis (joonis 10) kirjeldab, kuidas suhtlus simulatsioonisüsteemis agentide vahel saab staatuse õnnestunud või ebaõnnestunud. Õnnestunud suhtluse korral tekitatakse agentide vahele ühendus, et kasutajaliideses visuaalselt tähistada õnnestunud suhtlus ja selle partnerid. Samuti tõstetakse edukalt suhelnud agentidevahelist suhtlemise tõenäosust ning edukas suhtlemine ja suhtlemiseks kasutatud keel tähistatakse keelemaatriksis (*lang-matrix*) järgmiste iteratsioonide jaoks. Kui suhtlus kõigi kolme keele korral ebaõnnestub, siis vähendatakse Starteril tõenäosust suhelda ebaõnnestunud Partneriga, vähendades suhtluspartneri valiku tõenäosust Partneri valiku maatriksis (*my-matrix*).



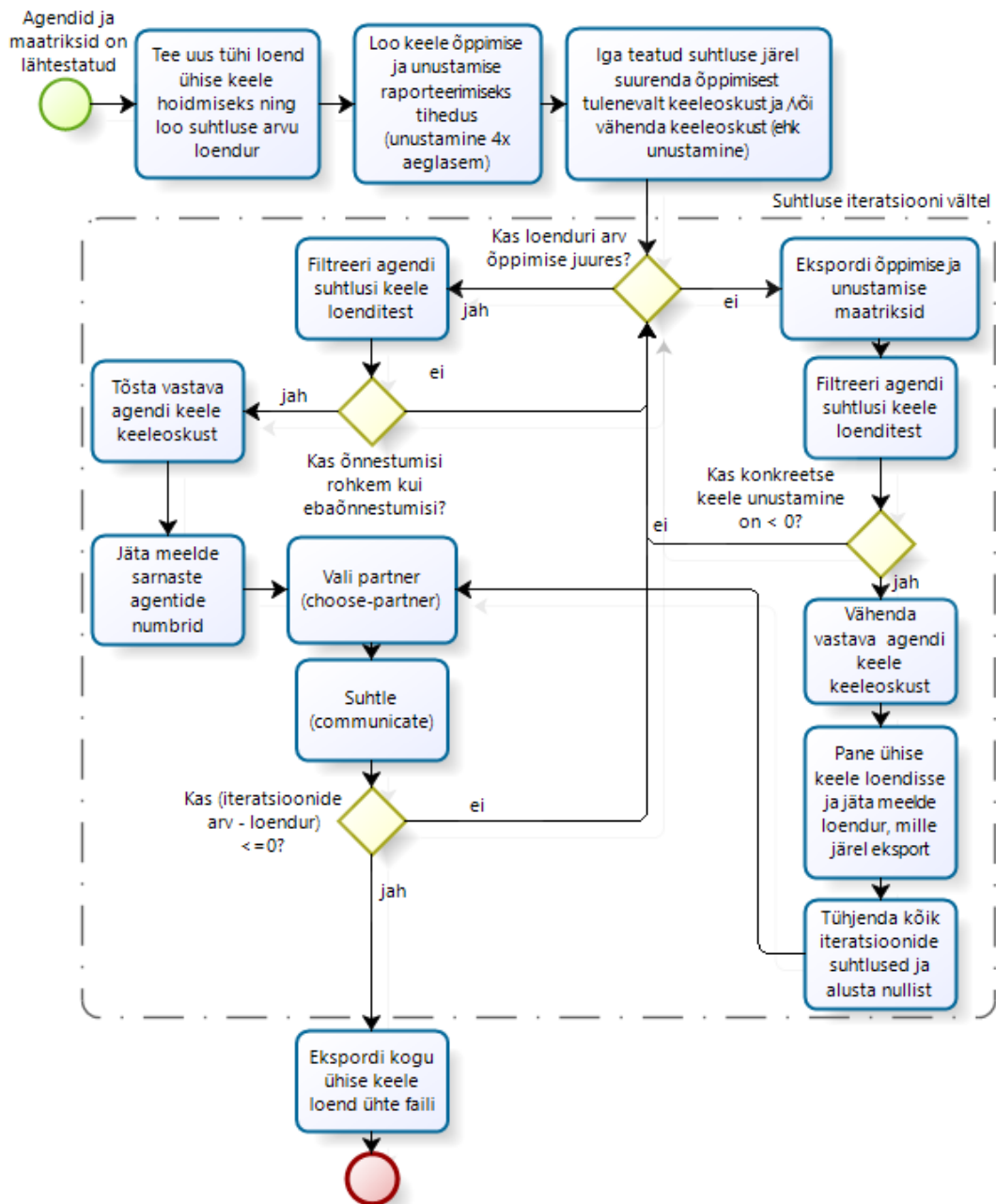
Joonis 10. Suhtlemise protsess.

Suhtlemise õnnestumist või ebaõnnestumist kontrollitakse keeleläve alusel, mida kirjeldab täpsemalt joonis 11.



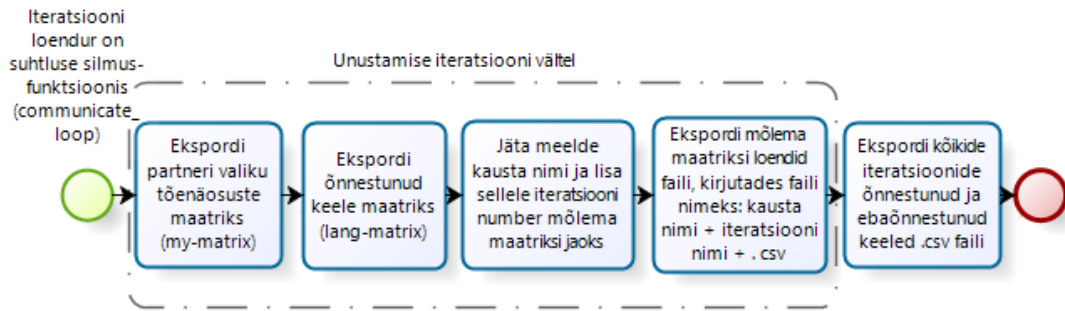
Joonis 11. Suhtluse kontrollimise protsess.

Joonis 12 illustreerib, kuidas simulatsioonisüsteemis toimub suhtlus paljude iteratsioonide vältel. Loenduri abil on seadistatud keele õppimise ja unustamise tsüklid. Õppimistsükli korral tõuseb keelesokus, kui keelt on kasutatud rohkem kui üks kord $\frac{n}{4}$ iteratsiooni jooksul vastava keelekasutaja korral, kus n tähistab unustamise iteratsioonis toimuvate suhtluste arvu. Keele unustamine ehk hääbumine on aeglasem kui keele õppimine ehk omandamine, mistõttu unustamise tsüklil on esialgu valitud toimuma 4 korda aeglasemalt kui õppimise tsüklil. Kui agent on keelt edukalt kasutanud n iteratsiooni jooksul suhtluses rohkem kui ebaedukalt, siis keelt õpitakse, mis tähendab, et keelesokust tõstetakse parameetri väärtuse võrra, mida saab simulatsioonisüsteemi kasutajaliidese abil muuta. Keelesokuse tase langeb vastava teise parameetri väärtuse võrra, mis on samuti kasutajaliidese abil muudetav, kui keelt on kasutatud ebaedukalt või kui seda keelt ei ole suhtluses kordagi unustamise iteratsiooni jooksul kasutatud. Vastav kontroll tehakse iga keelekasutaja kõikide keelte osas. Seega, kui suhtlus on simulatsioonisüsteemis tugevalt rahvuskeskne, siis nõrgeneb kõikidel agentidel võõrkeelteoskus.



Joonis 12. Suhtluse iteratsiooni protsess.

Kui suhtlemise iteratsioonid on jõudnud lõpule, siis eksporditakse (joonis 13) kogu ühise keele loend (*lang-matrix*), kasutatud keelte loendid (ehk õnnestunud ja ebaõnnestunud keelte loendid) ning keelekeskkonda kirjeldav tõenäosuste maatriks (*my-matrix*).



Joonis 13. Faili eksportimise protsess.

Faili eksportimine käivitatakse ka iga unustamise iteratsiooni järel, mil eksporditakse kõikide iteratsioonide nii õnnestunud kui ka ebaõnnestunud keeled vastavasse faili.

3.3.4 Lähtestamine

Simulatsioonisüsteem lähtestatakse küsitluse algandmete failist, millest loetakse sisse sisendandmete muutujad koos väärtustega ning need omistatakse agentide vastavatele muutujatele. Iga simulatsiooni alguses on sama seis, kuid kasutajaliideses saab muuta parameetreid ning nende väärtustega eksperimenteerides saab muuta simulatsiooni tulemusi. Iga üksikut agent'i iseloomustavad algselt järgmised muutujad ja nende võimalikud väärtused:

- Sugu – agent võib olla mees – 1 või naine – 2.
- Vanusegrupp – agendid on jaotatud järgmisse kuude vanusegruppi:
 - 15–24 – 1;
 - 25–34 – 2;
 - 35–49 – 3;
 - 50–64 – 4;
 - 65–74 – 5;
 - 75+ – 6.
- Keelekeskkond – valitud on järgmine statistiliselt tõhusaim viis jagada Eesti neljaks [11, p. 11]:
 - Mitmekeelne Tallinn – 1;
 - Venekeelse enamusega linnad – 2;
 - Eestikeelse enamusega linnad – 3;
 - Eestikeelsed linnad ja maapiirkonnad – 4 .
- Rahvus – jaotatud statistiliselt Eestis elavate elanike järgi:
 - Eestlane – 1;

- Venelane – 2;
- Ukrainlane – 3;
- Valgevenelane – 4;
- Mitu rahvust (näiteks eesti/vene) – 5;
- Muu – 12.
- Emakeel – jaotatud Eestis kahe põhikeele järgi:
 - Eesti – 1;
 - Vene – 2;
 - Eesti/vene – 3;
 - Eesti/muu – 4;
 - Vene/muu – 5;
 - Muu – 6.
- Keeleoskus (eesti, vene, inglise) – algandmed sisaldavad iga agendi kohta väärtusi antud kolme keele valdamise osas vahemikus 1 (ei valda üldse) kuni 7 (valdab vabalt).
- Üle aasta viibinud mujal – kas elanud, õppinud või töötanud Eesti-siseselt pikemalt (üle ühe aasta) mõnes teises praegusega võrreldes erinevas asulas: ei – 1 või jah – 2.
- Mobiilsus – kui tihti agent käib mujal Eestis, kaugemal oma elukohast (suurem võimalus suhelda erinevas keelekeskkonnas), väärtused vahemikus 1 (kaks või rohkem korda nädalas) kuni 7 (harvem kui korra aastas).
- Kodanikuaktiivsus – kui tihti käiakse üritustel, osaletakse koosolekutel: vahemikus 1 kuni 7.
- Akommodatsioon – kohanemine suhtluses (kas agent vastab alati suhtluspartneri keeles või suhtleb agent ainult oma kõrgema keeleoskusega keeles): 1 (kohaneb) kuni 7 (ei kohane).
- Etnosuletus – suhtlemine päritolu järgi (erinev või sama etniline taust): 1 (madal suletus) kuni 7 (kõrge suletus).

Nimetatud muutujate väärtused omistatakse simulatsioonisüsteemis vastavalt häälestamisele peatükis 4.1 ning katsetustele peatükis 4.2. Täpsemalt on seadistatud simulatsioonisüsteemi parameetrid koos kirjeldustega toodud lisas 1 lk 64. Katsetuste käigus (peatükis 4.2) muudetakse simulatsioonisüsteemi seadistusi, et analüüsida mudeli parameetrite mõju ja leida keelekeskkonda kõige paremini iseloomustav seadistus.

3.3.5 Sisendandmed

Eesti keelekeskkonna agendipõhise simulatsiooni koostamise lähteandmeteks on keeleteadlaste poolt 2016. aastal Tartu Ülikooli institutsionaalse uurimisteema “Eesti keele kestlikkus avatud maailmas“ raames läbi viidud üle-eestilise küsitluse tulemused. Nimetatud küsitluse abil uuriti Eesti elanike keeleoskust, keelevelikute eelistusi ja keelehoiakuid. Küsitlustulemuste failis on iga inimese kohta küsitluse andmed, mis on arvutatud mitme küsimuse keskmisena. Kuna küsimusi, mille põhjal sisendandmed on loodud, on ligikaudu 300, siis kasutatakse käesoleva töö raames teostatud simulatsioonisüsteemi puhul nende hulgast simulatsiooni koostamiseks vajalikke parameetreid.

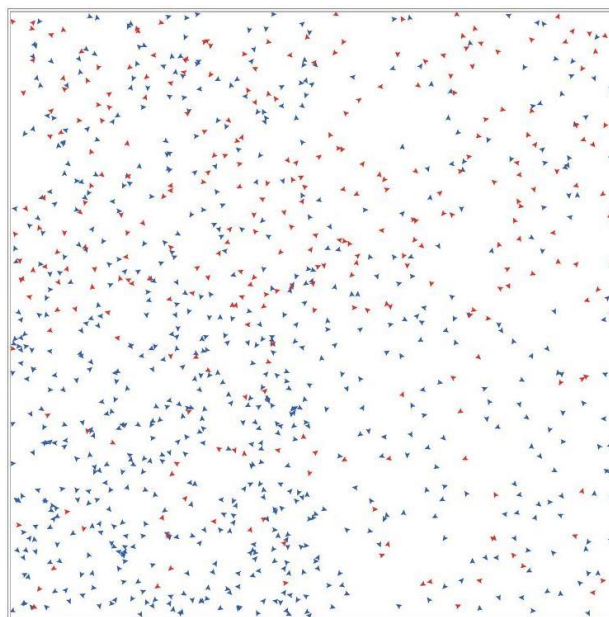
4 Simulatsioonisüsteemi valideerimine

Käesolevas peatükis häälestatakse ja määratakse simulatsioonide parameetreid, valideeritakse simulatsioonisüsteem koostöös keeleteadlase Martin Ehalaga ning analüüsitakse tekkinud võrgustiku struktuuri esialgseid mõjusid.

4.1 Simulatsiooni parameetrite häälestamine ja määramine

Selles peatükis määratakse kindlaks simulatsioonisüsteemi tööks vajalikud parameetrid ning kirjeldatakse nende väärtuste määramise põhimõtteid. Simulatsioonisüsteemile, mille põhjal parameetreid häälestatakse, viitab lisa 2 lk 65. Parameetrite häälestamiseks tehakse katseid, mida selles töös eraldi välja ei tooda, kuna need on ebakorrapärased ja katsetuste arv on suur.

Simulatsioonisüsteemi lähtestamisel luuakse suletud süsteem agentidest, mille omadused (näiteks keeleoskus ja käitumine) on saadud küsimustiku andmetest. Simulatsioonisüsteemis sisalduvad agendid on jaotatud keelekeskkondade järgi neljaks. Agentide eristamiseks kujutatakse neid (joonisel 14) ankeedi keele järgi kahe värviga – eestikeelse ankeedi valinuid sinisega ja venekeelse ankeedi valinuid punasega.



Joonis 14. Simulatsiooni algne vaade.

Lähtestamise järel saab simulatsioonisüsteemis suhtlust agentide vahel mitmeid kordi simuleerida. Suhtlemise õnnestumisel kujutatakse parema visuaalse ülevaate saamiseks ühendust vastava ühenduse loomiseks kasutatud õnnestunud keele järgi järgmiste värvidega: eesti keel on sinine, vene keel punane ning inglise keel roheline. Ühendused loovad võrgustiku, mis kirjeldab reaalse maailma sotsiaalvõrgustikku.

Koostöös keeleteadlase Martin Ehalaga selgitati simulatsiooni valideerimise jaoks piisav arv suhtlustsükleid ehk iteratsioone, et katsetada äärmuslikke juhtumeid kahe kõige olulisema parameetri puhul – etnosuletus ja keelelävi. Simulatsioonisüsteemi suhtlemise protsessi käivitatakse 10 000 korda, mille puhul enamik keele valiku väärtusi ja võrgustiku omadusi stabiliseeruvad. Simulatsiooni kordade järgi valiti algselt simulatsioonitulemuste raporteerimise tiheduseks (*raport-frequency*) üks raporteerimine iga 5 000 suhtluse järel. Läbiviidud simulatsioonide tulemusi kirjeldatakse peatükis 4.2.

Iga suhtluse iteratsiooni jaoks valitakse juhuslikult suhtluse algataja ning talle valitakse suhtluspartner vastavalt tõenäosusele, mis põhineb kasutajaliidese kaudu määratud partneri valiku mõjuteguritele. Mõjutegurite väärtused erinevad vastavalt sellele, kui olulised need on. Katsetuste alguses võiksid mõjutegurite väärtused olla vähemalt 0,1. Sama väärtusega 0,1 peaks olema ka juhuslikkuse tegur, mis annab võimaluse kõikide agentide vahelistele seostele ning muudab simulatsiooni dünaamilisemaks.

Kaks agenti alustavad omavahel suhtlemist ning nendevaheline suhtlemine kas õnnestub või ebaõnnestub. Suhtlemise õnnestumise määrab keelelävi (*communication-threshold*) ehk minimaalne keeleoskus suhtluseks. Selle parameetri võimalikud väärtused on 1–7, mis on tuletatud algandmete keeleoskuste skaalast. Kui valitud Partner on suhtluse Alustajast ehk Starterist erineva emakeelega, siis mõjutavad keeleläve ka teised parameetrid nagu näiteks akommodatsiooni (*accommodation-index*) ja etnilise suletuse (*ethnic-closure-index*) indeksid. Nimetatud indeksite väärtused korrutatakse läbi Alustaja vastavate kordajatega (mis on kasutajaliidese kaudu muudetavad), liidetakse omavahel ning saadud väärtusele liidetakse veel juurde ka meedia indeks (*media-index*), mille väärtus võib olla negatiivne või positiivne. Kuna meedia mõjutab suhtlemist ja keeleläve lineaarselt, siis meedia indeksi väärtus võiks jääda alla poole keeleläve väärtusest ehk vahemikku -2 kuni +2 ja võiks minimaalselt olla -0,1.

Kui suhtlus õnnestub, siis luuakse vastavate agentide vahele ühendus, mida visualiseeritakse NetLogo ekraanil selle suhtluskeele värviga, milles suhtlus õnnestus (eesti keel sinine, vene keel punane, inglise keel roheline). Lisaks suureneb suhtluse õnnestumisel õnnestunud Partneri edasise suhtlemise tõenäosus kasutajaliidese kaudu määratud õnnestunud suhtluse konstandi (*success-const*) võrra. Kui õnnestunud Partneril on sõpru, siis suurendatakse nende suhtlemise tõenäosust vastava mõjuteguri võrra (*triadic-closure-increment*), mille väärtus võiks suhtluspartneri valikut mõjutada rohkem kui ülejäänud partneri valiku parameetrid, aga olla väiksem kui üks ehk katsetustes võiks see olla 0,9. Nimetatud parameetri alusel on Alustajal ehk Starteril võimalik sõbruneda sõbra sõbraga (*triadic closure*). Kui suhtlus ebaõnnestub, siis ühendust ei looda, vaid vähendatakse ebaõnnestunud Partneri valiku tõenäosust kasutajaliidese kaudu esitatud konstandi (*failure-const*) võrra, mille väärtus on algselt sarnane õnnestumise konstandile.

Keele õppimine ja unustamine on seotud simulatsioonitulemuste raporteerimise tihedusega (*raport-frequency*), mille väärtus peab olema väiksem või võrdne suhtluse iteratsioonide arvuga. Kui keelt õpitakse (*lang-increment*) või unustatakse (*lang-decrement*), siis suurendatakse keeleoskust vastava kasutajaliidese kaudu esitatud parameetri väärtuse võrra, mis peaks mõjutama keeleoskust vähem kui pool keeleoskuse skaala ühest astmest ehk katsetustes võiks selle parameetri väärtuseks olla 0,35.

4.2 Valideerimise protseduur ja simulatsioonisüsteemi valideerimine

Valideerimisel saab hinnata, kas simulatsiooni väljund on ootuspärane vastavate algandmete oodatavatele tulemustele [27, p. 140]. Käesoleva lõputöö raames teostatud simulatsioonisüsteemi valideerimine viiakse läbi, võrreldes keelte kasutamise jaotust agentide keelekeskkonnas vastavalt küsimustiku andmetele [16]. Keelte kasutamise jaotust agentide keelekeskkonnas näitab viimase suhtluse iteratsiooni seis.

Nelja keelekeskkonna protsesside kõrvutamiseks saab eksporditud andmete põhjal võrrelda simulatsiooni alg- ja lõppseisus olevaid agentide parameetreid. Mudeli tulemuste analüüsimiseks vaadeldakse olulisemate parameetrite erinevust sisendi ja väljundi korral ning muutuse alusel tehakse järeldus simulatsiooni omaduste kohta [27, p. 157]. Agentide parameetrite võrdlemiseks tehakse simulatsioonisüsteemis neli simulatsiooni 10 000 suhtluse iteratsiooniga, võttes fookusesse kaks olulist parameetrit.

Sotsiaalse simulatsiooni ühe valideerimise tehnika kohaselt katsetatakse ekstreemseid olukordi parameetrite ebatõenäoliste väärtuste kombinatsioonidega, tavaliselt väga kõrgete või väga madalate väärtustega, et tõestada simulatsiooni loogilisust parameetrite ekstreemumite juures [27, p. 161]. Näiteks kui teatud parameetri mõju on nullilähedane, siis ei tohiks vastav parameeter simulatsioonisüsteemi agentidele mõju avaldada [27, p. 161]. Käesoleva magistritöö raames teostatud simulatsioonisüsteemis on neli simulatsiooni jaotatud kasutajaliidese kaudu määratavate parameetrite väärtuste järgi järgmiselt:

- mõlemate parameetrite väärtused on ühe otspunkti lähedal;
- esimese parameetri väärtus on teise otspunkti lähedal;
- teise parameetri väärtus on teise otspunkti lähedal;
- mõlemate parameetrite väärtused on teise otspunkti lähedal.

Ülevaatlikkuse saamiseks luuakse statistilised tabelid koos parameetrite näitude kirjeldusega iga simulatsiooni kohta. Igast tabelist on näha keeleoskuste keskmised väärtused enne ja pärast simulatsiooni iga keelekeskkonna kohta eraldi. Nende andmete põhjal saab teada, kas tulemused on sotsiolingvistiliselt usutavad ja millised reaalsed tingimused peaksid valitsema, et vastavad muutused leiaksid aset. Kui simulatsiooni eesmärgiks on simuleerida erinevaid tuleviku stsenaariume ja ennustused tõestavad testide tulemusi, siis võib eeldada, et simulatsioon annab sarnastel tingimustel usaldusväärse tulemuse [27, p. 157].

Katsetustes on olulisel kohal keelekeskkonna mõjutegur (*lang-environment-proportion*), mille väärtus määrab ära, kui palju suhtlust toimub samas keelekeskkonnas vaadeldavast neljast keelekeskkonnast. Realistlikkuse huvides peaks keelekeskkonna parameeter olema sellise väärtusega, et umbes 80% suhtlust toimuks samas keelekeskkonnas, sest inimesed ei hakka rohkem suhtlema teistest keelekeskkondadest pärit inimestega isegi siis, kui oluliselt suurendada teiste parameetrite mõju, näiteks kui võrd nad eelistavad suhelda samast soost inimestega [17]. Seega tuleb katsetustega välja asjaolu, et keelekeskkonna parameeter tuleb seadistada suurema väärtusega, mis muudab selle olulisemaks parameetriks. Kuna ülejäänud Partneri valiku mõjutegurid mõjutavad tulemust minimaalselt, siis tuleb vastavate parameetrite väärtused seada neutraalsesse asendisse (näiteks minimaalsele väärtusele 0,1) ja neid ei ole mõtet muuta. Nii saab kontrollitud olukorras analüüsida oluliste muutujate mõju simulatsiooni käigule.

Järgnevalt katsetati korraga kahe parameetriga, et uurida ja analüüsida nende omavahelist suhestamist ja nende mõju simulatsioonis esitatud keelekasutajatele. Valitud parameetriteks on keelelävi (*communication-threshold*) ja etnosuletus (*ethnic-closure-index*), mille väärtused seadistatakse kordamööda miinimum ja maksimum väärtuste lähedale.

Kõikide simulatsioonikatsete puhul on simulatsiooni alguses arvatud välja esialgsed keskmised keeleoskuste väärtused (tabelis 10), et nendega võrrelda katsete keskmisi tulemusi viimase simulatsiooni seisuga. Keeleoskuse väärtus on vahemikus 1 kuni 7 (peatükk 3.3.4).

Tabel 10. Esialgsed keeleoskuste keskmised väärtused.

| Keelekeskkond | Keskmine eesti keele oskus | Keskmine vene keele oskus | Keskmine inglise keele oskus |
|---------------|----------------------------|---------------------------|------------------------------|
| 1 | 5,640 | 5,463 | 3,743 |
| 2 | 4,227 | 6,401 | 2,576 |
| 3 | 6,282 | 4,685 | 3,376 |
| 4 | 6,608 | 4,312 | 2,971 |

Keelekeskkondade lõikes on kujutatud kolme vaadeldava keele keskmisi väärtusi, milles kajastuvad kõrgemad keskmised eesti keele osas neljandas keelekeskkonnas ehk eestikeelsetes linnades ja maapiirkondades, vene keele osas teises keelekeskkonnas ehk venekeelse enamusega linnades ning inglise keele osas hoopis esimeses keelekeskkonnas ehk mitmekeelses Tallinnas.

Esimese katse puhul on tabelis 11 etnosuletus (*ethnic-closure-index*) simulatsioonisüsteemi kasutajaliideses määratud 0,1 ning keelelävi (*communication-threshold*) 1,0, teise katse puhul (tabel 12) on etnosuletus 3,0 ning keelelävi 6,0, kolmanda katse (tabel 13) puhul on etnosuletus 0,1 ning keelelävi 6,0 ning neljanda katse (tabel 14) etnosuletus on 3,0 ning keelelävi 1,0.

Tabel 11. Katse 1 keskmised keeleoskuste väärtused.

| Keelekeskkond | Keskmine eesti keele oskus | Keskmine vene keele oskus | Keskmine inglise keele oskus |
|---------------|----------------------------|---------------------------|------------------------------|
| 1 | 5,217 | 4,866 | 3,043 |
| 2 | 3,586 | 6,099 | 1,876 |

| Keelekeskkond | Keskmine eesti keele oskus | Keskmine vene keele oskus | Keskmine inglise keele oskus |
|---------------|----------------------------|---------------------------|------------------------------|
| 3 | 6,109 | 4,013 | 2,676 |
| 4 | 6,579 | 3,632 | 2,288 |

Tabelis 11 kujutatud esimese katse puhul on kõrgeimad keskmised keeleoskused samad võrreldes esialgsete keskmiste väärtustega tabelis 10. Esimese katse puhul on suurim eesti keele oskuse muudatus, mis on keskmiselt vähenenud 0,64 võrra venekeelse enamusega linnades, vene keele oskus on kõige enam vähenenud keskmiselt 0,68 võrra eestikeelsetes linnades ja maapiirkondades ning inglise keel on vähenenud keskmiselt võrdselt 0,7 võrra keelekeskkondades 1–3 (seega kõige vähem on muutunud neljandas keelekeskkonnas ehk eestikeelsetes linnades ja maapiirkondades). Lisaks selgus esimese katse simulatsioonist, et õnnestunud suhtlusi samas keelekeskkonnas oli 6 015 ning ebaõnnestumisi 3 717, aga üle keelekeskkondade piiride oli õnnestunud suhtlusi 176 ning ebaõnnestumisi 92.

Tabel 12. Katse 2 keskmised keeleoskuste väärtused.

| Keelekeskkond | Keskmine eesti keele oskus | Keskmine vene keele oskus | Keskmine inglise keele oskus |
|---------------|----------------------------|---------------------------|------------------------------|
| 1 | 5,576 | 5,358 | 3,090 |
| 2 | 3,801 | 6,363 | 1,882 |
| 3 | 6,271 | 4,375 | 2,688 |
| 4 | 6,592 | 3,830 | 2,275 |

Tabelis 12 kujutatud teise katse puhul on sarnaselt esimesele katsele kõrgeimad keskmised keeleoskused samad võrreldes esialgsete keeleoskuste keskmiste väärtustega (tabelis 10). Teise katse puhul on tabelis 12 kõige rohkem eesti keele oskus vähenenud 0,43 võrra (aga vähem kui esimese katse puhul) venekeelse enamusega linnades, vene keele oskus on kõige enam vähenenud 0,48 võrra (ka vähem kui esimese katse puhul) eestikeelsetes linnades ja maapiirkondades ning inglise keel on kõige enam vähenenud 0,69 võrra eestikeelsetes linnades ja maapiirkondades, mis on ligikaudu sama palju kui esimese katse puhul. Teise katse puhul oli õnnestunud suhtlusi samas keelekeskkonnas 7 913 ja ebaõnnestumisi 1 842 ning üle keelekeskkonna piiride oli õnnestunud 199 ja ebaõnnestunud 46 suhtlust.

Tabel 13. Katse 3 keskmised keeleoskuste väärtused.

| Keelekeskkond | Keskmine eesti keele oskus | Keskmine vene keele oskus | Keskmine inglise keele oskus |
|---------------|----------------------------|---------------------------|------------------------------|
| 1 | 5,169 | 4,885 | 3,043 |
| 2 | 3,574 | 6,008 | 1,876 |
| 3 | 6,033 | 4,017 | 2,676 |
| 4 | 6,480 | 3,613 | 2,273 |

Kolmanda katse keskmised keeleoskuste väärtused on esitatud tabelis 13 ja neljanda katse keeleoskuste keskmised väärtused tabelis 14. Kolmanda ja neljanda katse kõrgemad keskmised väärtused on sarnaselt eelmistele katsetele saadud samade keelekeskkondade piires. Kolmanda katse puhul on tabelis 13 võrreldes esialgsete keeleoskuste keskmiste väärtustega (tabelis 10) eesti keele oskus kõige enam vähenenud teises keelekeskkonnas keskmiselt 0,65 võrra (enam kui esimese katse puhul), samuti on vene keele oskus kõige enam vähenenud neljandas keelekeskkonnas 0,69 võrra ning inglise keele oskus on vähenenud 0,7 võrra sarnaselt esimese katsega keelekeskkondades 1–3. Kolmanda katse puhul oli õnnestunud suhtlusi sama keelekeskkonna piires 5 544 ja ebaõnnestumisi 4 218 ning üle keelekeskkondade piiride õnnestus 137 ja ebaõnnestus 101 suhtlust.

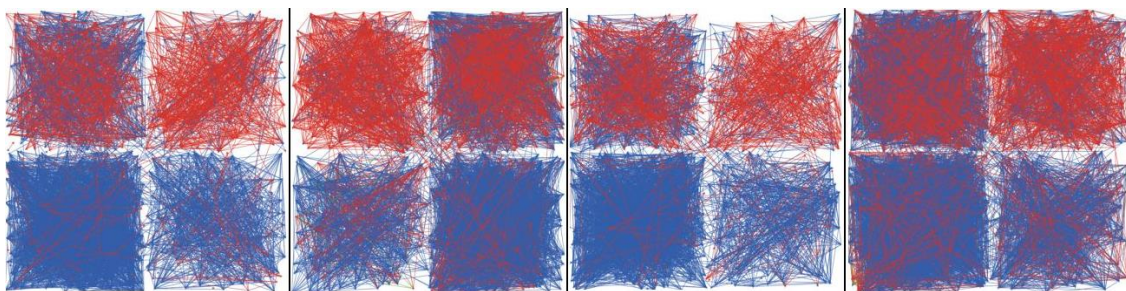
Tabel 14. Katse 4 keskmised keeleoskuste väärtused.

| Keelekeskkond | Keskmine eesti keele oskus | Keskmine vene keele oskus | Keskmine inglise keele oskus |
|---------------|----------------------------|---------------------------|------------------------------|
| 1 | 6,207 | 5,933 | 3,044 |
| 2 | 4,883 | 6,755 | 1,882 |
| 3 | 6,630 | 5,096 | 2,685 |
| 4 | 6,834 | 4,318 | 2,341 |

Neljanda katse puhul on tabelis 14 eesti ja vene keele oskused kõikide keelekeskkondade lõikes suurenenud võrreldes esialgsete keskmiste väärtustega (tabelis 10). Kõige rohkem on eesti keele oskus keskmiselt suurenenud 0,65 võrra teises keelekeskkonnas ehk venekeelse enamusega linnades ning vene keele oskus on kõige enam suurenenud 0,47 võrra esimeses keelekeskkonnas ehk mitmekeelses Tallinnas. Neljanda katse puhul on inglise keele oskus ainsa keelena vähenenud kõigis neljas keelekeskkonnas 0,69 võrra mitmekeelses Tallinnas, mis on sarnane teise katsega tabelis 12. Neljanda katse puhul

õnnestusid kõik suhtlused, millest 9 723 olid sama keelekeskkonna piires ning 277 üle keelekeskkondade piiride.

Nelja katse puhul moodustunud võrgustikud on kujutatud joonisel 15. Joonis 15 kujutab iga katse puhul nelja keelekeskkonda ruudukujulises võrgustikus, eraldades joonisel neli katset püstjoontega. Ühes katses olevad keelekeskkonnad on joonisel 15 kujutatud järgmiselt: esimene keelekeskkond on vasakul üleval nurgas, teine keelekeskkond on paremal üleval nurgas, kolmas keelekeskkond on all paremas nurgas ning neljas keelekeskkond on vasakul alumises nurgas.



Joonis 15. Katsete 1–4 võrgustiku struktuurid.

Joonisel 15 kujutatud võrgustikest on näha, et teises keelekeskkonnas (venekeelse enamusega linnades) on kõige enam punaseid ühendusi ehk suhtlus on õnnestunud vene keeles. Esimeses keelekeskkonnas (mitmekeelses Tallinnas) on nii eesti- kui ka venekeelseid õnnestunud suhtlusi. Kolmandas keelekeskkonnas (eestikeelse enamusega linnades) ja neljandas keelekeskkonnas (eestikeelsetes linnades ja maapiirkondades) on enim siniseid ühendusi ehk eestikeelseid. Teise katse puhul, kus etnosuletuse parameetri väärtuseks oli 3,0 ning keeleläve parameetri väärtuseks oli 6,0, on kolmanda ja neljanda keelekeskkonna puhul näha ühenduste muutust (joonisel 15 teine ruudukujuline võrgustik). Nimelt esimese, kolmanda ja neljanda katse puhul on neljandas keelekeskkonnas vähem ühendusi ehk õnnestunud suhtlusi ja teise katse puhul on kolmandas keelekeskkonnas vähem ühendusi, mis võib olla tingitud suhteliselt madalatest keeleoskustest.

Kõigi nelja katse puhul muutub inglise keele oskus esialgsete keeleoskuste keskmiste väärtustega võrreldes kõige rohkem. See on tingitud sellest, et inglise keelt emakeelena kõnelevaid agente simulatsioonisüsteemi algandmetes ei ole. Rahvuskeskse suhtluse puhul jääb inglise keel kasutusest kõrvale ning seetõttu ununeb see

simulatsioonisüsteemis pidevalt keele unustamise languse parameetri (*lang-decrement*) võrra.

Katsete 1–4 keskmiste keeleoskuste väärtuste täpsem analüüs on toodud peatükis 4.3. Teise sotsiaalsete simulatsioonide valideerimise tehnika kohaselt võrreldakse tulemusi muude vastavate simulatsioonidega, mis on koostatud sama probleemi uurimiseks [27, p. 163]. Peatükis 4.3 võrreldakse käesoleva magistritöö raames loodud simulatsioonisüsteemi katseid Andres Karjuse vastavate simulatsioonikatsetega.

Praegune simulatsioonisüsteem näitab, et keeleoskus paraneb vaid siis, kui suhtlus on etniliselt väga eraldatud, mida see tavaliselt ei ole, ning muudel juhtudel domineerib unustamine. See ei ole aga kooskõlas tegelikkusega, sest üldiselt inimeste keeleoskus on üsna püsiv ja eriti peaks see nii olema enamuskeele kõnelejate puhul. Simulatsiooni realistlikumaks muutmiseks tuleks seada keele õppimise ja unustamise suhteks 5:1, mis tähendab, et unustamine toimub viis korda aeglasemalt kui õppimine. Praegu toimub unustamine simulatsioonisüsteemis neli korda aeglasemalt kui õppimine. Kuna üks iteratsioon on tsükkel, mille jooksul toimub ka unustamine ja ühes iteratsioonis on simulatsioonisüsteemis katsete puhul määratud 5 000 suhtlust (*raport-frequency*), siis ühe unustamise tsükli jooksul toimub 5 õppimise tsüklit. Kirjeldatud muudatusest tulenevalt on tehtud veel neli katsetust, milles vaadeldakse keeleoskuse vähendamise parameetrit (*lang-decrement*), mis on otseselt seotud unustamisega, sest kui keel ununeb, siis vähendatakse keeleoskust just selle parameetri väärtuse võrra.

Eelnevate katsete 1–4 põhjal tundub, et kombinatsioon, kus etnosuletuse (*ethnic-closure-index*) väärtuseks on 0,1 ning keeleläve (*communication-threshold*) väärtuseks on 1,0, on kõige avatum keeleliste muutustele ning arvestades ajapiiranguga, et ei peaks sooritama väga palju iteratsioone muutuste nägemiseks, on keele unustamise sammu (*lang-decrement*) puhul mõistlik kasutada volatiilsuse sätet. Teiste muutujate puhul tuleb valida neutraalsed väärtused, et oleks võimalik analüüsida vaadeldava parameetri mõju simulatsiooni käigule. Seega uute katsete 5–8 puhul on teiste muutujate väärtused valitud järgmised: keele õppimise samm (*lang-increment*) on 0,35, etnosuletus 0,1, keelelävi 1,0, keelekeskkonna parameeter 40, meedia indeks -0,1. Keele unustamise sammu varieeritakse järgnevates katsetes 5–8 järgmiselt: katse 5 (tabel 15) puhul 0,15, katse 6 (tabel 16) puhul 0,001, katse 7 (tabel 17) puhul 0,25 ja katse 8 (tabel 18) puhul 0,05.

Tabel 15. Katse 5 keskmised keeleoskuste väärtused.

| Keelekeskkond | Keskmine eesti keele oskus | Keskmine vene keele oskus | Keskmine inglise keele oskus |
|---------------|----------------------------|---------------------------|------------------------------|
| 1 | 5,451 | 5,206 | 3,443 |
| 2 | 3,974 | 6,300 | 2,276 |
| 3 | 6,224 | 4,398 | 3,076 |
| 4 | 6,636 | 4,043 | 2,697 |

Vastavalt tabelile 15, viienda katse kõrgeimad keskmised keeleoskuste näitajad on samad, kui tabelis 10 esitatud esialgsed keeleoskuste keskmised väärtused. Viienda katse puhul on kõige enam vähenenud eesti keele oskus keskmiselt 0,25 võrra venekeelse enamusega linnades, vene keele oskus on kõige enam vähenenud keskmiselt 0,28 võrra eestikeelse enamusega linnades ning inglise keel on vähenenud keskmiselt võrdselt 0,3 võrra keelekeskkondades 1–3 (seega kõige vähem on see muutunud neljandas keelekeskkonnas ehk eestikeelsetes linnades ja maapiirkondades). Lisaks selgus simulatsioonisüsteemi kasutajaliidesest, et õnnestunud suhtlusi samas keelekeskkonnas oli 6 055 ja ebaõnnestumisi oli 3 697 ning üle keelekeskkondade piiride oli õnnestunud suhtlusi 150 ja ebaõnnestumisi 98.

Tabel 16. Katse 6 keskmised keeleoskuste väärtused.

| Keelekeskkond | Keskmine eesti keele oskus | Keskmine vene keele oskus | Keskmine inglise keele oskus |
|---------------|----------------------------|---------------------------|------------------------------|
| 1 | 5,651 | 5,465 | 3,741 |
| 2 | 4,229 | 6,440 | 2,574 |
| 3 | 6,319 | 4,694 | 3,374 |
| 4 | 6,678 | 4,322 | 2,980 |

Kuuenda katse puhul on tabelis 16 sarnaselt viiendale katsele kõrgeimad keskmised väärtused samad, mis olid esialgsed väärtused (tabelis 10). Kuuenda katse puhul on kõigis keelekeskkondades eesti ja vene keelt juurde õpitud, sest unustamise samm on väga väike. Inglise keelt on õpitud vaid neljandas keelekeskkonnas ehk eestikeelsetes linnades ja maapiirkondades, aga muudes keelekeskkondades on inglise keele oskus kahekordse unustamise sammu võrra väiksem. Kõige rohkem on eesti keele oskus suurenenud 0,07 võrra neljandas keelekeskkonnas ehk eestikeelsetes linnades ja maapiirkondades, vene keele oskus on kõige enam suurenenud teises keelekeskkonnas ehk venekeelse

enamusega linnades 0,039 võrra ning inglise keel on kõige rohkem suurenenud 0,009 võrra neljandas keelekeskkonnas ehk eestikeelsetes linnades ja maapiirkondades, mis on kordades vähem kui viienda katse puhul (tabelis 15). Kuuenda katse puhul oli õnnestunud suhtlusi sama keelekeskkonna piires 6 064 ja ebaõnnestumisi 3 690 ning üle keelekeskkonna piiride oli õnnestunud 148 ja ebaõnnestunud 98 suhtlust.

Tabel 17. Katse 7 keskmised keeleoskuste väärtused.

| Keelekeskkond | Keskmine eesti keele oskus | Keskmine vene keele oskus | Keskmine inglise keele oskus |
|---------------|----------------------------|---------------------------|------------------------------|
| 1 | 5,338 | 5,050 | 3,243 |
| 2 | 3,791 | 6,196 | 2,076 |
| 3 | 6,157 | 4,216 | 2,876 |
| 4 | 6,603 | 3,828 | 2,489 |

Seitsmenda katse keskmised keeleoskuste väärtused on esitatud tabelis 17 ja kaheksanda katse keeleoskuste keskmised väärtused tabelis 18. Seitsmenda katse kõrgemad keskmised väärtused on sarnaselt eelmistele katsetele saadud samade keelekeskkondade piirides. Seitsmenda katse puhul on võrreldes esialgsete keeleoskuste keskmiste väärtustega (tabelis 10) kõige enam vähenenud eesti keele oskus teises keelekeskkonnas keskmiselt 0,43 võrra (enam kui viienda katse puhul tabelis 15) ja samuti on vene keele oskus vähenenud neljandas keelekeskkonnas 0,48 võrra ning inglise keel on vähenenud 0,5 võrra sarnaselt viienda katsega keelekeskkondades 1–3. Seitsmenda katse puhul oli õnnestunud suhtlusi sama keelekeskkonna piires 6 036 ja ebaõnnestumisi oli 3 714 ning üle keelekeskkondade piiride õnnestus 153 ja ebaõnnestus 97 suhtlust.

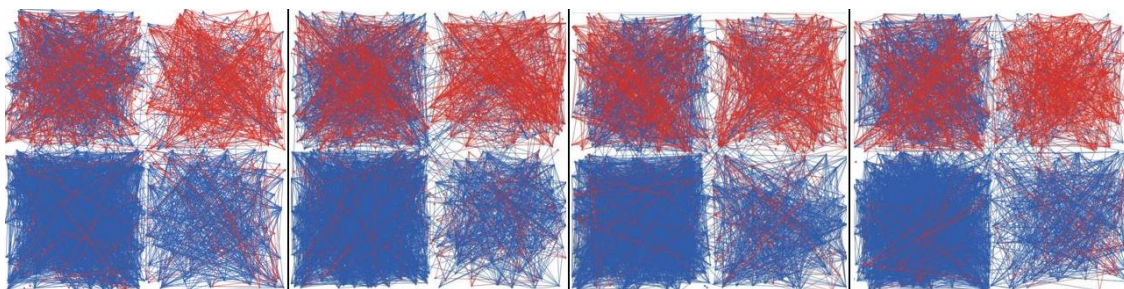
Tabel 18. Katse 8 keskmised keeleoskuste väärtused.

| Keelekeskkond | Keskmine eesti keele oskus | Keskmine vene keele oskus | Keskmine inglise keele oskus |
|---------------|----------------------------|---------------------------|------------------------------|
| 1 | 5,590 | 5,385 | 3,643 |
| 2 | 4,156 | 6,381 | 2,476 |
| 3 | 6,271 | 4,604 | 3,276 |
| 4 | 6,658 | 4,223 | 2,880 |

Tabelis 18 esitatud kaheksanda katse tulemuste osas on vene ja inglise keele oskused kõikide keelekeskkondade lõikes vähenenud võrreldes esialgsete keskmiste väärtustega

(tabelis 10) ning eesti keele puhul tõusis keeleoskus vaid neljandas keelekeskkonnas 0,05 võrra. Kõige rohkem on eesti keele oskus keskmiselt vähenenud 0,07 võrra teises keelekeskkonnas ehk venekeelse enamusega linnades, vene keele oskus on vähenenud 0,089 võrra neljandas keelekeskkonnas ehk eestikeelsetes linnades ja maapiirkondades ning inglise keele oskus on vähenenud keskmiselt võrdselt 0,1 võrra keelekeskkondades 1–3 (kõige vähem on see muutunud eestikeelsetes linnades ja maapiirkondades). Kaheksandal katsel õnnestus katsetest 5–8 kõige vähem suhtlusi ehk 5 993 suhtlust sama keelekeskkonna piires, aga ka ebaõnnestus kõige rohkem suhtlusi ehk 3 722 suhtlust. Kaheksandal katsel õnnestus 172 suhtlust üle keelekeskkondade piiride ja ebaõnnestus 113 suhtlust üle keelekeskkondade piiride.

Katsete 5–8 puhul moodustuvad võrgustikud on kujutatud joonisel 16. Joonis 16 kujutab iga katse puhul nelja keelekeskkonda ruudukujuliste võrgustikena, jaotades joonisel neli katset eraldi püstjoontega.



Joonis 16. Katsete 5–8 võrgustiku struktuurid.

Joonisel 16 kujutatud võrgustiku struktuurid kujutavad väga tõepäraselt suhtlusi nelja keelekeskkonna lõikes, kus kõige enam on vene keeles õnnestunud suhtlusi esimeses keelekeskkonnas (mitmekeelses Tallinnas) ja teises keelekeskkonnas (venekeelse enamusega linnades) ning enam eesti keeles õnnestunud suhtlusi just neljandas keelekeskkonnas (eestikeelsetes linnades ja maapiirkondades), aga ka kolmandas keelekeskkonnas (eestikeelse enamusega linnades).

Läbikukkunud kommunikatsioonide hulk katsete 5–8 puhul on ligikaudu 3 000. Keelte väiksem unustamine võib tugevalt mõjutada ebaõnnestunud suhtluste arvu just seetõttu, et etnosuletuse (*ethnic-closure-index*) väike väärtus suurendab ebaõnnestumisi, sest etnosuletuse väärtus 0,1 tähendab, et eelistatakse suhelda endast erinevaga ja siis ka ebaõnnestutakse rohkem. Sarnaselt on ebaõnnestumiste arv suur ka katsete 1 ja 3 puhul, kui etnosuletuse parameetri väärtuseks oli samuti valitud 0,1.

Katsete 5–8 järel on märgata peaaegu kõigi katsetuste puhul seda, et vaadeldavad kolm keelt ununevad suhteliselt ühtlaselt. Samas tegelikus elus ununevad vaid need keeled, mida ei kasutata ja see toimub aeglaselt. Seega tuleb unustamist simulatsioonisüsteemis veelkord keele õppimisest aeglasemaks muuta (näiteks 6 korda), et saavutada veelgi loomulikumaid tulemusi. Järelikult tasub simulatsioonisüsteemi parameetrite häälestamisega jätkata, kuni tulemustega jõuab selliste seadistusteni, mille põhjalt saab hakata katsetama realistlikke keelevahetusstrateegiaid. Selle jaoks tuleb lõplikus faasis teha arvatavasti 10 või 100 korda rohkem iteratsioone, kui selle lõputöö katsetustega on tehtud.

4.3 Võrgustiku struktuuri mõjude analüüs

Katsete 1 ja 3 tulemustest selgub, et etnosuletuse (*ethnic-closure-index*) madal väärtus toob kaasa suure ebaõnnestunud kommunikatsioonide arvu. Väga väike etnosuletuse väärtus, näiteks 0,1 tähendab, et agendid väldivad suhtlemist nendega, kes on sarnased. See seletab suurt läbikukkumiste arvu – kui agent püüab kogu aeg suhelda ainult muust rahvusest isikuga, siis on suur tõenäosus, et ühine keel puudub ja kommunikatsioon kukub läbi. Etnosuletuse kõrgem väärtus 3,0 aga tähendab omasuguste eelistamist ning see seletab, miks ebaõnnestumiste hulk väheneb. Ka Andres Karjuse simulatsioonisüsteemi [16], [28] põhjal selgus, et madal etnosuletus viitab erinevatega suhtlemisele, aga kõrge etnosuletuse väärtusega välditakse endast erinevaid ja suheldakse pigem omasugustega. Sotsiolingvistiliselt on fenomen, kus kõik väldivad rahvuskaaslastega suhtlemist ja püüavad suhelda teistega, üsna harv. Seda võib ette tulla vaid selliste vähemusrahva esindajate puhul, kes soovivad sulanduda enamuse hulka ja seetõttu väldivad teiste vähemusliikmetega suhtlemist.

Peatükis 4.2 kirjeldatud ebaõnnestunud kommunikatsioonide pingerea alusel (katse 3, katse 1, katse 2 ja katse 4) on näha, et etnosuletus on tugevam tegur ja keelelävi nõrgem tegur, mis on igati sotsiolingvistiliselt loogiline ning samuti sarnaneb Andres Karjuse katsetuste tulemustele [16]. Seega küsimus, kas suhelda ainult oma rahvuskaaslastega või ainult teistega, mõjutab oluliselt ebaõnnestunud suhtluste arvu. Mõnevõrra vähendab läbikukkumiste arvu see, kui suhtleja on avatud ja üritab suhelda ka siis, kui ta keeleoskus on väga nõrk.

Katsetuste tulemusena selgub, et olulisemad keelekasutajate omadused (keeleoskus erinevate keelte osas, etnosuletus, keelekeskkond) mõjutavad otseselt keelekasutajate omavahelist suhtlemist. Erinevates keeltes suhtlemist eelistavate keelekasutajate geograafiline jaotus keelekeskkonnas mõjutab keelekasutajate omadusi vahetult suhtluste õnnestumistega, sest inimese keelekeskkond määrab enamjaolt, kellega ja mis keeles saadakse suhelda. Tulemused kajastavad, et eesti ja vene keelt mitteoskavad keelekasutajad ei mõjuta neid keeli oskavate inimeste keelevalikuid, sest simulatsioonisüsteemis ei ole inglise emakeelega agente ning seetõttu ei ole inglise keeleoskus piisaval tasemel, et ületada keeleläve ja selle keelega õnnestunult suhelda. Vene keel mõjutab oluliselt peamiselt eesti keelt kõnelevat keelekeskkonda, sest keelekeskkondades, kus on rohkem vene emakeelega inimesi, õnnestub suhtlus enamjaolt just vene keeles. Inglise keel seevastu ei mõjuta niivõrd palju eelkõige eesti keelt kõnelevat keelekeskkonda, sest katsetuste käigus oli simulatsioonis märgata vaid üksikuid juhtumeid, mil suhtluskeeleks valiti inglise keel (näiteks neljandas keelekeskkonnas ehk eestikeelsetes linnades ja maapiirkondades). Ka eesti keel mõjutab olulisel määral peamiselt vene keelt kõnelevat keelekeskkonda, kui keelekeskkonnas on rohkem eestlasi (sellisel juhul suheldakse õnnestunult pigem eesti keeles), aga inglise keel ei mõjuta praktiliselt üldse vene emakeelega inimesi, kuna eelkõige vene keelt kõnelevas keelekeskkonnas ei suhelda inglise keeles edukalt. Katsetuste tulemustest selgus, et olulisemad tegurid, mis mõjutavad kõikides keelekeskkondades suhtlust, on etnosuletus ja keelelävi.

Tulevikus saavad keeleteadlased teostatud NetLogo simulatsiooni abil erinevate muutujate mõju ükshaaval katsetada ja analüüsida, millised on efektid. Kui üksikute muutujate mõju on selge ja realistlik, siis saab katsetada keerulisemate kombinatsioonidega. Niisugusel viisil on simuleeritud keelekeskkondades võimalik tekitada kompleksne muutus hoiakutes ning väljundi põhjal analüüsida tulemusi, kui simulatsioonisüsteemi parameetrid on seadistatud nii, et selle põhjal saavad keeleteadlased hakata katsetama realistlikke keelevahetusstrateegiaid.

Huvitav asjaolu antud töö algandmete faili puhul on see, et peaaegu kõik vastanutest eelistavad rääkida oma kõrgema keeleoskusega keeles ning väga suur osa on vastanud, et nad on nõus suhtlema ka teistes keeltes [16]. Samas on ka hulk vastajaid, kes on vastanud täiesti vastandlikult [16], mis tähendab, et nad valivad suhtluseks alati keele, mida kõik mõistavad, aga kui suhtluspartneri keeleoskus on kehv, siis nad ei eelista kõnelda keeles,

mida partner paremini valdab. Seega küsitluse andmed kajastavad ebastandardsust, mida on raske analüüsida ja simuleerida. Vastuste failist selgub veel, et akommodatsiooni puhul ei kajastu, millist keelt vastanu on nõus rääkima, vaid pigem see, kas ta on nõus suhtlema ka teistes keeltes [16]. Selgus keelte eelistuste osas teeks aga simulatsioonid ja nende analüüsimise lihtsamaks ning tulemused oleksid realistlikumad.

Teostatud simulatsioonisüsteemis esineb palju juhuslikkust, mis tähendab, et iga simulatsioon annab veidi erinevaid tulemusi, nagu ka Andres Karjuse simulatsioonisüsteemis [16]. Seetõttu on mõistlik itereerida simulatsiooni mitmeid tuhandeid kordi, et varieeruvus oleks minimaalne. Ka suhtluse algataja valitakse juhuslikult ning suhtluspartner valitakse talle potentsiaalsete partnerite hulgast. Sarnaselt võib agendi suhtluskatse vastuvõtmine olla pooleldi juhuslik, kuid see varieerub vastavalt keeleoskustele ja akommodatsiooni väärtustele, mis sõltuvad katsetes ette antud parameetrite väärtustest.

4.4 Simulatsiooni rakendamise ja edasiarendamise võimalused

Antud lõputöö simulatsioonisüsteemi valideerimine näitab, et edasisi eksperimente tehes on võimalik eesti keele tuleviku huvides seda kasutada näiteks keeleõppe ressursside suunamiseks, sest simulatsiooni abil on võimalik näha, kus on eesti keele mõju vähenemine (ehk unustamine) kiirem. Katsetuste keeleoskuste keskmine muutus näitab vaadeldavate keelekeskkondade piires võimalikke suhtlusi erinevate suhtlust iseloomustavate parameetrite väärtuste puhul. Eesti keele tuleviku huvides saaks paremini suunata keeleõpet just nendesse keelekeskkondadesse ja nende keelte osas, mis võivad selguda simulatsiooni katsetuste käigus. Selle abil saaks säästa keeleõppeks määratud raha keelekeskkondades, kus keeleoskus on niigi kõrge, aga samas saaks ka suunata keeleõppe ressursse just nendesse piirkondadesse, kus seda on enim vaja. Näiteks selgus teostatud simulatsioonisüsteemi põhjal tõsiasi, et inglise keele oskus vajaks enim järeleaitamist kõigis keelekeskkondades. Inglise keel annaks inimestele võimaluse end välismaal rakendada, mis laiendaks nende silmaringi või kasvõi võimaluse paremini suhelda välismaalastega, näiteks muukeelsete isikutega ja välisüliõpilastega.

Keeleteadlastel on võimalus simulatsioonisüsteemi abil eksperimenteerides ennustada vastavate parameetrite väärtuste korral Eestis tulevikus kasutatava kolme keele

keeleoskusi ning suhtluste õnnestumisi või ebaõnnestumisi ka üle keelekeskkondade piiride.

Teostatud simulatsioonisüsteemi jaoks olid algandmeteks läbiviidud sotsioloogilise küsitluse vastused. Üks oluline suund edasiminekaks oleks uurida lisaks eesti, vene ja inglise keelele ka Eesti elanike keeleoskusi teiste keelte osas nagu näiteks saksa keel ja soome keel. Nende keelte oskamine Eestis on üha suurenev tendents, sest nii seoses töö kui ka õpingutega puututakse kokku just nende keeltega. Tulevikus oleks huvitav jälgida, kuidas ka saksa ja soome keele oskajad mõjutavad suhtlust ja kas mõni eestlane valiks Eestis suhtluseks hoopis ühe nimetatud keeltest.

Simulatsioonisüsteemis on ligikaudu 1 000 küsitletud inimest jaotatud nende elukoha järgi nelja keelekeskkonda ning selle põhjal on tehtud ka antud lõputöö simulatsioonikatsetused ja analüüs. Samas oleks realistlikum, kui jaotus oleks väiksem, näiteks maakondade lõikes. Selline distributsioon annaks täpsemad tulemused keeleoskuste kohta ning parema võimaluse tulemuste analüüsimiseks ja nende tulemuste põhjal näiteks keeleõppe ressursside rakendamiseks.

Käesoleva magistritöö raames teostatud simulatsioonisüsteemis saab tulevikus modelleerida ka seadusandluse ja institutsioonid (näiteks koolid), et hinnata nende mõju Eesti keelekeskkonna arengule [16]. Samas täielikuks koolide analüüsimiseks peaks vastanute hulgas olema ka alla 15-aasta vanuseid inimesi. Hetkel puuduvad algandmete failist lapsed ning seega ei saa antud simulatsioonisüsteemi abil praegu vastuseid küsimustele, mis puudutavad kooli ja perekonna mõju keeleoskustele.

Lisaks on võimalik antud magistritöö simulatsiooni sisse viia ka põlvkondade vahetus, mis mõjutab oluliselt keele unustamist (näiteks eesti murdekeeled ja nende väljasuremine). Peale selle saaks teostatud simulatsioonisüsteemi abil uurida, kas ja kuidas uue keele õppimise alustamise lävendite väärtused, mis on esitatud agendi vastavate lokaalsete muutujate väärtustena, mõjutavad agentide jaotust keelekeskkonnas.

5 Kokkuvõte

Käesoleva magistritöö eesmärgiks oli kavandada ja koostada agendipõhine arvutisimulatsioon, mis võimaldaks keeleteadlastel ja muude valdkondade ekspertidel analüüsida erinevate Eestis kõneldavate keelte elujõulisust ja keelevelikute muutumist globaliseerumise kontekstis. Täiendavaks eesmärgiks oli uurida keelekasutajate omaduste mõju suhtlemisel ja keelekeskkonna mõju keelekasutajate omadustele ning uurida erinevate keelte mõju keelekeskkondadele. Eesmärkide saavutamiseks oli vaja kavandada simulatsioonisüsteem, järgides AOM-metoodikat ja ODD-metoodikat.

Töö koosneb neljast peatükist. Esimeses peatükis kirjeldatakse töö tausta, lahendatavaid probleeme, ülesande püstitust ja metoodikat ning antakse ülevaade tööst kui tervikust. Teises osas kirjeldatakse agentorienteeritud modelleerimise ja agendipõhise simulatsiooni mõisteid ning tuuakse välja simulatsioonisüsteemi kavandamise metoodikad. Kolmandas peatükis modelleeritakse probleemvaldkond vaatepunktide raamistiku analüüsi tasemel, kavandatakse probleemvaldkonna jaoks simulatsioonisüsteem vaatepunktide raamistiku disaini tasemel ning lõpuks kavandatakse detailselt ja teostatakse simulatsioonisüsteem vaatepunktide raamistiku simulatsiooni tasemel. Neljandas peatükis valideeritakse agendipõhise simulatsiooni käitumist erinevate simulatsiooniparameetrite väärtuste korral, valideerides valminud simulatsioonisüsteemi sisendandmete abil ning võrreldes antud simulatsioonisüsteemi tulemusi ühe teise sarnase simulatsioonisüsteemi tulemustega. Valideerimisel kinnitati ka keeleteadlase poolt süsteemi kasutatavust sotsiolingvistilisteks uuringuteks.

Magistritöö tulemusena valmis Eesti keelekeskkonna arvutipõhine mudel, mille abil saab analüüsida eesti keele dünaamikat ja püsijäämist ning erinevate keelte ristmõju suletud süsteemis. Samuti saavad keeleteadlased Eesti keelekeskkonna mudeli abil katsetada mitmesuguseid võimalikke tulevikustsenaariume. Töös püstitatud eesmärgid said simulatsioonisüsteemi kavandamise ja parameetrite väärtusi häälestades ning valideerimiskatsetuste tulemusel saavutatud. Simulatsioonisüsteemi kavandamisega saavutati ka kvaliteedieesmärk esitada adekvaatselt keelekasutajate omavahelist suhtlemist.

Käesoleva lõputöö raames kavandatud ja teostatud simulatsioonisüsteemi rakendamise abil saab edasiste eksperimentide käigus erinevate sisendparameetrite väärtustega katsetades tulemusi, mis võimaldavad näiteks keeleõppe ressursside täpsemat suunamist. Katsetades nelja keelekeskkonna lõikes mitmesuguste parameetrite erinevate väärtustega, saab hinnata erinevate keelte oskuste keskmisi muutusi.

Eesti keelekeskkonna arvutipõhise simulatsioonikeskkonna põhjal saab kujundada tulevikustsenaariume, mille käigus on simulatsiooni abil võimalik analüüsida keele mõjutegureid erinevate võimalike stsenaariumite korral.

Käesoleva magistritöö tulemusi rakendatakse 2017. aastal Tartu Ülikooli institutsionaalse uurimisteema “Eesti keele kestlikkus avatud maailmas“ raames. Tulevikus saab loodud simulatsioonisüsteemis täiendavalt uurida muude Eestis kõneldavate keelte oskusi, sest loodud simulatsioonisüsteemis on võimalik sisse lugeda ja agentidele omistada erinevate sotsioloogiliste küsitluste andmeid.

Kasutatud kirjandus

- [1] L. S. Sterling ja K. Taveter, *The Art of Agent-Oriented Modeling*, London: The MIT Press, 2009.
- [2] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, ... ja D. L. DeAngelis, „A standard protocol for describing individual-based and agent-based models,“ *Ecological Modelling*, kd. 198, nr 1, pp. 115-126, 2006.
- [3] V. Grimm, U. Berger, D. L. DeAngelis, J. G. Polhill, J. Giske ja S. F. Railsback, „The ODD protocol: a review and first update,“ *Ecological Modelling*, kd. 221, nr 23, pp. 2760-2768, 2010.
- [4] A. R. Hevner, S. T. March, J. Park ja S. Ram, „Design Science in Information Systems Research,“ *MIS Quarterly*, kd. 28, nr 1, pp. 75-105, 2004.
- [5] K. Taveter, *A Multi-Perspective Methodology for Agent-Oriented Business Modelling and Simulation*, Tallinn: TUT Press, 2004.
- [6] I. Shvartsman, K. Taveter, M. Parmak and M. Meriste, "Agent-oriented modelling for simulation of complex environments," in *International Multiconference on Computer Science and Information Technology*, Wisla, 2010.
- [7] T. ShaoPeng ja Z. Jun, „A Research on Multi Agent Modeling Language,“ *Procedia Engineering*, kd. 15, p. 1842 – 1847, 2011.
- [8] C. Macal and M. North, "Introductory tutorial: Agent-based modeling and simulation," in *Winter Simulation Conference*, Argonne, 2014.
- [9] P. L. Borrill ja L. Tesfatsion, „Agent-Based Modeling: The Right Mathematics for the Social Sciences?,“ Iowa State University, Iowa, 2010.
- [10] T. Miller, B. Lu, L. Sterling, G. Beydoun ja K. Taveter, „Requirements Elicitation and Specification Using the Agent Paradigm: The Case Study of an Aircraft Turnaround Simulator,“ *IEEE Transactions on Software Engineering*, kd. 40, nr 10, pp. 1007-1024, 2014.
- [11] M. Ehala, K. Koreinik, K. Praakli ja M. Siiner, „Kuidas uurida keele kestlikkust?,“ *Keel ja Kirjandus*, kd. 7, pp. 1-19, 2014.
- [12] X. Castelló, L. Loureiro-Porto ja M. S. Miguel, „Agent-based models of language competition,“ *International Journal of the Sociology of Language*, kd. 2013, nr 221, p. 21–51, 2013.
- [13] M. McPherson, L. Smith-Lovin ja J. M. Cook, „Birds of a feather: Homophily in social networks,“ *Annual Review of Sociology*, kd. 27, pp. 415-444, 2001.
- [14] T. H. Stark, „Understanding the Selection Bias: Social Network Processes and the Effect of Prejudice on the Avoidance of Outgroup Friends,“ *Social Psychology Quarterly*, kd. 78, nr 2, pp. 127-150, 2015.

- [15] K. Leetmaa, T. Tammaru ja D. B. Hess, „Preferences Toward Neighbor Ethnicity and Affluence: Evidence from an Inherited Dual Ethnic Context in Post-Soviet Tartu, Estonia,“ *Annals of the Association of American Geographers*, kd. 105, nr 1, pp. 162-182, 2015.
- [16] A. Karjus, *Testing an agent based model of language choice on real life multilingual data*, Edinburgh: Käsikirjaline materjal autori valduses, 2016.
- [17] M. Ehala ja K. Taveter, *Modeling I meeting*, Tallinn: Meeting memo 03.11, 2015.
- [18] S. M. Goodreau, J. A. Kitts ja M. Morris, „Birds of a Feather, or Friend of a Friend? Using Statistical Network Analysis to Investigate Adolescent Social Networks,“ *Demography*, kd. 46, nr 1, p. 103–125, 2009.
- [19] A. Wimmer ja K. Lewis, „Below and Beyond Racial Homophily. Erg Models of a Friendship Network Documented on Facebook,“ *American Journal of Sociology*, kd. 116, nr 2, p. 583–642, 2010.
- [20] G. Bianconi, R. K. Darst, J. Iacovacci ja S. Fortunato, „Triadic closure as a basic generating mechanism of communities in complex networks,“ *Physical Review E*, kd. 90, nr 4, 2014.
- [21] D. Cartwright ja F. Harary, „Structural balance: a generalization of Heider's theory,“ *Psychological Review*, kd. 63, nr 5, pp. 277-293, 1956.
- [22] R. Allan, „Survey of Agent Based Modelling and Simulation Tools,“ Science and Technology Facilities Council, Warrington, 2011.
- [23] I. Sakellariou, P. Kefalas and I. Stamatopoulou, "Enhancing NetLogo to Simulate BDI Communicating Agents," in *Artificial Intelligence: Theories, Models and Applications*, Syros, 2008.
- [24] P. Schreinemachers ja T. Berger, „An agent-based simulation model of human–environment interactions in agricultural systems,“ *Environmental Modelling & Software*, kd. 26, nr 7, p. 845–859, 2011.
- [25] U. Wilensky, „NetLogo,“ Center for Connected Learning and Computer-Based Modeling, 1999. [Võrgumaterjal]. Available: <http://ccl.northwestern.edu/netlogo/>. [Kasutatud 13 03 2017].
- [26] F. Santana, D. Nagata, M. Cursino, C. Barberato and S. Leal, "Using BPMN-based Business Processes in Requirements Engineering: the Case Study of Sustainable Design," in *International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE)*, Athens, 2016.
- [27] B. Edmonds ja R. Meyer, *Simulating Social Complexity - A Handbook*, Berlin: Springer, 2013.
- [28] A. Karjus, „Keelekeskkonnad,“ Tartu Ülikool, 2016. [Võrgumaterjal]. Available: http://kodu.ut.ee/~andres87/abm_proto.html. [Kasutatud 13 03 2017].

Lisa 1 – Simulatsioonisüsteemi seadistatud parameetrid

Tabel esitab valminud simulatsioonisüsteemi valideerimise järel seadistatud parameetrite väärtused ning nende kirjeldused.

| Parameeter | Väärtus | Kirjeldus |
|------------------------------------|----------------|---------------------------|
| <i>RUN-times</i> | 10000 | Iteratsioonide arv |
| <i>raport-frequency</i> | 5000 | Raporteerimise tihedus |
| <i>gender-proportion</i> | 0.1 | Soo mõjutegur |
| <i>age-group-proportion</i> | 0.1 | Vanusegrupi mõjutegur |
| <i>nationality-proportion</i> | 0.1 | Rahvuse mõjutegur |
| <i>native-lang-proportion</i> | 0.1 | Emakeele mõjutegur |
| <i>lang-environment-proportion</i> | 40 | Keelekeskkonna mõjutegur |
| <i>elsewhere-proportion</i> | 0.1 | Mujal viibimise mõjutegur |
| <i>often-away-proportion</i> | 0.1 | Mobiilsuse mõjutegur |
| <i>active-citizen-proportion</i> | 0.1 | Aktiivsuse mõjutegur |
| <i>random-proportion</i> | 0.1 | Juhuslikkuse mõjutegur |
| <i>triadic-closure-increment</i> | 0.9 | Sõbra sõbra mõjutegur |
| <i>lang-increment</i> | 0.35 | Keeleoskuse kasv |
| <i>lang-decrement</i> | 0.15 | Keeleoskuse langus |
| <i>success-const</i> | 0.1 | Õnnestumise konstant |
| <i>failure-const</i> | 0.1 | Ebaõnnestumise konstant |
| <i>accommodation-index</i> | 0.1 | Akommodatsiooni indeks |
| <i>ethnic-closure-index</i> | 0.1 | Etnosuletuse indeks |
| <i>media-index</i> | -1.0 | Meedia indeks |
| <i>communication-threshold</i> | 1.0 | Keelelävi |

Lisa 2 – Simulatsioonisüsteemi kood

Magistritöö tulemusel valminud simulatsioonisüsteemi kood on tööga eraldi kaasas oleval CD-plaadil, mis on lõputöö tagakaanele kinnitatud CD-tasku sees. CD-plaadil olevas kaustas on simulatsioonisüsteemi fail ning import ja eksport kaustad. Simulatsiooni avamiseks ja käivitamiseks on vaja NetLogo tarkvara versiooni 6.0 või sellest uuemat versiooni. Keelesimulatsiooni kasutajaliideses on kolm alternatiivset akent kasutajaliidese, info ja koodi jaoks. Vaikimisi avaneb kasutajaliidese aken, millel on näha liugureid suhtluspartneri valiku mõjutegurite, keele õppimise ja unustamise muutmise, suhtluspartneri valiku statistika muutmise ja keeleläve arvutuse jaoks. Veel sisaldab kasutajaliides kahte graafikut, mis näitavad infot simulatsiooni käigu kohta, õnnestunud ja ebaõnnestunud suhtlusi ning keelemuutuste arvu. Keelesimulatsioonis saab luua algseisu nupuga „*Setup*“, millega luuakse erinevate värvidega tähistatud keelekogukondade liikmed vastavalt simulatsiooni aluseks olevale küsimustikule. Suhtluskatsete tegemise algoritm hakkab tööle nupuga „*Communicate loop*“, mille puhul suhtluskatseid peegeldatakse erivärviliste joontega.