

TARTU ÜLIKOOL
MAAILMA KEELTE JA KULTUURIDE KOLLEDŽ

Krete Saak

**Tõlkemälu ja masintõlke vastete mõju kognitiivsele koormusele
tõlkimisprotsessis**

Magistripjekt

Juhendajad: Marju Taukar
Terje Loogus

Tartu 2019

Sisukord

Sissejuhatus	4
1. Masintõlke ja tõlkeabiprogrammide ajaloost	6
2. Masintõlke liigid	6
3. Tõlkeabiprogrammid	8
4. Pilgujälgimine ja pilgujälgijad	9
5. Tõlkeprotsessi uuringud	12
5.1 Meetodite ülevaade	12
5.2 Varasemad tõlkeprotsessi uuringud	13
6. Masintõlke järeltoimetamise uuringud	14
7. Tõlkeprotsessi uurimise katse	17
7.1 Osalejad	17
7.2 Tekst	18
7.3 Katse käik	19
7.4 Andmeanalüüs	21
7.5 Kasutatud kognitiivse koormuse näitajad	23
8. Tulemused	25
8.1 Töötlemisskiirus	26
8.2 Pupilli suuruse muutuse protsent	27
8.3 Fiksatsioonide arv lähteteksti sõna kohta	29
8.4 Silmitsemisaeg lähteteksti sõna kohta	30
8.5 Katse kestus	31
9. Tulemuste analüüs	32
9.1 Kognitiivne koormus	32
9.2 Erinevused professionaalide ja algajate vahel	33

9.3 Muud tähelepanekud	35
9.4 Tagasiside küsitluse vastused	36
Kokkuvõte	37
Resümee	39
Kasutatud kirjandus	41
Lisad	44
Lisa 1. Katses tõlgitud tekst	45
Lisa 2. Lähteteksti segmendid ja nende vasted	46
Lisa 3. Uuritava informeeritud nõusolekuleht	49
Lisa 4. Küsitlus	50
Lisa 5. Katseisiku P1 andmeanalüüsi tulemused segmentide kaupa	52

Sissejuhatus

Viimase kolmekümne aasta jooksul on maailmas aina enam uuritud seda, mis toimub tõlkimise ajal tõlkija peas. Krings (1986), üks esimesi tõlkeprotsessi uurijaid, kirjutas, et kuni me ei proovi paljastada seni musta kastina koheldud tõlkija peas toimuvat protsessi, ei saa me teada, mis moodustab tõlkepädevuse. Üle maailma on palju uuritud selle musta kasti sisu ning see pakub aina rohkem huvi globaliseeruvast maailmas, kus tõlge omab tähtsat rolli.

Viimastel aastakümnetel on hüppelise arengu teinud ka masintõlge ja tõlkeabiprogrammid. Aina enam kasutavad tõlkijad tõlkeabiprogramme ning aina enam peavad tõlkijad tegelema hoopis masintõlke järeltoimetamisega. Tõlkebürood arvestavad tõlkijale palka määrates tõlkemälu. Segmendi eest, millele on tõlkemälu vaste olemas, saab tõlkija vähem palka. Samuti ei ole nullist tõlkimine ja masintõlke järeltoimetamine sama hinnaga. Tõlkeprotsessi uuringutes on aina rohkem huvi pakkunud just nimelt nende tehnoloogiliste arengute tegelik mõju tõlkeprotsessile. Masintõlke kvaliteet ja seega ka sellest saadav abi sõltuvad aga tugevalt lähte- ja sihtkeelest, mistõttu ei saa teiste keelte uurimisest infot eestikeelse masintõlke kohta.

Eesti teadusmaastikul on tõlkeprotsessi ning tõlkemälu ja masintõlke vastete mõju tõlkeprotsessile uuritud vähe. Samas sõltub paljude Eesti tõlkijate palk nendest abivahenditest. Valisin oma magistritöö teemaks tõlkemälu ja masintõlke mõju tõlkija kognitiivsele koormusele ehk mentaalsele tööle, et panustada Eesti tõlkija n-ö musta kasti uurimisse. Töö idee põhineb O'Brieni (2007) artiklil, kus ta vaatas tõlkimist inglise-prantsuse ja inglise-saksa suunal.

Eesti tõlkeprotsessi uuringutesse panustamiseks viisin läbi katse, kus 20 tõlkijat tõlkisid lühikest teksti inglise-eesti suunal. Tõlkimisel kasutasid nad tõlkeabiprogrammi, mis sisaldas erineva täpsusega vastetega tõlkemälu ning osaliselt sisse lülitatud masintõlget. Katse toimus spetsiaalse arvuti taga, mis oli ühendatud pilgijälgijaga.

Siinses töös uurin, kuidas tõlkemälu osalised ja täpsed vasted ning masintõlge mõjutavad tõlkija kognitiivset koormust. Töö peamine eesmärk on järjestada kognitiivne koormus segmendi tõlkimisel tõlkemälu täpse vaste, osalise vaste või masintõlke olemasolul või ilma vasteta.

Lähemalt vaatan masintõlke järeltoimetamise kognitiivset koormust. O'Brieni (2007) uurimuse tulemused viitasid sellele, et järeltoimetamise kognitiivne koormus on sarnane

osalise vastega segmendi tõlkimise kognitiivsele koormusele. Kas see peab paika ka inglise-eesti tõlkesuunal?

Tõlkemälu mõju uurimiseks üritan leida vastust küsimusele, kas tõlkijad liiguvad täpsetest vastetest automaatselt edasi või kontrollivad/muudavad neid. Täpsed vasted peaksid tegema tõlkija elu lihtsamaks ja nende eest makstakse ka seetõttu vähem, kuid kontekstide erinevus võib muuta teksti sobimatuks. Sama lause ühest tekstist ei pruugi sobida teise. Kuid kas tõlkijad võtavad aja nende kontrollimiseks, kui aeg on tihtipeale just see, mida neil on vähe?

Selles uuringus osales kakskümmend katseisikut: kolmteist algajat ja seitse professionaali. Püüan välja selgitada, kas kogemus vähendab kognitiivset koormust ja kas töö peamise eesmärgina loodud järjestus on algajate ja professionaalide puhul sama.

Kuna tõlkeprotsessi uurimine Eestis on arengujärgus, siis vaatan lõpetuseks katseisikute tagasisidet ning annan mõned soovitusel edasisteks uuringuteks.

Töö algab väikese ülevaatega masintõlkest, tõlkeabiprogrammidest ja pilgijälgijast. Seejärel kirjutan lühidalt varasematest uurimustest, mis seostuvad selle tööga. Järgneb tõlkeprotsessi uurimiseks läbi viidud katse meetodi kirjeldus ning seejärel esitan tulemused ning analüüsin neid.

1. Masintõlke ja tõlkeabiprogrammide ajaloost

Idee kasutada masinaid tekstide tõlkimiseks sai alguse arvuti loomisest ehk juba 1940ndatel, mil kasutati esimesi arvuteid koodide murdmiseks. 1947. aastal kirjutas üks masintõlke arendamise eestvedajaid Warren Weaver, et keelt saab vaadata krüptoloogilisest aspektist. Ta vaatas keelt kui krüptitud versiooni teisest keelest. Weaveri mõtted tekitasid suurt entusiasmi – masintõlke võimalust hakati uurima ning selle arendamisse investeeriti suuri summasid. (Koehn, 2010; Nitzke, 2019; Trujillo, 1999)

USA valitsus moodustas loomulike keelte automaattõõtluse konsultatiivkomitee, mille ülesanne oli uurida, kui hästi masintõlge tegelikult töötas. Nende raport, mis avaldati 1966. aastal, vähendas massiliselt masintõlke uuringute investeringuid. Nad leidsid, et masintõlge ei tasu end ära, kuna selle järeltoimetamine on sama kallis kui teksti tõlkimine. See andis tagasilöögi masintõlke uuringutele USAs, kuid mujal arendamine jätkus. (Koehn, 2010; Nitzke, 2019)

Juba 1970ndatel suudeti luua esimesi töötavaid masintõlkesüsteeme – Météo süsteem ilmaennustuste tõlkimiseks ja Systran, mis oli algselt vene-inglise ja hiljem inglise-prantsuse süsteem. 1990ndatel võttis masintõlke arendamine uue suuna ning loodi tõlkeabiprogrammid, mis toetavad inimese tõlkeprotsessi. (Koehn, 2010; Nitzke, 2019; Trujillo, 1999)

Tõlkeabiprogrammidest on saanud tõlkijate igapäevased abivahendid ning masintõlge on nendesse integreeritud, kuid tõlkimine (või siis masintõlke järeltoimetamine) on siiski inimese teha. Praegu jätkub masintõlke arendamine paljudes ülikoolides ja ettevõtetes.

2. Masintõlke liigid

Masintõlge on teksti tõlkimine arvuti abil ühest loomulikust keelest teise (Hutchins, 1986). Masintõlke vajadus on peamiselt praktiline – väga paljud inimesed erinevatelt aladelt peavad lugema ja suhtlema keeltes, mida nad ei oska. Masintõlge muudab selle palju lihtsamaks.

Masintõlke süsteeme liigitatakse tõlkimise strateegia järgi järgmiselt.

- Otsetõlge oli esimene strateegia, mida kasutati. See süsteem otsib igale sõnale vaste sõnastikust ning kasutab talle teada olevaid reegleid, et teha vajalikud morfoloogilised

muudatud. Sellise süsteemi arendamine on väga keerukas ja ajakulukas, kuna iga keele ja keelesuuna reeglid tuleb eraldi sisestada. (Nitzke, 2019; Poibeau, 2017)

- Ülekandestrategie on otsetõlkest keerulisem, kuna see sisaldab ka süntaktilist analüüsi. See loob lähtetekstist süntaktilise struktuuri, mis kantakse üle vastavaks sihtkeele struktuuriks, millest luuakse tõlge. (Nitzke, 2019; Trujillo, 1999)
- Interlingvistiline süsteem kasutab vahekeelt ehk interlinguaalset esitust, mis annab edasi lähteteksti mõtte abstraktsel kujul. Interlinguaalse esituse saab tõlkida sihtkeelde lähtekeelest sõltumatult, mistõttu on uue keele lisamine lihtsam. Sellise interlinguaalse esituse, mis sobiks paljudele keeltele, leidmine on aga keeruline. (Nitzke, 2019)
- Näitepõhine süsteem kasutab mitmekeelset korpust, kust ta otsib lähtelausele kõige sarnasema lause ja seob selle teiste sarnaste lausetega. See süsteem töötati peamiselt välja Jaapanis 1980ndatel. (Nitzke, 2019)
- Viimastel kümnenditel on peamiselt uuritud ja arendatud statistilist süsteemi. See kasutab paralleelkorpust tõlkimise õppimiseks. Sihtteksti saamiseks koostab tarkvara n-ö treeningfaasis masinõppimise abil statistilise mudeli. Seda mudelit kasutab see hiljem tõlke genereerimiseks. (Goutte, Cancedda, Dymetman, & Foster, 2009; Nitzke, 2019)
- Kõige uuem on närvivõrgupõhine süsteem ehk neurotõlge. See kasutab samuti paralleelseid treeningkorpuseid, kuid loob suure ühise tõlke närvivõrgu. Iga keele jaoks on üks kooder ja dekodeer. Kooder loeb lähtelauset ja väljastab kindla pikkusega vektori, mida loeb dekodeer, mis seejärel väljastab sihtlause. (Nitzke, 2019)

Hetkel kasutatakse kõige rohkem statistilist masintõlget ja neurotõlget, mille tulemused on ka kõige paremad. Neurotõlge on küll väga uus süsteem, kuid sellel on palju potentsiaali.

Praegu on masintõlge internetis kõigile kättesaadav. Hetkel on masintõlkesüsteeme, mis kindlas valdkonnas ja kindlate keelte vahel annavad väga häid tulemusi, kuid enamik masintõlgitud tekste vajavad siiski inimese toimetust. Peamiselt on masintõlke eesmärk aga üldtõlge, mis ei pruugi olla keeleliselt korrektne, aga annab edasi tähtsaima informatsiooni (Koehn, 2010; Nitzke, 2019).

Suured tarkvara arendajad nagu IBM, Google ja Microsoft on välja tulnud oma masintõlke süsteemidega ning masintõlkega tegelevad ettevõtted nagu Systran muudavad oma

süsteeme aina paremaks. Igapäevaselt nähakse vaeva selle nimel, et ühel päeval ei oleks masintõlge ainult kasulik vaid ka hea. (Koehn, 2010)

3. Tõlkeabiprogrammid

Tõlkeabiprogrammid on tõlkija suurimad abilised. Praegu on turul palju tõlkeabiprogrammide pakkujaid, näiteks Memsources, memoQ ja SDL Trados.

Tõlkeabiprogrammidel on väga palju erinevaid funktsioone, mis tõlkijat aitavad. Funktsioonide kasutusmugavus ja hulk määravad tõlkeabiprogrammi väärtuse tõlkijale, kuid iga programmi peamine tööpõhimõte on sama. Tõlkija saab programmi sisestada tõlgitava faili, mille tõlkeabitarkvara muudab segmentideks. Üks segment koosneb ühest lausest, mida saab vaadata eraldiseisvana. Lisaks sellele hoiab programm ka teksti vormistust, paigaldades vajalikesse kohtadesse märgendeid. Nii saab tõlkija teha oma tööd vormistuse pärast muretsemata. Programmis on ühel küljel kastikeses lähtesegmendid ja nende kõrval (või all, olenevalt tõlkija harjumusest) algselt tühi kast ehk sihtsegmendi koht, kuhu tõlkija saab oma tõlke sisestada.

Miks aga üldse muuta teksti lauseteks? Tõlkeabiprogrammi süda on tõlkemälu. See salvestab iga varem tõlgitud lähte- ja sihtteksti segmendi. Kui tekstis (või järgmistes tekstides) tuleb ette sama lähtetekstiga segment, siis tõlkemälu on vaste olemas ning tõlkija ei pea seda uuesti tõlkima. Niimoodi saab ta aastate jooksul koostada andmebaasi oma tõlgetega. Tihtipeale antakse aga tarbetekstide tõlkijale koos lähtetekstiga ka tõlkemälu, mis sisaldab sarnaste tekstide varasemaid tõlkeid. Sel moel saab iga tõlget vajav ettevõtte garanteerida, et terminoloogia nende tekstides on ühtlane.

Tõlkemälu sisaldab erinevat sorti vasteid. Sellist vastet, mille lähtetekst on täpselt sama tõlgitava segmendi lähtetekstiga, nimetatakse täpselt vasteks ehk 100% vasteks. Kui ka lauset ümbritsevad laused on täpselt samad, mis eelmisel tõlkimisel, siis on tegemist konteksti vastega ehk 101% vastega. Kui tõlgitav tekst on väga sarnane varem tõlgitud segmendile, kuid seal on mõningad erinevused (näiteks märgendid või mõni sõna), siis programm näitab seda siiski tõlkijale ning näitab ka lähtetekstide erinevust. Sellist vastet nimetatakse osaliseks vasteks. Vaste protsent sõltub lähtelausete kattuvusest. Mida suurem on lähtelausete erinevus, seda väiksem on vaste protsent. Näiteks märgendite erinevuse korral on tavaliselt tegemist 99% vastega, kuid ühe sõna erinevusel on vaste protsent väiksem.

Joonisel 1 on näha Memsource'i tõlkeabiprogrammi tavaline tõlkevaade. Vasakul on lähtetekst ja paremal sihtteksti kast, kuhu on joonisel sisestatud täpsed, osalised ja masintõlke vasted. Tõlkimisala all on teksti eelvaade. Täiesti paremal üleval on kast, kuhu kuvatakse tõlkemälu vasted ja masintõlge (lähtetekst vasakul, keskel vaste protsent ja paremal varem tõlgitud sihttekst). Selle all on kast, kus kuvatakse osaliste vastete puhul lähteteksti erinevused. Punasega ja maha tõmmatult on märgitud see osa, mida praegu tõlgitavas lähtetekstis ei ole, ning rohelisega on märgistatud see, mis praeguses lähtetekstis on lisaks. Pildil on avatud katses kasutatud tekst.

The screenshot displays the Memsource Web Editor interface. At the top, there is a menu bar with options like 'Edit', 'Tools', 'Format', 'Document', and 'Help'. Below the menu is a toolbar with various icons for text editing and navigation. The main area is divided into several sections:

- Filter Source Text** and **Filter Target Text** input fields with a 'Clear Filter' button.
- A table with columns for '#', 'Source: en', and 'Target: et'. The table contains 17 rows of text, including safety instructions and warnings. Some rows are highlighted in orange, indicating differences or specific actions.
- Context Note** section showing a preview of the source text: 'FILE BELT SANDER INSTRUCTION MANUAL 1/8-3/4 in. x 18-24 in.'.
- Translation memory** section on the right, showing a list of previous translations with their scores (e.g., 83, 68, 76, 100, 99, 99).
- At the bottom, there is a status bar showing 'Segments: 44/0', 'Words: 533/0', 'Chars: 2867/0', '128/142', and 'File: TU projekt.docx'.

Joonis 1: Memsource'i tõlkeabiprogrammi tavaline tõlkevaade.

Tõlkeabiprogrammidel on lisaks veel palju erinevaid funktsioone, alustades terminibaasidest ja lõpetades kvaliteedikontrolli ning projektipakkidega. Selles töös omab aga suurimat tähtsust tõlkemälu ning ka masintõlge. Nimelt on võimalik ühendada tõlkeabiprogrammi masintõlkega ning sel juhul pakub see tõlkijale lisaks tõlkemälu vastetele ka masintõlgitud vastet.

4. Pilgijälgimine ja pilgijälgijad

Inimese pilgu liikumise ja selle tähenduse üle on arutletud juba Aristotelese ajast peale. Lihtne jälgimine viis väärvõimuseni, et silmad libisevad üle vaadatava ja peatuvad

liikumatult huvipakkuval. Alles 19. sajandi lõpus ja 20. sajandi alguses loodi esimesed pilgijälgijad ning saadi teada, kuidas silmad tegelikult liiguvad. (Conklin, Pellicer-Sanchez, & Carroll, 2018)

Arvati, et silmad liiguvad ühe sujuva liigutusega, kuid tegelikult liiguvad need üle vaadatava, tehes hüppeid ja lühikesi pause. Ühe hüppega läbitav vahemaa ja pauside pikkused on muutuvad, kuigi ilma pingutuseta jääb pausi pikkus alati väga lühikeseks. Inimene teeb 3-5 silmaliigutust sekundis (Holmqvist, Nyström, & Andersson, 2011). Seega ei näe inimene mitte liikuvat pilti vaid üksikute fikseeritud piltide seeriat, mis vaheldub kiirelt. (Conklin et al., 2018)

Silma hüppeid ehk liikumist nimetatakse sakaadiks ja silma peatumist nimetatakse fikseerimiseks, mis on tegelikult veidi eksitav nimetus, kuna silmas on alati kasvõi väike värin. Kui sakaad liigutab näiteks teksti lugedes pilku tagasi, siis seda nimetatakse regressiooniks. Lugedes või lihtsalt vaadates fikseerub pilk, et töödelda informatsiooni selles punktis, ning seejärel liigub järgmisesse punkti, mis sisaldab muud informatsiooni. Fiksatsioonil ajal töötleb kognitiivne süsteem visuaalset sisendit ja planeerib järgnevat liigutust. (Conklin et al., 2018; Holmqvist et al., 2011)

Fiksatsioon, sakaad ja regressioon toimuvad automaatselt ja seega võimaldab pilgijälgimine saada informatsiooni alateadliku tegevuse kohta. Pilgijälgimise süsteem võimaldab uurida, kuhu inimese pilk langeb, mitu korda see sellele kohale või alale langeb (fikseerimiste arv), kui kaua iga fikseerimine kestab (fikseerimise kestus) ning ka sakaadi pikkust ja kestust. Masinat, mis kõike eelnevat mõõdab, nimetatakse pilgijälgijaks. (Conklin et al., 2018)

Miks see pakub huvi? Inimene liigutab oma silmi selleks, et saada mingit kindlat objekti oma vaateväljas fookusesse selle selgeks nägemiseks. Enamiku ajast langeb pilk sellele, millele soovitakse kasvõi lühikeseks ajaks keskenduda. Seega saab pilgijälgimisel näha, mis inimese tähelepanu haaras ja mis talle huvi pakkus. (Duchowski, 2017)

Pilgu fikseerimise tuvastamine võimaldab uurida ka inimese mõttetegevust. Millele pilk fikseerub, selle töötlemisega aju ka hetkel tegeleb (Just & Carpenter, 1980). Keerulistes kohtades, mille töötlemine (ehk millest aru saamine) nõuab rohkem aega, fikseerub pilk pikemalt või rohkem kordi. Seega saab teksti keerukust uurida vaadates fikseerimiste pikkust ja hulka.

Lisaks fiksatsioonidele on mentaalse töö üks peamine mõõdupuu pupilli suuruse muutumine. Erinevad uuringud on näidanud, et mida rohkem inimene pingutab, seda rohkem tema pupill suureneb (Holmqvist et al., 2011). Pupilli suuruse muutust kutsuvad aga esile ka paljud muud faktorid ning seetõttu peab olema eksperimendi ülesehitus ning aparatuuri asukoht hästi läbi mõeldud. Näiteks valgus mõjutab pupilli suuruse muutust märgatavalt, mistõttu peab sellise uuringu tegemisel olema katse asukohas muutumatu valgustus. Pupilli suurust mõjutavad ka näiteks emotsioonid, väsimus, vanus, valu ja ravimid (Holmqvist et al., 2011). Seetõttu on soovitatav erinevate inimeste pupillimuutuste võrdlemisel kasutada suhtarvu.

Esimesed pilgujälgijad valmistati 19. sajandi lõpus ja need olid katsealustele väga ebamugavad. Näiteks tuimestati silm kokaiinilahusega ja asetati sellele rõngas, mis ühendati mehaanilise hoova külge. 20. sajandil hakati fotografeerima välise valgusallika peegeldust tsentraallohus, mis muutis uuringud palju mugavamaks. Kuni 70ndateni pidi iga teadlane endale ise masina ehitama, mis muutis uurimistöö aeglaseks, kuid iga teadlane teadis täpselt, mida tema masin võimaldab teha ja kuidas seda parimate tulemuste saamiseks kasutada. (Holmqvist et al., 2011)

Alates 70ndatest saab silmajälgijaid osta. Tänapäeval saab valida paljude erinevate mudelite vahel. Enamik neist kasutab videopõhist süsteemi ehk analüüsitakse videopilti jälgitavast silmast, mida valgustatakse infrapunase valgusallikaga. Silmajälgija salvestab pilgu x ja y koordinaadid ekraanil ning seejärel tarkvara ühendab need andmed eksperimendi andmetega, mille tulemusel saadakse silmade liikumise ja peatumise kohta andmeid huvipakkuvatel aladel. (Conklin et al., 2018)

Pilgu jälgimisel saadavat informatsiooni on niivõrd palju, et selle kohta on kirjutatud mitmeid raamatuid (eelpool tsiteeritud jm). Kuna tegemist on üsnagi uue ja pidevalt areneva valdkonnaga, siis mitmeid pilgu jälgimisel saadavaid andmeid ei osata veel tõlgendada. Olenevalt teadustöö eesmärgist tuleb valida muutujad, mis töö eesmärkidega kõige paremini sobivad.

5. Tõlkeprotsessi uuringud

5.1 Meetodite ülevaade

Holmes jagas kirjeldavad tõlkeuuringud kolmeks: tõlkefunktsiooni, -produkti ja -protsessi uuringud (Toury, 1995). Tõlkeprotsessi uuringute eesmärk on mõista tõlkija kognitiivset protsessi, mis on väga keeruline. Tõlkimine hõlmab ühes keeles teksti hoomamist ja selle esitamist teises keeles ning ka kahe keele vahel ümberlülitumise protsessi. Seetõttu tuginevad tõlkeprotsessi uuringud tugevalt erinevate valdkondade teoreetilistele lähenemistele. Neist peamised on lingvistika, psühholingvistika, kakskeelsuse ja teise keele omandamise uuringud ja kognitiivne psühholoogia. (Englund Dimitrova, 2010)

Tõlkimise kognitiivset protsessi ei saa uurida lihtsa vaatluse teel. Seetõttu seisnevad meetodid peamiselt erinevate vaatluste ja salvestuste andmete kogumises, millest saab teha järeldusi kognitiivse protsessi kohta. Tõlkeprotsessi uuringutele andis tõuke kognitiivpsühholoogide Ericssoni ja Simoni 1984. aastal avaldatud teos „Protocol analysis“, mis tutvustas introspektiivseid meetodeid, mis muutsid võimalikuks tõlkeprotsessi empiirilise uurimise. Kõige enam on kasutatud nendest meetoditest valjustimõtlemist, kus tõlkija tõlkimise ajal kirjeldab valjusti oma mõtteid. Teine palju kasutatud meetod on retrospektiivne intervjuu, kus tõlkija räägib pärast ülesande lahendamist oma tõlkeprotsessist. (Englund Dimitrova, 2010)

1990ndatel arenesid välja erinevad klahvivajutuste salvestamise programmid nagu Translog, mis salvestavad iga klahvivajutuse ja hiireliigutuse ning nende toimumise aja ning võimaldavad näha kõiki kustutusi, parandusi ja ka pause (Jakobsen, Mees, & Göpferich, 2012). Abimaterjalide kasutamise uurimisel on tavaline ekraani salvestamine ja osaleja jälgimine. Ekraani salvestamine võimaldab ka uurida tõlkija töökeskkonda ekraanil. Tähtsal kohal on ka pilgujälgimine. (Englund Dimitrova, 2010)

Tihti peale kasutatakse mitmeid meetodeid korraga, näiteks pilgujälgimine koos ekraani ja klahvivajutuste salvestamisega (nagu selles töös).

Tõlkeprotsessi uuringud on peamiselt olnud hüpoteese loovad, kuna uurimismeetodid annavad isegi lühikese teksti korral väga suure hulga andmeid, mistõttu kasutatakse paljudes uuringutes vähe katsealuseid ning lähte- ja sihttekste. Seetõttu ei saa teha laialdasi järeldusi. (Englund Dimitrova, 2010; O'Brien, 2009)

Lisaks suurele andmehulgale valmistab probleeme katse keskkond, mis ei pruugi samastuda igapäevase tõlkeskeskkonnaga (Englund Dimitrova, 2010; O'Brien, 2009). Näiteks igapäevaselt tõlkijad pigem ei mõtle valjusti, mistõttu võib see nõue muuta nende tavapärast tegevust. Tõlkijad teevad tavaliselt tööd üksinda ja seetõttu võib vaatleja juuresolek nende valikuid ja liigutusi muuta. Kuid aina arenev tehnoloogia ja tõlkimise muutumine arvutikeskseks vähendab selle probleemi ulatust.

5.2 Varasemad tõlkeprotsessi uuringud

Tõlkeprotsessi uuringute esimesed uurimisvaldkonnad olid tõlkimisel eettulevad probleemid ja nende lahendamine ning tõlkeprotsessi üldine põhiraamistik. Üks esimesi töid oli Hans-Peter Kringsi uurimus (Krings, 1986), kus ta kirjeldas kaheksa prantsuse keelt õppiva sakslase (ehk mitteprofessionaalide) tõlkeprotsessi. Ta analüüsis katsealuste abimaterjalide ja aja kasutust, tekkinud probleeme ning nende lahendusi.

Kui esimesed tööd tegelesid tõlkimise alustaladega, siis edaspidi keskenduti pigem kindlatele teemadele ja uurimisküsimustele. Selle töö raames on tähtsad kognitiivset koormust, tõlkeabitarkvara ja masintõlke järeltoimetamist puudutavad uurimused.

Juba 1964. aastal uurisid pilgujälgijaga kognitiivset koormust Hess ja Polt. Nad leidsid, et pupilli suurenemine sõltub ülesande raskusest. Mida raskem on ülesanne, seda rohkem pupill suureneb (Hess & Polt, 1964). Seda sama tulemust on hiljem tõestatud paljudes erinevate valdkondade uuringutes (Beatty & Lucero-Wagoner, 2000; Iqbal, Adamczyk, Zheng, & Bailey, 2005; Moresi, 2008). Beatty (1982) uuris suurt hulka eksperimentaalseid andmeid ja järeldas, et pupilli muutus on usaldusväärne mentaalse raskuse indikaator.

Sharmin, Spakov, Rähä, ja Lykke Jakobsen (2008) uurisid tõlkija pilku tõlkimisel ning leidsid, et teksti keerukus ei mõjutanud fiksatsioonide pikkust, aga fiksatsioonide arv oli keerulise teksti korral tunduvalt suurem kui kerge teksti juures. Ka Schnitzer ja Kowler (2006) leidsid, et keerulist teksti lugedes toimub rohkem regressioone, kuid ka sakaadidevaheline aeg on pikem. Dragsted (2012) uuris erinevate tõlkevariantide arvu seost kognitiivse koormuse näitajatega. Tema uuringu tulemused näitasid, et paljude tõlkevariantidega sõnadel toimunud fiksatsioonide arv, nende pikkus ja pauside pikkus olid suuremad kui väheste tõlkevariantidega sõnadel.

Selle magistr töö idee alus on Sharon O'Brieni 2007. aastal avaldatud artikkel „Eye-Tracking and Translation Memory Matches“ (O'Brien, 2007). Tema uuringu eesmärk oli välja töötada meetod kognitiivse koormuse mõõtmiseks tõlkeabitarkvaraga tõlkimisel. Neljal tõlkijal paluti tõlkida lühikest teksti. Kaks tõlkisid seda prantsuse keelde ja kaks saksa keelde. Lähtetekstiga oli kaasas ka tõlkemälu, mis sisaldas erinevaid vasteid, sealhulgas nii osalisi kui täpseid vasteid. Lisaks loodi mõned vasted masintõlkega ning osa teksti oli ka ilma vasteta. Uuringu tulemusel saadi teada, et pupilli suurenemine ja töötlemiskiirus (aeg, mis kulub ühe sõna tõlkimiseks) olid tugevalt seotud erinevate vastetega. Suurim oli tõlkija koormus ilma vasteta segmendi tõlkimisel ja vähim täpse vastega segmendi juures. Masintõlke ja osalise vastega segmentide koormus jäi samasse piirkonda.

Hiljem on O'Brien (2008) uurinud eraldi ka osaliste vastete keerukust. Ta avastas, et töötlemise kiirus kasvas koos vaste protsendi kasvuga, kuid pupilli suurenemise ja vaste protsendi kasvu vahel ei olnud sellist sõltuvust.

Alves ja Campos (2009) vaatlesid tõlkemälu ja ajanappuse mõju abimaterjalide kasutamisele. Nad leidsid, et tõlkemälu muudab tõlkijate käitumismustrit ja optimeerib väliste abivahendite kasutust (näiteks kasutatakse tõlkemälust otsingut), kuid tõlkemälu olemasolu ei vähenda tõlkija sisetunde tähtsust, vaid hoopis suurendab seda. Nende uuringu tulemused näitasid, et tõlkijad kontrollivad tõlkemälu vasteid, kui selleks on aega. Ajapuudusel aga suureneb tõlkemälu vastete usaldamine.

6. Masintõlke järeltoimetamise uuringud

Järeltoimetamine on masina toortõlke parandamine tõlkija poolt, järgides kindlaid juhtnööre ja kvaliteedinõudeid (O'Brien, 2011). Aina rohkem kasutatakse teksti tõlkimiseks masintõlget, mille väljund saadetakse tõlkijale järeltoimetamiseks. See erineb tavapärasest tõlkimisest, kuna tõlkija ei pea sihtteksti looma algusest, vaid tal on olemas võimalik tõlge. Kuid see ei ole ka tavaline toimetamine, sest esinevate vigade tüübid on täiesti erinevad. Näiteks masintõlkes leidub harva trükivigu, kuid seal võib olla süntaktilisi ja leksikaalseid vigu, mida esineb harva inimese tõlgitud tekstis. (Nitzke, 2019)

Järeltoimetamise uuringutes on kaks peamist küsimust: kui suurel määral on masintõlgitud vasted vastuvõetavad ning kui palju inimene peab pingutama, et selle vigu parandada. Koormust on kolme tüüpi: ajaline (kui palju aega kulub teksti järeltoimetamiseks),

tehniline (kui palju on vaja kustutada, lisada või ümber tõsta) ning kognitiivne (kui palju mõttetööd peab järeloimetaja tegema) (O'Brien, 2006; Rossetti & Gaspari, 2017).

Esimese järeloimetamise uuringu viis läbi Krings 1997. aastal, kes kasutas valjustimõtleme meetodit. Tema katsete tulemused näitasid muuhulgas, et järeloimetamine oli tavapärasest tõlkimisest 7–20% kiirem, tõlkimisel kasutati rohkem abimaterjale, järeloimetamisel oli vaja rohkem keskenduda lähtetekstile kui tõlkimisel ning koormus oli suurem keskmise kvaliteediga masintõlke järeloimetamisel kui halva kvaliteediga masintõlke puhul. Kokkuvõttes leidis Krings, et arvestades muutusi tähelepanu fookuses, nõuab järeloimetamine vähem kognitiivset pingutust, kuid üldiselt järeloimetamine kui protsess nõuab pigem rohkem kui vähem kognitiivset pingutust. (Nitzke, 2019)

Guerberof Arenas (2013) küsis 24 professionaalse tõlkija käest, mida nemad masintõlke järeloimetamisest arvavad. 40% vastanutest leidis, et järeloimetamine nõuab rohkem pingutust kui tõlkimine, 30% vastas, et pingutus on sama, ning 20% arvas, et pingutus on väiksem. Guerberof Arenas (2014) on uurinud ka tõlkija produktiivsust osaliste ja masintõlgitud vastete olemasolul. 24 professionaalset tõlkijat tõlkisid teksti nii, et nad ei teadnud, kas vaste on pärit tõlkemälust või masintõlgitud. Katse tulemusel sai järeldada, et masintõlke vaste ja osalise vaste, mille kattuvus on vahemikus 85–94%, toimetamiseks kulub tõlkijal keskmiselt sama palju aega ning see aeg on lühem, kui ilma vasteta segmendi tõlkimiseks kuluv aeg.

Carl, Dragsted, Elming, Hardt ja Jakobsen (2011) võrdlesid tõlkimist ja järeloimetamist inglise-taani suunal, kasutades klahvivajutuste salvestamist, pilgujälgimist ning kvaliteedi hindamist inimeste poolt. Järeloimetatud teksti peeti veidi paremaks tõlgitud tekstist. Huvitav oli see, et järeloimetamine võttis vaid veidi vähem aega kui tõlkimine. Tõlkimisele kulus keskmiselt 7 min 52 s, järeloimetamisele 7 min 35 s. Seega masintõlge säästis küll aega, aga mitte väga palju. Pilgujälgija andmed näitasid, et järeloimetamisel oli fiksatsioonide arv ja kogu fiksatsioonide pikkus sihttekstil tunduvalt suurem kui tõlkimise korral. Tõlkimisel olid aga fiksatsioonid lähtetekstil tunduvalt pikemad kui järeloimetamisel. Sellest võib oletada, et tõlkimisel on vajalik parem arusaamine lähtetekstist ning järeloimetamisel vaadatakse lähteteksti pigem kiirelt masintõlke kontrollimiseks.

Green, Heer ja Manning (2013) võrdlesid järeloimetamist ja tõlkimist inglise keelest araabia, prantsuse ja saksa keelde ning Plitt ja Masselot (2010) tegid sama inglise keelest prantsuse, itaalia, saksa ja hispaania keelde tõlkimisel või masintõlke järeloimetamisel.

Mõlema uuringu tulemused näitasid, et järeltoimetamine on tõlkimisest kiirem ning sihttekstis on vähem vigu ehk selle kvaliteet on parem.

Eestis on masintõlke kasutamisest kirjutanud oma magistritöö Katre Sepp (2017). Ta küsis 17 professionaalselt tõlkijalt nende arvamust masintõlke kohta, võrdles ajakulu tõlkimisel masintõlke abiga ja ilma ning võrdles ka saadud tõlgete kvaliteeti. Tõlkijate suhtumine masintõlkesse oli pigem negatiivne. Vaid 7 vastajal oli masintõlke funktsioon sisse lülitatud, masintõlke abi kasutati pigem harva ning nad tundsid, et sellest oli abi vaid vähestel kordadel, ning selle suurimateks puudusteks peeti seda, et masintõlke parandamine võtab kauem aega kui ise tõlkimine ning eesti keel on selle süsteemi jaoks liiga keeruline. Katse tulemused aga näitasid, et 7 juhul 10-st tõlkimise kiirus masintõlke aktiveerimisel kasvas ning keskmiselt tõlgiti 25% kiiremini. Tõlgete kvaliteedis suurt erinevust ei olnud.

Kokkuvõtvalt saab öelda, et varasemad uurimused näitavad pigem, et järeltoimetamine on n-õ nullist tõlkimisest kiirem.

7. Tõlkeprotsessi uurimise katse

Kasutasin pilgujälgijat Eyelink 1000 Plus 20 tõlkija pilgu jälgimiseks, kui nad tõlkisid Memsources'i tõlkeabiprogrammis 520-sõnalist lõiku kasutusjuhendist, kasutades tõlkemälu, milles oli umbes neljandikule tekstist täpsed vasted ning neljandikule osalised vasted. Lisaks sellele oli umbes neljandik tekstist masintõlgitud ning umbes neljandikule segmentidest vaste puudus. Peale katset kasutasin tarkvara Data Viewer (SR Research Ltd., 2017) vaatlusalade lisamiseks. Vaatlusalad on need piirkonnad, millest igaühe kohta eraldi soovib katse läbiviija teada saada, mis toimus, kui katseisiku pilk oli selles piirkonnas. Siinses katses olid vaatlusalad tõlkeabitarkvara segmentid. Iga vaatlusala kohta sain Data Vieweri abil andmed, mida edasi analüüsisin. Kognitiivse koormuse uurimise aluseks võtsin O'Brieni (2007) töös kasutatud töötlemiskiiruse ning pupilli muutuse. Lisaks neile vaatasin igas segmentis ka fiksatsioonide arvu lähteteksti sõna kohta ning silmitsemisaega lähteteksti sõna kohta.

Järgmistest peatükkides kirjeldan täpsemalt katses osalejaid, kasutatud teksti, katse käiku, andmeanalüüsi ja kognitiivseid näitajaid.

7.1 Osalejad

Katses osales 20 tõlkijat, kellest 13 olid algajad tõlkijad ning 7 professionaalid. Algajad tõlkijad olid Tartu Ülikooli tõlkeõpetuse magistriõppe tudengid, kellel kokkupuude tõlkimisega enne õpingute alustamist oli väike või üldse puudus. Neist kuus oli tõlkinud alla kahe aasta, kuus 2–5 aastat ning üks 5–10 aastat (see katseisik puutus nende aastate jooksul tõlkimisega kokku vähesel määral). Neli katseisikut veetsid iga nädal tõlkides alla 10 tunni, seitse 10–20 tundi ja kaks 20–40 tundi. Kõik katses osalenud algajad tõlkijad olid tuttavad tõlkeabiprogrammidega. Kolme jaoks ei olnud tuttav Memsources'i programm, kuid nad olid kasutanud teisi sarnaseid programme. Algajate tõlkijate hulgas oli üks, kes ei olnud varem kasutusjuhendeid tõlkinud.

Professionaalseid tõlkijaid osales katses seitse, kellest kaks olid tõlkinud 2–5 aastat, kaks 5–10 aastat ning kolm üle 10 aasta. Üks katseisik kulutas tõlkimisele iga nädal alla 10 tunni, viis isikut 20–40 tundi ning üks üle 40 tunni. Professionaalsetest tõlkijatest kolm olid tuttavad Memsources'i tõlkeabiprogrammiga, ülejäänud olid tuttavad mõne teise sarnase

programmiga. Kõik professionaalsed katses osalejad olid varasemalt kasutusjuhendeid tõlkinud, kuid neist ühegi jaoks ei olnud see peamine tõlgitav tekstiliik.

Iga katsealune sai endale koodi kujul A/Px, kus A tähistab algajat, P professionaali ning x numbrit. Katsealuse isikut ja tema koodi ei ole võimalik kuidagi kokku viia.

Katses osales küll 20 isikut, kuid arvestada sai 17 katseisiku andmeid. Katseisikute A2 ja P5 pupilliga kaotas pilgijälgija katse ajal kontakti. Mõlemal juhul oli võimalik läbi viia masina kalibreerimine ja valideerimine. Katse ajal ei leidnud pilgijälgija katseisiku A2 pupilli peale paari minuti möödumist enam üldse. Katseisiku P5 pupilli ei suutnud masin enam tuvastada siis, kui katseisik vaatas ekraani vasakut alumist neljandikku. Arvestusest tuli välja jätta ka katseisiku A3 andmed, kuna ta ei võtnud arvesse ei täpseid ega osalisi vasteid ega ka masintõlget, vaid kustutas sihtteksti kasti sisestatu ning tõlkis kõik ise, mistõttu erinesid tema pilgijälgimise andmed suurel määral kõigi teiste katseisikute omadest. Ülejäänud 17 katseisikut kasutasid tõlkemälu ja masintõlke abi ning nende andmetes oli näha sarnast tendentsi.

7.2 Tekst

Katses osalejad pidid tõlkima lühikese osa viil-lintlihvija kasutusjuhendi algusest (tekst tervikuna lisas 1). Teksti ja sellega kaasas oleva tõlkemälu sain tõlkebüroolt.

Algselt oli teksti pikkus 596 sõna, kuid lühendasin seda 520 sõnani. Teksti lühendasin selleks, et katse aega lühendada ja sellega ka andmetötluse hulka vähendada ning teksti katse jaoks sobivaks muuta. Eesmärk oli saada selline tekst, mis sisaldaks tõlkemälust täpsete vastetega ja osaliste vastetega segmente, masintõlgitud segmente ning ilma vasteta segmente.

Teksti ettevalmistamine toimus Memsourc'e'i tõlkeabiprogrammi veebiversioonis. Tegin igale katsealusele eraldi projekti, mille nimi oli nende kood, ning iga projekti sidusin samanimelise tõlkemäluga, millesse importisin tõlkebüroolt saadud tõlkemälu. Masintõlke vastete saamiseks kasutasin Memsourc'e'ga ühendatud masintõlget Microsoft Translate.

Tõlgitav tekst jagunes programmis 44 segmendiks. Kahe esimese segmendi (tõlkemälus osalised vasted) tõlkimine jäi andmete analüüsist välja, et vähendada olukorraga harjumise mõju andmetele. Ülejäänud 42 segmendist oli tõlkemälus täpne (100%) vaste 11-le ja osaline (65–99%) vaste 10 segmendile. Seejärel teostasid eeltõlkimise ehk programm

sisestas sihtteksti väljale kõik tõlkemälu vasted, mille vastavus oli üle 60%-i. Sisestasin masintõlke kümne segmendi sihtteksti väljale. Segmentide valikul lähtusin sellest, et masintõlgitud oleks nii lühikesi kui pikki segmente. Andmeanalüüsis kasutasin üheksat valitud segmentidest, sest üks segment kordus, mistõttu tõlkeprogramm kinnitas automaatselt selle esimese tõlke mõlemas segmendis ning enamus tõlkijaid jättis selle teisel korral vahele. Seejärel lülitasin masintõlke välja ning tõlkijate jaoks jäi tühjaks 11 segmendi sihtteksti kast. Tekst koos vastetega on lisas 2 ning pilt programmist koos sisestatud vastetega joonisel 2.

#	Source: en	Target: et	
1	FILE BELT SANDER INSTRUCTION MANUAL 1/8-3/4 in. x 18-24 in.	VIIL-LINTLIHVUJA KASUTUSJUHE 1/8-3/4" x 18-24"	X 99
2	Important Safety Information	Spetsiifiline ohutusteave	X 67
3	Please read, understand and follow all safety information contained in these instructions prior to the use of this tool.	Lugege enne selle tööriista kasutamist läbi kogu selles juhendis sisalduv ohutusteave, tehke see endale selgeks ja järgige seda. Hoidke see juhend edaspidiseks kasutamiseks alles.	X 100
4	Retain these instructions for future reference.	Neid juhiseid edasiseks kasutamiseks alles.	X MT
5	Intended Use		X
6	This pneumatic tool is intended for use in industrial locations, and used only by skilled, trained professionals in accordance with the instructions in this manual.	Viimane Pneumoseadmete on mõeldud kasutamiseks tööstusliku kohtades ja mida ainult kvalifitseeritud, koolitatud spetsialistid juhendis antud juhiste kohaselt.	X MT
7	This pneumatic tool is designed to be used with a contact arm and appropriate abrasive belts for sanding metals, wood, stone, plastics and other materials.		X
8	It should only be used for such sanding applications and within marked capacity and ratings.	Seda võib kasutada ainult selliste lihvimine rakendusi ja märgitud võimsuse ja hinnangud.	X MT
9	Only accessories specifically recommended by 3M should be used with this tool.		X
10	Use in any other manner or with other accessories could lead to unsafe operating conditions.	Muul viisil või muude lisaseadmete kasutamine võib põhjustada ohtlikke tingimusi.	X MT
11	Do not operate tool in water or in an excessively wet application.	Ärge kasutage tööriista vees või liiga niisked rakenduses.	X MT
12	Explanation of Signal Word Consequences	Ühesõnaliste teadete seletus	X 76

Translation memory: P8
Specific/Important Safety Information
 Context: No content
 Metadata
 Created: 8.2.2019 17:14
 Modified: 8.2.2019 17:14

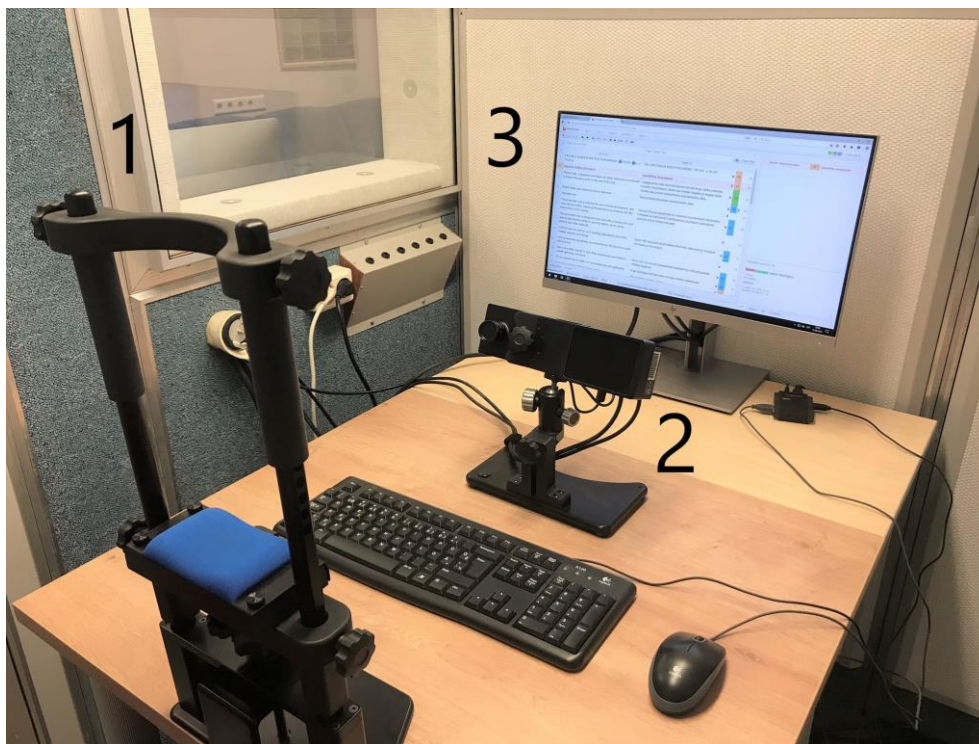
Segments: 44/0 Words: 533/0 Chars: 2867/0 28/25 Created/Modified: kretesaak/kretesaak File: TU projekt.docx

Joonis 2: Katseisikule avanenud vaade. Sulgesin elvaate, et katseisikud näeksid võimalikult palju teksti korraga, mis vähendas kerimise vajadust ning muutis lihtsamaks hiljem vaatlusalade lisamise.

7.3 Katse käik

Katses kasutasin SR Researchi pilgujälgijat EyeLink® 1000 Plus koos lauale paigutatud kaamera (Desktop Mount, joonisel 3 nr 2) ja peatoega (joonisel 3 nr 1). Masin jälgis ühte silma ning salvestussagedus oli 500 Hz. Katse aparatuur asus Tartu Ülikooli foneetika laboris ning selle paigutus on näha joonisel 3. Plaanitud katse tegemiseks ei sobinud tavaliselt kasutatud SR Researchi tarkvara Experiment Builder. Selle asemel kasutasin nende beetatarkvara Screen Recorder, mis salvestab ekraanil toimuvat ning pilgu liikumist ning seob need kaks kokku. Lisaks oli võimalik salvestada ka klahvivajutusi ja hiire liikumist ning määrata klahve, mille vajutamise hetked programm salvestab valitud nimega. Katsetel oli

sisse lülitatud klahvivajutuste salvestamine ning klahvi „Enter“ vajutused pidi programm salvestama nimega „Vaetus“.



Joonis 3: Katse aparatuur. Nr 1 on peatugi, nr 2 pilgujälgija kaamera ja nr 3 kuvar, millel tõlkimine toimus

Enne katsed lugesid katseisikud läbi informeeritud nõusolekulehe ja andsid allkirja, et nõustuvad katses osalemisega ja saadavate andmete kasutamisega edasistes uurimistöodes (nõusolekuleht on toodud lisa 3). Seejärel rääkisin neile täpsemalt katse käigust ja vajadusel tõlkeabitarkvarast. Soovitasin neil tõlkimisel segmentide vahel liikuda kasutades klahvi „Enter“ (tõlkeabiprogrammis liigub järgmise segmendi juurde) või klahvikombinatsiooni „Ctrl + Enter“ (kinnitab segmendi ja liigub järgmise juurde). Toonitasin ka, et katse ajal ei tohi lahkuda tõlkeabitarkvara lehelt ehk tarkvaraväliseid infoallikaid ei tohi kasutada. Infoallikate kasutamine ei olnud katse ajal lubatud, kuna see oleks toonud juurde palju muutujaid ning teinud andmeanalüüsi tunduvalt keerulisemaks. Muus osas soovitasin katseisikutel tunda end mugavalt ning tõlkida nii, nagu nad on harjunud.

Seejärel istus katseisik pilgujälgija ette ning panin paika peatoe ja tooli asendi. Valisin pilgu jälgimiseks silma, mida kaamera paremini nägi, panin kaamera fookusesse ning tegin pilgu kalibreerimise ja valideerimise. Nende käigus pidi katseisik jälgima ekraanile ilmuvaid täppe. Vajadusel vahetasin jälgitavat silma ning kalibreerisin ja valideerisin uuesti.

Kui katseks oli kõik valmis, siis panin tarkvara Screen Recorder salvestama ning katseisiku ees avanes Memsources'i lehekülge tõlgitava tekstiga (peatükis 7.2 joonis 2). Katseisik sai seejärel alustada tõlkimisega tavapärasel viisil.

Kui katseisik oli teksti ära tõlkinud ning soovi korral üle vaadanud, siis lõpetasin salvestamise ning palusin katseisikul täita taustaküsimustikku Google Forms keskkonnas, kus küsisin tema vanuse ja varasema tõlkekogemuse kohta ning kus ta sai anda tagasisidet katsele (küsimused on toodud lisa 4).

Pärast katset oli katseisiku tõlketööst MP4-formaadis video, mis sisaldas ekraanisalvestust koos pilgu liikumisega, EDF-fail, mis sisaldas sama videot ja mida sain edasi analüüsida SR Researchi tarkvaraga Data Viewer, ja tõlgitud tekst, mida ma selles töös ei käsitle.

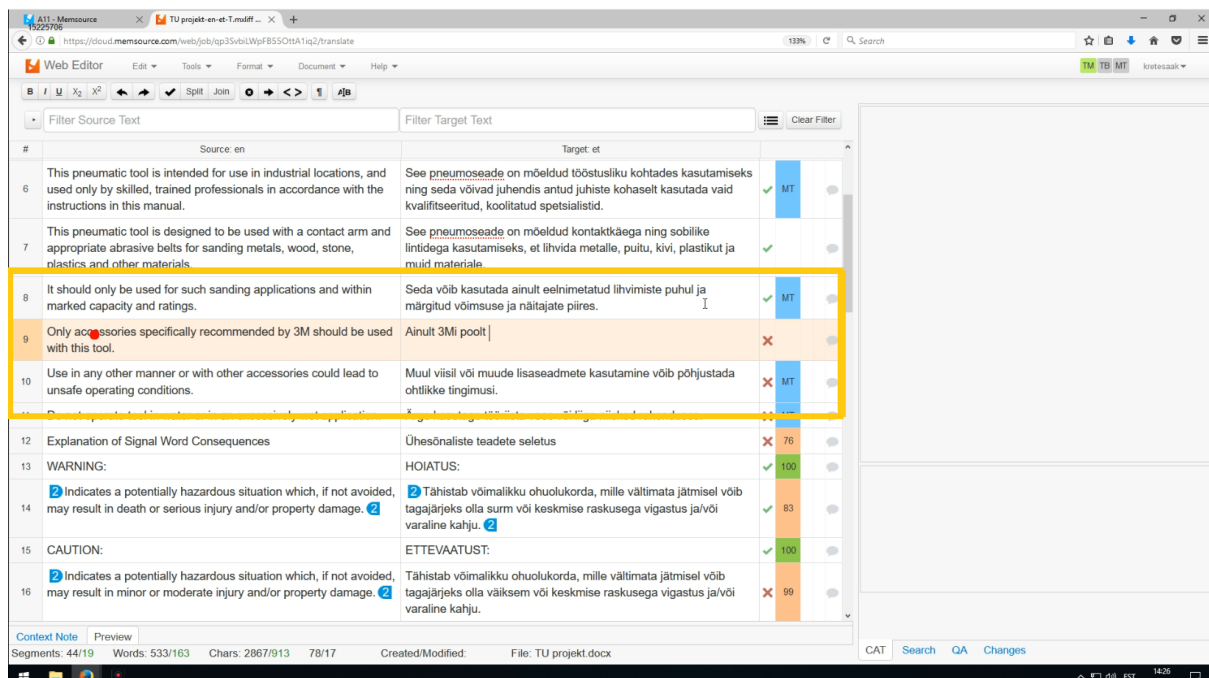
7.4 Andmeanalüüs

Katsete analüüsimiseks kasutasin tarkvara Data Viewer (SR Research Ltd., 2017) ning saadud andmeid analüüsisin Excelis.

Tavaliselt pannakse vaatlusalad ehk huvi pakkuvad alad paika katset üles ehitades. Tarkvara Screen Recorder kasutades tuleb aga vaatlusalad lisada hiljem. Selle katse jaoks oli see kõige sobilikum, kuid samas ka väga töömahukas lahendus. Lisasin käsitsi iga katseisiku iga segmendi jaoks dünaamilise(d) vaatlusala(d), mis sisaldas endas lähte- ja sihtteksti kasti ning kestis selle aja, mis katsealune seda tõlkides veetis.

Vaatlusalad olid ristküliku kujulised ning mõnevõrra suuremad lähte- ja sihtteksti kastidest, kuna videol oli näha, et pilku märkiv täpp tekkis mõne katseisiku puhul teksti kohale ja mõne puhul tekst alla (näide vaatlusalast on toodud joonisel 4). Vaatlusalad olid dünaamilised ehk sel hetkel, kui katseisik liikus järgmise segmendi juurde, vaatlusala kadus ning tulemusesse läks kirja selles piirkonnas toimunu just selles ajavahemikus. Vaatlusala ajalise alguse ja lõpu märkimiseks kasutasin klahvivajutuste salvestust. Kahjuks märguklahvi „Enter“ vajutamisel programm alati ei salvestanud liigutust tekstiga „Vaetus“. Õnneks olid vajalikud klahvivajutused siiski olemas, kuna ka klahvivajutuste salvestamine oli sisse lülitatud. Vaatasin video pealt umbkaudse aja, mil tõlkija uude segmenti suundus, seejärel

otsisin vajaliku klahvivajutuse üles ning sain selle täpse toimumishetke. Ühe segmendi vaatlusalala lõpuajaks sain selle hetke, mil katseisik sellest järgmise segmendi juurde liikus.



Joonis 4: Katseisiku A11 segmendi 9 vaatlusalala on märgitud oranži joonega. Punane täpp on pilgu asukoht.

Vaatlusalade lisamise tegid keeruliseks kerimine ning hiireklikkiga segmentide vahel liikumine. Kuna vaatlusalala on video peal ühes kindlas paigas, siis peale kerimist pidin looma uue vaatlusalala isegi siis, kui tõlkija jätkas tegelemist sama segmendiga. Kümne katse kohta olid olemas ka hiire liikumise, klõpsamise ja kerimise andmed ning nende toimumise ajad. Kuna hiirega seostuvate andmete hulk oli väga suur ning katseisikutel oli palutud liikuda segmentide vahel klahvivajutustega, siis algselt oli hiire jälgimine välja lülitatud. Kuid esimene andmeanalüüs näitas, et neid andmeid on siiski vaja ning järgnevatel katsetel lülitasin funktsiooni sisse. Nendel kümnel juhul sain ka kerimise ja klõpsamise hetked täpselt teada. Muudel juhtudel pidin videot vaadates jälgima täpsemalt, millal liigutus toimus või millal pilk liikus uue segmendi juurde. See oli andmeanalüüsi kõige aeganõudvam osa.

Paljudel juhtudel oli ühe segmendi jaoks vaja mitut vaatlusalala, kuna tõlkijad liikusid tagasi juba varem tõlgitud segmentide juurde. Programm võimaldas sellisel juhul kaks või rohkem vaatlusalala omavahel siduda, mis tegi edasise analüüsi lihtsamaks. Kuna vaatlusalade lisamine oli väga mahukas töö, siis jätsin välja need vaatlusalast läbi liikumised, kui midagi ei muudetud, välja arvatud esmakordsed liikumised 100% vaste juurde, millele järgnes

segmendi kinnitamine. Seega jäid vaatluse alt välja algne teksti vaatamine/lugemine ja hilisem üle lugemine. Selliste lugemiste lisamine oleks paljudel juhtudel olnud väga keeruline, sest lugedes oli võimalik mõtte liikumist ühelt segmendilt teisele täheldada vaid pilgu liikumisest, mistõttu samasuguste ajaliste segmentide lisamine oleks olnud väga keeruline. Kuna aga tuttava teksti kiirel üle lugemisel on mõttetöö tõenäoliselt tunduvalt väiksem, kui teksti tõlkimisel, siis ei oleks see palju tulemustesse lisanud.

Peale vaatlusalade lisamist lasin programmil Data Viewer koostada iga katseisiku kohta aruande, mis sisaldas järgmiseid andmeid:

- kogu katse kohta
 - kestus (ms)
 - keskmine pupilli suurus (suhteline ühik, ingl. k *arbitrary unit*)
- iga uuritava segmendi kohta
 - vaatlusala algusaeg (ms)
 - vaatlusala lõppaeg (ms)
 - fiksatsioonide arv
 - silmitsemisaeg (ms)
 - maksimaalne pupilli suurus (suhteline ühik)
 - minimaalne pupilli suurus (suhteline ühik)

Saadud andmeid analüüsisin edasi Excelis, et saada näitajad, mis võimaldavad kognitiivset koormust võrrelda.

7.5 Kasutatud kognitiivse koormuse näitajad

Esimene tulemustes käsitletav suurus on segmendi töötlemiskiirus. Töötlemiskiirus näitab, mitu sõna sekundis katsealune tõlgib. Mida suurem on saadud arv, seda kiiremini töötles katseisik selles segmendis sõnu ja seda suurem oli tema produktiivsus. Seda näitajat kasutas kognitiivse koormuse näitajana oma töös O'Brien (2007), Sepp (2017) kasutas samasugust näitajat tõlkimiskiiruste võrdlemiseks ning Plitt ja Masselot (2010) produktiivsuse näitajana. Ühes oma hilisemas artiklis seob O'Brien (2011) produktiivsuse ja kognitiivse koormuse. Ta leiab, et mida suurem on koormus, seda madalam on produktiivsus. Keerulisema teksti tõlkimine võtab kauem aega ja nõuab tõlkijalt suuremat koormust.

Töötlemiskiiruse arvutamiseks kasutasin kõigepealt vaatlusala algus- ja lõppaega, et saada segmendi tõlkimiseks kulunud aega. Keerulisem oli nende segmentidega, mille juurde tõlkija läks tagasi, kuna seal oli ühe segmendi kohta mitu vaatlusala ning Data Vieweri aruandes oli vaid esimese algus ja viimase lõpp. Programmis endas aga nägin kõigi vaatlusalade aegu ning sain kestuse arvutada neid kasutades. Töötlemiskiiruse arvutamiseks jagasin lähtesegmendis oleva sõnade arvu ajaga, mis kulus katseisikul segmendi tõlkimisele.

Kognitiivse koormuse uurimiseks vaatasin ka pupilli suuruse muutust. Peatükis 5.2 kirjutasin pikemalt uurimustest, kus on näidatud, et pupilli suuruse muutus ja kognitiivne koormus on omavahelises sõltuvuses (nt Beatty, 1982; Hess & Polt, 1964). Pupilli suurus ning selle muutumine on väga individuaalne (eriti arvestades seda, et selles katses pooled katseisikud kandsid prille või läätsesid, mis silma moonutavad) ning tihtipeale mõõdetakse seda suhtelistes ühikutes. Seetõttu kasutatakse tavaliselt pupilliandmete võrdlemiseks protsendilist muutust. O'Brien (2007) arvutas samateemalises töös protsendilist pupilli suuruse muutust selliselt, et leidis kõigepealt lugemisel keskmise pupillimuutuse baasjoone, seejärel keskmise pupillimuutuse iga segmendi tõlkimisel ning esitas protsendiliselt selle, kui palju rohkem pupilli suurus muutus tõlkimisel võrreldes baasjoonega.

Mina kasutasin pupilli suuruse muutuse protsenti (edaspidi: PSMP) arvutamiseks teistsugust lähenemist. Lahutasin kõigepealt segmendi maksimaalsest pupilli suurusest minimaalse. Nii sain iga katseisiku kohta iga segmendi tõlkimisel toimunud muutuse pupilli suuruses. Seda muutust võrdlesin iga katseisiku kogu katse keskmise pupilli suurusega, et võtta arvesse individuaalset pupillide erinevust. Ehk kui pupilli pindala muutus segmendis oli 500 ühikut ning kogu katse keskmine pupilli pindala oli 1500 ühikut, siis pupilli muutuse protsent oli $\frac{500}{1500} \times 100\% = 33,3\%$. See näitab, et selle segmendi tõlkimisel pupilli suurus muutus 1/3 pupilli keskmisest suurusest. Mida suurem see muutus on, seda suurem on tõlkimise kognitiivne koormus.

Kognitiivse koormuse näitajana käsitlen veel segmendi fiksatsioonide arvu ja silmitsemisaega ehk fiksatsioonide summaarset pikkust. Mõlemaid suuruseid on palju kasutatud kognitiivse koormuse ja teksti keerukuse uurimisel (vt peatükke 5.2 ja 6.). Fiksatsioonide arvu ja kognitiivset pinget on seostatud lugemisel (Just & Carpenter, 1980; Schnitzer & Kowler, 2006), tõlkimisel (Dragsted, 2012; Sharmin et al., 2008) ning seda on kasutatud järeltoimetamise uurimisel (Carl et al., 2011). Fiksatsioonide pikkust on samuti kasutatud sarnaselt (Carl et al., 2011; Dragsted, 2012).

Fiksatsioonide arvu ja summaarset pikkust mõjutab tugevalt teksti pikkus. Just ja Carpenter (1980) näitasid oma töös, et juba tavalisel lugemisel toimub pea igal sõnal fikseering. Kuna siinses töös lähteteksti sõnade arv segmendis varieerus 1-st kuni 25-ni, siis vaatlen igas segmendis neid suuruseid jagatuna lähteteksti sõnade arvuga. Lähteteksti segmendid sisaldasid ka tõlkeabitarkvarale omaseid vormistuse märgendeid, mida lugesin eraldi sõnaks, kuna ka nende lisamisega pidi tõlkija tegelema. Segmendi pikkuse arvestamiseks jagasin vaatlusala fikatsioonide arvu ja silmitsemisaja sõnade arvuga selle vaatlusala segmendi lähtetekstis.

Lisas 5 on näitena katseisiku P1 andmed ja nende põhjal arvutatud tulemused.

Järgmiseks leidsin iga katseisiku jaoks keskmised andmed iga vastetüübi kohta ning seejärel nende keskmiste pealt kokku üldised tulemused.

8. Tulemused

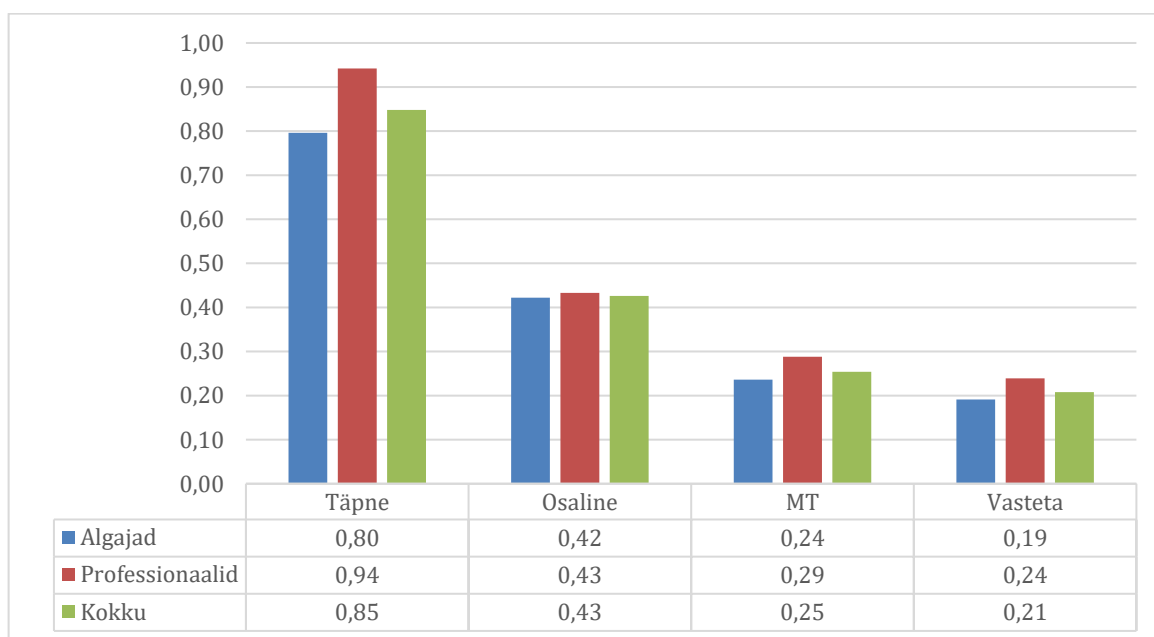
Tabelis 1 on kokkuvõtvalt kõik siin töös vaadeldavad kognitiivse koormuse näitajad iga vastetüübi jaoks. Järgnevates peatükkides vaatan neid kõiki lähemalt.

	Keskmine töötlemiskiirus (sõna/s)	Keskmine pupilli suuruse muutus %	Fiksatsioonide arv sõna kohta	Silmitsemisaeg sõna kohta (s)
Täpne	0,85	25,1	6,94	1,71
Osaline	0,43	36,2	9,71	2,69
MT	0,25	46,1	12,82	3,86
Vasteta	0,21	53,5	13,28	4,11
Pikkus	45,5 min			

Tabel 1: Kognitiivse koormuse näitajad iga vastetüübi jaoks

8.1 Töötlemiskiirus

Segmendi töötlemiskiirus näitab, mitu sõna sekundis katseisik selles segmendis tõlkis. Mida suurem arv, seda rohkem jõudis katseisik ühe sekundi jooksul läbi töödelda ehk seda kiiremini ta tõlkis ja seda vähem koormust tõlkimine nõudis. Joonisel 5 on esitatud keskmine töötlemiskiirus iga vastetüübi jaoks ehk seda tüüpi vastega segmentide töötlemiskiiruste keskmine. Eraldi tõin välja algajate ja professionaalide ning kõigi katseisikute keskmise töötlemiskiiruse.



Joonis 5: Katseisikute keskmine töötlemiskiirus (sõna/s) vastetüüpide kaupa

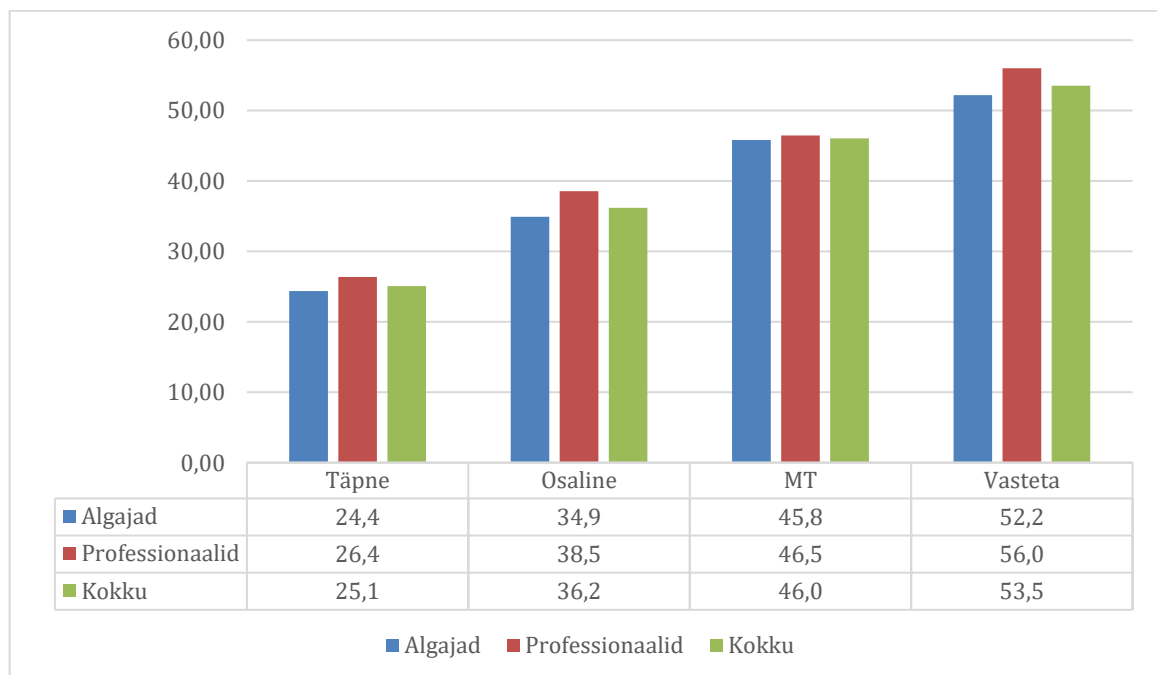
Kõigi rühmade jaoks oli kõige suurem täpse vaste töötlemiskiirus (algajad 0,80 sõna/s, professionaalid 0,94 sõna/s ja kokku 0,85 sõna/s). Kõigi katseisikute peale kokku oli sellest umbes kaks korda väiksem osalise vaste töötlemiskiirus, mis oli 0,43 sõna/s. Masintõlke järeltoimetamise keskmine töötlemiskiirus oli kõigi katseisikute peale 0,25 sõna/s ehk 42% väiksem osalise vaste töötlemiskiirusest. Kõige madalam töötlemiskiirus oli ilma vasteta segmentidel – 0,21 sõna/s ehk 16% väiksem järeltoimetamise töötlemiskiirusest.

Algajate ja professionaalide töötlemiskiirused vastetüüpide kaupa järjestusid samamoodi. Mõlemal juhul vähenes töötlemiskiirus järgmiselt: täpse vastega tõlkimine, osalise vastega tõlkimine, järeltoimetamine ja vasteta tõlkimine. Professionaalide töötlemiskiirus oli iga vastetüübi korral algajate omast väiksem – täpse vastega 18%, osalise vastega 2%, järeltoimetamisel 21% ja nullist tõlkimisel 26% väiksem.

17 osalejast 12 individuaalne töötlemiskiiruste järjestus vastetüüpide kaupa vastas üldisele järjestusele. Katseisikutel A5, A7 ja A13 olid järeltoimetamise ja nullist tõlkimise töötlemiskiirused väga sarnased ehk järeltoimetamine ja nullist tõlkimine nõudsid nende jaoks sama suurt pingutust. Katseisikul P3 olid väga sarnased osalise vastega segmentide ja masintõlgitud segmentide töötlemiskiirused (vastavalt 0,28 ja 0,26 sõna/s). Katseisiku P7 töötlemiskiirus nullist tõlkides oli suurem nii osalise vastega segmentide tõlkimise kiirusest kui ka järeltoimetamise töötlemiskiirusest (vastavalt 0,27; 0,25 ja 0,22 sõna/s) ehk tema jaoks nõudis vähem pingutust ise tõlkimine kui masintõlke või osalise vaste toimetamine.

8.2 Pupilli suuruse muutuse protsent

Pupilli suuruse muutus on seotud kognitiivse koormusega (vt peatükk 5.2). Mida keerulisem on tekst, seda rohkem pupill muutub. Vaatan segmentis toimunud pupilli muutust (maksimaalne suurus miinus minimaalne suurus) katseisiku keskmise pupillisuuruse suhtes. Joonisel 6 on toodud vastetüüpide kaupa keskmine PSMP seda tüüpi segmentides. Eraldi toin välja professionaalide ja algajate ning kõigi katseisikute keskmised PSMP-d.



Joonis 6: Katseisikute keskmine PSMP (%) vastetüüpide kaupa

Pupilli suurus muutus kõige rohkem ilma vasteta segmentide tõlkimisel. Kõigi katseisikute peale kokku oli keskmine PSMP nullist tõlkimisel 53,5%. Masintõlke järeltoimetamisel oli keskmine PSMP 46% ehk 7,5% väiksem nullist tõlkimisel toimunud keskmisest PSMP-st. Osalise vastega segmenti tõlkimisel oli PSMP veel 9,8% võrra väiksem ehk 36,2%. Nullist tõlkimise PSMP-st oli rohkem kui poole väiksem täpse vastega segmenti tõlkimise PSMP, mis oli 25,1%.

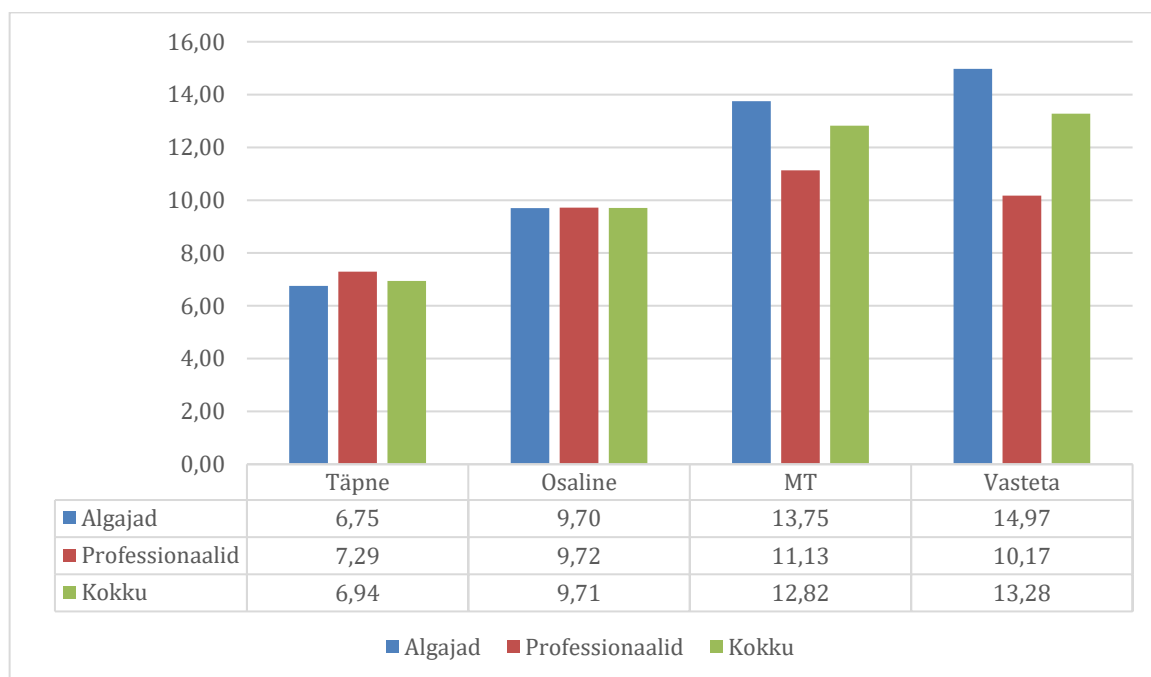
Mõlemas rühmas ja kõigi katseisikute peale kokku eraldusid teineteisest selgelt vastetüüpide PSMP-d. Iga vastetüübi PSMP oli professionaalide puhul veidi kõrgem algajate omast.

17 katseisikust 13 vastetüüpide keskmised PSMP-d järjestusid samamoodi nagu kõigi katsealuste keskmised PSMP-d. Kolme katseisiku (A10, A11 ja P7) PSMP-d vasteta segmenti tõlkimisel ja järeltoimetamisel olid lähedased ehk nende katseisikute jaoks nõudis järeltoimetamine sama palju pingutust kui nullist tõlkimine. Katseisiku A7 PSMP nullist tõlkimisel oli 59% ja järeltoimetamisel 64%, millest võib järeldada, et järeltoimetamine nõudis selle katseisiku jaoks suuremat pingutust kui nullist tõlkimine.

Vaadates individuaalseid pupilli suuruse muutuseid, siis katseisikute kõige suurem muutus pupilli suuruses toimus enamasti ilma vasteta segmenti tõlkimisel. Neljal juhul oli suurim muutus masintõlke järeltoimetamisel. Seitsme katseisiku pupill muutus kõige rohkem segmenti nr 7 tõlkimisel (lähtetekst: „This pneumatic tool is designed to be used with a contact arm and appropriate abrasive belts for sanding metals, wood, stone, plastics and other materials.“), millele puudus vaste. Katselindistustel oli näha, et katseisikutele valmistas probleemi väljendi „contact arm“ tõlkimine. Seda väljendit tekstis rohkem ei kasutatud ning tõenäoliselt oleks katseisikud soovinud kasutada väliseid allikaid konteksti sobiva vaste leidmiseks. Väikseim pupilli suuruse muutus toimus 16 katseisikul täpse vastega segmenti tõlkimisel. Ühel katseisikul oli väikseim muutus 99% vastega segmenti tõlkimisel (segment nr 17), mis koosnes ühest sõnast ning oli peaaegu identne eespool olnud segmentiga. Seitsme katseisiku pupill muutus kõige vähem segmenti nr 42 tõlkimisel, mis koosnes numbrist 1 ning sellele oli olemas ka täpne vaste, milleks oli 1.

8.3 Fiksatsioonide arv lähteteksti sõna kohta

Keerulistes kohtades, mille töötlemine (ehk millest aru saamine) nõuab rohkem aega, fikseerub pilk pikemalt või rohkem kordi (Just & Carpenter, 1980). Joonisel 7 on näidatud keskmine fikatsioonide arv lähteteksti sõna kohta vastetüüpide kaupa. Eraldi tõin välja professionaalide ja algajate ning kõigi katseisikute keskmised tulemused.



Joonis 7: Keskmine fikatsioonide arv lähteteksti sõna kohta vastetüüpide kaupa

Vaadates kõiki katseisikuid, on keskmiselt enim fikseerimisi lähteteksti sõna kohta ilma vasteta segmendi tõlkimisel – 13,28 fikseerimist. Masintõlke järeltoimetamisel oli keskmine fikseerimiste arv sõna kohta veidi väiksem – 12,82. Osaliste vastetega segmentide tõlkimisel oli näitaja 24% väiksem kui järeltoimetamisel ehk 9,71 fikseerimist. Kõige vähem fikseerimisi lähteteksti sõna kohta oli täpse vastega segmentide tõlkimisel – 6,94 fikseerimist.

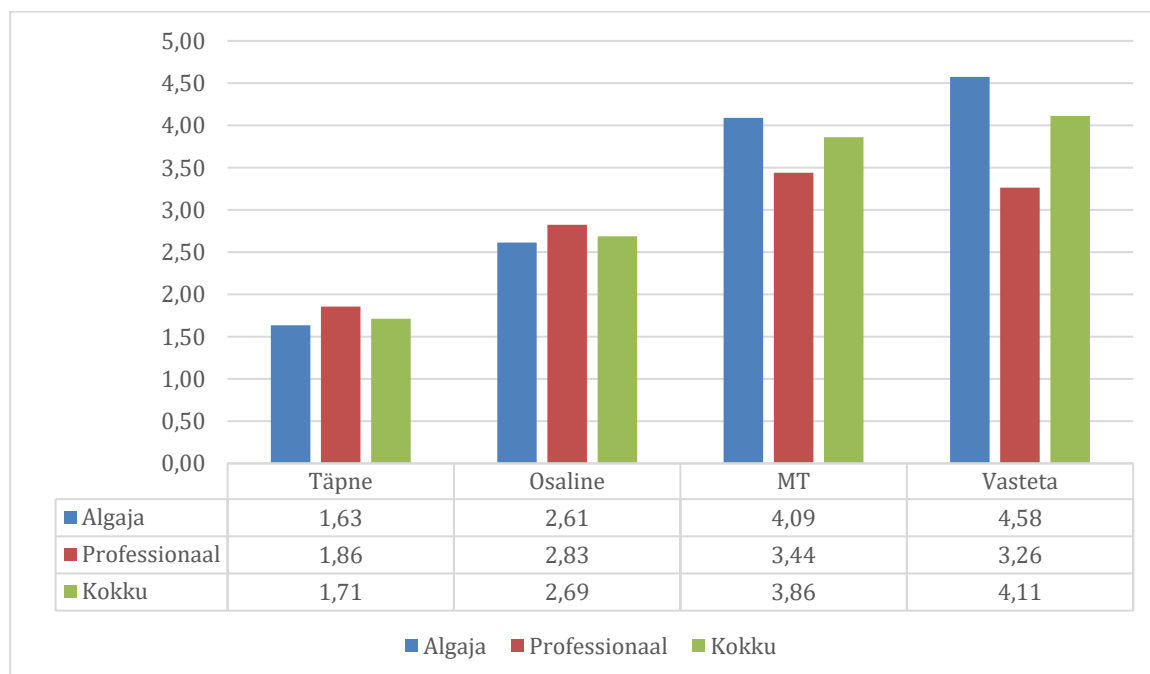
Algajate puhul oli nullist tõlkimisel rohkem fikseerimisi sõna kohta kui järeltoimetamisel. Järeltoimetamisel oli neid 8% vähem kui nullist tõlkimisel. Professionaalide puhul oli aga järeltoimetamisel fikseerimiste arv sõna kohta 9% suurem kui nullist tõlkimisel ehk professionaalide pilk peatus järeltoimetamisel keskmiselt rohkem kordi.

Individuaalselt järjestusid kuue katseisiku fikseerimiste arvud sõna kohta samasuguselt nagu kõigi katseisikute keskmised. Neljal juhul oli järjestus samasugune

professionaalide omaga ehk kõige vähem fiksatsoone oli täpse vastega segmendi tõlkimisel, järgnes osalise vastega segmendi tõlkimine, ilma vasteta segmendi tõlkimine ning enim fiksatsoone sõna kohta oli masintõlke järeltoimetamisel. Kolmel katseisikul oli nullist tõlkimisel vähem fiksatsoone kui osalise vastega segmendi tõlkimisel ja järeltoimetamisel, kahel juhul oli kõige vähem fiksatsoone osalise vastega segmendi tõlkimisel ning kahel juhul paigutus masintõlke järeltoimetamine osalise vastega ja nullist tõlkimisest ette.

8.4 Silmitsemisaeg lähteteksti sõna kohta

Silmitsemisaeg näitab, kui pikalt kokku pilk sellel segmendil peatus, kui katseisik seda tõlkis/järeltoimetas. Mida pikemalt pilk peatub, seda rohkem on vaja töödelda seda, millel pilk peatub (Just & Carpenter, 1980). Seda suurust vaatan samuti lähteteksti sõna kohta, kuna segmentide pikkused on väga erinevad (täpsem selgitus peatükis 7.5). Joonisel 8 on toodud vastetüüpide kaupa keskmine silmitsemisaeg lähteteksti sõna kohta. Eraldi tõin välja keskmise professionaalide ja algajate näitaja, lisaks kõigi katseisikute keskmisele.



Joonis 8: Keskmine silmitsemisaeg lähteteksti sõna kohta vastetüüpide kaupa (s)

Silmitsemisaeg lähteteksti sõna kohta oli kõige lühem täpse vaste tõlkimisel, mil keskmiselt peatus ühe sõna kohta katseisiku pilk 1,71 sekundiks. Osalise vastega segmenti tõlkimisel oli see aeg pea sekundi võrra pikem ehk 2,69 sekundit. Masintõlke järeltoimetamisel kasvas silmitsemisaeg sõna kohta 43% võrra ehk selleks oli 3,86 sekundit. Kõige pikem oli keskmine silmitsemisaeg lähteteksti sõna kohta ilma vasteta segmenti tõlkimisel, mil see oli 4,11 sekundit.

Samamoodi järjestusid vastetüübid, vaadates algajate keskmisi silmitsemisaegu sõna kohta. Nagu ka fiksatsioonide arvu puhul, järjestus silmitsemisaeg professionaalide puhul teisiti – järeltoimetamisel oli keskmine silmitsemisaeg sõna kohta 5% suurem võrreldes ilma vasteta tõlkimisega.

Individuaalseid tulemusi vaadates ühtis kaheksa katseisiku silmitsemisaegade järjestus vastetüüpide kaupa kõigi katseisikute keskmiste järjestusega. Nelja katseisiku tulemused järjestusid sarnaselt professionaalide omadega. Kolme katseisiku silmitsemisaeg sõna kohta oli lühem nullist tõlkides kui osalise vastega või masintõlgitud segmenti juures. Ühe katseisiku jaoks oli osalise vastega segmenti silmitsemisaeg sõna kohta kõige lühem ning ühe jaoks oli masintõlgitud segmentide keskmine väiksem osalise vastega segmentide omast.

8.5 Katse kestus

Lisaks kognitiivsetele näitajatele vaatan ma ka üldist katse kestust, mis hõlmab nii tekstiga tutvumiseks ja selle tõlkimiseks kui ka toimetamiseks ja üle lugemiseks kulunud aega. Tabelis 2 on kõigi katseisikute katse kestused minutites ning algajate ja professionaalide keskmine katseaeg.

	A1	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	Keskm.
Aeg (min)	33,8	48,4	48,8	51,4	48,2	47,0	50,2	56,8	53,7	25	55,4	47,2
	P1	P2	P3	P4	P6	P7						Keskm.
Aeg (min)	43,6	45,2	42,0	28,1	43,3	52,7						42,5

Tabel 2: Katse kestus katseisikute kaupa ning algajate ja professionaalide keskmised ajad

Professionaalide keskmine katse kestus oli 42,5 minutit ja algajate oma 47,2 minutit ehk professionaalid tõlkisid ja toimetasid teksti keskmiselt 4,7 minutit kiiremini. Professionaalide katse kestuse mediaan oli 43,4 min ja ulatus 28,1–52,7 min. Algajate katse kestuse mediaan oli 48,8 min ja ulatus 25–56,8 min. Kõige kiirem tõlkija oli A12, kellel kulus katse jaoks 25 minutit ning kõige aeglasem A10, kellel kulus 56,8 minutit. Seega oli nii kõige kiirem kui ka kõige aeglasem katseisik algaja.

Kahe katsegrupi keskmiseid kestuseid võrreldes ei tundu 4,7 min väga suure vahena, kuid vaadates kõiki tulemusi eraldi, saame kahe grupi erinevusest paremini aru. Algajate hulgas oli 5 ehk 45%, kelle katseaeg ületas 50 min. Professionaalide hulgas oli neid üks ehk 17%. Alla 40 min kulus katseks algajate hulgas kahel ehk 18% ja professionaalide hulgas ühel ehk 17%. Seega saab öelda, et professionaalidel kulus katsele vähem aega kui algajatel.

9. Tulemuste analüüs

9.1 Kognitiivne koormus

Selles töös vaatasin nelja kognitiivse koormuse näitajat: töötlemiskiirust, pupilli suuruse muutuse protsenti, fiksatsioonide arvu lähteteksti sõna kohta ja silmitsemisaega lähteteksti sõna kohta. Kõigi näitajate keskmised tulemused osutavad sellele, et kognitiivne koormus kasvab järgmiselt:

täpse vastega segmenti tõlkimine < osalise vastega segmenti tõlkimine < masintõlke järeltoimetamine < vasteta segmenti tõlkimine

Selles järjekorras vähenes töötlemiskiirus ehk keskmiselt jõudis katsealune tõlkida ühe sekundi jooksul vähem, PSMP kasvas ehk katseisik pingutas tõlkimisel aina rohkem, fiksatsioonide arv lähteteksti sõna kohta kasvas ehk katseisiku pilk fikseerus segmendis rohkem kordi ning silmitsemisaeg lähteteksti sõna kohta samuti kasvas ehk katseisikute pilk peatus segmendis pikemalt.

Samasuguse kognitiivse koormuse järjestuse sai oma uurimuses ka O'Brien (2007). Kuid tema uuringu tulemused näitasid, et inglise-prantsuse ja inglise-saksa suunal on kognitiivne koormus masintõlke toimetamisel sarnane osalise vastega segmenti tõlkimiseks

vajaliku pingutusega. Selle töö tulemused aga viitavad, et inglise-eesti suunal on masintõlke toimetamise kognitiivne koormus kõige lähedasem ilma vasteta segmenti tõlkimiseks vajaliku koormusega. Kuigi kaks tööd on tehtud 12-aastase vahega ning selle jooksul on masintõlge väga palju arenenud, siis töö tulemused viitavad siiski sellele, et väikesesse ja keerulisse eesti keelde masintõlkimine ei ole jõudnud suurte keelte vahelise masintõlkega samale tasemele.

Tulemused ühtivad ka Sepa (2017) magistr töö tulemustega. Ta leidis, et masintõlge kiirendab tõlkeprotsessi. Selles töös saadud töötlemiskiirused näitavad sama, kuigi järeldoimetamise ja vasteta tõlkimise töötlemiskiiruste vahe ei ole väga suur.

Teine huvitav tulemus on täpsete vastete tõlkimisel toimunud kognitiivne koormus. Täpsete vastete tõlkimisel kulus katseisikul umbes sekund iga sõna tõlkimiseks, pupill suurenes u 25%, sõna kohta toimus fiksatsioon umbes 7 ning silmitsemisaeg sõna kohta oli 1,71 sekundit. Need tulemused näitavad, et katseisikud ei liikunud automaatselt täpsest vastest edasi. Seda kinnitavad ka katse ekraanisalvestused. Pigem võtsid katseisikud aja lähte- ja sihtteksti lugemiseks ning mõned katseisikud isegi muutsid täpset vastet oma äranägemise järgi teksti sobivamaks. Siin võib tulemusi mõjutada ka see, et katseisikul ei olnud ajapiirangut. Tavaolukorras juhtub tihti, et tõlkimiseks ei ole piisavalt aega, ning siis võivad tulemused olla teistsugused.

Kokkuvõttes näitavad uurimuse tulemused, et tõlkemälu ja masintõlge on tõlkijale palju abiks, kuid isegi täpse vaste korral ei lähe tõlkija sellest kohe edasi, vaid ka neile kulub aega ja vaeva.

9.2 Erinevused professionaalide ja algajate vahel

Professionaalide ja algajate tulemustes olid mõned erinevused. Neist üks oli erinev ajakulu. Professionaalide katse kestus oli keskmiselt 4,7 min lühem. Sarnase tulemuseni jõudsid ka Jakobsen ja Jensen (2008), kes leidsid, et professionaalid loevad ja ka tõlgivad kiiremini.

Professionaalide töötlemiskiirus oli iga vastetüübi jaoks veidi kiirem algajate omast. See on seotud sellega, et nende katse kestus tervenisti oli lühem. Professionaalid jõudsid ühes sekundis tõlkida keskmiselt 0,48 sõna ja algajad 0,41 sõna ehk professionaalid tõlkisid 16%

rohkem ühes sekundis. Nende katse kestuste ulatus oli väiksem ehk kõigil professionaalidel kulus tekstiga tegelemiseks sarnasem aeg (kui kaks katseisikut arvestusest välja jätta, siis oli ulatus vaid 3,2 min). Kognitiivse koormuse näitajate juures oli tulemuste erinevus vastetüüpide kaupa väiksem kui algajate puhul, mis võib viidata sellele, et professionaalid ei toetu nii palju tõlkemälule ja masintõlkele kui algajad.

Pupilli suuruse muutus oli aga professionaalidel iga vastetüübi jaoks pisut suurem kui algajatel. See nagu viitaks suuremale koormusele. Siin võib aga üheks põhjuseks olla see, et 50% professionaalidest kandis prille ja 1/3 läätsesid ehk ainult üks katseisik ei kandnud kumbagi. Algajate seas aga 73% ei kandnud ei prille ega läätsesid. Kuna läätsed ja prilliläätsed on optilised vahendid, siis need moonutavad pupilli. Samas on see moonutus ühtlane kogu katse jooksul ning seega ei tohiks muuta suhtarvulist tulemust. Vahe võib aga olla ka juhuslik või tingitud väikesest katsealuste hulgast.

Väga huvitav oli professionaalide fiksatsioonide arv ja silmitsemisaeg lähteteksti sõna kohta. Kuigi pupilli suuruse muutus ja töötlemisaeg viitasid sellele, et nullist tõlkimine on järeltoimetamisest keerulisem, siis fiksatsioonide arv ja silmitsemisaeg sõna kohta näitasid vastupidist. Üheks võimalikuks seletuseks on see, et masintõlke järeltoimetamisel vaatab katseisik tõenäoliselt rohkem ekraani kui nullist tõlkimisel. Trükkimisel vaatavad paljud klaviatuuri ning enamik katses osalejaid seda ka tegi. Järeltoimetamisel kasutati aga palju sihtteksti kastis lauseosade ümber paigutamist ja vaid sõnalõppude muutmist, mistõttu ekraanilt vaadati eemale vähem. Kuid miks siis selline nähtus tekkis vaid professionaalidel? Ei saa ka öelda, et algajate jaoks oleks järeltoimetamine lihtsam olnud, kuna nende fiksatsioonide arv ja silmitsemisaeg sõna kohta olid kõrgemad professionaalide omadest. Võib olla oli algajate hulgas rohkem neid, kes trükkimisel üldse klaviatuuri ei vaata. Võimalik põhjus võib olla ka see, et järeltoimetamisel on vaja tutvuda kahe tundmatu tekstiga, kuid sama kehtib ka osalise vaste ja täpse vaste puhul. Üks võimalus on ka see, et vahe põhjustas kogemus või see on tingitud tekstist. See on üks küsimus, mis vajaks põhjalikumalt uurimist.

9.3 Muud tähelepanekud

Selles peatükis toon välja mõned huvitavad tähelepanekud tõlkijate tööst, mida ma aga sihilikult ei uurinud.

Kuigi selles töös ma ei käsitle tõlkekvaliteeti, siis ekraanisalvestusi vaadates märkasin huvitavat viga. Segmendil 12, mille lähtetekst oli „Explanation of Signal Word Consequences“, oli tõlkemälus osaline vaste „Ühesõnaliste teadete seletus“. Osalise vaste lähtetekst, mille tõlkimisel see tõlkemälusse salvestus, oli „Explanation of Single Word Consequences“. Seega erinesid varem tõlgitud lause ja katses tõlgitav lause ühe sõna võrra. Kõigist katseisikutest üheksa ehk 53% muutsid lähteteksti katis olnud osalist vastet. Ülejäänud kaheksa katseisikut, kelle hulgas oli viis algajat (ehk pea pooled) ja kolm professionaali (ehk pooled professionaalidest), jätsid sõnaühendi „signal word“ vasteks „ühesõnaline“. Kuna need lähtetekstide sõnad, „signal“ ja „single“, on niivõrd sarnased, siis siin võis juhtuda selline asi, et katseisiku peas, nähes vastet „ühesõnaline“, tekkis kohe seos sõnaga „single“ ning ta ei näinudki, et tegelikult on tõlgitavas lauses tähed teises järjekorras. Internetiavarustes ringlevad tekstid, kus sõnade esimesed ja viimased tähed on õiged, kuid ülejäänud on vales järjekorras. Inimene suudab siiski selle teksti kokku klapitada. Selline viga on tingitud tõlkemälust, kuna seda seost ei saaks tekkida nullist tõlkimisel. Seega on siin välja tulnud üks tõlkemälu kasutamise probleem, mida aga tõenäoliselt tuleb ette harva.

Tõlkimisel valmistasid paljudele katseisikutele probleeme tõlkeprogrammi märgendid, mis hoiavad vormistust. Paljude jaoks, kes seda konkreetset programmi väga tihti ei kasuta, ei olnud märgendite saamine sihtteksti kerge. Ekraanisalvestustelt oli näha, et nende märgendite kopeerimine, mis olid segmendi alguses, oli väga keeruline, sest valikkuse tekkis millegi pärast kogu tekst. Memsources'is on märgendite lisamise nupp umbes selline: <>. See seostub küll märgendi kujutamisega tekstifailis (nagu lisas 2 näha), kuid programmis endas on raske seda seost näha. Seetõttu arvan, et edasistes uuringutes oleks hea märgendid välja jätta. Siinses katses olid märgendid juba esimeses segmendis, mida tulemuste arvutamisel ma ei arvestanud, mistõttu arvasin, et katseisik saab kohe alguses endale selgeks teha, kuidas neid lisada. Tegelikult aga tegelesid paljud katseisikud märgendite lisamisega hiljem, alustasid seda tegevust mujalt või siis sai see läbivalt neile takistuseks. Kuid kuna pupilli suuruse muutuste järgi ei sisaldanud katseisikute jaoks kõige keerulisemad segmendid märgendeid, siis leian, et see ei mõjuta terviklikult selle töö tulemusi väga palju.

9.4 Tagasiside küsitluse vastused

Peale katses osalemist vastasid katseisikud mõnele küsimusele ning said anda tagasisidet (küsimused on lisas 4). Küsisin neilt, kuidas nad tundsid end füüsiliselt ja vaimselt ning kas nad saavutasid tavapärase tõlkimisrütmi.

Kõigist kahekümnest katseisikust kaheksa kirjutas, et neil oli füüsiliselt ebamugav. Üks mainis, et peatugi oli ebamugav, kaks tõid välja, et pead oleks tahtnud liigutada, ja ühel väsis istumisest selg ära. Kolm katseisikut mainisid, et nad tundsid end vaimselt ebakindlat. Neli tõi välja, et ebamugavust tekitas väliste allikate abi puudumine. Üks katseisik leidis, et nii vaimselt kui füüsiliselt polnud midagi viga, kuus tundis end hästi ja üks väga hästi.

Tõlkerütmi saavutamise küsimusele vastas 11 katsealust, et neid häiris see, et nad ei saanud abivahendeid kasutada. Kaks leidsid, et nad tõlkisid tänu sellele kiiremini. Üks arvas, et ta tõlkis seetõttu aeglasemalt. Kaks katseisikut saavutasid oma tavapärase tõlkerütmi, kolm saavutas selle enam vähem. Lisaks toodi segavate faktoritena välja ka järgmist: masintõlge oli võõras, tõlkeabiprogramm oli võõras, kogu katseolukord mõjutas ja katse tegija soovitusel (nt hoia aktiivsena see segment, millega tegeled) segasid.

Seega soovitan edaspidi ehitada katset üles nii, et saaks ka abivahendeid kasutada, ning kaaluda katse teostamist ilma peatoeta.

Kokkuvõte

Aina rohkem viiakse maailmas läbi tõlkeprotsessi ja masintõlke järeltoimetamise uuringuid. Need kaks valdkonda on tugevalt seotud tõlkijate ja seega ka kogu tõlketööstuse igapäevatööga. Teades, mis tõlkijale rohkem koormust valmistab, saab panustada just nende valdkondade arendamisesse ja maksta tõlkijatele ka väärlist tasu.

Varasemates uuringutes on leitud, et tõlkemälu lihtsustab ja kiirendab tõlkija tööd (Alves & Campos, 2009; O'Brien, 2007) ning ka masintõlke olemasolu muudab tõlkimist kiiremaks (Green et al., 2013; Guerberof Arenas, 2014; Plitt & Masselot, 2010; Sepp, 2017).

Selle magistritöö käigus salvestasin kahekümne tõlkija tõlkeprotsessi, millest 18 andmeid saab kasutada edasistes uuringutes ning 17 andmeid kasutasin selles uuringus. Kogusin nende tööst ekraani-, pilgu- ja klahvivajutuste salvestuse ning ka tõlgitud tekstid.

Vaatlesin lähemalt kognitiivset koormust tõlkemälu ja masintõlke abiga tõlkimisel. Töös kasutasin nelja erinevat kognitiivse koormuse näitajat ning kõigi tulemused kinnitasid üksteist. Katsete tulemused näitasid, et inglise keelest eesti keelde tõlkimisel kasvab kognitiivne koormust järgnevalt:

$$\text{täpse vastega segmendi tõlkimine} < \text{osalise vastega segmendi tõlkimine} < \text{masintõlke järeltoimetamine} < \text{vasteta segmendi tõlkimine}$$

Sellest järjestusest saab järeldada, et tõlkemälu ja masintõlge teevad tõesti tõlkimise lihtsamaks. Täpse vastega segmendi tõlkimise kõik kognitiivse koormuse näitajad viitasid vähemalt poole väiksemale koormusele võrreldes ilma vasteta segmendi tõlkimisega. Osalise vastega segmendi kognitiivne koormus oli mõnevõrra suurem täpse vastega segmendi omast, kuid siiski palju väiksem vasteta segmendi tõlkimise kognitiivsest koormusest. Seega on põhjendatud tõlkemälu vastete järgi tõlke hinna arvutamine.

Lähemalt vaatasin järeltoimetamise kognitiivset koormust. Kui O'Brien (2007) leidis, et järeltoimetamise kognitiivne koormus on sarnane osalise vastega segmendi tõlkimise koormusele, siis inglise-eesti suunal selle töö andmetel see nii ei ole. Katse andmed näitavad, et järeltoimetamise kognitiivne koormus on kõige lähedasem nullist tõlkimisele. Siit võib järeldada, et katses kasutatud inglise-eesti masintõlge ei ole veel jõudnud sellele tasemele, mis oli 2007. aastal suurte keelte vahel, mistõttu ei saa pakkuda masintõlke toimetamise eest inglise-eesti suunal väga palju madalamat hinda kui nullist tõlkimise eest.

Nii videolt kui arvandmetest on näha, et tõlkijad ei liigu täpsete vastetega segmentidest automaatselt edasi. Mitmed tõlkijad tegid täpsetes vastetes muudatusi ning kõik lugesid neid. Seega isegi täpse vaste olemasolu nõuab tõlkijalt tööd, mistõttu peaks nende segmentide tõlkimine olema kindlasti tasustatud.

Professionaalide ja algajate peamine erinevus oli see, et professionaalid tõlkisid algajatest kiiremini ning ka nende töötlemiskiirus oli suurem. Lisaks oli neil fiksatsioonide arv suurim ja silmitsemisaeg pikim just järeltoimetamisel, kuid töötlemiskiirus ja pupilli suurenemine olid sarnaselt algajatele suurimad ilma vasteta segmendi tõlkimisel. Ka siin saaks edasi uurida, mis täpselt põhjustas rohkem fiksatsioone ja pikema silmitsemisaja. Kasutades neid samu andmeid, on võimalik uurida täpselt, kuhu professionaalid ja algajad vaatasid, ning võib olla on seal peidus vastus.

Selle töö üks suur panus on Eesti tõlkijate tõlkeprotsessi andmekogum, mida saab ka edasi uurida, kuna vaatlusalad saab lisada pärast katset. Näiteks saab uurida pilgu jaotumist lähte- ja sihtteksti vahel, kindlate sõnade või väljendite tõlkimist, eraldi masintõlke järeltoimetamist või just tõlkemälu vastetega segmentide tõlkimist. Kuna olemas on ka tõlgitud tekstid, siis saab ka nende kvaliteeti hinnata, kas siis masintõlke ja nullist tõlkimise vaatenurgast või näiteks algajate ja professionaalide võrdluses. Lisaks saab videolt uurida, milliseid tõlkeabiprogrammi funktsioone või milliseid tõlketehnikaid tõlkijad kasutavad.

Katseisikute tagasisidet ja minu tähelepanekuid arvestades annan mõned soovitusel edasisteks katseteks. Kuna tõlkeabiprogramm ei olnud kõigile tuttav, siis soovitan edaspidi kasutada teksti, kus puuduvad märgendid, kuna nendega tegelemine võib tulemusi mõjutada. Kõige suurem probleem tõlkijate endi jaoks oli väliste allikate abi puudumine, mistõttu soovitan edaspidi võimalusel seda võimalust neile pakkuda. Kuna ka asend oli paljude jaoks probleem, siis võiks võimalusel peatuge mitte kasutada.

Magistritöö teema valikul soovisin panustada tõlkija n-ö musta kasti sisu uurimisse ehk tõlkeprotsessi uuringutesse. Kuna Eesti kontekstis ja eesti keelde tõlkimisel on seda tehtud vähe, siis usun, et see töö oli vajalik, ning arvan, et seda on paljudel huvitav lugeda.

Lõpetuseks tänan oma juhendajaid, kes mind aitasid ja minusse uskusid. Suure tänu on ära teeninud ka Reelika Saar, kes aitas mind tehniliselt ja tänu kellele oli mul nii palju katseisikuid. Suured tänud ka Tartu Ülikooli foneetika laborist Anton Malmile ja Pärtel Lippusele, kes vastasid mu paljudele küsimustele ning mind nõu ja jõuga igati aitasid.

Resümee

In this master's theses I investigated the cognitive effort of translation process with the help of translation memory and machine translation. The main objective was to compare the cognitive load that different match types placed on translators. In addition, I took a closer look at the cognitive effort of post-editing and exact matches and compared the cognitive effort of professional translators and beginners.

To investigate translation process, I conducted an experiment using eye tracking technology, key-logging and screen capturing. The translation process of 20 translators, 7 professionals and 13 beginners was recorded and the data of 17 recordings was used in this study. I investigated cognitive effort using processing speed, pupil dilation, fixation count and dwell time as indicators.

The results suggest that cognitive load increases in the following order:

$$\text{exact match segment} < \text{fuzzy match segment} < \text{machine translation} < \text{no match segment}$$

All indicators demonstrated that translating with the help of an exact match needed no more than half the effort that was needed with no match segments. A fuzzy match put more cognitive load on translators than an exact match, but still considerably less than no match. Thus, the results demonstrate that translation memory indeed decreases cognitive load and therefore it can be taken into consideration when calculating the price of a translation.

O'Brien (2007) suggested that post-editing and fuzzy matches put similar cognitive load on translators. The results of this master's theses do not suggest that. This research demonstrates that the cognitive load of post-editing is closest to translation from scratch. This suggests the price for post-editing machine translation from English to Estonian should be similar to the price for translation.

The cognitive effort for exact match segments was also investigated. The results and the screen captures indicate that translators do not confirm exact match segments automatically. They read them through and even make changes where they consider them necessary. Thus, exact matches put cognitive load on translators and therefore they should receive pay for checking and modifying them.

The main difference between professionals and beginners was the time it took to complete the translation. Professionals translated faster and their processing speeds for each match type were higher meaning they processed more words in one second. The fixation count and dwell time for professionals were highest when post-editing, but pupil dilation and processing speed indicate that cognitive effort was highest when translating from scratch. This is something that needs further investigation.

This work also contributed to the translation process research in Estonia by collecting data that can be used for further research. New interest areas can be created to investigate other research questions like fixations on source and target text, translation of certain words, post-editing or fuzzy matches. Using key-logging and screen capture data translation techniques and quality can be studied. All of this and a lot more can be researched using the same files and the number of participants can be increased with more experiments.

Feedback from participants suggests that the use of external sources is important in the translation process and so if possible, it should be provided. Many participants found the head support uncomfortable and so if possible, I suggest not to use it.

With this thesis I was hoping to peek inside the black box of translators and learn more about their translation process. The results of this study are interesting and hopefully a good reading and inspiration for many others.

Kasutatud kirjandus

- Alves, F., & Campos, T. L. (2009). Translation technology in time: Investigating the impact of translation memory systems and time pressure on types of internal and external support. *Copenhagen Studies in Language*, 37, 191–218.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91(2), 276–292. <https://doi.org/10.1037/0033-2909.91.2.276>
- Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. *Handbook of psychophysiology*, 2nd ed. (lk 142–162). New York, NY, US: Cambridge University Press.
- Carl, M., Dragsted, B., Elming, J., Hardt, D., & Jakobsen, A. L. (2011). The process of post-editing: a pilot study. *Copenhagen Studies in Language*, 41, 131–142.
- Conklin, K., Pellicer-Sanchez, A., & Carroll, G. (2018). *Eye-tracking: a guide for applied linguistics research*. Cambridge ; New York, NY: Cambridge University Press.
- Dragsted, B. (2012). Indicators of difficulty in translation — Correlating product and process data. *Across Languages and Cultures*, 13(1), 81–98. <https://doi.org/10.1556/Acr.13.2012.1.5>
- Duchowski, A. T. (2017). *Eye tracking methodology: theory and practice* (Third edition). Cham: Springer.
- Englund Dimitrova, B. (2010). Translation process. Y. Gambier & L. van Doorslaer (Toim), *Handbook of Translation Studies* (Kd 1, lk 406–411). <https://doi.org/10.1075/hts.1.tra6>
- Goutte, C., Cancedda, N., Dymetman, M., & Foster, G. (2009). *Learning Machine Translation*. Salvestatud <https://books.google.ee/books?id=vaAPdPTtJ2IC>
- Green, S., Heer, J., & Manning, C. D. (2013). The efficacy of human post-editing for language translation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, 439. <https://doi.org/10.1145/2470654.2470718>
- Guerberof Arenas, A. (2013). What do professional translators think about post-editing? *The Journal of Specialized Translation*, (19).
- Guerberof Arenas, A. (2014). Correlations between productivity and quality when post-editing in a professional context. *Machine Translation*, 28(3–4), 165–186. <https://doi.org/10.1007/s10590-014-9155-y>

- Hess, E. H., & Polt, J. M. (1964). Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science*, *143*(3611), 1190–1192. <https://doi.org/10.1126/science.143.3611.1190>
- Holmqvist, K., Nyström, M., & Andersson, R. (2011). *Eye Tracking: a comprehensive guide to methods and measures*. Oxford: OUP Oxford.
- Hutchins, W. J. (1986). *Machine translation: past, present, future*. Chichester: Horwood [u.a.].
- Iqbal, S. T., Adamczyk, P. D., Zheng, X. S., & Bailey, B. P. (2005). Towards an index of opportunity: understanding changes in mental workload during task execution. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '05*, 311. <https://doi.org/10.1145/1054972.1055016>
- Jakobsen, A. L., & Jensen, K. T. H. (2008). Eye Movement Behaviour Across Four Different Types of Reading Task. *Copenhagen Studies in Language*, *36*, 103–124.
- Jakobsen, A. L., Mees, I. M., & Göpferich, S. (2012). *Behind the mind: methods, models and results in translation process research*. Frederiksberg: Samfundslitteratur Press.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Koehn, P. (2010). *Statistical machine translation*. Cambridge ; New York: Cambridge University Press.
- Krings, H. P. (1986). Translation problems and translation strategies of advanced German learners of French (L2). *Interlingual and intercultural communication*, 263–276.
- Moresi, S. (2008). Pupil dilation in response preparation. *International Journal of Psychophysiology*, *67*(2), 124–130. <https://doi.org/10.1016/j.ijpsycho.2007.10.011>
- Nitzke, J. (2019). *Problem solving activities in post-editing and translation from scratch: a multi-method study*. Salvestatud <https://www.doabooks.org/doab?func=fulltext&uiLanguage=en&rid=33067>
- O'Brien, S. (2006). Pauses as Indicators of Cognitive Effort in Post-editing Machine Translation Output. *Across Languages and Cultures*, *7*(1), 1–21. <https://doi.org/10.1556/Acr.7.2006.1.1>
- O'Brien, S. (2007). EYE-TRACKING AND TRANSLATION MEMORY MATCHES. *Perspectives: Studies in Translatology*, *14*(3), 185–205.

- O'Brien, S. (2008). Processing fuzzy matches in translation memory tools: An eye-tracking analysis. *Copenhagen Studies in Language*, 36, 79–102.
- O'Brien, S. (2009). Eye tracking in translation process research: methodological challenges and solutions. *Methodology, technology and innovation in translation process research*, 38, 251–266.
- O'Brien, S. (2011). Towards predicting post-editing productivity. *Machine Translation*, 25(3), 197–215. <https://doi.org/10.1007/s10590-011-9096-7>
- Plitt, M., & Masselot, F. (2010). A Productivity Test of Statistical Machine Translation Post-Editing in a Typical Localisation Context. *The Prague Bulletin of Mathematical Linguistics*, 93(1). <https://doi.org/10.2478/v10108-010-0010-x>
- Poibeau, T. (2017). *Machine translation*. Cambridge, Massachusetts: The MIT Press.
- Rossetti, A., & Gaspari, F. (2017). Modelling The Analysis Of Translation Memory Use And Post-Editing Of Raw Machine Translation Output: A Pilot Study Of Trainee Translators' Perceptions Of Difficulty And Time Effectiveness. *Zenodo*. <https://doi.org/10.5281/zenodo.1090952>
- Schnitzer, B. S., & Kowler, E. (2006). Eye movements during multiple readings of the same text. *Vision Research*, 46(10), 1611–1632. <https://doi.org/10.1016/j.visres.2005.09.023>
- Sepp, K. (2017). *Masintõlke järetoimetamine* (Magistritöö). Tallinna Ülikool, Humanitaarteaduste instituut, Tallinn.
- Sharmin, S., Spakov, O., Rähä, K.-J., & Lykke Jakobsen, A. (2008). Where on the screen do translation students look while translating, and for how long? *Copenhagen Studies in Language*, 36, 31–51.
- SR Research Ltd. (2017). EyeLink Data Viewer 3.1.97 [Computer software]. Mississauga, Ontario, Canada: SR Research Ltd.
- Toury, G. (1995). *Descriptive Translation Studies and Beyond*. Salvestatud <https://books.google.ee/books?id=aNHENtMoqEIC>
- Trujillo, A. (1999). *Translation Engines: Techniques for Machine Translation*. (R. J. Paul, P. J. Thomas, & J. Kuljis, Toim). Salvestatud <http://public.ebib.com/choice/publicfullrecord.aspx?p=3073816>

Lisad

Lisa 1. Katses tõlgitud tekst

FILE BELT SANDER INSTRUCTION MANUAL $\frac{1}{8}$ - $\frac{3}{4}$ in. x 18-24 in.

Important Safety Information

Please read, understand and follow all safety information contained in these instructions prior to the use of this tool. Retain these instructions for future reference.

Intended Use

This pneumatic tool is intended for use in industrial locations, and used only by skilled, trained professionals in accordance with the instructions in this manual. This pneumatic tool is designed to be used with a contact arm and appropriate abrasive belts for sanding metals, wood, stone, plastics and other materials. It should only be used for such sanding applications and within marked capacity and ratings. Only accessories specifically recommended by 3M should be used with this tool. Use in any other manner or with other accessories could lead to unsafe operating conditions.

Do not operate tool in water or in an excessively wet application.

Explanation of Signal Word Consequences

- WARNING:** Indicates a potentially hazardous situation which, if not avoided, may result in death or serious injury and/or property damage.
- CAUTION:** Indicates a potentially hazardous situation which, if not avoided, may result in minor or moderate injury and/or property damage.

WARNING

To reduce the risks associated with impact from abrasive product, or tool breakup, sharp edges, hazardous pressure, rupture, vibration and noise:

- Read, understand and follow the safety information contained in these instructions prior to the use of this tool. Retain these instructions for future reference.
- Only personnel who are properly trained should be allowed to service this tool.
- Practice safety requirements. Work alert, have proper attire, and do not operate tools under the influence of alcohol or drugs.
- Operators and other personnel must always wear protection for eyes, ears, and respiratory protection when in the work area or while operating this product. Follow your employer's safety policy for PPEs and/or ANSI Z87.1 or local/national standards for eyewear and other personal protective equipment requirements.
- Wear protective apparel, taking into consideration the type of work being done.
- Proper eye protection must be worn at all times.
- Tool shall not be operated in the presence of bystanders.
- If you notice any abnormal noise or vibration when operating the tool, immediately discontinue its use and inspect for worn or damaged components. Correct or replace the suspect component. If abnormal noise or vibration still exists, return the tool to 3M for repair or replacement. Refer to warranty instructions.
- Never operate this tool without all safety features in place and in proper working order.
- Never over-ride or disable the safety features of the start-stop control such that it is in the on position.
- Prior to use inspect abrasive product and accessories for possible damage. If damaged, replace with new abrasive product and accessories available from 3M.
- Never allow this tool to be used by children or other untrained people.
- Use only with mounting hardware recommended by 3M; check with 3M for mounting hardware requirements.
- Prior to use, ensure guard is oriented to protect the operator from flying fragments and is properly secured.
- If the tool is jammed, shut off the tool, disconnect the air hose and ease it free. Ensure the abrasive product is correctly secured.

Lisa 2. Lähteteksti segmendid ja nende vasted

ID	#	Source	Target	
	1	FILE BELT SANDER INSTRUCTION MANUAL {1>1/8-3/4 <1}in. x 18-24 in.	VIIL-LINTLIHVIA KASUTUSJUHEND 1/8–3/4“ x 18–24“	99
	2	Important Safety Information	Spetsiifiline ohutusteave	66
	3	Please read, understand and follow all safety information contained in these instructions prior to the use of this tool.	Lugege enne selle tööriista kasutamist läbi kogu selles juhendis sisalduv ohutusteave, tehke see endale selgeks ja järgige seda. Hoidke see juhend edaspidiseks kasutamiseks alles.	100
	4	Retain these instructions for future reference.	Neid juhiseid edasiseks kasutamiseks alles.	MT
	5	Intended Use		
	6	This pneumatic tool is intended for use in industrial locations, and used only by skilled, trained professionals in accordance with the instructions in this manual.	Viimane Pneumoseadmete on mõeldud kasutamiseks tööstusliku kohtades ja mida ainult kvalifitseeritud, koolitatud spetsialistid juhendis antud juhiste kohaselt.	MT
	7	This pneumatic tool is designed to be used with a contact arm and appropriate abrasive belts for sanding metals, wood, stone, plastics and other materials.		
	8	It should only be used for such sanding applications and within marked capacity and ratings.	Seda võib kasutada ainult selliste lihvimine rakendusi ja märgitud võimsuse ja hinnangud.	MT
	9	Only accessories specifically recommended by 3M should be used with this tool.		
	10	Use in any other manner or with other accessories could lead to unsafe operating conditions.	Muul viisil või muude lisaseadmete kasutamine võib põhjustada ohtlikke tingimusi.	MT
	11	Do not operate tool in water or in an excessively wet application.	Ärge kasutage tööriista vees või liiga niisked rakenduses.	MT
	12	Explanation of Signal Word Consequences	Ühesõnaliste teadete seletus	75
	13	WARNING:	HOIATUS:	100
	14	{2>Indicates a potentially hazardous situation which, if not avoided, may result in death or serious injury and/or property damage.<2}	Tähistab võimalikku ohuolukorda, mille vältimata jätmisel võib tagajärjeks olla väiksem või keskmise raskusega vigastus ja/või varaline kahju.	82
	15	CAUTION:	ETTEVAATUST:	100
	16	{2>Indicates a potentially hazardous situation which, if not avoided, may result in minor or moderate injury and/or property damage.<2}	Tähistab võimalikku ohuolukorda, mille vältimata jätmisel võib tagajärjeks olla väiksem või keskmise raskusega vigastus ja/või varaline kahju.	99
	17	WARNING	HOIATUS:	99
	18	To reduce the risks associated with impact from abrasive product, or tool breakup, sharp	Toimige järgmiselt, et vähendada riske, mis kaasnevad abrasiivvahendilt	100

	edges, hazardous pressure, rupture, vibration and noise:	saadud löögi või tööriista purunemisega või on tingitud teravatest servadest, ohtlikust rõhust, rebenemisest, vibratsioonist ja mürast.	
19	Read, understand and follow the safety information contained in these instructions prior to the use of this tool.	Lugege enne selle tööriista kasutamist läbi kogu selles juhendis sisalduv ohutusteave, tehke see endale selgeks ja järgige seda. Hoidke see juhend edaspidiseks kasutamiseks alles.	87
20	Retain these instructions for future reference.		MT
21	Only personnel who are properly trained should be allowed to service this tool.		
22	Practice safety requirements.	Täitke ohutusnõudeid.	100
23	Work alert, have proper attire, and do not operate tools under the influence of alcohol or drugs.	Täitke ohutusnõudeid. Olge töötamise ajal tähelepanelik, kandke nõuetekohast rõivastust ja ärge kasutage tööriista alkoholijoobes või ravimite mõju all.	76
24	•{1} <1} Operators and other personnel must always wear protection for eyes, ears, and respiratory protection when in the work area or while operating this product.		
25	Follow your employer's safety policy for PPEs and/or ANSI Z87.1 or local/national standards for eyewear and other personal protective equipment requirements.	Järgige oma tööandja poliitika PPEs ja/või ANSI Z87.1 või prillid ja muude isikukaitsevahendite nõuete/Rahvuslikud standardid.	MT
26	Wear protective apparel, taking into consideration the type of work being done.	Kandke tehtava töö iseloomule vastavaid kaitserõivaid.	100
27	Proper eye protection must be worn at all times.	Nõuetekohaseid kaitseprille tuleb kasutada kogu aeg.	100
28	Tool shall not be operated in the presence of bystanders.	Tööriista ei tohi käitada kõrvaliste isikute juuresolekul.	100
29	•{1} <1} If you notice any abnormal noise or vibration when operating the tool, immediately discontinue its use and inspect for worn or damaged components.		
30	Correct or replace the suspect component.	Parandada või asendada kahtlustatava osa.	MT
31	If abnormal noise or vibration still exists, return the tool to 3M for repair or replacement.		
32	Refer to warranty instructions.	Vaadake juhiseid garantii.	MT
33	Never operate this tool without all safety features in place and in proper working order.	Ärge mingil juhul käitage seda tööriista, kui kõik kaitsed ja ohutusvahendid pole paigas ega nõuetekohaselt töökorras.	84
34	Never over-ride or disable the safety features of the start-stop control such that it is in the on position.		
35	•{1} <1} Prior to use inspect abrasive product and accessories for possible damage.		
36	If damaged, replace with new abrasive	Kahjustuste korral asendage	100

	product and accessories available from 3M.	abrasiivvahend ja lisatarvikud uute 3M-i pakutavate toodetega.	
37	Never allow this tool to be used by children or other untrained people.	Ärge mingil juhul lubage seda tööriista kasutada lastel või väljaõppimata isikutel.	86
38	Use only with mounting hardware recommended by 3M; check with 3M for mounting hardware requirements.	Kasutage üksnes 3M-i soovitatud kinnitusdetaile; küsige 3M-ist kinnitusdetaile puudutavate nõuete kohta.	100
39	Prior to use, ensure guard is oriented to protect the operator from flying fragments and is properly secured.	Enne kasutamist, tagada on suunatud killud lendavad ettevõtja kaitsmine ja nõuetekohane kaitstus.	MT
40	If the tool is jammed, shut off the tool, disconnect the air hose and ease it free.	Kui tööriist kinni kiilub, lülitage see välja ja vabastage ettevaatlikult.	65
41	Ensure the abrasive product is correctly secured.		
42	1	1	100
43	3M {1>'8 <1}in. to {2>3<2}{3>/4 <3}in.		
44	File Belt Sander	22 000 p/min viil-lintlihvija	80

Lisa 3. Uuritava informeeritud nõusolekuleht

Uuritava informeeritud nõusolekuleht

Hea huviline,

Kutsun Teid osalema tõlkimiskatsesse, mille teema on masintõlke ja tõlkemälu kasutamine ning tõlkimisega kaasnev kognitiivne koormus. Uurimuses osalemine on vabatahtlik ning Teil on õigus sellest igal hetkel loobuda. Et saaksite anda osalemiseks informeeritud nõusoleku, selgitatakse järgnevalt täpsemalt mõõtmiste eesmärgid ning protseduure.

Tõlkimisprotsessi kohta saab infot eeskätt tõlkija silmade liikumisest ekraanil. See tähendab, et kui tõlkija tõlgib, salvestan tema tegevust pilgujälgimisseadme abil. Uuringu käigus palun Teil tõlkida etteantud teksti Memsources'i tõlkeabiprogrammis pilgujälgimisseadmega ühendatud arvuti taga. Silmade liikumise salvestamine toimub arvuti ekraani ette paigaldatud seadmega ning tõlkimine ei erine olulisel määral tavaliselt tõlkimisolukorrast. Katse ei kesta üle kahe tunni ning igal hetkel on võimalik teha puhkepaus.

Peale katset palun Teil täita lühikese taustaküsimustiku, kus küsin Teie vanuse ja varasema tõlkekogemuse kohta ning kus saate anda tagasisidet katse kohta.

Kõik uuringu käigus kogutud andmed on konfidentsiaalsed. Uuringus kasutame ainult Teile omastatud katsekoodi „A/Px“ (A – algaja, P – professionaal, x tähistab numbrit) ning Teie nime ja katsekoodi ei viida kuskil kokku. Neid analüüsitakse anonüümsel kujul ega avaldata Teile loata isikut tuvastada võimaldaval kujul mitte kellelegi. Lisainfo saamiseks pöörduge uurija poole, helistage +372 53061214 või saatke e-kiri aadressil KreteSaak@gmail.com. Uuringu igas etapis on Teil õigus ja võimalus küsida täiendavaid küsimusi katse läbiviijalt. Uuringus osalemine on vabatahtlik ja Teil on õigus uuring igal hetkel katkestada.

Kui olete valmis uuringus osalema, kinnitage järgnevalt oma nõusolekut.

Mina, _____, olen nõus

- osalema eelpool nimetatud uuringus, _____

- andmete anonüümse kasutamisega teadusotstarbel. _____

allkirjad

Uurija

allkiri

kuupäev

Krete Saak

Lisa 4. Küsitlus

Tõlkeprotsessi uuringud

Ankeet tõlkeprotsessi uuringutes osalejale.

* Required

Kood (küsi uuringu läbiviijalt) *

Vanus *

Kas kandsite prille/läätsi katse ajal?

- Kandsin prille
- Kandsin läätsi
- Ei kandnud kumbagi

Olen tõlkimisega tegelenud *

- alla kahe aasta
- 2–5 aastat
- 5–10 aastat
- üle 10 aasta

Tõlgin nädalas *

- 40 tundi või rohkem
- 20–40 tundi
- 10–20 tundi
- alla 10 tunni

Tõlkeprogrammide varasem kasutus *

- Olen tuttav Memsources'i desktopi versiooniga.
- Olen tuttav Memsources'i veebiversiooniga.
- Ei ole tuttav Memsources'i tõlkeabiprogrammiga, kuid olen tuttav mõne muu sarnase programmiga.
- Ei ole tuttav ühegi tõlkeabitarkvaraga.

Masintõlke varasem kasutus *

- Kasutan tõlkeabiprogrammi automaatselt masintõlget.
- Kasutan masintõlget muul viisil.
- Ei kasuta masintõlget üldse.

Kokkupuude kasutusjuhendite tõlkimisega *

- Tõlgin peamiselt kasutusjuhendeid.
- Tõlgin teiste hulgas ka kasutusjuhendeid.
- Olen kasutusjuhendeid tõlkinud harva.
- Ei ole kasutusjuhendit tõlkinud.

Kuidas tundsite end uuringu käigus? (füüsiliselt ja vaimselt) *

Kas saavutasite oma tavalise tõlkimisrütmi? Kui ei, siis mis segas? *

Kui soovite veel midagi öelda uuringu kohta, siis siin saab seda teha.

Lisa 5. Katseisiku P1 andmeanalüüsi tulemused segmentide kaupa

Kood	Vaatlusala (VA) nr	Segmenti nr	Tüüp	Sõnade arv	VA algus (ms)	VA lõpp (ms)	VA aeg (ms)	Töötlemiskirrus (sõna/s)	Fiksatsioonide arv	Fiksatsioonide arv sõna kohta	Silmitsemisaeg (ms)	Silmitsemisaeg sõna kohta (s)	Max pupilli suurus	Min pupilli suurus	Pupilli muutus	Pupilli muutuse %
P1	1	3_1	100	19	111040	141887	30847	0,616	110	5,79	24889	1310	1764	1329	435	30,04
P1	2	4_1	MT	6	141887	917160	27487	0,218	49	8,17	20255	3376	2162	1442	720	49,72
P1	3	5_1	no	2	162534	181469	18935	0,106	39	19,50	13134	6567	2256	1499	757	52,28
P1	4	6_1	MT	25	181469	2137304	126855	0,197	321	12,84	107968	4319	1999	1235	764	52,76
P1	5	7_1	no	25	277320	2153610	159187	0,157	338	13,52	101447	4058	1973	1180	793	54,76
P1	6	8_1	MT	15	420201	2184016	80531	0,186	233	15,53	67645	4510	2072	1181	891	61,53
P1	7	9_1	no	12	470326	2213543	86734	0,138	220	18,33	64554	5380	2216	1201	1015	70,09
P1	8	10_1	MT	15	521676	2230094	48149	0,312	129	8,60	37977	2532	1748	1202	546	37,70
P1	9	11_1	MT	12	553274	585024	31750	0,378	65	5,42	26842	2237	1963	1235	728	50,27
P1	10	12_1	75	5	585024	603208	18184	0,275	52	10,40	14344	2869	1881	1348	533	36,81
P1	11	13_1	100	1	603208	604720	1512	0,661	9	9,00	1246	1246	1452	1398	54	3,73
P1	12	14_1	82	20	604720	2279475	146360	0,137	308	15,40	93916	4696	1666	1168	498	34,39
P1	13	15_1	100	1	716578	719250	2672	0,374	10	10,00	2374	2374	1606	1366	240	16,57
P1	14	16_1	99	20	719250	2303786	52621	0,380	126	6,30	39019	1951	1523	1172	351	24,24
P1	15	17_1	99	1	747560	750904	3344	0,299	7	7,00	1366	1366	1566	1457	109	7,53
P1	16	18_1	100	21	750904	2325257	143464	0,146	440	20,95	120180	5723	1824	1181	643	44,40
P1	17	19_1	87	18	875321	2335721	38136	0,472	133	7,39	31272	1737	1868	1175	693	47,86
P1	18	20_1	Kordus	6	902993	907720	4727	1,269	11	1,83	4233	706	1656	1457	199	13,74

P1	19	21_1	no	13	917160	2349584	39798	0,327	103	7,92	32544	2503	1887	1136	751	51,86
P1	20	22_1	100	3	946358	951238	4880	0,615	14	4,67	2728	909	1400	1185	215	14,85
P1	21	23_1	76	17	951238	987100	35862	0,474	107	6,29	29264	1721	1797	1154	643	44,40
P1	22	24_1	no	25	987100	1156332	111709	0,224	251	10,04	93552	3742	1893	1304	589	40,67
P1	23	25_1	MT	21	1090122	2407261	151966	0,138	434	20,67	126225	6011	1824	1157	667	46,06
P1	24	26_1	100	12	1209617	1223856	14239	0,843	54	4,50	11787	982	1659	1349	310	21,41
P1	25	27_1	100	9	1223856	2420700	49518	0,182	122	13,56	41993	4666	1781	1132	649	44,82
P1	26	28_1	100	10	1264751	1269918	5167	1,935	21	2,10	4412	441	1675	1338	337	23,27
P1	27	29_1	no	24	1269918	2445235	89723	0,267	182	7,58	75175	3132	1932	1324	608	41,99
P1	28	30_1	MT	6	1354242	1377689	23447	0,256	53	8,83	20512	3419	1689	1214	475	32,80
P1	29	31_1	no	16	1377689	2465402	69052	0,232	152	9,50	52361	3273	1853	1123	730	50,41
P1	30	32_1	MT	4	1431686	1442046	10360	0,386	25	6,25	7582	1896	1863	1288	575	39,71
P1	31	33_1	84	15	1442046	1501051	59005	0,254	160	10,67	51527	3435	1865	1233	632	43,64
P1	32	34_1	no	19	1501051	2539390	198381	0,096	504	26,53	165643	8718	2017	1146	871	60,15
P1	33	35_1	no	12	1632564	1692121	53916	0,223	117	9,75	46936	3911	2034	1184	850	58,70
P1	34	36_1	100	12	1670425	1718928	20567	0,583	69	5,75	16678	1390	1561	1187	374	25,83
P1	35	37_1	86	13	1718928	2559422	27121	0,479	75	5,77	22536	1734	1729	1042	687	47,44
P1	36	38_1	100	15	1738713	1762470	23757	0,631	68	4,53	19711	1314	1554	1070	484	33,42
P1	37	39_1	MT	18	1762470	1872432	109962	0,164	262	14,56	94882	5271	1759	1178	581	40,12
P1	38	40_1	65	17	1872432	1906462	34030	0,500	100	5,88	26380	1552	1650	1174	476	32,87
P1	39	41_1	no	7	1906462	1927549	21087	0,332	44	6,29	18067	2581	1733	1372	361	24,93
P1	40	42_1	100	1	1927549	1929165	1616	0,619	4	4,00	1147	1147	1607	1499	108	7,46
P1	41	43_1	no	6	1929165	2066462	93795	0,064	166	27,67	64729	10788	1852	1178	674	46,54
P1	42	44_1	80	3	2039008	2057647	18639	0,161	36	12,00	13573	4524	1852	1380	472	32,59

Olen magistriprojekti kirjutanud iseseisvalt. Kõigile töös kasutatud teiste autorite töödele, põhimõtteliste seisukohtadele ning muudest allikaist pärinevatele andmetele on viidatud.

Autor: Krete Saak

20.05.2019

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Krete Saak,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose

Tõlkemälu ja masintõlke vastete mõju kognitiivsele koormusele tõlkimisprotsessis,

mille juhendajad on Terje Loogus ja Marju Taukar,

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Krete Saak
20.05.2019