

Niklas Nieminen

VENÄJÄN KIELEN VERBIASPEKTIN MERKITYKSEN VÄLITTYMINEN SUOMENKIELISIIN KONEKÄÄNNÖKSIIN

Informaatioteknologian ja viestinnän tiedekunta
Pro gradu -tutkielma
Elokuu 2024

TIIVISTELMÄ

Nieminen, Niklas: Venäjän kielen verbiaspektin merkityksen välittyminen suomenkielisiin konekäännöksiin
Pro gradu -tutkielma
Tampereen yliopisto
Monikielisen viestinnän ja käännöstieteen maisteriohjelma, venäjän kääntämisen ja tulkkauksen opintosuunta
Elokuu 2024

Tässä tutkielmassa tarkastellaan venäjän kielen verbiaspektin merkityksen välittymistä suomenkielisissä konekäännöksissä. Konekääntäminen on hyvin ajankohtainen aihe käännöstieteessä, sillä konekääntämisen käyttö on lisääntynyt huomattavasti. Konekäännösten laatu on myös viime vuosina parantunut merkittävästi. Venäjän verbiaspekti on usein ongelmallinen asia suomalaisopiskelijoille, joten siksi on kiinnostavaa selvittää, onko se ongelmallinen asia myös konekääntimelle. Aihe on mielenkiintoinen, koska aspektin ilmaisemisessa on huomattavia eroja suomen ja venäjän kielten välillä.

Keskeisenä tutkimuskysymyksenä on se, miten konekäännin (Google-kääntäjä) selviää venäjän kielen verbien aspektien merkityksen välittämisestä suomen kieleen. Hypoteesina on se, että konekäännin selviytyy hyvin vain osasta käännöksiä. Hypoteesi perustuu siihen, että konekääntimet tekevät lähes aina virheitä, joten siksi on syytä olettaa, että näin tulee käymään myös tässä tutkimuksessa. Samaan aikaan kuitenkin oletan, että myös onnistuneita käännösratkaisuja ilmenee. Aspektin lisäksi tutkitaan myös konekäännösten laatua.

Tutkielman teoreettisessa osassa esitellään konekääntämistä yleisesti sekä neuroverkkokääntämistä. Lisäksi esitellään aspekti kielitieteellisenä käsitteenä sekä keinot, millä tavoin aspekti ilmenee suomen ja venäjän kielissä. Aspektin ilmaisukeinoja sekä verbien aikamuotoja suomen ja venäjän kielten välillä vertaillaan keskenään. Lisäksi käsitellään rinnakkaiskorpusen käyttöä käännöstieteellisessä tutkimuksessa.

Tutkimusmateriaali koostuu venäjänkielisestä Wikipediasta kerätyistä henkilöartikkeleista sekä näiden Google-kääntäjällä tehdyistä suomenkielisistä konekäännöksistä. Artikkeleista ja niiden konekäännöksistä on koottu rinnakkaiskorpus, jonka avulla tekstejä on vertailtu. Korpuksen kokoamisessa on käytetty apuna ohjelmia LF Aligner ja Sketch Engine, jota käytettiin myös korpusaineiston analysointiin. Tutkielman analyysiosassa käydään läpi kymmenen keskeistä venäjän kielen verbiä ja esitetään näiden verbien aspekteihin liittyviä esimerkkejä teksteistä ja käännöksistä.

Tapaustutkimuksissa käy ilmi, että konekäännin selviää valittujen venäjän verbien aspektin merkityksen välittämisestä varsin hyvin. Peräti yli 98 prosentissa tapauksista aspektin merkitys välittyi käännökseen hyvin. Epäselviä tapauksia sekä tapauksia, joissa aspektin merkitys välittyi vain osittain tai ei ollenkaan, oli vain muutama. Aspektin merkityksen välittymisen ohella on tarkasteltu lisäksi valittuja verbejä sisältävien virkkeiden konekäännösten laatua. Enemmistö (57 prosenttia) näistä käännöksistä on laadultaan hyviä, mutta heikkojen ja käyttökelvottomien käännösten osuus (43 prosenttia) on myös merkittävä. Aspektin merkityksen välittyminen hyvin ei siis vielä tarkoita sitä, että koko virke olisi käännöslaadultaan niin ikään hyvä.

Avainsanat: konekääntäminen, neuroverkkokääntäminen, aspekti, aspektuaalisuus, rinnakkaiskorpus
Tämän julkaisun alkuperäisyys on tarkastettu Turnitin Originality Check -ohjelmalla.

SISÄLLYS

1 JOHDANTO	1
1.1 Tutkimusaihe	1
1.2 Aiempi tutkimus	2
1.3 Tutkimuskysymys ja hypoteesi	3
1.4 Aineisto ja metodi	3
1.5 Tutkielman rakenne	4
2 KONEKÄÄNTÄMINEN JA ASPEKTI	6
2.1 Konekääntäminen ja käännösten koheesio	6
2.1.1 Konekääntäminen	6
2.1.2 Neuroverkkokääntäminen	9
2.1.3 Google-kääntäjä	11
2.1.4 Käännösten koheesio	12
2.2 Aspekti	13
2.2.1 Aspekti kielitieteellisenä käsitteenä	13
2.2.2 Aspekti suomen kielessä	14
2.2.3 Aspekti venäjän kielessä	16
2.2.4 Yhteenveto	19
2.3 Aspekti ja verbin aikamuodot suomen ja venäjän välillä	20
3 VERBIEN JA KÄÄNNÖSTEN TUTKIMUS	25
3.1 Tutkimusmetodi ja -materiaali	25
3.1.1 Korpus ja rinnakkaiskorpus	25
3.1.2 Sketch Engine	27
3.1.3 Tutkimusmateriaali	28
3.2 Verbit ja aspektiparit	29
3.2.1 Verbit ja aspektiparit teksteissä: yleisiä tendenssejä	29
3.2.2 Venäjänkielisten verbien käännösvastineet	32

3.2.3 Tutkittavien verbien ja aspektiparien valinta	36
3.3 Tapaustutkimukset	38
3.3.1 выигрывать / выиграть	39
3.3.2 выпускать / выпустить	41
3.3.3 играть / сыграть	43
3.3.4 начинать / начать	46
3.3.5 писать / написать	48
3.4 Analyysin yhteenveto	49
4 LOPPUPÄÄTELMÄT	53
LÄHTEET	57
Aineistolähteet	57
Kirjallisuuslähteet	57
LITTEET	61
Liite 1: Luettelo tutkimusmateriaalin artikkeleista	61
Liite 2: Esimerkki artikkelista ja konekäännöksestä	63
VENÄJÄNKIELINEN LYHENNELMÄ	I
ПЕРЕДАЧА ЗНАЧЕНИЯ ВИДА РУССКОГО ГЛАГОЛА В МАШИННОМ ПЕРЕВОДЕ НА ФИНСКИЙ ЯЗЫК	I
Введение	i
Машинный перевод и аспектуальность	ii
Метод и материал исследования	vi
Основные результаты исследования	viii
Заключение	xi

1 JOHDANTO

1.1 Tutkimusaihe

Konekääntäminen on tällä hetkellä hyvin ajankohtainen aihe käänntieteessä, sillä konekääntämisen käyttö lisääntyy koko ajan kovaa vauhtia. Konekäännösten laatu on parantunut huomattavasti viime vuosina neuroverkkokääntämisen nopean kehittymisen myötä. Siksi myös ammattikäntäjät hyödyntävät nykyisin konekäännöksiä työssään. Nykyisin konekääntämistä käyttää arvioiden mukaan maailmassa ainakin miljardi ihmistä, joista vain pieni osa on ammattikäntäjiä (Nurminen 2021, 23).

Konekääntimiä käytettäessä on kuitenkin syytä kiinnittää huomiota asioihin, jotka saattavat tuottaa hankaluuksia. On hyvä tarkastella esimerkiksi sitä, pyrkiikö konekäännin välittämään lähdetekstin rakenteita myös kohdetekstiin, vai pystyykö käännin tekemään kohdetekstistä kyseisen kielen rakenteiden mukaisen. Huolimatta siitä, pystyykö konekäännin edellä mainittuun asiaan vai ei, pitää muistaa, että täydellisiä käännöksiä ei kuitenkaan lopulta ole olemassa. Siksi haluankin selvittää, kuinka hyvin konekäännin selviää kääntämisestä jonkin tietyn ilmiön osalta. Tutkimuskohteekseni olen valinnut venäjän verbien aspektien ilmaiseman merkityksen välittymisen suomenkieliseen konekäännökseen.

Aspektilla tarkoitetaan erilaisia tilanteen keston ilmaisemiseen liittyviä seikkoja (VISK § 1498). Eri kielissä on erilaisia tapoja aspektin ilmaisemiseen. Venäjässä aspekti ilmenee systemaattisesti kieliopillisin keinoin verbeissä, kun taas suomessa aspektia ei voi automaattisesti tunnistaa eri kieliopillisten muotojen perusteella (Nurminen 2015, 14). Aspektin ilmaiseminen eroaa siis huomattavasti suomen ja venäjän kielten välillä. Venäjän kielen aspektijärjestelmä tuottaa usein suomalaisille venäjän kielen opiskelijoille haasteita. Siksi venäjästä kääntäminen suomeen voi myös olla hankalaa. Jos tiettyjen ilmiöiden kääntäminen saattaa olla ihmiselle haastavaa, niin on mielenkiintoista tutkia, onko se haastavaa myös konekääntimelle. Suomen ja venäjän väliset erot aspektin ilmaisemisessa tekevät tutkimuskohteestani mielestäni myös kiinnostavan. Siksi olen tutkielmassani myös vertailut aspektin ilmaisemista näiden kielten välillä sekä esittänyt osittaisia vastaavuuksia eri keinojen välillä.

1.2 Aiempi tutkimus

Konekääntäminen sekä aspektit ja aspektuaalisuus ovat kaikki laajasti tutkittuja aiheita, joten niistä on olemassa runsaasti aiempaa tutkimusta. Konekääntämisen teoria perustuu tässä tutkimuksessa pääosin Lynne Bowkerin ja Jairo Buitrago Ciron (2019) teokseen. Neuroverkkokääntämisen osalta teoria perustuu suurimmaksi osaksi Philipp Koehnin (2020) teokseen. Neuroverkkokääntämistä sekä konekääntämistä yleisemmin käsitellään myös Dorothy Kennyn (2022) toimittamassa teoksessa.

Aspektia ja aspektuaalisuutta yleisellä tasolla ovat tutkineet teoksissaan Bernard Comrie (1976) sekä Östen Dahl (1985). Venäjän kielen aspektia ja aikamuotoja käsitellään laajasti A. V. Bondarkon (1971) teoksessa. Lisäksi teoriaosan koostamiseen on käytetty myös kotimaisia lähteitä liittyen venäjän aspektiin. Näiden avulla saadaan myös näkökulmaa siihen, miten venäjän aspektista puhutaan suomalaisille venäjän opiskelijoille. Suomen kielen aspektia käsitellään muun muassa Ebba Askosen (2001) ja Salla Nurmisen (2015) väitöskirjoissa. Heistä Askonen käsittelee aspektia objektin sijanvalinnan näkökulmasta, Nurminen tiettyjen verbien näkökulmasta. Keskeisenä lähteenä suomen aspektin ja kieliopin osalta on myös Ison suomen kieliopin verkkoversio (VISK).

Suomen ja venäjän aspektuaalisuutta yhdessä on tutkinut Hannu Tommola (1986) väitöskirjassaan. Hänen väitöskirjansa muodostaa vahvan perustan tässä tutkimuksessa esitellyille vertailuille suomen ja venäjän aspektuaalisuudesta. Venäjän ja suomen aspektuaalisuuden ilmaisemisen samankaltaisuuksia ja eroja käsitellään lisäksi John Dingleyn (2003) artikkelissa sekä Östen Dahlin ja Fred Karlssonin (1976) artikkelissa. Myös Arto Mustajoki ja Helena Niemensivu (1990) käsittelevät teoksessaan venäjän aspektia suomalaisesta näkökulmasta.

Aspektia ja konekääntämistä yhdessä on sen sijaan tutkittu vähemmän. Barbara Sonnenhauserin ja Robert Zangenfeindin (2016) artikkelissa käsitellään aspektia ja konekääntämistä venäjästä englantiin. Tutkimuksessaan he tutkivat, miten venäjän aspektiin liittyviä asioita ilmaistaan englanninkielisissä käännöksissä. Näiden tietojen perusteella he muodostivat sääntöjä ja malleja, joiden pohjalta venäjän aspektiin liittyvää konekääntämistä olisi mahdollista kehittää. Heidän tutkimuksensa kohdistui siis sääntöpohjaisen konekääntämisen malliin. Tässä tutkimuksessa sen sijaan käytetään neuroverkkokääntämisen malliin pohjautuvaa konekääntämistä. Jo aiemmassa artikkelissaan Zangenfeind ja Sonnenhauser (2014) totesivat, että

venäjän aspekti on monimutkainen ongelma konekääntimelle, kun käännetään kieleen, jossa aspektia ei ilmaista samalla tavalla kuin venäjässä.

Tutkimuksessa käytettävään korpusmetodiin liittyviä seikkoja on myös syytä tarkastella. Rinnakkaiskorpusen käyttöä käännoistieteellisessä tutkimuksessa käsitellään etenkin Mikhail Mikhailovin ja Robert Cooperin (2016) teoksessa. Tutkimuksessa korpuksen kokoamiseen sekä aineiston analysointiin käytettävää korpusohjelmaa Sketch Engine käsitellään Adam Kilgarriffin yms. (2014) kirjoittamassa artikkelissa.

1.3 Tutkimuskysymys ja hypoteesi

Tutkimuskysymyksenä tässä tutkimuksessa on se, miten konekäännin (Google-kääntäjä) selviää venäjän kielen verbien aspektien merkityksen välittämisestä suomen kieleen. Olen päättänyt käyttää tutkimuksessa Google-kääntäjää, koska se on tällä hetkellä yksi tunnetuimmista ja suosituimmista konekääntimistä. Lisäksi Google-kääntäjän toiminta perustuu nykyisin neuroverkkokääntämiseen, joten Google-kääntäjä on osa konekäännintä sukupolvea. Neuroverkkoihin perustuvilla kääntimillä saadaan parhaat tulokset. Toissijaisesti tutkimuksessa kiinnitetään huomiota myös tutkittavien verbien aikamuotoihin sekä yleisesti konekäännösten laatuun.

Hypoteesina tässä tutkimuksessa on se, että konekäännin selviytyy hyvin vain osasta käännöksistä. Hypoteesini perustuu siihen, että konekääntimet tekevät lähes aina virheitä, joten siksi on syytä olettaa, että näin tulee käymään myös tässä tutkimuksessa. Samaan aikaan kuitenkin oletan, että myös onnistuneita käännösratkaisuja ilmenee aineistoa analysoitaessa. Oletus on se, että aspektin merkityksen välittäminen tuottaa jonkin verran haastetta konekääntimelle.

1.4 Aineisto ja metodi

Tutkielman tutkimusmateriaalina on kokoelma venäjänkielisiä Wikipedia-artikkeleita, joista on tehty lisäksi suomenkieliset konekäännökset Google-kääntäjää käyttämällä. Tutkimusmateriaalin artikkeleista ja niiden konekäännöksistä on koottu rinnakkaiskorpus käyttäen apuna ohjelmia LF Aligner sekä Sketch Engine. Näistä jälkimmäistä käytetään myös korpusaineiston analysointiin.

Aineistoa analysoidaan siten, että ensin tutkitaan, mitkä verbit esiintyvät usein teksteissä. Venäjänkielisille verbeille etsitään niiden aspektiparit. Tämän jälkeen näille verbeille määritetään rinnakkaiskorpuksen avulla käänösvastineet. Kun käänösvastineet on saatu selville, on aika valita ne verbit ja aspektiparit, jotka pääsevät mukaan lopullisiin tapaustutkimuksiin. Näissä tapaustutkimuksissa tutkitaan varsinaista tutkimusaihetta eli venäjän aspektin merkityksen välittymistä suomenkielisessä konekäännöksessä. Analyysin apuna käytetään tilastoja sekä käänösesimerkkejä korpuksessa mukana olevista teksteistä. Käänösesimerkit luokitellaan niin aspektin merkityksen välittymisen sekä käänöksen laadun perusteella.

1.5 Tutkielman rakenne

Tutkielmani koostuu neljästä luvusta. Ensimmäisessä luvussa eli johdannossa olen esitellyt tutkimusaiheen ja perustellut sen ajankohtaisuuden. Olen tuonut esille aihepiiriin liittyvää aiempaa tutkimusta sekä omat tavoitteeni ja tutkimuskysymykseni. Myös tutkimuksen metodi ja tutkimusaineisto on esitelty lyhyesti.

Toisessa luvussa esittelen teorian liittyen tutkimusaiheeseen. Luku alkaa konekääntämisestä ja neuroverkkokääntämisestä kertovilla kappaleilla, joilla pyritään luomaan helposti ymmärrettävä yleiskuva näistä aiheista. Lisäksi luvussa käsitellään Google-kääntäjää sekä käänösten koheesiota. Tämän jälkeen esittelen aspektin käsitteen niin yleisestä kuin myös sekä suomen että venäjän kielten näkökulmasta. Aiheen yhteydessä käsitellään myös verbien aikamuotoja kummankin kielten osalta. Lopuksi olen tehnyt yhteenvedon, jossa käsitellään aspektin ilmaisemisen sekä verbien aikamuotojen yhtäläisyyksiä ja eroja suomen ja venäjän kielten välillä.

Kolmas luku eli analyysiosa alkaa tutkimusmetodin ja -materiaalin esittelyllä. Seuraavana luvussa esitellään aineistosta löytyviä verbejä ja aspektipareja. Rajatulle aspektiparien joukolle etsitään käänösvastineet, minkä pohjalta lopulta suoritetaan lopullinen tutkittavien verbien ja aspektiparien valinta varsinaisia tapaustutkimuksia varten. Tapaustutkimuksissa verbejä ja aspekteja analysoidaan tilastojen ja tekstiesimerkkien kautta. Luku päättyy analyysin yhteenvedon.

Neljäs ja viimeinen luku sisältää loppupäätelmät, joissa vielä tiivistetään tutkielman tarkoitus sekä käydään läpi vastaukset tutkimuskysymyksiin. Lisäksi luvussa pohditaan tutkielman

luotettavuutta, kattavuutta sekä jatkotutkimusmahdollisuuksia. Varsinaisen tutkielman lisäksi lukujen jälkeen mukana on myös lähdeluettelo, liitteet sekä tutkielman venäjänkielinen lyhennelmä, johon on koottu tutkielman keskeisimmät asiat sekä tärkeimmät yksityiskohdat tutkimuksen lopputuloksista.

2 KONEKÄÄNTÄMINEN JA ASPEKTI

2.1 Konekääntäminen ja käännösten koheesio

2.1.1 Konekääntäminen

Konekääntämisellä tarkoitetaan tekstin automaattista kääntämistä luonnollisesta kielestä toiseen tietokonesovellusta käyttäen sekä tutkimusaluetta, jolla tietokoneolingvistit yrittävät löytää keinoja tällaiseen kääntämiseen. Konekääntämistä kuvaillaan erittäin vaikeaksi tehtäväksi, koska luonnolliset kielet ovat monimutkaisia. Lauseita voidaan tulkita monella eri tapaa ja kielten kieliopilliset rakenteet voivat erota toisistaan huomattavasti. Hankaluudestaan huolimatta konekääntäminen on yksi ensimmäisistä aloista, joihin tietokoneohjelmia on kehitetty. (ISO 2017, 1; Bowker & Buitrago Ciro 2019, 37.)

Matemaatikko Warren Weaver esitteli konekääntämisen idean yleisölle vuonna 1949. Idea perustui toisen maailmansodan aikaisiin kryptografian tekniikoihin. Seuraavan vuosikymmenen aikana konekääntämistä tutkittiin paljon. Tutkijat uskoivat vahvasti siihen, että pian tietokoneet pystyisivät tuottamaan korkealaatuisia käännöksiä. Kuitenkin 1960-luvun alkuun mennessä oli tullut selväksi, että kryptografian tekniikat olivat liian rajattuja ja kääntäminen liian monimutkaista koneille. Konekääntämisen tutkimus hiipui hetkellisesti. Tutkimus kuitenkin jatkui, ja 1970-luvulla kehitettiin esimerkiksi tiettyjen erikoisalojen teksteille sopivia kääntimiä. (Bowker & Buitrago Ciro 2019, 37–38.) Automaattisen kääntämisen järjestelmiä käytettiin tuohon aikaan muun muassa puolustuksen ja hallinnon organisaatioissa sekä kansainvälisissä järjestöissä (Kenny 2022, 32).

Kiinnostus konekääntämistä kohtaan alkoi jälleen kasvaa 1980-luvulla. Tietokoneavusteisen kääntämisen ohjelmiin, kuten käännoäsmuistiohjelmiin, alettiin kiinnittää huomiota 1990-luvulla. (Bowker & Buitrago Ciro 2019, 38.) Käännoäsmuistin toiminta perustuu siihen, että se tallentaa kääntäjän tekemät käännoäset, jotta niitä voisi hyödyntää uudelleen (Kenny 2022, 30). Nykyään konekääntäminen on integroitu käytettäväksi esimerkiksi käännoäsmuistiohjelmissa (Bowker & Buitrago Ciro 2019, 39). Tietokoneavusteisen kääntämisen ohjelmat eivät siis itsessään tee käännoästyötä kääntäjän puolesta, vaan ne ovat vain apuväline. Vuonna 1997 ensimmäinen ilmainen internetissä toimiva konekäännoäsin, Babel Fish, avautui kuluttajien käytettäväksi (Kenny 2022, 32). Nykyään maksuttomia internetissä toimivia konekääntimiä on

olemassa useita, Google-kääntäjän lisäksi muun muassa Microsoft Translate, DeepL, Reverso sekä erilaisten sosiaalisen median palveluiden konekääntimet (Dorst ym. 2022, 55).

Konekääntämisen historiasta on eroteltavissa erilaisia kehitysvaiheita. Viime vuosituhanen puolella hallitsevin konekääntämisen malli oli sääntöpohjainen konekääntäminen (rule-based machine translation, RBMT) (Bowker & Buitrago-Cirio 2019, 39). Malli perustuu siihen, että ohjelmalle syötetään lista kaikista lähde- ja kohdekielen sanoista sekä säännöistä, miten sanat voidaan yhdistää järkevasti. Mallin kehittäminen oli kuitenkin todella kallista, koska sääntöjen kirjoittamiseen joka kieliparille tarvittiin päteviä lingvistejä. Lisäksi kaikkien mahdollisten sääntöjen kirjoittaminen on yksinkertaisesti mahdotonta. (Kenny 2022, 35.) Säännöt ovat myös sitä monimutkaisempia, mitä monimutkaisempia käännettävien kielten rakenteet ovat. Kielen monitulkintaisuus ja huonosti kirjoitetut lähdetekstit aiheuttavat mallissa myös paljon ongelmia. (Bowker & Buitrago-Cirio 2019, 40–41.) Säännöt voivat myös olla ristiriitaisia tai ne toimivat vain osassa tapauksista.

Nykyisen vuosituhanen puolella hallitsevaksi konekääntämisen malliksi alkoi nousta tilastollinen konekääntäminen (statistical machine translation, SMT). Se säilytti asemansa 2010-luvun puoliväliin saakka. (Kenny 2022, 37.) Tilastollinen konekääntäminen perustuu vahvasti rinnakkaiskorpuksen käyttöön sekä todennäköisyyslaskentaan. Tilastolliseen konekääntämiseen perustuva käännin jakaa virkkeet lyhyisiin lausekkeisiin, joille käännin etsii kohdekielen käännösvastineet. Todennäköisyyslaskentaa hyödyntäen käännin laskee lausekkeelle todennäköisimmän käännösvaihtoehdon kyseisessä kontekstissa. Käännin hyödyntää toiminnassaan rinnakkaiskorpuksia. Jotta tilastollinen konekääntäminen toimisi mahdollisimman hyvin, niin korpuksen pitäisi olla mahdollisimman suuria. Kaikkien mahdollisten sanojen ja ilmausten sisällyttäminen korpuksiin olisi kuitenkin työlästä ja lähes mahdotonta. Lisäksi jos korpuksen aineisto liittyy vahvasti esimerkiksi johonkin erikoisalaan, niin tällöin käännin ei välttämättä sovellu kovin hyvin muiden alojen aineiston kääntämiseen. Korpuksia pitää myös kerätä erikseen jokaiselle kieliparille. Pienempien kielten välillä sopivia rinnakkaistekstejä on hankala löytää. (Bowker & Buitrago-Cirio 2019, 43–44.) Tilastollinen konekääntäminen kuitenkin osoitti sen, että datan pohjalta oppivat konekääntimet toimivat paremmin. Tilastollista konekääntämistä voidaan käyttää edelleen esimerkiksi uusien käänninprojektien alkuvaiheissa. (Kenny 2022, 38.)

Viime vuosina hallitsevaksi konekääntämisen malliksi on noussut neuroverkkokääntäminen (neural machine translation, NMT). Neuroverkkokääntämisen toiminta perustuu

neuroverkkoihin. Ne käsittelevät informaatiota tavalla, joka on samankaltainen eliöiden hermoston toiminnan kanssa. Ihmisten lailla neuroverkot oppivat asioita. (Bowker & Buitrago-Ciro 2019, 44–45.) Neuroverkot koostuvat lukuisista neuroneista, jotka ovat yhteydessä toisiinsa. Neuronit sisältävät tietoja, joita käännin yhdistelee ja täten luo sopivia käännöksiä. Myös neuroverkkokääntimen algoritmia koulutetaan rinnakkaisdatalla. (Kenny 2022, 40.) Tässä tutkimuksessa konekäännösten tekemiseen käytettävän Google-kääntäjän toiminta perustuu nykyisin neuroverkkokääntämiseen (mts. 39). Neuroverkkokääntämistä käsitellään tarkemmin seuraavassa alaluvussa.

Konekääntämistä voidaan käyttää moniin eri tarkoituksiin. Yksi tapa on käyttää niin sanottuja raakakonekäännöksiä siten, että halutaan ymmärtää tietty määrä tekstin sisällöstä jotakin tiettyä tarkoitusta varten. Raakakonekäännöksellä tarkoitetaan konekäännöksen tuottamaa tekstiä, jota ei ole editoitu mitenkään. Lukijat siis tietävät lukevansa koneen tuottamaa käännoästä, joten he suhtautuvat tekstiin eri tavalla kuin ihmisen kirjoittamaan tekstiin. (Nurminen 2021, 30.) Konekäännöksiä voidaan käyttää myös ihmisten väliseen kommunikointiin (mts. 40–41). Monet käyttävät konekääntämistä silloinkin, kun he osaavat lähdetekstin kieltä joko hyvin tai vähän (mts. 89). Esimerkiksi opiskelijat saattavat käyttää konekääntämistä apuna lukiessaan vieraskielisiä artikkeleita ja kirjoittaessaan tekstejä vieraalla kielellä. Konekääntämistä käytetään esimerkiksi korvaamaan sanakirjan käyttöä. (Dorst ym. 2022, 59–60.)

Konekäännösten laatua parannetaan usein käyttämällä jälkieditointia. Jälkieditoinnilla tarkoitetaan konekäännösten editointia ja korjaamista (ISO 2017, 2). Jälkieditointi voi olla joko kevyttä tai perusteellista riippuen käännoäksen käyttötarkoituksesta (mts. 5). Perusteellisen jälkieditoinnin tavoitteena on saada tehtyä tasoltaan ihmiskäännöistä vastaava teksti. On suositeltavaa, että jälkieditoinnissa hyödynnetään mahdollisimman paljon konekääntimen tuottamaa tekstiä. (mts. 8.)

On arvioitu, että tämän tutkielman kirjoittamisen aikana konekääntämisellä on noin miljardi käyttäjää. Konekäännösten käyttäjät voidaan jakaa kahteen ryhmään: ammattikäntäjiin ja niihin, jotka käyttävät raakakonekäännöksiä vieraskielisten tekstien ymmärtämiseen. Jälkieditointia tekeviä kääntäjiä kutsutaan usein jälkieditoijiksi. Jälkieditoijia voidaan pitää tärkeinä konekääntimien käyttäjinä, sillä heille käännoästeknologia on tuttua, ja he osaavat työskennellä raakakonekäännösten kanssa. Heidän näkemyksiään voidaan myös hyödyntää konekääntämisen kehittämisessä. Heidän osuutensa kaikista konekääntämisen käyttäjistä on kuitenkin vain murto-osan. Ylivoimaisesti suurin osa, arvion mukaan jopa 99,97 prosenttia

käyttäjistä, kuuluu siis ryhmään, joka käyttää konekääntämistä pelkästään tekstien ymmärtämiseen. (Nurminen 2021, 23–24.)

2.1.2 Neuroverkkokääntäminen

Neuroverkkokääntämistä tutkittiin jo 1980- ja 1990-luvuilla. Jotkin 1990-luvulla esitellyt neuroverkkokääntämisen mallit ovat hyvin samankaltaisia nykyisten mallien kanssa. Dataa mallien kouluttamiseen ei kuitenkaan vielä tuolloin ollut saatavilla tarpeeksi. Tietokoneiden resurssit eivät myöskään vielä riittäneet. 2000-luvulla esiteltiin malleja, joissa neuroverkkoja oli yhdistelty tilastollisen konekääntämisen menetelmiin. 2010-luvulla alettiin kehittää yhä enemmän malleja, jotka perustuivat pelkästään neuroverkkoihin. Erilaisten, kehittyneempien mallien myötä neuroverkkokääntäminen alkoi nousta konekääntämisen hallitsevaksi alueeksi. Tällä hetkellä selvästi suurin osa konekääntämisen tutkimuksesta keskittyy neuroverkkokääntämiseen. (Koehn 2020, 39–40.)

Neuroverkkokääntämisen toiminta perustuu keinotekoisiiin neuroverkkoihin, jotka jäljittelevät ihmisaivojen toimintaa (Kenny 2022, 40). Neuroverkoissa on neuroneiksi kutsuttuja yksiköitä, jotka ovat yhteydessä toisiinsa. Neuronit reagoivat ja aktivoituvat muilta neuroneilta ja neuronien välisiltä yhteyksiltä saamiensa ärsykkeiden mukaan. (Forcada 2017, 292–293.) Ensimmäisissä neuroverkoissa oli vain yksi käsittelykerros (Koehn 2020, 31). Nykyisin neuroverkkoja hyödynnetään laajasti erilaisissa tekoälysovelluksissa. Syväoppimisessa käytetään monikerroksisia neuroverkkoja. (mts. 33.) Neuroverkot käsittelevät syötettyjä tietoja niin sanotuissa piilokerroksissa sen sijaan, että ne vain suoraan tulostaisivat tietoa pelkästään syötetyn tiedon pohjalta (mts. 68). Syötetty ja tuotettu tieto ovat ulkokerroksissa (Kenny 2022, 42).

Neuroverkkokääntämisen malleissa sanoja edustavat vektorit. Kenny (2022, 41) antaa esimerkin, jossa englanninkielinen sana *apple* ("omena") on esitetty muodossa [1,20; 2,80; 6,10]. Vektorit sen sijaan edustavat sanojen välisiä suhteita. Tietokone oppii määrittämään sopivat vektorit sanoille koulutusmateriaalin perusteella. Samankaltaisissa yhteyksissä esiintyville sanoille tietokoneet määrittelevät myös samankaltaisia vektoreja. Tietokoneen oppimia sanoja edustavia vektoreita kutsutaan niin sanotuiksi sanan upotuksiksi (word embedding), joita voi olla useissa eri tasoissa neuroverkon sisällä. Vektoreita voidaan näin ollen käyttää laskutoimituksiin kuten matematiikassa. (mts. 41–42.) Yleensä kuvittelemme asioita kolmiulotteisesti, mutta sanoja edustaville vektoreille tarvitaan yleensä enemmän ulottuvuuksia

(Forcada 2017, 295). Kun ulottuvuuksia on enemmän, voidaan sanoja sijoittaa kauemmas tai lähemmäs toisistaan tarkemmin. Ulottuvuuksia voi olla jopa satoja. (Pérez-Ortiz ym. 2022, 151.)

Neuroverkkokääntämisen toiminnassa hyödynnetään niin sanottua enkoodaus-dekoodausmallia. Enkoodauksessa neuroverkko käsittelee syötetyn lähdetekstin ja dekoodauksessa neuroverkko tuottaa käsittelyn pohjalta käännöksen. (Koehn 2020, 125.) Enkooderi laskee vektorit tekstin jokaiselle sanalle. Dekooderi ottaa huomioon näiden vektorien lisäksi myös kohdetekstin sanoille lasketut vektorit. (Pérez-Ortiz ym. 2022, 156–157.) Dekooderi laskee todennäköisyydet sana kerrallaan kaikille mahdollisille vaihtoehdoille ja määrittää tämän perusteella, mikä sana sopii parhaiten tiettyyn kohtaan lauseessa (Forcada 2017, 298).

Neuroverkkokääntäminen on korpuspohjaista konekääntämistä, sillä neuroverkkokääntimien kouluttamiseen käytetään suuria rinnakkaiskorpuksia. Vaikka myös tilastollinen konekääntäminen pohjautuu rinnakkaiskorpuksiin, nämä mallit eroavat kuitenkin toisistaan teknisesti. (Forcada 2017, 292.) Neuroverkkojen kouluttamisella tarkoitetaan prosessia, joka määrittää neuronien välisten suhteiden painoarvot. Erilaisten algoritmien avulla neuroverkkoa koulutetaan siten, että verkko pystyy suoriutumaan uusien tekstien kääntämisestä mahdollisimman hyvin olemassa olevaa dataa hyödyntäen. (Pérez-Ortiz ym. 2022, 148–149.) Kouluttaminen on usein vaativaa, ja monimutkaisimpien mallien kouluttaminen voi kestää jopa kuukausia (Forcada 2017, 295).

Neuroverkkokääntämistä pidetään yleisesti tähän asti parhaana konekääntämisen mallina. Neuroverkkokääntimet ottavat kokonaiset lauseet hyvin huomioon, ja ne selviytyvät edeltäjiään paremmin hankalimmistakin rakenteista. (Kenny 2022, 43.) Esimerkiksi tilastollisen konekääntämisen toiminta perustuu enemmän fraaseihin, kun taas neuroverkot huomioivat paremmin yksittäiset sanat (Forcada 2017, 301). Neuroverkkokääntimien, kuten kaikkien muidenkin mallien, ongelmana on se, että ne eivät huomioi kovin hyvin lauseiden yhteyttä toisiinsa esimerkiksi pronomien viittaussuhteiden osalta. Kääntimet voivat myös tuottaa sujuvaa, mutta sisällöltään virheellistä tekstiä. Joskus kääntimet voivat jopa keksiä sanoja, joita ei oikeasti ole olemassa. (Kenny 2022, 43.) Ongelmia voi tulla vastaan myös harvinaisempien sanojen kohdalla (Koehn 2020, 297). Erityisesti jälkieditoijien on oltava tarkkana, koska virheellistä käännöstä on hankalampi huomata, jos lause kuitenkin kuulostaa järkevältä (Forcada 2017, 303). Neuroverkkokääntimien kouluttaminen vaatii edeltäjiään enemmän aikaa

ja resursseja (Kenny 2022, 44). Vahingoittunut tai virheellinen rinnakkaisdata vaikuttaa negatiivisesti käännosten laatuun (Koehn 2020, 299).

2.1.3 Google-kääntäjä

Google-kääntäjä (Google Translate) on teknologiayritys Googlen kehittämä konekäännin. Google-kääntäjä julkaistiin suuren yleisön käytettäväksi vuonna 2006 (Turovsky 2016b). Toukokuussa 2022 Google-kääntäjä oli käytettävissä 133 eri kielellä (Caswell 2022). Kääntimen vuonna 2010 julkaistu Android-sovellus saavutti miljardin latauskerran rajan maaliskuussa 2021 (Pitman 2021). Huhtikuussa 2016 Google ilmoitti, että kääntimellä on yli 500 miljoonaa käyttäjää (Turovsky 2016b). Kennyn (2022, 44) mukaan Google-kääntäjä on kaikkein tunnetuin konekäännin. Myös Dorstin ym. (2022, 55–56) tutkimuksessa havaittiin, että Google-kääntäjä oli tutkimukseen osallistuneiden keskuudessa ylivoimaisesti tunnetuin ja käytetyin konekäännin. Tavallisten tekstien lisäksi kääntimellä voi kääntää myös kuvissa olevia tekstejä tekstintunnistusta hyödyntäen. Puheentunnistuksen avulla myös puheen kääntäminen on mahdollista. (Pitman 2021.)

Ensimmäisten kymmenen vuoden ajan Google-kääntäjän toiminta perustui tilastolliseen konekääntämiseen. Marraskuussa 2016 Google ilmoitti siirtyvänsä neuroverkkokääntämiseen. Muutosta perusteltiin sillä, että neuroverkkokääntämiseen perustuva käännin pystyy kääntämään kokonaisia lauseita paremmin ja käännin oppii itse ajan mittaan tekemään parempia käännöksiä. (Turovsky 2016a). Käyttöönottoa edeltäneessä tutkimuksessa havaittiin, että Googlen neuroverkkokäännin (Google Neural Machine Translation, GNMT) paransi käännösten laatua verrattuna aiempaan tilastolliseen konekääntämiseen perustuvaan kääntimeen. Googlen käännin pystyy myös kääntämään kahden kielen välillä, vaikka käännintä ei olisi opetettuakaan toimimaan näiden kielten välillä, jos molempia kieliä voidaan kääntää jollekin yhteiselle kolmannelle kielelle. (Schuster ym. 2016.) Tällaista yhteistä välikieltä kutsutaan pivot-kieleksi (Tieteen termipankki). Kesäkuussa 2020 Google kertoi, että käännösten laatu on parantunut vuosi vuodelta uusien tekniikoiden ja mallien myötä. Erityisesti edistystä on tullut vähäresurssisten kielten kääntämisessä. Laatu on kuitenkin vielä kaukana täydellisestä. Ongelmia on vielä esimerkiksi tiettyjen aihealueiden teksteissä, liian sananmukaisissa käännöksissä sekä murteellisten tai puhekielisten tekstien kääntämisessä. (Caswell & Liang 2020.)

2.1.4 Käännösten koheesio

Koheesiolla tarkoitetaan sellaisia tekstin ominaisuuksia, jotka muodostavat yhteyksiä tekstin eri osien välille. Erilaiset koheesion keinot muodostavat sidoksia sanojen ja virkkeiden välille. Sidokset luovat siis koheesiota. Näin ollen koheesiota kutsutaan myös sidosteisuudeksi. Esimerkiksi aiemmin mainittuun asiaan voidaan myöhemmin tekstissä viitata pronomiinilla. Tällöin asiaa ilmaisevan sanan (substantiivin) ja pronominin välille on muodostunut sidos. Koheesio voi olla sekä kieliopillinen että semanttinen ominaisuus. Edellä mainittu esimerkki substantiivin ja pronominin välisestä sidoksesta edustaa semanttista ominaisuutta, koska sanat viittaavat samaan asiaan. Semanttista koheesiota voidaan luoda myös esimerkiksi käyttämällä synonyymejä, antonyymejä tai ala- ja yläkäsitteitä. Kieliopillisesta koheesiosta on kyse esimerkiksi silloin, kun samaan asiaan viitataan saman sanan eri sijamuodoilla. Tällöin viittausten muodostamaa ketjua muokataan nimenomaan kieliopillisin keinoin. Kieliopilliseksi koheesioksi luetaan myös lauseiden yhdistäminen konjunktioita käyttämällä. (Hiippala 2015, 7–9.)

Koheesion käsitteeseen liittyy myös läheisesti käsite koherenssi. Koherenssilla tarkoitetaan ominaisuuksia, jotka tuovat tekstissä esitetyt väittämät yhteen. Ajatellaan, että tekstit eivät yleensä koostu vain satunnaisista virkkeistä. (Hiippala 2015, 9–10.) Koherenssin tehtävänä on siis luoda loogisesti etenevä kokonaisuus toisiinsa liittyvistä asioista (Menzel ym. 2017, 1). Sekä koheesion että koherenssin ominaisuuksia tarvitaan, jotta teksti voi olla ymmärrettävä ja yhtenäinen kokonaisuus (Hiippala 2015, 9). Koheesio on osa koherenssia (Sim Smith & Specia 2017, 133). Erilaiset keinot ja tavat, joiden myötä koherenssi voidaan saavuttaa, vaihtelevat kielittäin (mts. 138).

Tekstejä käännettäessä on etsittävä lähdetekstissä esiintyvien asioiden viittaussuhteet. Näiden ilmaisemiseen käännöksessä tarvitaan kohdekielen koheesion luomisen keinoja. Tämä koskee luonnollisesti niin ihmisen kuin koneen tekemiä käännöksiä. (Menzel ym. 2017, 2.) Koherentista lähdetekstistä pitäisi ihanteellisesti saada tulokseksi koherentti käännös. Ihmiskääntäjälle tämä on konekäännintä helpompaa, koska ihminen voi keskittyä tekstiin kokonaisuutena. Ihmiselle koheesion varmistaminen voi olla myös intuitiivista. Konekääntimet sen sijaan kääntävät usein tekstejä vain yksittäinen lause kerrallaan, eivätkä ne pysty havaitsemaan koheesioon ja koherenssiin liittyviä piirteitä. Konekääntimet eivät välttämättä ota kontekstia huomioon, joten käännös saattaa olla asiayhteyden sopimaton. (Sim Smith & Specia 2017, 131–133.)

Konekäännösten koheesiota on tutkittu jonkin verran. Lapshinova-Koltunski (2017, 106) ei ole aiemmissa tutkimuksissaan havainnut merkittäviä eroja ihmisen ja konekääntimen tekemien käännösten koheesioon liittyvissä ominaisuuksissa, vaikka konekäännösten laatu ei ollut verrattavissa ihmiskäännöksiin. Sim Smithin ja Specian (2017, 132) mukaan asiaa koskevien tutkimusten tulokset ovat olleet vaihtelevia. Joissakin tutkimuksissa konekääntimien on havaittu pärjäävän hyvin koheesioon liittyvien ominaisuuksien kanssa, joissakin sen sijaan ei. Tutkimuksissa on käytetty sekä rinnakkaisia ihmis- ja konekäännöksiä että erillisiä ihmiskäännöksiä vertailumateriaalina (Lapshinova-Koltunski 2017, 112). Sim Smith ja Specia (2017, 131) itse tutkivat sitä, miten koherenssi voidaan taata konekäännöksissä.

2.2 Aspekti

2.2.1 Aspekti kielitieteellisenä käsitteenä

Aspektin käsitteelle on olemassa kielitieteessä useita eri määritelmiä, jotka vaihtelevat yleensä sen mukaan, tarkastellaanko aspektia merkitys- vai muotokategoriana. Nurminen (2015, 14) määrittelee aspektin seuraavasti: "[A]spekti on merkityskategoria, joka ilmaisee lauseen kuvaaman asiointilan ja siinä tapahtuvan tekemisen kestoa eli ajassa etenemistä." Ison suomen kieliopin (VISK § 1498) mukaan aspektilla tarkoitetaan "lauseen rakenteen eri tasoilla ilmeneviä tilanteen keston ilmaisemiseen liittyviä seikkoja". Myös VISK pitää aspektin käsitettä laajasisältöisenä ja moniselitteisenä. Comrien (1976, 3) mukaan aspektit ovat erilaisia tapoja tarkastella tilanteen sisäistä kestoa. Edellä mainitut määritelmät eroavat toisistaan jonkin verran, mutta kaikille niille yhteistä on se, että aspekti liittyy jollain tavalla ilmaistun tilanteen keston ilmaisemiseen. Aspektin ilmaisemisessa on eroavaisuuksia eri kielten välillä. Joissakin kielissä, kuten venäjässä, aspekti ilmenee systemaattisesti kieliopillisin keinoin. Sen sijaan esimerkiksi suomen kielessä aspektia ei voi systemaattisesti tunnistaa muotojen perusteella. (Nurminen 2015, 14.) Tällöin aspektia on käytännössä tutkittava nimenomaan merkityskategoriana.

Lauseen aspektilla tarkoitetaan lauseen merkityksen sisäistä aikarakennetta. Lauseessa kuvailtu tilanne voi olla joko rajattu eli perfektiivinen tai rajaamaton eli imperfektiivinen. Rajatussa tilanteessa on kyse tilanteesta, jolla on alku ja loppu tai se on muutoin saatettu loppuun. Rajaamaton tilanne sen sijaan tarkoittaa tilannetta, joka on tai on ollut käynnissä, on keskeneräinen tai ei ole johtanut tulokseen. Myös toistuvaa toimintaa on tapana kuvata

imperfektiivisellä aspektilla. (VISK § 1498.) Aspekti voi näin ollen kuvata puhujan näkökulmaa tilanteeseen (Kirvesmäki ym. 2021, 141). Rajatusta aspektista voidaan puhua myös terminatiivisena aspektina ja rajaamattomasta aspektista kursiivisena tai duratiivisena aspektina (VISK § 1500).

Aspektista voidaan puhua toiminnan keston ilmaisemisen lisäksi myös toiminnan resultatiivisuuden eli tulokseen johtavuuden näkökulmasta. Aspekti on resultatiivinen, jos verbin kuvaama toiminta aiheuttaa muutoksen objektin kohteessa. Aspekti on irresultatiivinen, jos muutosta ei aiheudu. (Askonen 2001, 38; Hakulinen & Karlsson 1979, 183–185.) Resultatiivisuutta käsitellään usein etenkin lauseen objektin sijanvaihtelun yhteydessä. Voidaan puhua myös resultatiivisista verbeistä (esimerkiksi *löytää* ja *unohtaa*). Irresultatiivisina verbeinä pidetään muun muassa tunnetta ja edestakaista liikettä kuvaavat verbit (esimerkiksi *rakastaa* ja *heiluttaa*), joiden kuvaama toiminta ei yleensä johda mihinkään tulokseen. (VISK § 1500.)

Aspektia tarkastellaan usein yhdessä tempuksen eli aikamuotojen kanssa (Nurminen 2015, 52). Tempus ilmaisee verbien tarkoittaman tilanteen aikaa suhteessa siihen, millä hetkellä asiasta puhutaan. Tempus sisältää siis viittauksen puhujan nykyhetkeen ja kyse on tilanteen ulkoisesta ajasta. Aspektissa on sen sijaan kyse tilanteen sisäisestä ajasta, jolloin se kuvaa tilannetta tapahtumahetkellä. (VISK § 1523.) Aspektin ja tempuksen erottaminen toisistaan voi kuitenkin joissain tapauksissa olla epäselvää (Dahl 1985, 25). Vaikeuksia voi aiheuttaa esimerkiksi se, että aspektin ja tempuksen kuvaamiseen käytetään samankaltaisia käsitteitä. Aspektista puhutaan usein imperfektiivisenä tai perfektiivisenä, kun taas tempuksen kuvaamiseen käytetään esimerkiksi termejä imperfekti ja perfekti. Näistä perfektiä saatetaan käsitellä myös aspektina. (Nurminen 2015, 53–54.) Comrien (1976, 52) mukaan perfekti ilmaisee sen, että jo tapahtunut tilanne on edelleen relevantti puhehetkellä. Siksi hän käsitteleeikin omassa tutkimuksessaan perfektiä aspektina. Suomen kielessä perfektiä on perinteisesti pidetty aikamuotona (Nurminen 2015, 55).

2.2.2 Aspekti suomen kielessä

Suomen kielessä aspektia voidaan ilmaista muun muassa lauseen objektin sijalla, verbien aikamuodoilla sekä erilaisilla verbeillä ja niiden johdoksilla. Askosen (2001, 39) mukaan objektin sijanvalinta on suomen kielen tärkein morfologinen keino aspektin ilmaisemiseksi. Juva (2019, 128) pitää tätä suomen kielen erityispiirteenä. Keskeisiä sijamuotoja aspektin

ilmaisemisessa ovat partitiivi ja akkusatiivi. Partitiivissa oleva osaobjekti kuvaa kesken olevaa toimintaa, kun taas akkusatiivissa oleva kokonaisobjekti ilmaisee lopullista, päättynyttä tai tulokseen johtanutta tekoa. (Askonen 2001, 39; Juva 2019, 128.) Näin ollen siis partitiivi on irresultatiivisen ja akkusatiivi resultatiivisen aspektin tunnus (Hakulinen & Karlsson 1979, 183). Esimerkiksi ilmaus "luin sanomalehden" kertoo sen, että sanomalehti on luettu alusta loppuun saakka, mutta ilmaus "luin sanomalehteä" ei kerro mitään siitä, kuinka paljon, mihin asti ja mistä kohtaa sanomalehteä on luettu.

On syytä muistuttaa, että suomessa akkusatiivi on lähinnä syntaktinen käsite, sillä vain persoonapronomineilla (*minut, sinut, hänet* jne.) ja kysymyspronominilla *kuka* (*kenet*) on morfologinen t-päätteinen akkusatiivi. Muissa tapauksissa akkusatiivilla on tarkoitettu muita objektin sijoja kuin partitiivia. Voidaan puhua esimerkiksi genetiiviakkusatiivista ("söin omenan") ja nominatiiviakkusatiivista ("syön omenat"). (VISK § 1221 & 1226.) Objektia, joka on nominatiivissa, genetiivissä tai t-päätteisessä akkusatiivissa, kutsutaan totaaliobjektiksi. Totaaliobjektia käytetään aspektiltaan rajatussa lauseissa, kun taas partitiiviobjektia käytetään aspektiltaan rajaamattomissa lauseissa. Kielteisissä lauseissa objekti on kuitenkin pääsääntöisesti partitiivissa. (VISK § 930.) Aspektiltaan rajatussa lauseessa esiintyvää verbiä voidaan kutsua rajaavaksi ja rajaamattomassa lauseessa esiintyvää verbiä rajapakoiseksi verbiksi. Lisäksi on myös olemassa rajahakuisia verbejä, jotka voivat esiintyä aspektiltaan kummanlaisessa tahansa lauseessa joko totaali- tai partitiiviobjektin kanssa. (VISK § 1508.) Objektin sijanvalintaan vaikuttaa myös kvantitatiivinen määräisyys. Totaaliobjekti on kvantitatiivisesti määräinen ("söin jäätelön"), kun taas partitiiviobjekti on epämääräinen ("söin jäätelöä"). (VISK § 931.)

Aspektia voi suomen kielessä ilmaista myös aikamuodoilla eli tempuksilla. Suomen kielessä on neljä tempusta: preesens, imperfekti eli preteriti, perfekti ja pluskvamperfekti. (Nurminen 2015, 19.) Menneistä aikamuodoista imperfekti ja perfekti sisältävät päätyneisyyden ja jatkuvuuden ulottuvuuden. Imperfektillä viitataan menneeseen aikaan sijoittuvaan tapahtumaan, kun taas perfektillä viitataan tilanteeseen, joka edelleen jatkuu tai se vaikuttaa nykyhetkeen. (Juva 2019, 130.) Imperfektin ilmaisema tilanne on siis päättynyt puhehetkeen mennessä (VISK § 1530). Perfekti on sen sijaan osittain sekä menneen että menemättömän ajan aikamuoto. Päättynyttä tilannetta kuvaava perfekti ilmaisee tilanteen vaikutusta tai tulosta. Aiemmin alkanutta, käynnissä olevaa tilannetta kuvaava perfekti ilmaisee tilanteen jatkuvuutta. (VISK § 1534.)

Suomen kielessä verbin aspektiin voi vaikuttaa myös verbijohtimilla (Askonen 2001, 40). Muuntelujohdokset vaikuttavat verbin kuvaamaan toimintaan, tapahtumaan tai tilaan. Johdokset kertovat esimerkiksi tilanteen rajaamattomuudesta, kestosta, jatkuvuudesta, toistuvuudesta tai hetkellisyydestä. (VISK § 351.) Esimerkiksi ilmauksissa "heittää palloa" ja "heitellä palloa" on nähtävissä ero tilanteen kestossa. Ensimmäisen esimerkin kohdalla kyse on yhdestä kerrasta, kun taas jälkimmäisessä esimerkissä on kyse pidempään jatkuvasta, toistuvasta tapahtumasta. Verbin *heitellä* kaltaisia verbejä kutsutaan frekventatiivisiksi verbeiksi, joita voidaan muodostaa lisäämällä verbiin frekventatiivinen johdos. Tyypillisesti objekti on partitiivissa ja aspekti on rajaamaton, mutta totaaliobjektilla voidaan kuvata rajattua aspektia (esim. "lueskelin lehteä", "lueskelin lehden"). Momentaanijohdoksilla kuvataan tyypillisesti hetkellisiä tapahtumia, joten aspekti on luontaisesti rajattu. (VISK § 352.) Merkityksessä korostuu kerrallisuus ja lyhytkestoisuus. Verbejä voidaan muodostaa muun muassa *-Ahta*-johtimella (esimerkiksi *istua: istahtaa, nukkua: nukahtaa, tippua: tipahtaa*). (VISK § 368.)

2.2.3 Aspekti venäjän kielessä

Venäjän kielen verbijärjestelmässä aspekti (вид) on kieliopillinen kategoria. Aspekteja on kaksi. Ajallisesti rajattua toimintaa kuvaa perfektiiivinen aspekti (совершенный вид), kun taas päättymätöntä, jatkuvaa, toistuvaa tai tavanomaista toimintaa kuvaa imperfektiivinen aspekti (несовершенный вид). (Kirvesmäki ym. 2021, 141.) Useimmilla venäjän kielen verbeillä on kaksi aspektia (Wade & Gillespie 2011, 268). Samaa leksikaalista merkitystä ilmaisevat perfektiiivisen ja imperfektiivisen aspektin verbit muodostavat aspektiparin (видовая пара) (Kirvesmäki ym. 2021, 141). Esimerkiksi verbin *lukea* imperfektiivinen aspekti on *читать* ja perfektiiivinen aspekti *прочитать* (Alestalo & Hämäläinen 2017, 97). Kieltä opiskelevan voi olla aluksi vaikea ymmärtää, että venäjässä on usein olemassa kaksi verbiä, joilla on melkein sama merkitys. Ongelmia tuottavat myös aspektiparien muodostaminen ja ylipäätään aspektien käyttö. (Mustajoki & Niemensivu 1990, 7.)

On kuitenkin olemassa joitakin verbejä, joilla ei ole aspektiparia. Parittomat verbit voivat olla aspektiltaan joko imperfektiivisiä tai perfektiiivisiä. (Wade & Gillespie 2011, 269–270.) Parittomia verbejä kutsutaan yksiaspektisiksi verbeiksi (одновидовые глаголы). Imperfektiiivisiä yksiaspektisiä verbejä ovat muun muassa kaikki liikettä osoittavat verbit, joissa ei ole etuliitteitä, kuten *носить* (*kantaa, viedä, tuoda*). Imperfektiiivisiä ovat myös jotkin

pysyviä tiloja tai ominaisuuksia kuvaavat verbit, kuten *быть* (*olla*), *весить* (*painaa*) ja *находиться* (*sijaita*). Perfektiivisiä yksiaspektisiä verbejä ovat muun muassa toiminnan alkua tai tietyn ajanjakson kestävää toimintaa osoittavat verbit, kuten *засмеяться* (*purskahtaa nauruun*) ja *поспать* (*nukkua jonkin aikaa*). (Kirvesmäki ym. 2021, 148–149).

On myös olemassa verbejä, joilla on sekä imperfektiivisen että perfektiivisen aspektin merkitykset (Wade & Gillespie 2011, 270). Tällaisia verbejä kutsutaan kaksiaspektisiksi verbeiksi (двувидовые глаголы). Kaksiaspektisten verbien aspektia ei voi tunnistaa pelkästä verbimuodosta, vaan sen merkitys pitää tulkita kontekstin mukaan. (Kirvesmäki ym. 2021, 150.) Esimerkiksi tilanteen kesto on kaksiaspektisten verbien kohdalla ilmaistava jollakin muulla tavalla, kuten adverbeilla, koska pelkän verbimuodon perusteella ei voida tällöin päätellä, onko kyse rajatusta vai rajattomasta toiminnasta. Kaksiaspektisiä verbejä ovat muun muassa *атаковать* (*hyökätä*), *использовать* (*käyttää*) ja *исследовать* (*tutkia*). (Wade & Gillespie 2011, 271).

Aspektipareja on mahdollista muodostaa monella eri tavalla. Ensimmäinen tapa on muodostaa (yleensä) imperfektiivisen aspektin verbistä perfektiivisen aspektin verbi lisäämällä etuliite. Esimerkiksi imperfektiivisen aspektin verbin *писать* (*kirjoittaa*) perfektiivinen aspekti *написать* saadaan lisäämällä etuliite *на-*. (Kirvesmäki ym. 2021, 143). On kuitenkin hyvä huomioda se, että jos verbin leksikaalinen merkitys muuttuu etuliitteen lisäämisen myötä, kyse ei ole enää aspektiparista. Esimerkiksi perfektiivisen aspektin verbillä *почитать* on merkitys *lukea jonkin aikaa*, joten sen leksikaalinen merkitys on eri kuin verbin *читать* (*lukea*). (mts. 145–146.)

Toinen tapa muodostaa aspektipareja on erilaiset suffiksit. Esimerkiksi verbin *бросать* imperfektiivisessä aspektissa *бросать* on suffiksi *-а-* ja perfektiivisessä aspektissa *бросить* on suffiksi *-и-*. Suffiksi voi myös kadota kokonaan, jolloin imperfektiivisen aspektin verbistä saadaan perfektiivisen aspektin verbi. Verbin *уношать* imperfektiivinen aspekti on *забывать* ja perfektiivinen aspekti *забыть*. Joissain tapauksissa aspektiparien kirjoitusasu voi olla täysin sama, mutta imperfektiivisen aspektin verbissä suffiksi on painollinen ja perfektiivisen aspektin verbissä painoton (esimerkiksi *leikata*-verbi: *отрезать* ja *отрёзать*). Kolmas vaihtoehto on se, että aspektiparin verbit eroavat sekä etuliitteen että suffiksin osalta. Esimerkiksi verbin *ошибать* imperfektiivinen aspekti on *покупать* ja perfektiivinen aspekti *купить*. (Kirvesmäki ym. 2021, 143–145).

Neljäs tapa aspektiparien muodostamiseen on verbien vartalonmuutos. Esimerkiksi verbin *valita* imperfektiivinen aspekti on *выбирать* ja perfektiivinen aspekti *выбрать*. Perfektiivisessä aspektissa verbin vartalosta on hävinnyt и-kirjain. Viides tapa on muodostaa aspektiparin verbit täysin eri sanoista. Esimerkiksi verbin *puhua* imperfektiivinen aspekti on *говорить* ja perfektiivinen aspekti *сказать*. (Kirvesmäki ym. 2021, 145).

Venäjän kielessä verbeillä on kolme aikamuotoa: preesens (настоящее время), preteriti (прошедшее время) ja futuuri (будущее время). Imperfektiivisen aspektin verbit esiintyvät kaikissa aikamuodoissa, kun taas perfektiivisen aspektin verbit esiintyvät vain preteritissä ja futuurissa. Perfektiiviset verbit eivät voi esiintyä preesensissä, koska niillä ei kuvata puhehetkellä tapahtuvaa toimintaa. (Bondarko 1971, 50.) Preesens osoittaa muun muassa parhaillaan tapahtuvaa toimintaa, pysyvää asiantilaa, toistuvaa tai tavanomaista toimintaa, kykyä tai taitoa suorittaa toiminta tai puhehetkellä edelleen jatkuvaa menneisyydessä alkanutta toimintaa (Kirvesmäki ym. 2021, 165).

Taulukko 1. Venäjän kielen aspektien ja aikamuotojen suhde sekä imperfektiivisen aspektin verbin *читать* ja perfektiivisen aspektin verbin *прочитать* taivutus persoonan, suvun ja luvun mukaan (Mustajoki & Niemensivu 1990, 14; Kirvesmäki ym. 2021, 157, 166).

	imperfektiivinen aspekti	perfektiivinen aspekti
futuuri	буду читать, будешь читать, будет читать, будем читать, будете читать, будут читать	прочитаю, прочитаешь, прочитает, прочитаем, прочитаете, прочитают
preesens	читаю, читаешь, читает, читаем, читаете, читают	–
preteriti	читал, читала, читало, читали	прочитал, прочитала, прочитало, прочитали

Perfektiiviset verbit taipuvat muodollisesti imperfektiivisten verbien preesensin tavoin, mutta tällöin aikamuoto on futuuri (Kirvesmäki ym. 2021, 172). Otetaan esimerkiksi imperfektiivinen verbi *смотреть* ja perfektiivinen *посмотреть* (*katsoa*). Taipunut muoto *смотрю* on preesensissä, mutta *посмотрю* futuurissa. (mts. 155, 172). Imperfektiiviset verbit taipuvat futuurissa *быть*-verbin avulla (*буду смотреть*). Imperfektiivisten verbien futuuria kutsutaan liittofutuuriksi (будущее сложное) ja perfektiivisten verbien futuuria yksinkertaiseksi

futuuriksi (будущее простое). (mts. 172.) Preteritin taivutus on molemmilla aspekteilla samanlainen (смотрел, посмотрел) (mts. 166).

Preteritissä imperfektiivistä aspektia käytetään muun muassa silloin, kun huomio kiinnittyy toimintaan, sen keston, sen tulokselliseen yrittämiseen tai siihen, onko toiminta ylipäätään edes tapahtunut. Perfektiivistä aspektia käytetään kuvaamaan tapahtunutta, loppuun suoritettua tai ajallisesti rajoitettua tapahtumaa. Sitä käytetään myös osoittamaan toiminnan alkua tai odottamatonta toimintaa. Futuurin merkitykset ovat pitkälti samankaltaisia. Imperfektiivisten verbien liittofutuurissa tuleva tapahtuma voi olla jatkuvaa tai toistuvaa eikä toiminnan suorittaminen loppuun ole olennaista. Perfektiivisten verbien yksinkertaista futuuria käytetään kuvaamaan loppuun suoritettavaa, kerran tapahtuvaa ja alkavaa toimintaa. (Kirvesmäki ym. 2021, 169–173.)

2.2.4 Yhteenveto

Kuten edellisten alalukujen perusteella voidaan päätellä, on suomen ja venäjän kielten välillä havaittavissa merkittäviä eroja aspektin ilmaisemisen keinoissa. Venäjässä aspekti on kieliopillinen kategoria, kun taas suomessa aspektia ei voi systemaattisesti tunnistaa verbimuotojen perusteella. Eroavaisuuksista huolimatta aspektin ilmaisemisen keinoista voidaan kuitenkin vetää myös yhteisiä johtopäätöksiä ja määritellä ainakin osittaisia vastaavuuksia keinoissa.

Aiemmin on todettu, että lauseen aspektilla voidaan kuvailla rajattua tai rajaamatonta tilannetta. Venäjässä ajallisesti rajattua toimintaa ilmaistaan perfektiivisen aspektin verbillä ja rajaamatonta toimintaa imperfektiivisen aspektin verbillä. Suomessa rajatun ja rajallisen toiminnan ero ilmaistaan lauseen objektin sijanvaihtelulla. Yleensä rajatuissa lauseissa käytetään totaaliobjektia ja rajaamattomissa lauseissa partitiiviobjektia. Tilanteen keston voidaan suomessa vaikuttaa myös verbijohtimilla.

Aspektilla voidaan ilmaista myös resultatiivisuutta eli tulokseen johtamista. Venäjässä tulokseen johtanutta toimintaa kuvataan perfektiivisen aspektin verbillä, kun taas suomessa sitä ilmaisee akkusatiivissa oleva kokonaisobjekti. Keskenpäistä toimintaa ilmaistaan venäjässä imperfektiivisen aspektin verbillä, suomessa partitiivissa olevalla osaobjektilla.

Taulukko 2. Suomen ja venäjän kielten aikamuotojen väliset suhteet yksinkertaistettuna (Alestalo & Hämäläinen 2017, 96).

	suomi	venäjä
tuleva aika	preesens	futuuri
nykyaika		preesens
mennyt aika	imperfekti perfekti pluskvamperfekti	preteriti

Verbien aikamuotojen kohdalla on syytä huomata, että aikamuodoissa ei ole täyttä vastaavuutta suomen ja venäjän kielten välillä. Molemmille kielille yhteistä on preesens. Venäjässä on erikseen futuuri, kun taas suomessa tulevaa toimintaa ilmaistaan preesensillä. Menneiden ajan muodoissa on sen sijaan merkittävä ero. Venäjässä on vain yksi menneen ajan aikamuoto, preteriti. Suomessa niitä on kolme: imperfekti, perfekti ja pluskvamperfekti. Näistä etenkin imperfektillä ja perfektillä on merkitystä, kun ilmaistaan tilanteen jatkuvuutta, vaikutusta tai tulosta. Sekä imperfektillä että perfektillä voidaan ilmaista päättynyttä toimintaa, jota venäjässä vastaa perfektiivinen aspekti. Perfektillä voidaan ilmaista myös edelleen jatkuvaa toimintaa, jota venäjässä vastaa imperfektiivinen aspekti. Perfektillä saatetaan joissain tapauksissa ilmaista jopa tulevaa tapahtumaa, jolloin se vastaisi futuuria. Vastaavuudet toimivat kuitenkin hyvin vain helpoimmissa tapauksissa, koska aikamuodoissa ja aspektin ilmaisemisessa ei ole suoraa vastaavuutta kielten välillä. Tämän takia venäjänkielisillä saattaa olla usein vaikeuksia suomen kielen perfektin kanssa (ks. esim. Kachelina 2014). Aspektit ovatkin venäjässä yksi keino ilmaista merkityseroja, jotka suomessa ilmenevät eri aikamuotojen avulla (Alestalo & Hämäläinen 2017, 96).

2.3 Aspekti ja verbin aikamuodot suomen ja venäjän välillä

Tässä luvussa esitellään keskeisimpiä vastaavuuksia ja eroavaisuuksia, joita on havaittavissa suomen ja venäjän kielten aspektin ilmaisemisen ja verbin aikamuotojen välillä. Kaikkia erikoistapauksia ei ole kuitenkaan mahdollista käsitellä erikseen, joten tarkoituksena on vain luoda yleiskuva aiheesta. Lisäksi on syytä huomioida se, että vastaavuudet eivät välttämättä toimi kaikissa mahdollisissa tapauksissa, kuten jo edellisessä luvussa todettiin.

Suomen partitiiviobjekti vastaa usein venäjän imperfektiivistä aspektia, kun taas akkusatiiviobjekti vastaa usein perfektiivistä aspektia (Dahl & Karlsson 1976, 29).

- (1) Он строил дом.
Hän rakensi taloa.
- (2) Он построил дом.
Hän rakensi talon.

Esimerkissä 2 on selvää, että talo on saatu valmiiksi. Esimerkin 1 lauseet eivät sen sijaan vielä kerro sitä, missä vaiheessa talon rakentaminen on. Eroja voidaan havainnollistaa tarkemmin yhdyslauseilla (Dahl & Karlsson 1976, 31).

- (3) Он строил дом, и он сейчас готов.
Hän rakensi taloa, ja nyt se on valmis.
- (4) Он строил дом, но он ещё не готов.
Hän rakensi taloa, mutta se ei ole vielä valmis.
- (5) Он построил дом, и он сейчас готов.
Hän rakensi talon, ja nyt se on valmis.
- (6) *Он построил дом, но он ещё не готов.
*Hän rakensi talon, mutta se ei ole vielä valmis.

Esimerkkien 3 ja 4 imperfektiiviset lauseet virkkeiden alussa ovat mahdollisia molemmissa esimerkeissä, koska imperfektiivinen lause ei kerro päätöstilän saavuttamisesta. Päätöstilana voidaan pitää sitä, että on olemassa talo. Sen sijaan esimerkin 6 virkkeet ovat kieliopillisesti väärin, koska ensimmäinen lause ilmaisee talon olevan valmis, mutta jälkimmäinen lause kieltää päätöstilän saavuttamisen. Esimerkin 6 virkkeet ovat siis todenvastaisia. (Dahl & Karlsson 1976, 30–31.)

Kielteisissä lauseissa objekti on suomessa kuitenkin pääsääntöisesti partitiivissa. Dahlin & Karlssonin (1976, 38–39) mukaan venäjässä kieltolauseen aspekti on useimmiten perfektiivinen, kuten esimerkissä 7. Esimerkki kuvaa kesken jäänyttä toimintaa. Myös imperfektiivinen aspekti on mahdollinen, kuten esimerkissä 8, mutta tällöin verbin kuvaamaa toimintaa ei ole koskaan aloitettukaan.

- (7) Он не построил дом.
Hän ei rakentanut taloa.
- (8) Он не строил дома.
Hän ei rakentanut taloa.

Aina suomen objektin sijanvalinnan ja venäjän verbiaspektin välillä ei kuitenkaan ole edellä esiteltyä vastaavuutta. Esimerkissä 9 venäjänkielisessä lauseessa verbin aspekti on perfektiivinen, mutta suomenkielisessä lauseessa objekti on partitiivissa. Jos suomenkieliseen lauseeseen vaihtaisi tilalle akkusatiiviobjektin, niin lauseen merkitys muuttuisi siten, että susi on ammuttu kuoliaaksi. (Dahl & Karlsson 1976, 34.)

- (9) Он выстрелил в волка.
Hän ampui sutta.

Epämääräisyyttä ilmaistaan suomessa partitiiviobjektilla. Esimerkissä 10 suomenkielisen lauseen objekti on partitiivissa, kun taas venäjänkielisessä lauseessa verbin aspekti on perfektiivinen. (Dingley 2003, 6.)

- (10) Пекка выпил холодной воды.
Pekka joi kylmää vettä.

Suomen ja venäjän preesensit vastaavat luonnollisesti toisiaan erilaisissa merkityksissä. Venäjän futuurin sisältävät ilmaukset käännetään suomeen tavallisesti preesensillä. Liitofutuurissa mahdollisia ovat myös erilaiset verbiliitot, joissa БЫТЬ-verbi on suomennettu jollakin muulla tavalla kuin olla-verbillä, kuten seuraavassa esimerkissä. (Tommola 1986, 42–43.)

- (11) Ты будешь пять лет тянуть.
Sinä saat puurtaa viisi vuotta.

Suomen imperfekti taas vastaa yleensä venäjän preteritiä molemmissa aspekteissa, mutta joissain tapauksissa myös preesensia (vain imperfektiivisessä aspektissa), koska venäjäksi samanaikaisuutta ilmaistaan tavallisesti preesensillä, kun taas suomeksi imperfektillä (Tommola 1986, 44–46).

- (12) Он понял, что Петров не слушает его.
Hän tajusi, ettei Petrov kuunnellut häntä.

Myös suomen perfekti vastaa pääsääntöisesti venäjän preteritiä molemmissa aspekteissa, mutta myös preesens (imperfektiivisessä aspektissa) ja yksinkertainen futuuri (perfektiivisessä aspektissa) ovat mahdollisia. Perfektiivinen yksinkertainen futuuri voidaan kääntää perfektillä, jos toiminnan hetki edeltää tulevaisuudessa olevaa viitekohtaa. Imperfektiivisen aspektin preesens voi tulla kyseeseen silloin, kun kuvaillaan käynnissä olevaa toimintaa. (Tommola 1986, 46–49.)

- (13) Ты пойдешь играть в футбол, когда сделаешь уроки.
Lähdet pelaamaan jalkapalloa vasta, kun olet tehnyt läksyksi.
(14) Две тыщи лет он присутствует среди людей.
Kaksi tuhatta vuotta hän on ollut ihmisten parissa.

Kootaan edellä esitetyt vastaavuudet ja havainnot vielä taulukoiksi. Jälleen kerran on syytä muistuttaa, että nämä vastaavuudet pätevät useissa tapauksissa, mutta eivät läheskään kaikissa mahdollisissa tapauksissa. Siksi näihin vastaavuuksiin tulee suhtautua varauksella, eikä näitä voi tulkita sataprosenttisen varmoina ohjeina. Objektien ja aspektien oikeanlainen käyttö tulee

aina miettiä tarkkaan tapauskohtaisesti. Ensin esitellään suomen objektin sijanvalinnan ja venäjän verbiaspektien välillä vallitsevat osittaiset vastaavuudet.

Taulukko 3. Suomen objektin sijanvalinnan ja venäjän verbiaspektin osittaiset vastaavuudet.

	venäjässä imperfektiivinen aspekti	venäjässä perfektiivinen aspekti
suomessa partitiiviobjekti myönteisessä lauseessa	usein	
suomessa partitiiviobjekti kielteisessä lauseessa		usein
suomessa partitiiviobjekti ilmaisemassa epämääräisyyttä	kumpi tahansa	
suomessa akkusatiiviobjekti		usein
suomessa objektin sijan tietyt leksikaaliset merkitykset	kumpi tahansa	

Pääsääntöisesti siis suomen akkusatiiviobjektia vastaa usein venäjässä perfektiivinen aspekti. Suomen partitiiviobjekti vastaa pääosin venäjässä imperfektiivistä aspektia, mutta poikkeustapauksiakin on. Lisäksi erilaisten tiettyjen leksikaalisten merkitysten kohdalla vastaava aspekti voi olla kumpi tahansa. Seuraavaan taulukkoon on vielä koottu suomen ja venäjän aikamuotojen vastaavuudet suhteessa aspekteihin.

Taulukko 4. Suomen kielen aikamuotojen osittaiset vastaavuudet venäjän kielen aikamuotojen ja aspektien kanssa.

	suomen preesens	suomen imperfekti	suomen perfekti
venäjän yksinkertainen futuuri	perfektiivinen		perfektiivinen
venäjän liittofutuuri	imperfektiivinen		
venäjän preesens	imperfektiivinen	imperfektiivinen	imperfektiivinen
venäjän preteriti		imperfektiivinen/perfektiivinen	imperfektiivinen/perfektiivinen

Kuten jo aiemmin on todettu, niin venäjän yksinkertaisella futuurilla on ainoastaan perfektiivinen aspekti ja liittofutuurilla ainoastaan imperfektiivinen aspekti. Preesensissä vain imperfektiivinen aspekti on mahdollinen, kun taas preteritissä molemmat aspektit ovat mahdollisia.

3 VERBIEN JA KÄÄNNÖSTEN TUTKIMUS

3.1 Tutkimusmetodi ja -materiaali

3.1.1 Korpus ja rinnakkaiskorpus

Korpus on kokoelma tekstejä, jotka on kerätty kielitieteellistä tutkimuskäyttöä varten. Tekstit valitaan yleensä erilaisten ennakkoon määriteltyjen kriteerien perusteella ja niiden tarkoitus on edustaa kieltä mahdollisimman laajasti. (Sinclair 2004, 23). Tekstejä on kuitenkin saatavilla käytännössä rajattomasti, joten korpuksset eivät välttämättä edusta kaikkia kielen ilmiöitä. Lisäksi kieli muuttuu koko ajan. Siksi korpusta pitääkin käsitellä ikään kuin näytteenä kohdekielen aineistosta. (mts. 6). Tutkijat voivat koota itse omia korpuksiaan. Vaihtoehtoisesti on myös mahdollista käyttää erilaisia valmiita jo olemassa olevia korpuksia. (Laippala & Palander-Collin 2020, 461.)

Käännöstieteellisessä tutkimuksessa käytettäviä korpuksia ovat yleensä rinnakkaiskorpuksset sekä verrannolliset korpuksset. Rinnakkaiskorpuksset ovat monikielisiä korpuksia, joissa on tekstejä jollakin lähdekielellä sekä näiden tekstien käännöksiä yhdellä tai useammalla kielellä. Rinnakkaiskorpuksset voivat siis olla kaksi- tai monikielisiä. Kaksikielinen rinnakkaiskorpus voi sisältää tekstejä käännettynä molemmissa kielisuunnissa tai sitten korpus voi olla yksisuuntainen. Verrannolliset korpuksset sen sijaan sisältävät tekstejä, jotka ovat esimerkiksi tekstilajiltaan tai aihepiiriltään sopivia vertailtavaksi, mutta eivät ole toistensa käännöksiä. (Tiittula & Kolehmainen 2020, 955–956.) Rinnakkaiskorpuksia kutsutaan myös käännöskorpuksiksi (Marco 2019, 40). Tässä tutkimuksessa käytetään kaksikielistä rinnakkaiskorpusta, joka on yksisuuntainen, sillä kaikki korpuksen lähdetekstit ovat venäjänkielisiä, ja ne on käännetty suomeksi.

On olemassa paljon erilaisia valmiita korpusaineistoja. Esimerkiksi yleiskorpuksset kuvaavat kieltä yleisesti ja ne kootaan siten, että ne edustaisivat kieltä mahdollisimman laajasti. Muun muassa eri maiden kansalliskorpuksset ovat tällaisia. Lisäksi on olemassa erilaisia korpuksia, jotka keskittyvät johonkin tiettyyn aihealueeseen, kuten tiettyinä aikakautena julkaistuihin teksteihin tai puhuttuun kieleen. (Laippala & Palander-Collin 2020, 465–466.) Jos kuitenkin tutkimukselle sopivaa valmista aineistoa ei ole, tutkija voi koota itse oman korpuksensa. Korpuksen luomisessa on kuitenkin huomioitava monia eri seikkoja, kuten korpuksen

edustavuus, koko ja käyttötarkoitus. (mts. 469.) Nykyisten tietokoneiden ja ohjelmistojen myötä pienten korpusten kokoaminen on mahdollista lyhyessäkään ajassa (Mikhailov & Cooper 2016, 19). Tässä tutkimuksessa käytetään itse koottua rinnakkaiskorpusta, joka sisältää sekä venäjän- että suomenkielisiä tekstejä.

Korpuksen koko pitäisi päättää ensimmäisenä. Koko riippuu projektiin käytettävissä olevasta ajasta. (Mikhailov & Cooper 2016, 22.) Esimerkiksi sanastotyössä suuri korpus on hyödyksi, mutta tarkempaan tutkimusaiheisiin riittää usein pienempikin määrä aineistoa. Pohdittavaa voikin aiheuttaa se, onko suurempi korpus aina pienempää parempi tai se, onko korpuksen koko riittävä. Valtavan suurta aineistoa voi olla hankala käsitellä, kun taas pienemmän korpuksen käsittely voi onnistua helposti valmiilla korpusohjelmilla. (Laippala & Palander-Collin 2020, 470–471.) Tässä tutkimuksessa on syytä käyttää korpusta, joka ei ole kooltaan liian suuri, koska tämän tutkielman laajuus ei riitä liian suuren aineiston käsittelemiseen. Tutkittavia verbejä voidaan valita vain rajallinen määrä, joten siksi pienempikin aineisto riittää. Kymmenien tuhansien sanojen kokoiset korpuksat kummallakin kielellä ovat riittävän laajat tämän tutkielman tarpeita varten. Tällöin yleisimmät verbit esiintyvät teksteissä jo useita kertoja, jolloin tutkittavia tapauksiakin on saatavilla tarpeeksi.

Rinnakkaiskorpuksen kokoaminen voi viedä paljon aikaa, sillä esimerkiksi tekstien kohdistaminen voi olla hidasta työtä (Mikhailov & Cooper 2016, 22). Kohdistamisella tarkoitetaan sitä, että käännöksen segmentit on yhdistetty vastaamaan lähdetekstin osia. Kohdistaminen on tärkeää, sillä ilman sitä rinnakkaiskorpusta ei oikeastaan voisi edes kunnolla käyttää, koska tällöin käännösvastineiden etsimisestä tulisi erittäin työlästä, eikä korpus enää vastaisi täydellisesti tarkoitustaan. Kohdistamiseen voidaan käyttää siihen suunniteltuja tietokoneohjelmia, jotka etsivät teksteistä samankaltaisuuksia ja yhdistävät vastaavat kohdat toisiinsa automaattisesti. Ohjelma jakaa tekstien kohdat segmenteiksi. (mts. 29–30.) Tässä tutkimuksessa venäjän- ja suomenkielisten tekstien kohdistamiseen on käytetty ohjelmaa nimeltä LF Aligner. Tekstien kohdistus sujui tässä projektissa yleensä ongelmitta. Automaattinen kohdistus toimi melko hyvin, mutta lähes aina kohdistuksessa oli joitakin kohtia, jotka piti korjata manuaalisesti. Ohjelmassa voi tallentaa kohdistuksen tuloksen suoraan tnx-tiedostoksi (Translation Memory eXchange), jonka voi viedä sellaisenaan korpusohjelmaan. Tiedostoja voi nimensä mukaisesti käyttää myös käännösmuistiohjelmissa, joten kohdistamalla tekstejä voidaan myös luoda käännösmuisteja jo olemassa olevista teksteistä ja niiden käännöksistä.

3.1.2 Sketch Engine

Tässä tutkimuksessa käytettävä korpusohjelma on Sketch Engine. Ohjelman käyttöliittymää voi käyttää joko valmiiden korpusien tutkimiseen tai sitten ohjelman avulla voi myös koota omia korpuksia (Kilgarriff ym. 2014, 7). Kuten aiemmin on todettu, tässä tutkimuksessa on koottu oma venäjän- ja suomenkielisiä tekstejä sisältävä rinnakkaiskorpus juuri erityisesti tätä tutkimusta varten. Korpuksen kokoamisen lisäksi ohjelmaa käytetään myös korpusaineiston analysointiin ja käsittelyyn. Korpusohjelmassa on mahdollista käyttää useita erilaisia toimintoja, joita voidaan hyödyntää aineistoa analysoitaessa.

Aivan aluksi tulee valita tutkittavat verbit korpusaineistosta. Verbit valitaan venäjänkielisestä aineistosta, jotta voidaan tutkia niiden suomenkielisiä käännösvastineita. On syytä keskittyä niihin verbeihin, jotka ovat teksteissä yleisiä, koska tällöin materiaalia ja täten esimerkkejä on olemassa enemmän. Sketch Engine -ohjelmassa on Wordlist-toiminto, jonka avulla voidaan saada korpuksessa esiintyvien sanojen frekvenssilistat. Tällöin on mahdollista saada tietoa siitä, kuinka monta kertaa eri sanat esiintyvät korpuksen teksteissä. Frekvenssilistoja on yleensä kahdenlaisia: lemmoihiin sekä saneisiin perustuvia listoja (Mikhailov & Cooper 2016, 52). Tässä tutkimuksessa on järkevää käyttää lemmoihiin perustuvaa listaa, koska tutkimuskohdetta ei ole rajattu esimerkiksi verbien taivutusmuotojen perusteella.

Word Sketch -toiminnolla voidaan tarkastella sitä, miten jokin sana käyttäytyy kieliopillisesti. Toiminnon avulla voidaan tutkia esimerkiksi sitä, millaisia objekteja jonkin sanan kanssa käytetään. (Kilgarriff ym. 2014, 9.) Tämä toiminto on tämän tutkimuksen kannalta hyödyllinen, koska kuten on aiemmin todettu, aspektin ilmaisemisen yksi tärkeä keino suomen kielessä on objektin sijanvalinta. Asiaa voidaan tarkastella tarkemmin Concordance-toiminnolla (konkordanssi), jonka avulla voidaan katsoa, millaisissa konteksteissa jokin sana esiintyy (mts. 10). Kontekstin avulla on mahdollista saada enemmän tietoa esimerkiksi siitä, miksi lauseessa on käytetty tiettyä aspektia. Tässä tutkimuksessa käytetään rinnakkaiskorpusta, joten siksi tutkimuksessa käytetään erityisesti Parallel Concordance -toimintoa (rinnakkaiskonkordanssi). Tällöin on mahdollista tehdä haku toisella korpuksen kielistä ja tämän jälkeen nähdä konkordanssit molemmilla kielillä vierekkäin (mts. 23). Näin ollen käännösvastineiden tutkiminen on helppoa, kun lähde- ja kohdekielistä aineistoa voi kätevästi vertailla samassa näkymässä. Jotta toiminnosta olisi mahdollisimman paljon hyötyä, on tärkeää, että tekstien kohdistus on tehty huolellisesti.

3.1.3 Tutkimusmateriaali

Tämän tutkimuksen tutkimusmateriaalina on rajattu valikoima henkilöartikkeleita venäjänkielisestä Wikipediasta sekä näiden artikkeleiden suomenkieliset konekäännökset. Artikkelit on valittu Wikipediasta hyödyntäen Satunnainen artikkeli -toimintoa, kuitenkin siten, että artikkelit olisivat pituudeltaan riittäviä, jotta niissä olisi mahdollisimman paljon sopivaa materiaalia tutkimusta varten. Artikkeleista on poistettu niissä ollutta tutkimuksen kannalta epärelevanttia tietoa, kuten muun muassa kuvia, kuvatekstejä, luetteloita, sisällysluetteloita ja taulukoita. Artikkeleiden suomenkieliset konekäännökset on tehty Google-kääntäjällä. Artikkelit on kerätty kahtena päivänä tammikuussa 2023. Myös konekäännökset on tuotettu samana ajankohtana. Luettelo tutkimusmateriaaliin sisältyvistä artikkeleista on tämän tutkielman liitteessä 1. Sen lisäksi liitteessä 2 on esimerkki yhdestä tutkimusmateriaaliin sisältyvästä artikkelista sekä kyseisestä artikkelista tehdystä konekäännöksestä. Artikkelit ja konekäännökset on koottu rinnakkaiskorpuksiksi korpusohjelma Sketch Enginen avulla. Tutkimuskäyttöön kerättyä materiaalia myös analysoidaan ja käsitellään kyseistä korpusohjelmaa käyttäen. Kaiken kaikkiaan aineistossa on yhteensä 33 venäjänkielistä Wikipedia-artikkelia.

Korpuksessa esiintyviin sanoihin viitataan usein termeillä sane ja lemma. Sane tarkoittaa sananmuodon esiintymää tekstissä. Sane voi olla missä tahansa taivutusmuodossa (esimerkiksi *talossa* ja *pihalla*). Lemma sen sijaan tarkoittaa sanan niin sanottua sitaattimuotoa eli sanakirjamuotoa (esimerkiksi *talo* ja *piha*). (Kyröläinen & Laippala 2020, 500.) Alkuperäiset venäjänkieliset artikkelit sisältävän korpuksen koko on 54 519 sanetta ja suomenkieliset konekäännökset sisältävän korpuksen koko on 45 298 sanetta. Lähes kymmenen tuhannen saneen ero saneiden määrässä selittynee kielten erilaisilla rakenteilla. Venäjässä prepositioita käytetään enemmän, kun taas suomessa käytetään enemmän pidempiä sanoja sekä jälkiliitteitä, jotka kirjoitetaan sanaan kiinni (Mikhailov 2003, 166–167). Kyse voi olla myös siitä, että neuroverkkokääntimet saattavat joskus ikävästi jättää joitakin sanoja tai kohtia kokonaan kääntämättä.

3.2 Verbit ja aspektiparit

3.2.1 Verbit ja aspektiparit teksteissä: yleisiä tendenssejä

Aloitetaan korpuksessa esiintyvien verbien tutkiminen tarkastelemalla korpusohjelmassa frekvenssilistoja. Tarkastellaan ensin venäjänkielistä korpusta ja verbien esiintymistä yleisellä tasolla sekä etsitään verbien aspektipareja.

Taulukko 5. Venäjänkielisen korpuksen verbien frekvenssilistan 20 yleisintä verbiä esiintymiskertoineen.

1. быть 689	11. родиться 45
2. стать 209	12. принимать 44
3. работать 88	13. иметь 41
4. получить 80	14. сделать 40
5. выйти 80	15. выходить 39
6. выступать 57	16. участвовать 39
7. играть 52	17. написать 39
8. получать 51	18. мочь 39
9. становиться 49	19. смочь 36
10. сыграть 46	20. являться 35

Ylivoimaisesti yleisin venäjänkielisessä korpuksessa esiintyvä verbi on *быть* (689 kertaa). Kyseistä verbiä ei yleensä käytetä preesensissä (Kirvesmäki ym. 2021, 257). Konkordanssia tarkastelemalla huomataankin, että lähes kaikki verbin esiintymät ovat joko preteriti- tai futuurimuotoja. Verbi on imperfektiivinen ja sillä ei ole perfektiivistä aspektiparia (mts. 149). Toiseksi eniten venäjänkielisessä korpuksessa esiintyvä verbi on *стать* (209 kertaa). Se on perfektiivinen verbi, jolla merkityksen mukaan on aspektipari tai sitten ei. Merkityksessä *alkaa* verbillä ei ole aspektiparia. Sen sijaan merkityksessä *tulla joksikin* sillä on imperfektiivinen aspektipari *становиться*. (mts. 150). Näistä verbeistä *становиться* on 49 esiintymällään frekvenssilistan sijalla yhdeksän.

Kolmannella sijalla korpuksen frekvenssilistalla on verbi *работать* (88 kertaa). Verbi on imperfektiivinen ja sen perfektiivinen aspektipari on *поработать*, joka ei kuitenkaan esiinny korpuksessa kertaakaan. Neljännellä sijalla listalla on perfektiivinen verbi *получить* (80 kertaa), jonka imperfektiivinen aspektipari *получать* on listalla kahdeksantena (51 kertaa). Kärkikymmenikössä on edustettuna myös yksi aspektipari lisää, sillä seitsemännellä sijalla on imperfektiivinen verbi *играть* (52 kertaa) ja kymmenennellä sijalla on perfektiivinen verbi *сыграть* (46 kertaa). Viidennellä sijalla listalla on liikeverbi *выйти* (80 kertaa). Verbi on perfektiivinen ja sen imperfektiivinen aspektipari on *выходить* (39 kertaa). Kuudennella sijalla listalla on imperfektiivinen verbi *выступать* (57 kertaa), jonka perfektiivinen aspektipari on *выступить* (23 kertaa).

Aineiston kaikkein yleisimpien verbien tutkimisen jälkeen jatkoin frekvenssilistan huolellisempaa tarkastelua aina sijaan 50 asti. Kaiken kaikkiaan näiden listalla esiintyvien verbien joukosta löytyi yhteensä viisitoista aspektiparia. Tässä tutkimuksessa on tarkoitus vertailla sitä, miten aspektien välittämä merkitys ilmenee suomenkielisessä käännöksessä. Siksi on tarkoituksenmukaista valita tutkittavat verbit siten, että niillä on olemassa aspektipari. Aspektiparittomat sekä kaksiaspektiset verbit on siis syytä pudottaa pois tutkimuksesta tässä vaiheessa selvyuden vuoksi, koska tällaisten verbien vertailu saattaisi olla liian hankalaa tämän tutkielman laajuus huomioon ottaen. Kaksiaspektisten verbien kohdalla aspektien erojen esiintuominen saattaisi myös olla hankalampaa. Listalla oli myös muutama verbi, joiden imperfektiiviset tai perfektiiviset aspektiparit eivät olleet viidenkymmenen yleisimmän verbin joukossa ja jotka olivat siten harvinaisempia koko aineistossa. Päätin olla valitsematta näitä aspektipareja mukaan, koska on mielestäni tärkeää, että korpuksessa olisi riittävästi verbien molempien aspektien esiintymiä, jotta tutkittavaa materiaaliakin olisi niin ikään todennäköisemmin riittävästi.

Taulukko 6. Aspektiparit, joiden verbit olivat venäjänkielisen korpuksen verbien frekvenssilistalla 50 yleisimmän verbin joukossa sekä näiden verbien esiintymiskerrat korpuksessa.

Imperfektiivinen aspekti	Perfektiivinen aspekti
становиться 49	стать 209
получать 51	получить 80
выходить 39	выйти 80
выступать 57	выступить 23
играть 52	сыграть 46
принимать 44	принять 26
писать 27	написать 39
мочь 39	смочь 36
занимать 32	занять 26
начинать 30	начать 26
выигрывать 29	выиграть 28
выпускать 26	выпустить 28
проходить 24	пройти 22
проводить 17	провести 32
позволять 13	позволить 21

Kaikkia viittätoista aspektiparia ei kuitenkaan kannata valita mukaan tutkielman varsinaisiin tapaustutkimuksiin, koska muutoin tutkittavien konkordanssien määrä olisi liian suuri tämän tutkielman laajuus huomioon ottaen. Kun käänösvastineet verbeille ovat selvillä, voidaan tarkemmin pohtia sitä, mitkä aspektiparit kannattaa valita mukaan vertailtaviksi sekä varsinaisen tutkimuksen kohteiksi.

Tarkastellaan seuraavaksi vielä suomenkielisen korpuksen verbien frekvenssilistaa. Vaikka tutkittavia verbejä ei valitakaan suomenkielisen korpuksen perusteella, niin suomenkielisten verbien frekvenssilistaa katsomalla voidaan saada jo mahdollisesti alustavaa tietoa siitä,

millaisia käännösvastineita venäjänkielisen korpuksen yleisimmille verbeille suomenkielisestä korpuksesta voidaan löytää.

Taulukko 7. Suomenkielisen korpuksen verbien frekvenssilistan 20 yleisintä verbiä esiintymiskertoineen.

1. olla 929	11. osallistua 79
2. ei 256	12. päästä 73
3. tulla 221	13. alkaa 62
4. saada 153	14. pitää 61
5. julkaista 144	15. kirjoittaa 60
6. voittaa 127	16. mennä 55
7. tehdä 106	17. lukea 52
8. pelata 89	18. muuttaa 52
9. lähteä 87	19. antaa 52
10. työskennellä 80	20. aloittaa 51

Myös suomenkielisten verbien frekvenssilistalla ensimmäisenä on *olla*-verbi. Listalla toisena on kieltosana *ei*, joka suomen kielessä lasketaan verbiksi. Silmämääräisesti tarkastellen näyttäisi kuitenkin siltä, että suomenkielisten verbien listalla on monia merkitykseltään venäjänkielisten verbien listan sisältöä vastaavia verbejä. Venäjänkielisten verbien suomenkielisiä käännösvastineita korpuksessa tarkastellaan tarkemmin seuraavassa alaluvussa. Käännösvastineita tarkastelemalla voin myös mahdollisesti saada vahvistuksen alustavalle havainnolleni, jonka esitin edellä.

3.2.2 Venäjänkielisten verbien käännösvastineet

Venäjänkielisten verbien suomenkielisten käännösvastineiden etsimiseen ja tutkimiseen käytän Sketch Engine -ohjelman toimintoa Parallel Concordance (rinnakkaiskonkordanssi). Tarkoitukseni on etsiä suomenkieliset käännösvastineet edellisessä alaluvussa esitellyille venäjänkielisille verbeille, jotka kuuluvat valittujen viidentoista aspektiparin joukkoon. Toiminnon avulla voidaan tarkastella samanaikaisesti verbien esiintymistä alkuperäisissä venäjänkielisissä teksteissä sekä tekstien suomenkielisissä konekäännöksissä. Tällöin on

mahdollista nähdä suoraan, miten konekäännin on kääntänyt teksteistä näitä tiettyjä verbejä sisältävät kohdat. Käännösvastineita tutkittaessa on hyvä pitää mielessä se, että konekäännin on altis virheille, joten kaikki koneen tuottamat vastineet eivät välttämättä ole oikeita. Kävin läpi yksitellen kaikki aspektiparien sisältämät 30 verbiä ja keräsin korpusaineistosta kaikki löytämäni erilaiset käännösvastineet, joita konekäännin oli tuottamissaan käännöksissään käyttänyt. Kyseiset suomenkieliset käännösvastineet on esitelty seuraavassa taulukossa.

Taulukko 8. Venäjänkieliset verbit ja niiden suomenkieliset käännösvastineet rinnakkaiskorpuksessa.

Aspektipari	Käännösvastineet
становиться / стать	tulla (joksikin), olla, käydä, tehdä
получать / получить	saada, vastaanottaa, voittaa, levitä, loukkaantua, palkita, ansaita, saavuttaa
выходить / выйти	julkaista, menettää, päästä, saavuttaa, erota, nousta, astua, ilmestyä, jäädä, mennä (naimisiin), edetä
выступать / выступить	esiintyä, puhua, toimia, kilpailla, vastustaa, kannattaa, pelata
играть / сыграть	soittaa, pelata, näytellä
принимать / принять	hyväksyä, vastaanottaa, osallistua, ottaa, päättää, päästä, saada, palkata, houkutella, omaksua, esiintyä, vannoa
писать / написать	kirjoittaa, säveltää
мочь / смочь	voida, osata, saada, pystyä, päästä
занимать / занять	olla, viedä, lainata, toimia, sijoittua, saavuttaa, voittaa, nousta
начинать / начать	aloittaa, alkaa
выигрывать / выиграть	voittaa
выпускать / выпустить	ilmoittaa, julkaista, päästää, vapauttaa
проходить / пройти	läpikäydä, tapahtua, pitää, palvella, päästä, mennä, kulkea, työstää, käydä, äänittää, läpäistä, suorittaa, sujua, puolustaa
проводить / провести	antaa, pitää, tehdä, viettää, toteuttaa, suorittaa, pelata, johtaa
позволять / позволить	antaa mahdollisuus, mahdollistaa, voida

Kuten taulukosta voidaan havaita, osalle verbeistä on olemassa varsin selkeät ja yksiselitteiset käännösvastineet, kun taas joillekin verbeille erilaisia käännösvastineita löytyi monenlaisia

kontekstista riippuen. Kaikkein selkein tapaus oli aspektipari *выигрывать / выиграть*, joka esiintyi rinnakkaiskorpuksessa ainoastaan merkityksessä *voittaa*. Selkeä oli myös aspektiparin *начинать / начать* merkitys *aloittaa* tai *alkaa*. Aspektiparilla *играть / сыграть* oli yleisesti tunnetut merkitykset *soittaa*, *pelata* ja *näytellä*. Verbit *писать* ja *написать* esiintyivät tavallisen merkityksen *kirjoittaa* lisäksi erityisesti musiikkiin liittyvissä henkilöartikkeleissa merkityksessä *säveltää*.

Verbeille *становиться* ja *стать* yleisimmät merkitykset korpusaineistossa olivat *tulla joksikin* ja *olla*. Näitä verbejä voidaankin käyttää kopulaverbinä *быть*-verbin (*olla*) tavoin (Kirvesmäki ym. 2021, 266). Verbeillä *мочь* ja *сметь* selvästi yleisin merkitys oli *voida*. Lisäksi näillä verbeillä yleensä ilmaistiin teksteissä jonkinlaista mahdollisuutta tehdä jotakin. Usein nämä verbit esiintyvät yhdessä toisen verbin infinitiivimuodon kanssa (mts. 194). Mahdollisuuden ilmaiseminen liittyi aineistossa vahvasti myös aspektipariin *позволять / позволить*.

Aspektipari *принимать / принять* osoittautui verbejä tarkasteltaessa varsin monimerkityksiseksi. Suomenkieliset käännösvastineet riippuivat hyvin paljon siitä, minkä substantiivin kanssa kyseiset verbit esiintyivät alkuperäisissä venäjänkielisisä teksteissä. Ilmiössä saattaa olla kyse siitä, että venäjän kielessä verbin ja substantiivin sisältävät rakenteet ovat tyypillisempiä, kun taas suomen kielessä tämänkaltainen niin sanottu substantiivitus on epätavallisempi. Esimerkki toisesta samantyyppisestä aspektiparista oli *занимать / занять*. Aspektiparille *получать / получить* yleisin merkitys oli *saada*, mutta korpukselta löytyi paljon muitakin vastineita, jotka olivat vahvasti riippuvaisia kontekstista, jossa kyseiset verbit esiintyivät. Verbien *выступать* ja *выступить* yleisin käännösvastine korpusaineistossa oli *esiintyä*. Sen sijaan verbien *выпускать* ja *выпустить* yleisimmät käännösvastineet olivat *julkaista* ja *päästää*.

Kolme taulukossa olevaa aspektiparia koostuu verbeistä, jotka luokitellaan venäjässä kieliopillisesti liikeverbeiksi. Liikeverbeillä tarkoitetaan tiettyä liikettä ilmaisevia verbejä, jotka poikkeavat kieliopillisesti muista venäjän kielen verbeistä (Kirvesmäki ym. 2021, 240). Liikeverbejä käytetään venäjässä usein myös kuvaannollisessa merkityksessä (mts. 252). Tällöin liikeverbin suomenkielisenä käännösvastineena voi hyvinkin olla verbi, jota ei suomessa yleensä yhdistetä liikkeen kuvaamiseen. Suurimmassa osassa rinnakkaiskorpuksen tapauksista liikeverbit eivät näin ollen varsinaisesti kuvanneetkaan konkreettista liikettä, vaan jotakin muuta toimintaa. Esimerkiksi aspektiparin *выходить / выйти* verbit esiintyivät usein

merkityksessä *julkaista* puhuttaessa musiikin julkaisemisesta. Kaikkien kolmen aspektiparin verbeille löytyi aineistosta useita erilaisia käännösvastineita, mikä kertoo siitä, kuinka monipuolisesti liikeverbejä käytetään venäjässä erilaisten tapahtumien ja toimintojen kuvaamiseen. Aspektiparien *проходить / пройти* sekä *проводить / провести* verbit esiintyivät korpuksessa myös monenlaisissa merkityksissä. Liikeverbejä tarkasteltaessa konteksti on siis syytä ottaa tarkasti huomioon, koska merkitykset voivat vaihdella huomattavasti sen mukaan.

Vertaillaan seuraavaksi suomenkielisiä käännösvastineita suomenkielisen korpuksen frekvenssilistalla korkeimmilla sijoilla oleviin verbeihin. Näin ollen voidaan saada vahvistus aiemmalle havainnolle siitä, onko venäjänkielisestä korpusaineistosta valittujen yleisten verbien sekä suomenkielisen korpusaineiston yleisimpien verbien merkitysten välillä olemassa jonkinlainen yhteys.

Taulukko 9. Suomenkieliset verbien käännösvastineet, jotka ovat myös suomenkielisen korpuksen verbien frekvenssilistan 20 yleisimmän verbin joukossa.

olla	päästä
tulla	alkaa
saada	pitää
julkaista	kirjoittaa
voittaa	mennä
tehdä	antaa
pelata	aloittaa
osallistua	

Kuten taulukosta voidaan havaita, on suomenkielisen korpuksen frekvenssilistan kahdenkymmenen yleisimmän verbin joukossa peräti viisitoista sellaista verbiä, jotka löytyivät myös valittujen aspektiparien verbien suomenkielisten käännösvastineiden joukosta. Tämä vahvistaa aiempaa silmämääräistä havaintoa, jonka mukaan sekä yleisimpien suomen- että venäjänkielisten verbien joukossa on runsaasti merkitykseltään samankaltaisia verbejä. Tämä on luonnollisestikin hyvin odotettua, koska käännösten on tarkoitus vastata merkitykseltään

alkuperäisiä tekstejä. Sitä ei kuitenkaan tässä vaiheessa voida vielä tietää varmaksi, onko konekäännin tehnyt työnsä oikein. Tämä selviää tutkimalla aineistoa ja aineistoesimerkkejä tarkemmin.

3.2.3 Tutkittavien verbien ja aspektiparien valinta

Tutkimuksessa on nyt päästy siihen vaiheeseen, että on aika valita aspektiparit ja verbit, jotka etenevät varsinaisiin tapaustutkimuksiin. Edellisessä alaluvussa käsiteltiin sitä, millaisia verbejä ja aspektipareja venäjänkielisestä korpusaineistosta on mahdollista löytää. Aineistosta poimittiin kaikkiaan viisitoista aspektiparia, joiden sisältämien verbien suomenkielisiä käännösvastineita etsittiin rinnakkaiskorpuksista. Jo aiemmin tutkimuksesta rajattiin pois muun muassa sellaisia verbejä, jotka ovat aspektiparittomia tai kaksiaspektisiä. Jotta tutkittavien konkordanssien määrä olisi kohtuullinen, on syytä rajata tutkittavien aspektiparien ja verbien määrää entisestään. On hyödyllistä valita tutkielman tapaustutkimuksia varten sellaisia aspektipareja, joiden verbien merkitykset ovat selvät ja joiden avulla saadaan selkeästi tuotua eri verbien välisiä eroja esille.

Viidentoista aspektiparin joukossa on kolme aspektiparia, joiden verbit ovat venäjässä liikeverbejä. Näitä ovat aspektiparit *выходить / выйти, проходить / пройти* sekä *проводить / провести*. Kuten jo aiemmin todettiin, liikeverbit eroavat tietyiltä ominaisuuksiltaan kieliopillisesti venäjän kielen muista verbeistä. Niitä voidaan myös liikkeen kuvaamisen lisäksi käyttää kuvaannollisissa merkityksissä, mikä saattaa tuoda esiin hankaluuksia näitä verbejä tutkittaessa. Esimerkiksi verbeille *проходить* ja *пройти* löytyi korpusaineistosta toistakymmentä erilaista käännösvastinetta. Samoin myös kahden muun aspektiparin liikeverbeille löytyi useita eri merkityksiä. Koska näiden liikeverbien merkitykset saattavat vaihdella hyvin merkittävästi kontekstin mukaan, on mielestäni parempi, että nämä kyseiset verbit ja aspektiparit jätetään tällä kertaa tapaustutkimusten ulkopuolelle.

Tarkasteltujen aspektiparien joukossa on myös muita pareja, joiden sisältämät verbit osoittautuivat niin ikään monimerkityksisiksi. Muun muassa verbeille *принимать* ja *принять* löytyi useita käännösvastineita niitä tarkasteltaessa. Tällaisia aspektipareja olivat myös esimerkiksi *занимать / занять, получать / получить* sekä *выступать / выступить*. On tarkoituksenmukaisempaa keskittyä verbeihin, joille on korpusaineistossa olemassa selkeämmät käännösvastineet.

Aspektipari *становиться / стать* on muutoin varsin selkeä merkityksiltään, mutta näiden verbien kohdalla mahdollisena ongelmana on verbin *стать* sekoittuminen vastaavanlaiseseen aspektiparittomaan verbiin, jolla on eri merkitys. Selvyyden vuoksi on parempi jättää tämä aspektipari pois tutkimuksesta, koska korpusohjelma ei välttämättä osaa erottaa näitä kahta erimerkityksistä verbiä aineistosta. Verbit *мочь* ja *смочь* toimivat usein toisen verbin yhteydessä apuverbin tavoin. On parempi tutkia verbejä, jotka esiintyvät pääosin itsenäisesti. Päätän jättää tutkimuksesta pois myös toisen mahdollisuutta ilmaisevan verbiparin *позволять / позволить*.

Kymmenen aspektiparia on nyt karsittu pois joukosta, joten jäljellä on vielä viisi aspektiparia. Nämä aspektiparit osoittautuivatkin kaikkein selvimmiksi tapauksiksi, koska niiden sisältämille verbeille löytyi huomattavasti rajatumpi määrä käänkösvastineita. Esimerkiksi verbeille *выигрывать* ja *выиграть* ainoa aineistosta löytynyt käänkösvastine oli *voittaa*. Myös aspektiparit *играть / сыграть*, *писать / написать*, *начинать / начать* sekä *выпускать / выпустить* olivat selkeämmästä päästä tapauksia, koska niillä oli vain muutamia erilaisia käänkösvastineita. Näin ollen oli lopulta oikeastaan varsin selvää, että juuri nämä aspektiparit ja verbit ovat mukana lopullisissa tapaustutkimuksissa.

Taulukko 10. Tapaustutkimukseen valitut aspektiparit sekä niiden sisältämien verbien korpuksessa esiintyvät käänkösvastineet.

Aspektipari	Käänkösvastineet korpuksessa
выигрывать / выиграть	voittaa
выпускать / выпустить	ilmoittaa, julkaista, päästää, vapauttaa
играть / сыграть	soittaa, pelata, näyttää
начинать / начать	aloittaa, alkaa
писать / написать	kirjoittaa, säveltää

Tutkitaan seuraavaksi aspektiparien sisältämien verbien esiintymiskertojen määrää venäjänkielisessä korpusaineistossa.

Taulukko 11. Aspektiparien sisältämien verbien esiintymiskertojen määrät venäjänkielisessä korpuksessa.

Imperfektiivinen aspekti	lkm.	Perfektiivinen aspekti	lkm.	Yhteensä
выигрывать	29	выиграть	28	57
выпускать	26	выпустить	28	54
играть	52	сыграть	46	98
начинать	30	начать	26	56
писать	27	написать	39	66
	164		167	331

Kuten taulukosta voidaan nähdä, näistä aspektipareista eniten korpuksessa esiintyy *играть / сыграть*. Toiseksi eniten esiintymiä on aspektiparilla *писать / написать*. Loput kolme aspektiparia ovat varsin tasaisissa lukemissa. Mielenkiintoista on huomata myös se, että jokaisen aspektiparin kohdalla imperfektiivisen aspektin ja perfektiivisen aspektin verbien lukumäärät ovat hyvin lähellä toisiaan. Imperfektiivisen aspektin verbit esiintyvät aineistossa yhteensä 164 kertaa ja perfektiivisen aspektin verbit niukasti useammin, 167 kertaa. Määrien puolesta aspektiparien valinta osoittautui mielestäni hyväksi, koska molemmat aspektit ovat aineistossa edustettuina tasaisesti. Verbien esiintymiä on myös ylipäätään kohtuullinen määrä tarkasteltavaksi. Tutkittavia tapauksia on siis kaikkiaan yhteensä 331 kappaletta.

3.3 Tapaustutkimukset

Kaikkien viiden aspektiparin osalta on tarkasteltu korpuksen kaikki segmentit, joissa kyseisten aspektiparien sisältämät verbit esiintyvät. Segmentit on luokiteltu neljään ryhmään aspektin merkityksen välittymisen perusteella. Ensimmäisen ryhmän muodostavat segmentit, joissa aspektin merkitys välittyy hyvin käännökseen. Toisen ryhmän muodostavat segmentit, joissa aspektin merkitys välittyy vain osittain. Kolmannen ryhmän muodostavat segmentit, joissa aspektin merkitys ei välity ollenkaan. Neljännen ryhmän muodostavat epäselvät tapaukset, joiden kohdalla ei voida sanoa varmaksi, välittyykö aspektin merkitys vai ei. Käännössegmentit on luokiteltu myös konekäännöksen laadun perusteella kolmeen eri ryhmään: hyviin, heikkoihin ja käyttökelvottomiin. Aspektiparikohtaisissa alaluvuissa esiteltävät

esimerkkitapaukset on valittu siten, että ne edustaisivat mahdollisimman laajasti eri ryhmiin luokiteltuja tapauksia. Alalukujen kohdalla on annettu myös tiedot eri ryhmiin luokiteltujen segmenttien määrästä niin aspektin merkityksen välittymisen kuin käännosten laadun osalta. Jos käännoksen laatua ei ole erikseen kommentoitu esimerkkien kohdalla, on käänнос tällöin luokiteltu laadultaan hyväksi.

3.3.1 **выигрывать / выиграть**

Imperfektiivisen aspektin verbi *выигрывать* esiintyy aineistossa yhteensä 29 kertaa, ja perfektiivisen aspektin verbi *выиграть* esiintyy 28 kertaa. Korpusaineistosta löytyi molemmille verbeille ainoastaan yksi käänносvastine: *voittaa*.

Tarkastellaan ensin imperfektiivistä verbiä *выигрывать*.

- (15) Ближайшие несколько лет южноафриканка постепенно привыкает к соперничеству на новом уровне, выигрывает несколько титулов на соревнованиях второстепенного тура
- (16) Seuraavien vuosien aikana eteläafrikkalainen tottuu vähitellen kilpailuun uudella tasolla, voittaa useita titteleitä toisen kierroksen kilpailuissa

Tutkittava verbi on molemmissa esimerkeissä preesensissä. Suomenkielisessä lauseessa objekti on partitiivissa, joten imperfektiivisen aspektin merkitys välittyy eteenpäin.

- (17) Из менения в лучшую стороны стали происходить по ходу сезона-1996: Кётцер выигрывает сразу тринадцать матчей в рамках основных сеток турниров Большого шлема (в том числе побывав в полуфинале Australian Open), а также начинает более стабильно играть менее значимые соревнования.
- (18) Muutoksia parempaan alkoi tapahtua kauden 1996 aikana: Koetzer voittaa kolmetoista ottelua kerralla Grand Slam -turnausten pääarvona (mukaan lukien Australian välierissä Open), ja alkaa myös pelata vähemmän merkittäviä kilpailuja johdonmukaisemmin.

Edellä on esimerkki heikosta käännoksestä, joka on ilmaisuiltaan hieman epäluontevalta kuulostava. Tutkittava verbi on molemmilla kielillä preesensissä. Suomenkielisessä lauseessa objekti on jälleen partitiivissa, joten tässäkin tapauksessa aspektin merkitys välittyy käännoksen heikommasta laadusta huolimatta.

- (19) Во второй половине сезона чешка выигрывает свой второй турнир WTA -- в Форест-Хилсе и дошла до полуфинала на 75-тысячнике в Пуатье.
- (20) Kauden toisella puoliskolla tšekki voitti toisen WTA-turnauksensa Forest Hillsissä ja pääsi semifinaaliin 75 000:lla Poitiersissa.

Venäjänkielisessä virkkeessä tutkittavan verbin aikamuotona on preesens, mutta suomenkielisessä käännoksessa aikamuotona on imperfekti. Tässä tapauksessa kyseessä lienee niin sanottu historiallinen preesens. Se tarkoittaa sitä, että preesensia käytetään viittaamaan menneisiin tapahtumiin (VISK § 1529). Imperfektin käyttö käännoksessa ei siis varsinaisesti

ole väärin, tällöin vain tehokeinona käytetty historiallinen preesens katoaa. Aspektin välittymiseen tällä kuitenkin on vaikutusta. Venäjänkielisessä lauseessa aspekti on imperfektiivinen. Suomenkielisessä lauseessa objekti on akkusatiivissa, jolloin aspekti on resultatiivinen. Tässä esimerkissä alkuperäinen aspektin merkitys ei siis välity.

Edellä esitelty tapaus oli imperfektiivisen verbin kohdalla ainoa, jossa aspektin merkitys ei välittynyt. Kaikissa muussa 28 tapauksessa aspektin merkitys siis välittyi. Käännöksistä 19 luokiteltiin hyväksi ja 10 heikoiksi.

Tarkastellaan seuraavaksi perfektiivistä verbiä *выиграть*.

- (21) В начале апреля -- на зелёном грунте Хилтон-Хеда -- Аманде удалось выиграть самый статусный турнир своей одиночной карьеры.
- (22) Huhtikuun alussa Hilton Headin vihreällä maaperällä Amanda onnistui voittamaan yksinpeliuransa arvostetuimman turnauksen.

Venäjänkielisessä esimerkissä tutkittava verbi esiintyy yhdessä pääverbin *удаться* kanssa, suomenkielisessä esimerkissä verbin *onnistua* kanssa. Myös *удаться* on aspektiltaan perfektiivinen verbi. Suomenkielisessä esimerkissä lauseen objekti on akkusatiivissa, joten aspekti on resultatiivinen. Perfektiivinen aspekti ja resultatiivinen aspekti vastaavat toisiaan, joten aspektin merkitys välittyy.

- (23) Во время грунтового сезона Шафаржова выиграла домашний 100-тысячник из цикла ITF в Праге.
- (24) Savikauden aikana Shafarova voitti Prahassa ITF -kierroksen 100 000. kotikierroksen.

Erikoiset käännösvastineet tekevät tästä käännöksestä käyttökeltottoman, koska lauseen alkuperäinen ajatus ei välity kunnolla eteenpäin. Aspektin välittymiseen tämä ei kuitenkaan vaikuta. Tässäkin esimerkissä suomenkielisen lauseen aspekti on resultatiivinen, koska lauseen objekti on akkusatiivissa.

- (25) В 2014 году Луция продолжила активно играть парные состязания -- основной напарницей теперь была Главачкова, но и без неё время от времени удавалось добиться неплохих результатов: в январе чешка выиграла турнир в Сиднее вместе с Тимеей Бабош, попутно переиграв альянсы Пешке / Среботник и Эррани / Винчи.
- (26) Vuonna 2014 Luzia jatkoi aktiivisesti nelinpelien pelaamista - Glavachkova oli nyt pääkumppani, mutta jopa ilman häntä, hän onnistui aika ajoin saavuttamaan hyviä tuloksia: tammikuussa tšekki voitti turnauksen Sydneyssä yhdessä Timea Babosin kanssa samanaikaisesti.

Vaikka suomenkielinen käännös näyttää laadultaan pääosin hyvältä, on se kuitenkin jouduttu luokittelemaan käyttökeltottomaksi, sillä käännin on jättänyt alkuperäisen tekstin viimeisen lauseen sisällön kokonaan pois käännöksestä. Tutkittavan verbin aspektiin tällä virheellä ei

kuitenkaan ole vaikutusta. Käännöksessä lauseen aspekti on resultatiivinen, mikä vastaa venäjän perfektiivistä aspektia.

Aspektin merkitys välittyi hyvin kaikissa 28 tapauksessa. Käännösten laatu oli parempi kuin imperfektiivisen verbin esimerkkien käännösten kohdalla, sillä peräti 21 käännöstä luokiteltiin hyviksi ja vain neljä heikoksi. Toisaalta kolme perfektiivisen verbin sisältänyttä käännöstä luokiteltiin käyttökelttomiksi.

3.3.2 выпускать / выпустить

Imperfektiivisen aspektin verbi *выпускать* esiintyy aineistossa yhteensä 26 kertaa, ja perfektiivisen aspektin verbi *выпустить* esiintyy 28 kertaa. Aineistosta näille verbeille löytyi kaikkiaan neljä eri verbin merkityksistä riippuvaa käännösvastinetta: *ilmoittaa*, *julkaista*, *päästä* ja *vapauttaa*.

Aloitetaan tarkastelemalla imperfektiivistä verbiä *выпускать*.

- (27) Работал на Украине, в Харьковском театре оперы и балета имени Н. В. Лысенко (1929 -- 1931 главным режиссёром и главным балетмейстером), выпуская полнометражные спектакли и балеты, среди которых <<Днепрострой>> А. Бородина.
- (28) Hän työskenteli Ukrainassa N. V. Lysenkon (1929-1931 pääohjaajana ja pääkoreografina) nimetyssä Harkovin ooppera- ja balettiteatterissa ja julkaisi täyspitkiä esityksiä ja baletteja, mukaan lukien A. Borodin Dneprostroy.

Esimerkin lähdetekstissä verbin muotona on gerundin preesens, kun taas käännöksessä aikamuotona on imperfekti. Verbi *работать* on kuitenkin preteritissä, joten gerundimuoto viittaa myös menneeseen aikaan. Siksi se on tässä yhteydessä luontevaa kääntää suomeksi imperfektillä. Näin ollen aspektin merkityskin välittyy. Käännös on laadultaan tosin heikko, koska se mukailee liian tarkkaan alkuperäistekstiä.

- (29) Это про новые группы и исполнителей, которые выпускают свой первый сингл и не получают оплаты за его успех.
- (30) Se kertoo uusista bändeistä ja artisteista, jotka julkaisevat ensimmäisen singlensä eivätkä saa palkkaa sen menestyksestä.

Tässä esimerkissä aikamuoto on preesens sekä lähdetekstissä että käännöksessä. Tällöin tyypillisesti aspekti onkin imperfektiivinen.

Otetaan seuraavaksi mukaan tarkasteluun myös perfektiivinen verbi *выпустить*.

- (31) В 1964 году тренер команды получил указание перед игрой в Горьком против команды ГАЗ, проходившей в рамках чемпионата завода ЗИЛ, не выпускать на поле Стрельцова, однако в первом тайме болельщики устроили массовые беспорядки и угрожали даже поджечь стадион, если не выпустят Стрельцова.
- (32) Vuonna 1964 joukkueen valmentajaa ohjeistettiin ennen Gorkin ottelua GAZ-joukkuetta vastaan, joka pidettiin osana ZIL-tehtaan mestaruutta, että Streltsovia ei saa päästää kentälle, vaan ensimmäisellä puoliajalla. fanit järjestivät mellakoita ja uhkasivat jopa sytyttää stadionin tuleen, jos ei Streltsov vapautetaan.

Tässä esimerkissä on mukana sekä imperfektiivinen että perfektiivinen verbi. Imperfektiivinen verbi on käännetty suomeksi verbillä *päästää*, kun taas perfektiivinen verbi on käännetty verbillä *vapauttaa*. Käännös on luokiteltu heikoksi, koska käännös ei kuulosta täysin luontevalta. Pitkä virke on ilmeisestikin tuottanut hankaluuksia konekääntimelle. Imperfektiivisen verbin sisältävältä kohdalta käännös on kuitenkin onnistunut kelvosti. Perfektiivinen verbi on sen sijaan kääntynyt suomeksi kieliopillisesti virheelliseen muotoon, sillä oikea muoto olisi *ei vapauteta*. Käännös voisi olla tässä yhteydessä parempikin, koska virkkeessä ei varsinaisesti puhuta vapauttamisesta, vaan mahdollisuuden antamisesta pelaajalle tai pelaajan kentälle päästämisestä. Aspektin merkityksen välittymiseen käännösvirheellä on myös merkitystä osittain, koska objektin sija ei sovi käännökseen. Kielteiseen lauseeseen sopisi partitiiviobjekti *Streltsovia*, jolloin vastaavuus venäjän perfektiivisen aspektin kanssa olisi olemassa.

- (33) В эмиграции выпустил перевод <<Дивана>> Гёте (Париж, 1932).
- (34) Maanpaossa hän julkaisi käännöksen Goethen Divanista (Pariisi, 1932).

Alkutekstin lauseessa perfektiivinen verbi on preteritissä, ja suomenkielisessä lauseessa objekti on akkusatiivissa. Näiden vastaavuus on siis hyvin esillä tässä esimerkissä.

- (35) В финале против Югославии Стрельцов не играл по решению тренера Гавриила Качалина, который выпустил на позицию Стрельцова Никиту Симоняна.
- (36) Finaalissa Jugoslaviaa vastaan Streltsov ei pelannut valmentajan Gavriil Kachalinin päätöksellä, joka vapautti Nikita Simonyanin Streltsovin asemaan.

Käännös on hieman epäluonteva, joten luokittelin sen heikoksi. Aspektin merkityksen vastaavuus on kuitenkin olemassa, sillä tässäkin tapauksessa perfektiivinen verbi on preteritissä, ja suomenkielisessä lauseessa on akkusatiiviobjekti.

- (37) Тейлор Свифт выпустила свой третий студийный альбом, Speak Now, в октябре 2010 года.
- (38) Taylor Swift julkaisee kolmannen studioalbuminsa, Speak Nyt, lokakuussa 2010.

Tämä esimerkki on mielenkiintoinen, sillä alkutekstin virkkeessä verbi on preteritissä, mutta konekäännöksessä verbi on preesensissä. Siksi luokittelin tämän käännöksen heikoksi, koska mielestäni tässä pitäisi olla suomeksi imperfekti. Englanninkielinen nimi keskellä virkettä

näyttäisi myös hieman hämmentäneen konekääntimen toimintaa. Väärä aikamuoto ei kuitenkaan tässä tapauksessa vaikuta aspektin merkityksen välittymiseen. Venäjänkielisessä virkkeessä aspekti on perfektiivinen, ja suomenkielisessä lauseessa objekti on akkusatiivissa.

- (39) В феврале 2021 года Свифт и рок-группа Haim выпустили совместный сингл <<Gasoline>> из расширенной версии альбома Women in Music Pt.
(40) Helmikuussa 2021 Swift ja rockbändi Haim julkaisivat yhteistyösinglen " Gasoline" albumin Women laajennetulta versiolta. sisään Musiikki Pt .

Englanninkielinen nimi on mitä ilmeisemmin aiheuttanut sen, että konekäännös on sellaisenaan käyttökelvoton, mutta tässäkin esimerkissä tutkittavan verbin aspektin merkitys on välittynyt käännökseen hyvin.

- (41) В конце ноября Big Machine Records выпустили плейлист Reputation Stadium Tour для стриминговых сервисов.
(42) Marraskuun lopun suuri kone Records julkaisee Maine- soittolista Stadion Kierros suoratoistopalveluihin.

Tässäkin esimerkissä käännös on käyttökelvoton englanninkielisten nimien kääntymisen myötä. Aspektin merkityksen sanoisin välittyvän tässä tapauksessa vain osittain. Käännöksessä objekti on nominatiivissa, mikä on kieliopillisesti väärin tässä yhteydessä. Jostain syystä tässäkin preteritissä oleva verbi on käännetty preesensissä olevalla verbillä.

Imperfektiivisen aspektin verbin *выпускать* kohdalla kaikissa tarkastelluissa tapauksissa aspektin merkitys välittyi hyvin. Käännöksistä 14 luokiteltiin hyväksi, 12 heikoiksi. Täysin käyttökeltottomia käännöksiä ei siis aineistosta kyseisen verbin kohdalla löytynyt. Perfektiivisen aspektin verbin *выпустить* kohdalla aspektin merkitys välittyi niin ikään hyvin, sillä 26 tapauksessa merkitys välittyi hyvin. Edellä esitellyissä kahdessa tapauksessa vastaavuudet olivat vain osittaiset, koska käännökset olivat jääneet puutteellisiksi. Käännösten laatu sen sijaan osoittautui heikommaksi kuin imperfektiivisen verbin kohdalla, sillä vain viisi käännöstä luokiteltiin hyväksi. Heikkoja käännöksiä oli peräti 19 ja käyttökeltottomia neljä.

3.3.3 играть / сыграть

Imperfektiivisen aspektin verbi *играть* esiintyy aineistossa yhteensä 52 kertaa, ja perfektiivisen aspektin verbi *сыграть* esiintyy 46 kertaa. Näille verbeille löytyi aineistosta yhteensä kolme eri verbin merkityksistä riippuvaa käännösvastinetta: *soittaa, pelata ja näyttellä*.

Tarkastellaan ensimmäisenä imperfektiivistä verbiä *играть*.

- (43) Ольга родилась в спортивной семье -- её мама увлекалась волейболом, а папа играл в бильярд и был мастером спорта в этой игре.
- (44) Olga syntyi urheiluperheeseen - hänen äitinsä rakasti lentopalloa, ja hänen isänsä pelasi biljardia ja oli urheilun mestari tässä pelissä.

Tutkittavan verbin aikamuoto on venäjänkielisessä esimerkissä preteriti ja suomenkielisessä imperfekti. Objekti on suomenkielisessä esimerkissä partitiivissa, joten lauseessa on venäjän imperfektiivistä aspektia vastaava irresultatiivinen aspekti.

- (45) В частности, великий феодосийский маринист сообщил Глинке три татарских напева, наиграв их на скрипке (художник владел инструментом и часто играл на творческих вечерах композитора).
- (46) Erityisesti suuri Feodosian merimaalari kertoi Glinkalle kolme tatarilaista melodiaa soittaen niitä viululla (taiteilija omisti instrumentin ja soitti usein säveltäjän luovilla iltoilla).

Venäjänkielisessä esimerkissä imperfektiivinen aspekti on selkeä valinta, koska ennen tutkittavaa verbiä on sana *часто* (usein). Usein toistuvaa toimintaa kuvataan imperfektiivisellä aspektilla. Suomenkielisessä esimerkissä verbin yhteydessä ei ole objektiä, mutta sana *usein* välittää ajatuksen toistuvasta toiminnasta, joten määrittelin lauseen aspektin irresultatiiviseksi. Käännöksen olen luokitellut heikoksi, koska esimerkiksi sanan *ilta* taivutus on mennyt kääntimellä väärin.

- (47) В последние годы парной карьеры Аманда достаточно часто играла в дуэте с американкой Лори Макнил: девушки трижды доходили до финалов турниров WTA и взяли два титула.
- (48) Kaksinpeliuransa viimeisinä vuosina Amanda soitti usein duetossa amerikkalaisen Laurie McNeilin kanssa: tytöt pääsivät kolme kertaa WTA -turnausten finaaliin ja voittivat kaksi titteliä.

Esimerkin käännös on muuten sujuva, mutta huomio kiinnittyy nyt itse tutkittavan verbin virheelliseen käännösvastineeseen. Tässä käännöksessä pitäisi käyttää verbin *soitti* sijasta verbiä *pelasi*. Myös tässä esimerkissä verbien yhteydessä esiintyvät sanat *часто* ja *usein*. Virheellinen käännösvastine ei tässä tapauksessa kuitenkaan vaikuta aspektin merkityksen välittymiseen.

- (49) Выступала в 2005 -- 2007 годах в Германии за команду <<Анрохт> >, позже играла за французский <<Мондевилль> >.
- (50) Hän pelasi Saksassa vuosina 2005-2007 Anrochtin joukkueessa, myöhemmin ranskalaisessa Mondevillessä.

Tässä tapauksessa käännös on sen sijaan oikein hyvä, mutta huomio kiinnittyy nyt verbien määrään. Venäjänkielisessä esimerkissä on kaksi verbiä, kun taas suomenkielisessä vain yksi. Konekäännin on kääntänyt verbin *выступала* sanalla *pelasi*. Jälkimmäinen verbi *играла* on siis periaatteessa kadonnut suomennoksesta. Toisaalta asian voi tulkita myös siten, että ensimmäinen verbi onkin jäänyt pois ja suomennoksessa on mukana vain jälkimmäinen verbi.

Verbin *выступать* voisi tässä tapauksessa suomentaa esimerkiksi verbillä *edustaa* (joukkuetta). Käännös on kuitenkin sujuva ilmankin tätä verbiä. Koska en voi tietää kääntimen logiikkaa käännöksen taustalla, en osaa sanoa, välittyykö aspektin merkitys tutkittavan verbin kohdalla. Siksi tämä tapaus jää nyt epäselväksi.

Edellinen epäselvä tapaus jäi imperfektiivisen aspektin verbin kohdalla ainoaksi laatuun, sillä kaikissa muissa 51 tapauksessa aspektin merkitys välittyi hyvin käännöksiin. Kaikkiaan 33 käännöstä luokiteltiin hyväksi ja 19 heikoiksi.

Siirrytään seuraavaksi tarkastelemaan perfektiivisen aspektin verbiä *сыграть*.

- (51) Зимой 2005 года вызывался в юношескую сборную Украины до 17 лет и сыграл два матча -- против Турции и Германии.
- (52) Talvella 2005 hänet kutsuttiin Ukrainan alle 17-vuotiaiden nuorten maajoukkueeseen ja hän pelasi kaksi ottelua - Turkkiä ja Saksaa vastaan.

Suomen kielioppi vaatii nyt lukusanan yhteydessä partitiiviobjektin *ottelua*, vaikka usein venäjän perfektiivistä aspektia vastaa suomessa akkusatiiviobjekti. Vaikka objekti onkin nyt partitiivissa, on kuitenkin selvää, että mainitut ottelut on jo pelattu loppuun ja täten ne ovat jo mennyttä elämää. Kyse on siis suomenkielisessä esimerkissä resultatiivisesta aspektista, joka vastaa venäjän perfektiivistä aspektia.

- (53) Роль, которую готовился исполнить Михаил Вавич сыграл уже другой актёр.
- (54) Roolin, jota Mikhail Vavich valmistautui näyttelemään, näytteli jo toinen näyttelijä.

Tässä esimerkissä tutkittava verbi esiintyy merkityksessä *näytellä*. Tässä suomenkielisessä esimerkissä objekti *roolin* on nyt akkusatiivissa. Tapaus on siis nyt edellistäkin esimerkkiä selkeämpi. Suomenkielisen virkkeen aspekti on resultatiivinen, joten vastaavuus venäjän perfektiivisen aspektin kanssa on jälleen olemassa.

- (55) В дальнейшем, уже как актёрский дуэт, Михаил Вавич и Татьяна Павлова сыграли в целом ряде отечественных и зарубежных кинофильмов.
- (56) Tulevaisuudessa Mikhail Vavich ja Tatjana Pavlova soittivat jo näyttelevänä duetona useissa kotimaisissa ja ulkomaisissa elokuvissa.

Jälleen kerran konekäännin on käyttänyt väärää käännösvastinetta. Soittamisen sijaan kyse on tietenkin näyttelemisestä. Aspektin merkityksen välittymiseen virheellä ei tälläkään kertaa ole kuitenkaan vaikutusta. Käännöksen laatuun virheellä on sen sijaan tietenkin merkitystä, sillä tämä on luokiteltu heikoksi käännökseksi.

Kaikissa 46 tutkitussa perfektiivisen aspektin verbin tapauksessa aspektin merkitys välittyi hyvin. Käännöksistä 22 oli hyviä ja 24 heikkoja, joten heikkojen käännösten osuus oli nyt suurempi kuin imperfektiivisen aspektin verbin kohdalla.

3.3.4 начинать / начать

Imperfektiivisen aspektin verbi *начинать* esiintyy aineistossa yhteensä 30 kertaa, ja perfektiivisen aspektin verbi *начать* esiintyy 26 kertaa. Aineistosta löytyi näille verbeille kaksi eri verbin merkityksistä riippuvaa käännösvastinetta: *aloittaa* ja *alkaa*.

Otetaan ensin tarkasteltavaksi imperfektiivinen verbi *начинать*.

- (57) Помимо работы на театральной и концертной сцене Михаил Вавич начинает сниматься в кино.
(58) Teatteri- ja konserttilavalla työskentelyn lisäksi Mikhail Vavich alkaa näytellä elokuvissa.

Molemmissa esimerkeissä tutkittavan verbin aikamuotona on preesens. Lauseissa kerrotaan toiminnan alkavan, mutta toiminnalle ei määritetä tässä loppua. Siispä suomenkielisen lauseen aspektin on oltava irresultatiivinen, mikä vastaa tässä tapauksessa venäjän imperfektiivistä aspektia.

- (59) К марту 1996 года, закончив работу над романом в соответствии с договором издательства <<Локид>>, он начинает писать два новых романа, сначала <<Звёзды -- холодные игрушки>>, а затем и <<Лабиринт отражений>>.
(60) Maaliskuuhun 1996 mennessä, saatuaan päätökseen romaanin työskentelyn Lokid-julkaisusopimuksen mukaisesti, hän alkoi kirjoittaa kahta uutta romaania, ensin Stars are Cold Toys ja sitten Reflection Labyrinth.

Tässä esimerkissä tutkittavien verbien aikamuodot eroavat. Venäjänkielisessä esimerkissä aikamuotona on preesens, mutta suomenkielisessä esimerkissä imperfekti. Käännökseen tai aspektin merkityksen välittymiseen tällä ei kuitenkaan ole merkitystä, koska kyse on joka tapauksessa menneestä tapahtumasta. Objekti on suomenkielisessä lauseessa partitiivissa, joten lauseessa on irresultatiivinen aspekti, joten vastaavuus venäjän aspektiin on olemassa. Teosten nimet ovat kääntyneet englanniksi, joten käännös on siksi luokiteltu heikoksi.

Aspektin merkitys välittyi hyvin kaikissa 30 tarkastellussa tapauksessa. Käännöksistä 21 luokiteltiin hyväksi, yhdeksän heikoiksi.

Tarkastellaan seuraavaksi perfektiivisen aspektin verbiä *начать*.

- (61) Также президент Казахстана поручил министерству иностранных дел начать процедуру присоединения ко Второму Факультативному протоколу к Международному пакту о гражданских и политических правах.
- (62) Kazakstanin presidentti kehotti myös ulkoministeriötä aloittamaan menettelyn kansalaisyhteisönsä ja poliittisia oikeuksia koskevan kansainvälisen yleissopimuksen toiseen valinnaiseen pöytäkirjaan liittymiseksi.

Näissä esimerkeissä tutkittava verbi esiintyy yhdessä toisen verbin kanssa. Venäjänkielisessä lauseessa *начать* esiintyy yhdessä verbin *поручил* kanssa, ja suomenkielisessä lauseessa *aloittamaan* esiintyy yhdessä verbin *kehotti* kanssa. Venäjän verbi *поручить* on myös aspektiltaan perfektiivinen. Suomenkielisessä lauseessa objekti *menettelyn* on akkusatiivissa. Näin ollen suomenkielisessä lauseessa aspekti on resultatiivinen, mikä vastaa tässä tapauksessa venäjän perfektiiivistä aspektia.

- (63) Основной удар наносила армия Шермана, начавшая в мае вторжение в Джорджию.
- (64) Pääiskun antoi Shermanin armeija, joka aloitti hyökkäyksen Georgiaan toukokuussa.

Venäjänkielisessä esimerkissä verbi on partisiippimuodossa. Tarkemmin sanottuna kyseessä on aktiivin partisiipin preteriti. Kyseistä partisiippia voidaan käyttää relatiivilauseiden korvaamiseen (Kirvesmäki ym. 2021, 227). Konekäännin onkin kääntänyt kyseisen partisiipin käyttäen relatiivilauseetta. Aikamuoto on siis venäjänkielisessä esimerkissä preteriti ja suomenkielisessä imperfekti. Suomenkielisessä esimerkissä objekti on akkusatiivissa, joten suomenkielisen lauseen aspekti on resultatiivinen.

- (65) Начала карьеру игрока в настольный теннис в местном клубе, выступала за команду Югославии.
- (66) Hän aloitti uransa pöytätennispelaajana paikallisessa seurassa, pelasi Jugoslavian joukkueelle.

Tämän esimerkin käännöksen olen luokitellut heikoksi, koska lauseet tarvitsisivat suomessa väliin konjunktion. Aspektin kannalta tapaus on kuitenkin selkeä, sillä suomenkielisessä lauseessa objekti on akkusatiivissa, joten lauseen aspekti on resultatiivinen.

- (67) Уже в начале карьеры Тейлор Свифт начали называть кумиром для подростков, а в дальнейшем поп-иконкой и королевой поп-музыки.
- (68) Uransa alkuvaiheessa Taylor Swiftiä kehuttiin teini-idoliksi ja myöhemmin pop-ikoniksi ja popin kuningattareksi.

Tässä on sen sijaan tapaus, jota sanoisin epäselväksi. Konekäännin on ohittanut tutkittavan verbin kokonaan ja kääntänyt ilmaisun *начали называть* yhdellä verbillä *kehuttiin*. Suomennos ei ole mielestäni myöskään täysin onnistunut, sillä verbin *называть* merkitys on enemmänkin *nimittää* tai *kutsua*, ei *kehua*. Jos molemmat verbit haluttaisiin säilyttää käännöksessä, olisi parempi suomennos esimerkiksi *alettiin kutsua*. Aspektin merkityksen

välittymistä on hankala kommentoida, kun alkuperäinen tutkittava verbi puuttuu käännöksestä. Kaiken lisäksi tutkittavasta verbistä poiketen *называть* on aspektiltaan imperfektiivinen verbi.

Edellä esitellyn epäselvän tapauksen lisäksi muita vastaavanlaisia tapauksia ei ilmennyt, vaan kaikissa lopuissa 25 tapauksessa aspektin merkitys välittyi hyvin. Käännöksistä 16 luokiteltiin hyväksi, 10 heikoiksi.

3.3.5 писать / написать

Imperfektiivisen aspektin verbi *писать* esiintyy aineistossa yhteensä 27 kertaa, ja perfektiivisen aspektin verbi *написать* esiintyy 39 kertaa. Näille kahdelle verbille löytyi aineistosta kaksi eri verbin merkityksistä riippuvaa käännösvastinetta: *kirjoittaa* ja *säveltää*.

Tarkastellaan ensin imperfektiivistä verbiä *писать*.

- (69) <<Я пишу песни, и мой голос -- всего лишь способ донести эти тексты>>
(70) "Kirjoitan kappaleita, ja ääneni on vain tapa saada nuo sanat perille."

Sekä venäjän- että suomenkielisessä esimerkissä tutkittavan verbin aikamuotona on preesens. Suomenkielisessä lauseessa objekti on partitiivissa, joten kyseessä on venäjän imperfektiivistä aspektia vastaava irresultatiivinen aspekti.

- (71) Наряду с работой в опере, Кавос писал музыку и к балетам
(72) Оopperatyön ohella Cavos sävelsi musiikkia myös baletteihin

Tässä esimerkissä tutkittava verbi esiintyy hieman harvinaisemmassa merkityksessä *säveltää*. Verbien aikamuotona on venäjänkielisessä lauseessa preteriti ja suomenkielisessä imperfekti, mutta aspektit ovat samat kuin edellisessäkin esimerkissä. Objekti on suomenkielisessä lauseessa partitiivissa, joten lauseen aspekti on irresultatiivinen, mikä vastaa venäjän imperfektiivistä aspektia.

- (73) В этот же период он пишет такие значимые произведения, как повесть <<Атомный сон>>, рассказы <<Л -- значит люди>>, <<Дорога на Веллесберг>> и <<Мой папа -- антибиотик>>.
(74) Samana aikana hän kirjoitti sellaisia merkittäviä teoksia kuin tarina "Atomic Dream", tarinat "L tarkoittaa ihmisiä", "Tie Wellesbergiin" ja "Isäni on antibiootti".

Teosten nimien kohdalla käännin on tehnyt käännöksen laatua heikentäviä asioita. Aspektin merkityksen välittymiseen tällä ei kuitenkaan ole merkitystä. Suomenkielisen lauseen objekti on jälleen partitiivissa, joten lauseen aspekti on irresultatiivinen. Venäjänkielisessä esimerkissä

tutkittavan verbin aikamuotona on preesens, mutta suomenkielisessä esimerkissä aikamuotona on imperfekti. Kyse lienee jälleen jo aiemmin mainitusta historiallisesta preesensistä.

Tutkittuja tapauksia oli imperfektiivisen aspektin verbille kaikkiaan 27, joista kaikissa aspektin merkitys välittyi käännökseen. Hyviä käännöksiä oli 16 ja heikkoja 11.

Siirrytään seuraavaksi tarkastelemaan perfektiivistä verbiä *написать*.

- (75) Песню для Лазарева написал Филипп Киркоров в соавторстве с греческим композитором Димитрисом Контопулосом и американской певицей Шерон Вон.
(76) Laulun Lazareville on kirjoittanut Philip Kirkorov yhteistyössä kreikkalaisen säveltäjän Dimitris Kontopoulосin ja amerikkalaisen laulajan Sharon Vaughnin kanssa.

Venäjänkielisessä lauseessa tutkittavan verbin aikamuotona on preteriti, ja suomenkielisessä lauseessa perfekti. Suomenkielisen lauseen objekti on akkusatiivissa, joten lauseen aspekti resultatiivinen, mikä vastaa venäjän perfektiivistä aspektia.

- (77) Весной 1840 года композитор постоянно навещал Керн и именно тогда написал романс <<Я помню чудное мгновенье>> на стихи Пушкина, посвятив его дочери той, кому поэт адресовал эти стихи.
(78) Keväällä 1840 säveltäjä vieraili jatkuvasti Kernissä ja silloin hän kirjoitti Pushkinin runoihin perustuvan romanssin "Muistan ihanan hetken" ja omisti sen sen tyttärelle, jolle runoilija puhui nämä runot.

Käännöksen loppuosa on mennyt hieman pieleen kääntimeltä, mutta tutkittavan verbin sisältävään osaan virheillä ei ole vaikutusta. Venäjäksi aikamuotona on preteriti ja suomeksi imperfekti. Suomenkielisen lauseen aspekti on tässäkin tapauksessa resultatiivinen, koska lauseessa on partitiiviobjekti.

Edellä käsitelty verbipari osoittautui tässä tutkimuksessa varsin selkeäksi tapaukseksi, sillä myös perfektiivisen verbin kohdalla aspektin merkitys välittyi kaikissa tutkitussa 39 tapauksessa. Käännösten laatu oli perfektiivisen verbien sisältävien virkkeiden kohdalla hieman heikompi kuin imperfektiivisen verbin kohdalla. Heikkoja käännöksiä oli enemmän, yhteensä kaikkiaan 22, kun taas hyviä käännöksiä oli yhteensä 17.

3.4 Analyysin yhteenveto

Kaikki tapaustutkimukset on nyt suoritettu, joten on aika tehdä vielä yhteenveto niiden tuloksista. Seuraaviin taulukoihin on koottu tutkimustulokset liittyen luokitteluihin aspektin merkityksen välittymisen sekä konekäännösten laadun osalta. Aloitetaan tutkimustulosten käsittely aspektin merkityksen välittymisen tuloksista.

Taulukko 12. Aspektin merkityksen välittyminen verbeittäin.

	Välittyy hyvin	Välittyy osittain	Ei välity	Epäselvä
1. выигрывать	28		1	
выиграть	28			
2. выпускать	26			
выпустить	26	2		
3. играть	51			1
сыграть	46			
4. начинать	30			
начать	25			1
5. писать	27			
написать	39			
YHTEENSÄ	326	2	1	2
Prosentteina	98,5 %	0,6 %	0,3 %	0,6 %

Yhteensä 331 tapauksesta siis peräti 326 tapauksessa aspektin merkitys välittyi hyvin konekäännökseen. Muunlaisia tapauksia oli yhteensä vain viisi. Niistä kahdessa aspektin merkitys välittyi osittain ja yhdessä ei ollenkaan. Kaksi tapausta luokiteltiin epäselviksi. Prosentteina muunlaisten tapauksien osuus oli vain 1,5 prosenttia kaikista tapauksista, kun taas hyvin välittyneiden osuus on näin ollen 98,5 prosenttia tutkituista tapauksista.

Taulukko 13. Aspektin merkityksen välittyminen imperfektiivisten ja perfektiivisten verbien osalta.

	Välittyy hyvin	Välittyy osittain	Ei välity	Epäselvä
Imperfektiiviset verbit	162	0	1	1
Prosentteina	98,8 %	0,0 %	0,6 %	0,6 %
Perfektiiviset verbit	164	2	0	1
Prosentteina	98,2 %	1,2 %	0,0 %	0,6 %

Kun tarkastellaan erikseen imperfektiivisen ja perfektiivisten verbien tapauksia, huomataan pieniä eroja tuloksissa näiden verbiryhmien välillä. Imperfektiivisten verbien joukossa oli suhteellisesti niukasti enemmän tapauksia, joissa aspektin merkitys välittyi hyvin. Molemmat tapaukset, joissa aspektin merkitys välittyi osittain, olivat perfektiivisten verbien joukossa. Nämäkin tapaukset olivat molemmat saman verbin kohdalla. Ainoa tapaus, jossa aspektin merkitys ei välittynyt, oli imperfektiivisten verbien joukossa. Epäselviä tapauksia oli yksi sekä imperfektiivisten että perfektiivisten verbien joukossa.

Tarkastellaan seuraavaksi tuloksia segmenttien konekäännösten laadun luokittelun osalta.

Taulukko 14. Konekäännösten laatu lähdetekstin segmentin sisältämän verbin mukaan.

	Hyvä	Heikko	Käyttökelvoton
1. выигрывать	19	10	
выиграть	21	4	3
2. выпускать	14	12	
выпустить	5	19	4
3. играть	33	19	
сыграть	22	24	
4. начинать	21	9	
начать	16	10	
5. писать	16	11	
написать	22	17	
YHTEENSÄ	189	135	7
Prosentteina	57,1 %	40,8 %	2,1 %

Suurin osa (57,1 %) konekäännöksistä luokiteltiin siis laadultaan hyviksi. Heikkojen käännösten osuus (40,8 %) on kuitenkin myös merkittävä. Käyttökelvottomien käännösten osuus jäi vain pariin prosenttiin (2,1 %). Kahdeksan verbin kohdalla hyvien käännösten osuus oli heikkojen osuutta suurempi. Verbin *сыграть* sisältävien segmenttien käännöksistä niukka

enemmistö luokiteltiin heikoiksi. Merkittävän poikkeuksen tilastoon tekee kuitenkin perfektiiäinen verbi *выпустить*, jonka kohdalla ylivoimainen enemmistö käännöksistä luokiteltiin heikoiksi. Myös käyttökeltottomia käännöksiä oli eniten kyseisen verbin kohdalla. On kuitenkin syytä muistaa, että konekäännöksen heikolla laadulla ei ollut välttämättä mitään tekemistä lähdetekstin sisältämän verbin kanssa, vaan heikko laatu saattoi myös johtua muista syistä.

Taulukko 15. Konekäännösten laatu imperfektiivisiä ja perfektiiäisiä verbejä sisältäneiden segmenttien käännösten osalta.

	Hyvä	Heikko	Käyttökeltvoton
Imperfektiiviset verbit	103	61	0
Prosentteina	62,8 %	37,2 %	0,0 %
Perfektiiäiset verbit	86	74	7
Prosentteina	51,5 %	44,3 %	4,2 %

Kun tarkastellaan erikseen imperfektiivisten ja perfektiiäisten verbien osuuksia, huomataan jälleen eroavaisuuksia luokitusten määrissä. Imperfektiiäisen verbin sisältäneiden segmenttien käännöksistä 62,8 prosenttia luokiteltiin hyväksi, kun taas perfektiiäisen verbin sisältävien segmenttien osalta hyväksi luokiteltiin vain 51,5 prosenttia käännöksistä. Vastaavasti heikkojen käännösten osuus oli imperfektiivisten verbien kohdalla (37,2 %) pienempi kuin perfektiiäisten verbien kohdalla (44,3 %). Kaikki käyttökeltvottomat käännökset olivat segmenteille, jotka sisälsivät perfektiiäisiä verbejä. Tällaisia tapauksia oli yhteensä seitsemän, ja ne keskittyivät vain kahden eri verbin kohdalle. Tulosten perusteella näyttää siis siltä, että aspektiltaan imperfektiiviset verbit selviytyivät tutkimuksessa paremmin sekä aspektin merkityksen välittymisen osalta että konekäännösten laadun osalta.

4 LOPPUPÄÄTELMÄT

Tässä tutkielmassa olen tarkastellut venäjän kielen verbiaspektin merkityksen välittymistä suomenkielisissä konekäännöksissä. Tavoitteena oli selvittää, kuinka hyvin konekäännin selviytyy merkityksen välittämisestä. Olen esitellyt aspektin käsitteen ja sen, millä tavoilla aspekti ilmenee venäjän ja suomen kielissä. Olen myös vertaillut aspektin ilmaisemista näissä kielissä sekä esittänyt vastaavuuksia ja eroavaisuuksia, joita kyseisten ilmaisemiskeinojen välillä on. Näitä tietoja on hyödynnetty tutkielmassa suoritetuissa tapaustutkimuksissa, kun aspektin merkityksen välittymistä on arvioitu. Myös konekääntämisen teoriaa on käsitelty tutkielmassa.

Tutkimus toteutettiin korpusmetodia käyttämällä. Kokosin rinnakkaiskorpuksen, joka sisälsi venäjänkielisiä artikkeleita sekä niiden suomenkielisiä konekäännöksiä. Tutkin ensin yleisesti teksteissä esiintyviä verbejä ja niiden käännösvastineita korpustyökalujen avulla. Lopulta valitsin tutkittavat verbit aineiston joukosta. Tutkittavia verbejä oli yhteensä kymmenen, ja ne muodostivat kaikkiaan viisi aspektiparia. Tapaustutkimuksissa olen esimerkkien avulla esitellyt, miten konekäännin suoriutuu aspektin merkityksen välittämisestä. Tapaukset luokiteltiin neljään eri kategoriaan: tapauksiin, joissa aspektin merkitys välittyi hyvin, osittain tai ei ollenkaan sekä tapauksiin, jotka olivat asian osalta epäselviä. Samalla tarkasteltiin myös konekäännösten laatua.

Tutkimuskysymyksenä tässä tutkimuksessa oli se, miten konekäännin (Google-kääntäjä) selviää venäjän kielen verbien aspektien merkityksen välittämisestä suomen kieleen. Tapaustutkimusten jälkeen on varsin selvää, että konekäännin selviytyi tästä tehtävästä erittäin hyvin, sillä peräti yli 98 prosentissa tapauksista aspektin merkitys välittyi hyvin. Epäselviä tapauksia sekä tapauksia, joissa merkitys välittyi vain osittain tai ei ollenkaan, oli yli kolmensadan tapauksen joukossa vain muutama. Aspektin lisäksi tapaustutkimuksissa on tutkittu myös verbien aikamuotoja sekä mukana olleiden konekäännösten laatua.

Hypoteesina tässä tutkimuksessa oli se, että konekäännin selviytyy hyvin vain osasta käännöksiä. Tätä hypoteesia voidaan tarkastella kahdella eri mittarilla, aspektin merkityksen välittymisen ja konekäännösten laadun perusteella. Aspektin merkityksen välittymisen osalta hypoteesi ei täysin pitänyt paikkaansa, koska niin ylivoimainen enemmistö tapauksista oli onnistuneita. Oletus oli se, että aspektin merkityksen välittäminen tuottaisi jonkin verran haastetta konekääntimelle. Haasteiden ja hankaluuksien määrä jäikin reilusti pienemmäksi kuin

odotin. Konekäännösten laadun perusteella tilanne on sen sijaan tasaisempi. Laadultaan hyvien käännösten osuus (57,1 %) oli suurempi kuin heikkojen (40,8 %) ja käyttökeltottomien (2,1 %) osuudet. Onnistuneiden käännösratkaisujen määrä ei siis ole merkittävästi muunlaisten ratkaisujen osuutta suurempi. Lisäksi on syytä muistaa, että hyväksi käännöksiksi luokitellut segmentit eivät välttämättä olleet parhaita mahdollisia käännöksiä. Ihmiskääntäjä saattaisi pystyä vielä parempaan jälkeen. Tutkimustulos käännösten laadun osalta oli kuitenkin varsin odotettu. Ilahduttavaa oli nähdä se, kuinka käyttökeltottomien käännösten osuus jäi varsin pieneksi.

Vaikka tässä tutkimuksessa konekäännin selviytyikin oikein hyvin aspektin merkityksen välittämisestä, ei tuloksia voi kuitenkaan yleistää koskemaan ilmiötä laajemmin. Nämä tulokset koskevat lopulta kuitenkin vain valittuja verbejä, valittua materiaalia ja valittua konekäännintä. Lopullisissa tapaustutkimuksissa oli mukana ainoastaan kymmenen verbiä. Jos tutkimukseen olisi valittu eri verbit, olisivat tuloksetkin voineet olla täysin erilaiset. Valitsin verbit lähinnä sillä perusteella, että niiden käännösvastineet rinnakkaiskorpuksessa olivat selkeitä. Muiden verbien kohdalla olisi saattanut tulla vastaan monimutkaisempia ja -tulkintaisempia tapauksia. Tämän tutkimuksen perusteella näyttäisi kuitenkin siltä, että konekääntimen toiminta tutkittavan aiheen kohdalla vaikuttaa lupaavalta.

Tutkimusmateriaali koostui rajatusta valikoimasta venäjänkielisiä Wikipedia-artikkeleita sekä näiden suomenkielisistä konekäännöksistä. Tästä valinnasta voidaan olla montaa mieltä. Wikipediasta oli helppoa etsiä tekstejä tutkimusta varten, koska artikkeleita on valtava määrä. Toisaalta Wikipedia on avoin tietosanakirja ja sinne voi kirjoittaa käytännössä kuka tahansa. Artikkelien tekstin laatu voi siis vaihdella huomattavasti artikkelien välillä tai jopa artikkelien sisällä, koska artikkelia on vuosien saatossa saattanut olla kirjoittamassa suuri määrä ihmisiä, joiden kirjoitustaidot ja -tyyli eroavat toisistaan. Erilaisilla teksteillä olisi mahdollisesti saavutettu täysin erilaisia tuloksia.

Konekäännösten tekemiseen käytettiin tässä tutkimuksessa Google-kääntäjää. Kyseinen konekäännin valittiin siksi, että sen toiminta perustuu neuroverkkokääntämiseen, joka edustaa uutta sukupolvea konekääntämisessä. Venäjänkielisten artikkeleiden konekäännökset tuotettiin kääntimellä tammikuussa 2023. Näin ollen tutkimukset heijastavat sitä, millaista jälkeä käännin juuri tuona ajankohtana tuotti. Koska konekääntimiä kehitetään jatkuvasti, on hyvin mahdollista, että esimerkiksi tämän tutkielman julkaisuhetkellä konekäännösten laatu saattaisi

olla vielä parempi. Konekääntimiä koskevien tutkimusten kohdalla niiden ajankohta on otettava huomioon, kun tutkimustuloksia tarkastellaan.

Verbiaspektin merkityksen välittymisen lisäksi tässä tutkielmassa tarkasteltiin myös tapaustutkimuksissa mukana olleiden konekäännösten laatua. Päädyin arvioimaan laatua kolmiportaisella asteikolla: hyvä, heikko ja käyttökelvoton. Laadun arviointi on kuitenkin lopulta varsin subjektiivista toimintaa. Jonkun toisen ihmisen suorittamana arvioinnin tulokset olisivat voineet poiketa omista tuloksistani. Jokaisella meistä on kuitenkin omat rajansa sille, millaista laatua konekääntimeltä kelpuuttaa. Siksi tämän tutkimuksen luokitteluihinkaan ei voi suhtautua yhtenä totuutena eikä luokittelujen perusteella voi tehdä liikaa johtopäätöksiä siitä, millaista laatua käännin todellisuudessa pystyy tuottamaan.

Jatkotutkimusmahdollisuuksia on tutkielman aiheelle kuitenkin olemassa lukuisia. Jo pelkästään tämän tutkielman pohjalta voisi jatkaa samalla metodilla ja saman materiaalin parissa. Tässä tutkielmassa käytiin läpi kuitenkin vain viisi aspektiparia eli kymmenen verbiä, joten aspektipareja riittäisi tutkittavaksi vielä suuri määrä. Tosin tutkimusmateriaalin ollessa hyvin rajattu saattaisi suurelle osalle verbeistä olla vain pieni määrä tapauksia tutkittavaksi. Aspektien lisäksi voitaisiin myös antaa lisää huomiota verbien aikamuotojen tutkimiselle, joilla saattaa olla merkitystä aspektin tai tilanteen keston kannalta. Suomen kielessä mennyttä aikaa voidaan ilmaista kolmella aikamuodolla, kun taas venäjän kielessä vain yhdellä, joten konekääntimen on pystyttävä valitsemaan sopiva aikamuoto käännösvastineeseen.

Tässä tutkimuksessa käytettiin konekääntimenä Google-kääntäjää. Luonnollisesti Google-kääntäjän sijaan voitaisiin valita jokin toinen konekäännin ja tarkastella, millaisia tuloksia sillä saataisiin. Voitaisiin myös käyttää kahta tai jopa useampaa konekäännintä rinnakkain ja vertailla, miten tulokset eroaisivat eri käänninten välillä. Konekääntimen kehitystä olisi toki mahdollista myös tarkastella toteuttamalla sama tutkimus uudelleen jonain myöhempänä ajankohtana.

Tutkimusmateriaalin kokoa olisi tietenkin myös mahdollista kasvattaa. Artikkeleita voitaisiin ottaa korpukseen mukaan nykyistä selvästi suurempi määrä. Tutkimusmateriaali voitaisiin myös kokonaan korvata jollakin toisella ja tutkia joitakin aivan erilaisia ja erityyppisiä tekstejä. Laajemmassa tutkimuksessa olisi mahdollista käyttää esimerkiksi suuria rinnakkaiskorpuksia, jotka voisivat sisältää laajasti erityyppisiä tekstejä ja niiden käännöksiä. Jos tekstit olisivat laajasti otettu erilaisista lähteistä, niiden perusteella voitaisiin saada tuloksia, jotka olisivat

helpommin yleistettävissä. Voisi olla mielenkiintoista tutkia muun muassa sitä, miten konekäännin selviytyisi esimerkiksi kaunokirjallisuuden kääntämisestä ja miten aspektien merkitys kyseisen tyyppisessä tekstissä välittyisi. Kaunokirjallisen tekstit kuitenkin eroavat tyyllillisesti huomattavasti asiateksteistä, joita esimerkiksi tässä tutkimuksessa käytetyt henkilöartikkelit ovat. Konekäännösten tuottaminen ja rinnakkaiscorpusten kokoaminen on kuitenkin aikaa vievää toimintaa, joten laajempaa tutkimusta ei kannata lähteä suorittamaan kovin vähäisillä resursseilla.

LÄHTEET

Aineistolähteet

Tutkimusmateriaalina käytetyt artikkelit ovat listattuna liitteessä 1.

Kirjallisuuslähteet

Alestalo, Marjatta ja Eija Hämmäläinen 2017. *Venäjän kielioppi: venäjää aikuisille. 7.*, korjattu painos. Helsinki: Finn Lectura.

Askonen, Ebba 2001. *Objektin aspektuaalinen sijanvalinta*. Oulun yliopisto, väitöskirja.

Bondarko A. V. 1971. *Vid i vremâ russkogo glagola (značenie i upotreblenie)*. M.: Prosvešenie.

Bowker, Lynne, and Jairo Buitrago Ciro 2019. *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Bingley: Emerald Publishing Limited.

Caswell, Isaac, and Bowen Liang 2020. Recent Advances in Google Translate. <https://ai.googleblog.com/2020/06/recent-advances-in-google-translate.html>. (3.1.2023)

Caswell, Isaac 2022. Google Translate learns 24 new languages. <https://blog.google/products/translate/24-new-languages/>. (3.1.2023)

Comrie, Bernard 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Dahl, Östen ja Fred Karlsson 1976. Verbien aspektit ja objektin sijanmerkintä: vertailua suomen ja venäjän välillä. *Sananjalka* 18 (1), 28–52.

Dahl, Östen 1985. *Tense and Aspect Systems*. Oxford: Blackwell.

Dingley, John 2003. The Implementation of Aspect in Russian and Finnish: Similarities and Differences. *Canadian Slavonic Papers* 45 (1-2), 1–10.

Dorst, Aletta G., Susana Valdez, and Heather Bouman 2022. Machine translation in the multilingual classroom: how, when and why do humanities students at a Dutch university use machine translation? *Translation and Translanguaging in Multilingual Contexts* 8 (1), 49–66.

Forcada, Mikel L. 2017. Making sense of neural machine translation. *Translation Spaces* 6:2, 291–309.

Hakulinen, Auli ja Fred Karlsson 1979. *Nykysuomen lauseoppia*. Helsinki: Suomalaisen Kirjallisuuden Seura.

- Hiippala, Tuomo 2015. Tekstilingvistiikkaa kääntäjille.
<https://www.mv.helsinki.fi/home/thiippal/publications/2015-tekstilingvistiikkaa.pdf>.
 (13.3.2023)
- ISO 2017. *Translation services – Post-editing of machine translation output – Requirements* (ISO 18587:2017(E)). Geneva: International Organization for Standardization.
- Juva, Kersti 2019. *Löytöretki suomeen*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Kachelina, Natalia 2014. *Suomen kielen perfektin semanttiset tyypit venäjänkielisen käännöskorpuksen pohjalta*. Tampereen yliopisto, pro gradu -tutkielma.
<https://urn.fi/URN:NBN:fi:uta-201405131409>.
- Kenny, Dorothy 2022. Human and machine translation. In Dorothy Kenny (ed.), *Machine translation for everyone: Empowering users in the age of artificial intelligence*. Berlin: Language Science Press, 23–49.
- Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel 2014. The Sketch Engine: ten years on. *Lexicography* (Berlin) 1 (1): 7–36.
- Kirvesmäki, Arja, Hannes Viimaranta ja Julia Ekman 2021. *Venäjän kielioppi*. Helsinki: Helsingin yliopisto.
- Koehn, Philipp 2020. *Neural Machine Translation*. Cambridge: Cambridge University Press.
- Kyröläinen, Aki-Juhani ja Veronika Laippala 2020. Määrällinen korpuslingvistiikka. Teoksessa Milla Luodonpää-Manni, Markus Hamunen, Reetta Konstenius, Matti Miestamo, Urpo Nikanne ja Kaius Sinnemäki (toim.), *Kielentutkimuksen menetelmiä I–IV*. Helsinki: Suomalaisen Kirjallisuuden Seura, 487–524.
- Laippala, Veronika ja Minna Palander-Collin 2020. Korpusaineistot. Teoksessa Milla Luodonpää-Manni, Markus Hamunen, Reetta Konstenius, Matti Miestamo, Urpo Nikanne ja Kaius Sinnemäki (toim.), *Kielentutkimuksen menetelmiä I–IV*. Helsinki: Suomalaisen Kirjallisuuden Seura, 460–486.
- Lapshinova-Koltunski, Ekaterina 2017. Cohesion and translation variation: Corpus-based analysis of translation varieties. In Katrin Menzel, Ekaterina Lapshinova-Koltunski & Kerstin Kunz (ed.), *New perspectives on cohesion and coherence: Implications for translation*. Berlin: Language Science Press, 106–130.
- Marco, Josep 2019. Living with parallel corpora: The potentials and limitations of their use in translation research. In Irene Doval Reixa, and María Teresa Sánchez Nieto (ed.), *Parallel Corpora for Contrastive and Translation Studies: New Resources and Applications*. Amsterdam: John Benjamins Publishing Company, 39–56.
- Menzel, Katrin, Ekaterina Lapshinova-Koltunski, and Kerstin Kunz 2017. Cohesion and coherence in multilingual contexts. In Katrin Menzel, Ekaterina Lapshinova-Koltunski & Kerstin Kunz (ed.), *New perspectives on cohesion and coherence: Implications for translation*. Berlin: Language Science Press, 1–10.

- Mikhailov, Mikhail 2003. *Parallel'nye korpusa hudožestvennyh tekstov: principy sostavleniâ i vozmožnosti primeneniâ v lingvističeskikh i perevodovedčeskikh issledovaniâh na primere russko-finskogo parallel'nogo korpusa hudožestvennyh tekstov*. Universitet Tampere, dissertaciâ.
- Mikhailov, Mikhail, and Robert Cooper 2016. *Corpus Linguistics for Translation and Contrastive Studies: A Guide for Research*. London: Routledge.
- Mustajoki, Arto ja Helena Niemensivu 1990. *Venäjän kielen vaikeuksia*. Helsinki: Yliopistopaino.
- Nurminen, Mary 2021. *Investigating the Influence of Context in the Use and Reception of Raw Machine Translation*. Tampere University, Ph.D. dissertation.
- Nurminen, Salla 2015. *Lauseen aspektin ilmaiseminen suomen kielessä. Tutkimus olla-, tehdä- ja tulla-verbillisistä tekstilauseista käyttökonteksteissaan*. Turun yliopisto, väitöskirja.
- Pérez-Ortiz, Juan Antonio, Mikel L. Forcada, and Felipe Sánchez-Martínez 2022. How neural machine translation works. In Dorothy Kenny (ed.), *Machine translation for everyone: Empowering users in the age of artificial intelligence*. Berlin: Language Science Press, 141–164.
- Pitman, Jeff 2021. Google Translate: One billion installs, one billion stories. <https://blog.google/products/translate/one-billion-installs/>. (3.1.2023)
- Schuster, Mike, Melvin Johnson, and Nikhil Thorat 2016. Zero-Shot Translation with Google's Multilingual Neural Machine Translation System. <https://ai.googleblog.com/2016/11/zero-shot-translation-with-googles.html>. (3.1.2023)
- Sim Smith, Karin, and Lucia Specia 2017. Examining lexical coherence in a multilingual setting. In Katrin Menzel, Ekaterina Lapshinova-Koltunski & Kerstin Kunz (ed.), *New perspectives on cohesion and coherence: Implications for translation*. Berlin: Language Science Press, 131–150.
- Sinclair, John 2004. Corpus and Text — Basic Principles. In Martin Wynne (ed.), *Developing Linguistic Corpora: A Guide to Good Practice*. AHDS, 5–24.
- Sonnenhauser, Barbara, and Robert Zangenfeind 2016. Not by chance. Russian aspect in rule-based machine translation. *Russian Linguistics* 40(3), 199–213.
- Tieteen termipankki. Nimitys: pivot-kieli. <https://tieteentermipankki.fi/wiki/Nimitys:pivot-kieli>. (3.1.2023)
- Tiittula, Liisa ja Leena Kolehmainen 2020. Käännöstiede ja sen menetelmät. Teoksessa Milla Luodonpää-Manni, Markus Hamunen, Reetta Konstenius, Matti Miestamo, Urpo Nikanne ja Kaius Sinnemäki (toim.), *Kielentutkimuksen menetelmiä I–IV*. Helsinki: Suomalaisen Kirjallisuuden Seura, 942–983.
- Tommola, Hannu 1986. *Aspektual'nost' v finskom i russkom âzykah*. Helsinki: Neuvostoliittoinstituutti.

Turovsky, Barak 2016a. Found in translation: More accurate, fluent sentences in Google Translate. <https://blog.google/products/translate/found-translation-more-accurate-fluent-sentences-google-translate/>. (3.1.2023)

Turovsky, Barak 2016b. Ten years of Google Translate. <https://www.blog.google/products/translate/ten-years-of-google-translate/>. (3.1.2023)

VISK = Auli Hakulinen, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen ja Irja Alho 2004. Iso suomen kielioppi. Helsinki: Suomalaisen Kirjallisuuden Seura. Verkko-versio. Saatavissa: <http://scripta.kotus.fi/visk>, URN:ISBN:978-952-5446-35-7. (18.12.2022)

Wade, Terence, and David Gillespie 2011. *A Comprehensive Russian Grammar*. 3rd ed., rev. and expanded. Malden, MA: Wiley-Blackwell.

Zangenfeind, Robert, and Barbara Sonnenhauser 2014. Russian verbal aspect and machine translation. *Dialog-21*, Bekasovo, 4 June 2014 - 8 June 2014, 743–752.

LIITTEET

Liite 1: Luettelo tutkimusmateriaalin artikkeleista

Kaikki artikkelit ovat peräisin venäjänkielisestä Wikipediasta (<https://ru.wikipedia.org/>).

Asfendarova, Gul'sum Džafarovna.

https://ru.wikipedia.org/wiki/Асфендиарова,_Гульсум_Джафаровна. (6.1.2023)

Bobylëv, Boris Andreevič (žurnalist).

[https://ru.wikipedia.org/wiki/Бобылёв,_Борис_Андреевич_\(журналист\)](https://ru.wikipedia.org/wiki/Бобылёв,_Борис_Андреевич_(журналист)). (6.1.2023)

Danin, Dmitrij Âkovlevič. https://ru.wikipedia.org/wiki/Данин,_Дмитрий_Яковлевич. (6.1.2023)

Èdel', Ivan Vasil'evič. https://ru.wikipedia.org/wiki/Эдель,_Иван_Васильевич. (6.1.2023)

Evelinov, Boris Efimovič. https://ru.wikipedia.org/wiki/Евелинов,_Борис_Ефимович. (6.1.2023)

Foregger, Nikolaj Mihajlovič.

https://ru.wikipedia.org/wiki/Фореггер,_Николай_Михайлович. (6.1.2023)

Glinka, Mihail Ivanovič. https://ru.wikipedia.org/wiki/Глинка,_Михаил_Иванович. (29.1.2023)

Kavos, Katerino Al'bertovič. https://ru.wikipedia.org/wiki/Кавос,_Катерино_Альбертович. (29.1.2023)

Keržakov, Mihail Anatol'evič.

https://ru.wikipedia.org/wiki/Кержаков,_Михаил_Анатольевич. (6.1.2023)

Kětcer, Amanda. https://ru.wikipedia.org/wiki/Кётцер,_Аманда. (6.1.2023)

Kravčenko, Grigorij Panteleevič.

https://ru.wikipedia.org/wiki/Кравченко,_Григорий_Пантелеевич. (6.1.2023)

Lazarev, Sergej Vâčeslavovič.

https://ru.wikipedia.org/wiki/Лазарев,_Сергей_Вячеславович. (6.1.2023)

Lebedev, Sergej Ivanovič (biolog).

[https://ru.wikipedia.org/wiki/Лебедев,_Сергей_Иванович_\(биолог\)](https://ru.wikipedia.org/wiki/Лебедев,_Сергей_Иванович_(биолог)). (6.1.2023)

Linkol'n, Avraam. https://ru.wikipedia.org/wiki/Линкольн,_Авраам. (6.1.2023)

Lopyrin, Anatolij Ivanovič. https://ru.wikipedia.org/wiki/Лопырин,_Анатолий_Иванович. (6.1.2023)

Luginbûl', Bernhard. https://ru.wikipedia.org/wiki/Лугинбюль,_Бернхард. (6.1.2023)

- Luk'ânenko, Sergej Vasil'evič. https://ru.wikipedia.org/wiki/Лукьяненко,_Сергей_Васильевич. (6.1.2023)
- Magomaev, Muslim Magometovič. https://ru.wikipedia.org/wiki/Магомаев,_Муслим_Магометович. (29.1.2023)
- Milovanova, Ol'ga Ūr'evna. https://ru.wikipedia.org/wiki/Милованова,_Ольга_Юрьевна. (6.1.2023)
- Müller, Karl Vil'gel'm. https://ru.wikipedia.org/wiki/Мюллер,_Карл_Вильгельм. (6.1.2023)
- Nepogodov, Dmitrij Mihajlovič. https://ru.wikipedia.org/wiki/Непогодов,_Дмитрий_Михайлович. (6.1.2023)
- Paovič, Sandra. https://ru.wikipedia.org/wiki/Паович,_Сандра. (6.1.2023)
- Šafaržova, Luciâ. https://ru.wikipedia.org/wiki/Шафаржова,_Луция. (29.1.2023)
- Šipov, Ivan Pavlovič (1865). [https://ru.wikipedia.org/wiki/Шипов,_Иван_Павлович_\(1865\)](https://ru.wikipedia.org/wiki/Шипов,_Иван_Павлович_(1865)). (6.1.2023)
- Skalozubov, Nikolaj Lukič. https://ru.wikipedia.org/wiki/Скалозубов,_Николай_Лукич. (6.1.2023)
- Strel'cov, Èduard Anatol'evič. https://ru.wikipedia.org/wiki/Стрельцов,_Эдуард_Анатольевич. (29.1.2023)
- Swift, Tejlor. https://ru.wikipedia.org/wiki/Свифт,_Тейлор. (6.1.2023)
- Thorževskij, Ivan Ivanovič. https://ru.wikipedia.org/wiki/Тхоржевский,_Иван_Иванович. (6.1.2023)
- Tokaev, Kasym-Žomart Kemelevič. https://ru.wikipedia.org/wiki/Токаев,_Касым-Жомарт_Кемелевич. (6.1.2023)
- Topalov, Vladislav Mihajlovič. https://ru.wikipedia.org/wiki/Топалов,_Владислав_Михайлович. (6.1.2023)
- Vavič, Mihail Ivanovič. https://ru.wikipedia.org/wiki/Вавич,_Михаил_Иванович. (6.1.2023)
- Vehov, Nikolaj Kuz'mič. https://ru.wikipedia.org/wiki/Вехов,_Николай_Кузьмич. (6.1.2023)
- Volkov, Svâtoslav Semënovič. https://ru.wikipedia.org/wiki/Волков,_Святослав_Семёнович. (6.1.2023)

Liite 2: Esimerkki artikkelista ja konekäännöksestä

Milovanova, Ol'ga Ūr'evna. https://ru.wikipedia.org/wiki/Милованова,_Ольга_Юрьевна.
(6.1.2023)

Alkuperäinen venäjänkielinen artikkeli	Konekäännös suomeksi
<p>Ольга Юрьевна Милова́нова (род. 14 сентября 1986 года в Ленинграде) — российская профессиональная бильярдистка, мастер спорта России международного класса, четырёхкратная чемпионка мира в различных дисциплинах русского бильярда.</p>	<p>Olga Jurievna Milovanova (s. 14. syyskuuta 1986 Leningradissa) on venäläinen ammattibiljardipelaaja, Venäjän kansainvälisen luokan mestari, neljäkertainen maailmanmestari venäläisen biljardin eri lajeissa.</p>
<p>Биография и карьера</p>	<p>Elämäkerta ja ura</p>
<p>Ольга родилась в спортивной семье — её мама увлекалась волейболом, а папа играл в бильярд и был мастером спорта в этой игре. Впервые девочка попробовала играть в 6–7 лет, некоторое время ходила в бильярдную секцию. В детские годы Ольга также активно занималась водным поло, но в 1999 году, в возрасте тринадцати лет сделала решающий выбор в пользу бильярда, а ещё через несколько лет решила заниматься этим видом спорта профессионально. Участвовать в различных турнирах Милованова начала с 2004 года.</p>	<p>Olga syntyi urheiluperheeseen - hänen äitinsä rakasti lentopalloa, ja hänen isänsä pelasi biljardia ja oli urheilun mestari tässä pelissä. Ensimmäistä kertaa tyttö yritti pelata 6–7-vuotiaana, jonkin aikaa hän meni biljardiosastolle. Lapsuudessa Olga oli myös aktiivisesti mukana vesipallossa, mutta vuonna 1999, 13-vuotiaana, hän teki ratkaisevan valinnan biljardin hyväksi, ja muutamaa vuotta myöhemmin hän päätti pelata tätä urheilua ammattimaisesti. Milovanova aloitti osallistumisen erilaisiin turnauksiin vuonna 2004.</p>

<p>Первый успех Ольги на крупных турнирах национального и международного уровня среди женщин — победа на чемпионате России и финал Кубка Европы соответственно в 2007 году. В 2008 году она вышла в полуфинал чемпионата мира в свободной пирамиде, а в 2009 победила на этом же турнире и получила звание мастера спорта международного класса. Тогда же Милованова начала заниматься тренерской деятельностью (и продолжает до сих пор).</p>	<p>Olgan ensimmäinen menestys suurissa kansallisissa ja kansainvälisissä naisten turnauksissa oli voitto Venäjän mestaruudesta ja Euroopan Cupin finaalista vuonna 2007. Vuonna 2008 hän saavutti vapaan pyramidin MM-kisojen välieriin, ja vuonna 2009 hän voitti saman turnauksen ja sai kansainvälisen urheilumestarin tittelin. Sitten Milovanova alkoi harjoittaa valmennusta (ja jatkaa tähän päivään asti).</p>
<p>В последующие годы Ольга оставалась одним из сильнейших игроков в русский бильярд среди женщин и выиграла, среди прочего, ещё три чемпионата мира (последний — в 2016 году), чемпионат Европы, командный чемпионат мира (дважды — в 2016 и 2017 годах), Кубок Кремля (в 2011 и 2014 годах), а также неоднократно становилась чемпионкой России и до 2019 года продолжала показывать стабильно высокие результаты на крупнейших соревнованиях. Благодаря четырём победам на чемпионатах мира в настоящее время Ольга Милованова является самым титулованным игроком в пирамиду среди женщин после Дианы Мироновой.</p>	<p>Seuraavina vuosina Olga pysyi yhtenä vahvimista pelaajista Venäjän biljardissa naisten keskuudessa ja voitti muun muassa kolme muuta maailmanmestaruutta (viimeinen vuonna 2016), Euroopan mestaruuden, joukkueiden maailmanmestaruuden (kahdesti vuosina 2016 ja 2017), Kremlin Cupin (vuonna 2011 ja 2014), ja hänestä tuli myös toistuvasti Venäjän mestari ja vuoteen 2019 asti hän jatkoi jatkuvasti korkeita tuloksia suurissa kilpailuissa. Neljän maailmanmestaruusvoiton ansiosta Olga Milovanova on tällä hetkellä Diana Mironovan jälkeen arvostetuin naispyramidipelaaja.</p>

В 2009 году окончила Санкт-Петербургский государственный лесотехнический университет по специальности «бухгалтер». Не замужем.	Vuonna 2009 hän valmistui Pietarin valtion metsätekniikan yliopistosta kirjanpitäjän tutkinnon. Naimaton.
--	---

VENÄJÄNKIELINEN LYHENNELMÄ

ПЕРЕДАЧА ЗНАЧЕНИЯ ВИДА РУССКОГО ГЛАГОЛА В МАШИННОМ ПЕРЕВОДЕ НА ФИНСКИЙ ЯЗЫК

Введение

Машинный перевод (МП) в настоящее время является очень актуальной темой в переводоведении, поскольку использование МП быстро растет. Качество перевода значительно улучшилось, особенно благодаря переводу с использованием нейронных сетей. Машинным переводом пользуются как в профессиональном переводе, так и для повседневных нужд.

Однако машинные переводчики пока не способны выдавать идеальный текст. Машинный переводчик также может попытаться сохранить в переводе грамматику исходного текста. Некоторые явления языка могут быть трудными для машинных переводчиков. Поэтому нашей целью в данной работе является изучение одного важного явления – аспектуальности.

В настоящей работе мы рассматриваем передачу значения грамматического вида русского глагола в машинном переводе на финский язык. В разных языках существуют разные способы выражения вида. В русском языке вид систематически появляется в глаголах грамматическими средствами. В финском же языке грамматического вида нет, а особенности протекания действия выражается другими средствами (Nurminen 2015, 14). Поэтому между русским и финским языками существует значительная разница в выражении аспектуальности. Вид глагола русского языка часто является трудной темой для финских студентов, изучающих русский язык. Поэтому интересно выяснить, сложна ли эта тема и для машинного переводчика. В данной работе мы сравниваем, как выражение вида различается в этих языках. Кроме того, мы обсуждаем частичные соответствия, существующие между этими средствами.

Исследовательский вопрос данной работы заключается в том, чтобы выяснить, как машинному переводчику удастся передать на финском языке значение вида русских глаголов. Мы выбрали Google Переводчик (Google Translate), потому что это один из самых популярных машинных переводчиков в настоящее время. Кроме того, его работа в настоящее время основана на нейронных сетях. Гипотеза исследования заключается в том, что машинный переводчик хорошо справляется лишь с некоторыми переводами. Машинные переводчики по-прежнему часто допускают ошибки, поэтому естественно предположить, что это произойдет и в наших исследованиях.

Исследовательским материалом для данной работы послужил сборник статей русскоязычной Википедии и финские машинные переводы этих статей, сделанные автором настоящей работы. Из вышеуказанных материалов был составлен параллельный корпус. Содержание корпуса анализируется с помощью корпусной программы Sketch Engine. Предметом изучения являются русские глаголы в текстах и их соответствия в финских переводах.

В настоящем реферате мы кратко представляем теорию машинного перевода и аспектуальности. Кроме того, мы более подробно представляем материал и метод исследования. Однако наибольшее внимание мы уделяем результатам исследований. Мы рассматриваем как примеры текстов, так и статистику изучаемых глаголов. Наконец, мы рассмотрим надежность и полноту исследования, а также возможности дальнейших исследований.

Машинный перевод и аспектуальность

Машинный перевод означает автоматический перевод текста с одного естественного языка на другой с помощью компьютера (ISO 2017, 1). Ранние исследования проводил математик Уоррен Уивер в 1949 году, но из-за различных проблем развитие машинного перевода замедлилось в 1960-х годах. С 1980-х годов интерес снова возрос, а в 1990-х годах были разработаны программы автоматизированного перевода, такие как программы памяти переводов (Bowker & Buitrago Ciro 2019, 37–38).

Этапами развития в истории машинного перевода были машинный перевод на основе правил (RBMT), статистический машинный перевод (SMT) и, совсем недавно, нейронный машинный перевод (NMT). RBMT был основан на грамматических правилах,

SMT основан на вычислении вероятностей и параллельных корпусах, а NMT использует нейронные сети. Сегодня доминируют переводчики на основе нейронных сетей, такие как Google Переводчик (Bowker & Vuitrago-Cirio 2019, 39, 44–45; Kenny 2022, 37).

Машинные переводы широко используются либо в качестве неотредактированного перевода для обеспечения понимания, либо при общении между людьми, например, в социальных сетях (Nurminen 2021, 30, 40–41). Качество перевода может быть улучшено путем постредактирования МП (ISO 2017, 2). Машинный перевод чаще всего используется людьми, которые используют неотредактированный перевод для понимания текстов на иностранном языке, в то время как доля профессиональных переводчиков и постредакторов небольшая (Nurminen 2021, 23–24).

Работы по созданию нейронного машинного перевода начались уже в 1980-х и 1990-х годах, но только в 2010-х годах он стал доминирующей моделью с достаточными данными и вычислительной мощностью. Сегодня большинство разработок в области машинного перевода сосредоточено на нейронном машинном переводе. (Koehn 2020, 39–40.) Нейронный машинный перевод основан на искусственных нейронных сетях, имитирующих функционирование человеческого мозга, где слова представлены в виде векторов. Глубокое обучение использует многослойные нейронные сети. (Kenny 2022, 33, 40–41.) Нейронный машинный перевод считается лучшей моделью машинного перевода на сегодняшний день. Нейросетевые переводчики способны выполнять переводы высокого качества, но они также могут допускать ошибки. Они не обязательно хорошо учитывают связи между предложениями (там же, 43).

Когезия и когерентность имеют решающее значение для понятности и единства текста. Когезия образует связи между словами и предложениями в тексте как грамматически, так и семантически. Например, использование местоимений отсылает к ранее упомянутым вещам, создавая семантическую связность. Когерентность же связана с логическим единством и последовательностью утверждений, представленных в тексте. Как когезия, так и когерентность важны как для человеческого, так и для машинного перевода, гарантируя, что перевод сохранит понятность и контекст исходного текста (Niippala 2015, 7–9).

Вид или аспект (аспектуальность) — лингвистическое понятие, связанное с выражением длительности ситуации в языке. Его часто делят на ограниченную (совершенную) и

неограниченную (несовершенную) формы, которые относятся к завершению или неполноте ситуации. (VISK § 1498.) Вид также может описывать результативность деятельности, т. е. изменение, которое она вызывает в цели (Askonen 2001, 38). Взгляд на ситуацию может различаться в зависимости от структуры и значения языка, что делает вид многомерным понятием в лингвистике. Вид часто рассматривается вместе с временем глагола (Nurminen 2015, 52).

В русском языке вид — это грамматическая категория, разделяющая глаголы на две основные категории: совершенный вид и несовершенный вид. Совершенный вид описывает ограниченное во времени действие, которое было завершено, тогда как несовершенный вид относится к непрерывности, повторению или незавершенности. Большинство русских глаголов имеют оба вида, образуя видовую пару. Использование и понимание видов может быть сложной задачей для изучающего русский язык, поскольку одно и то же лексическое значение может иметь два разных вида. (Kirvesmäki *ум.* 2021, 141.) Кроме того, есть глаголы, не имеющие аналога в другом аспекте, и их называют одновидовыми глаголами. С другой стороны, есть также двувидовые глаголы, которые имеют значения как совершенного, так и несовершенного вида. Эти глаголы нельзя распознать только по форме глагола, но вид необходимо интерпретировать с помощью контекста (Wade & Gillespie 2011, 269–270).

Видовые пары могут быть сформированы разными способами, например, с использованием префикса, суффикса или изменения основы. Например, от глагола *писать* глагол совершенного вида *написать* образуется путем добавления приставки *на-*. Однако важно отметить, что добавление приставки может также изменить лексическое значение глагола, что отличает его от исходной видовой пары. (Kirvesmäki *ум.* 2021, 143–145.) Что касается времен, русские глаголы несовершенного вида имеют форму всех трех времен (настоящего, прошедшего, будущего), тогда как у глаголов совершенного вида только две временные формы: прошедшего и будущего времени. У глаголов совершенного вида нет формы настоящего времени, потому что они не описывают действие, происходящее в момент речи. (Бондарко 1971, 50.)

В финском языке вид выражается множеством различных способов, например, выбором падежа объекта, времен глаголов, а также различных глаголов и их производных. Выбор падежа объекта — важнейший способ выражения вида: партитив обозначает незавершенное действие, а генетив описывает действие, завершившееся или приведшее

к результату. (Askonen 2001, 39; Juva 2019, 128.) Выбор падежа объекта также важен для разграничения вида: тотальный объект ограничивает вид, тогда как разделительный объект придает предложению открытый, неограниченный тон (VISK § 930).

В финском языке четыре времени: настоящее, имперфект, перфект и плюсквамперфект (Nurminen 2015, 19). Они также выражают вид и временные отношения события. Имперфект относится к прошлому событию, а перфект описывает ситуацию, которая продолжается или влияет на настоящее (Juva 2019, 130). Производные глаголов также можно использовать для аспектной модификации значения глагола (Askonen 2001, 40). Таким образом, в финском языке вид выражается различными способами с помощью падежа объекта, времен и глагольных производных, что придает предложениям разные значимые тона и контексты в зависимости от того, как конструируется видовое значение.

Между русским и финским языками имеются существенные различия в выразительном виде. В русском языке вид — грамматическая категория, в которой совершенный вид обозначает ограниченное и ведущее к результату действие, а несовершенный вид выражает неограниченное и незавершенное действие. В финском языке соответствующая разница в значении передается через выбор падежа объекта: тотальный объект часто относится к ограниченному действию, тогда как разделительный объект часто описывает неограниченное действие. Что касается времен глаголов, то финский и русский языки также существенно различаются. Настоящее время – общее для двух языков время, но в финском языке события в будущем также выражаются настоящим временем, а в русском языке отдельно существует будущее время. Что касается прошедшего времени, то в русском языке есть только одна форма, а в финском три формы: имперфект, перфект и плюсквамперфект. С помощью этих форм в финском языке выражаются продолжительность, действие и результат ситуации, что частично соответствует задаче русского вида, но прямого соответствия нет. Подводя итог, можно констатировать, что хотя средства выражения вида в финском и русском языках различаются, имеются и частичные сходства. Оба языка стремятся выразить разные значения и временные отношения событий, но используемые средства и грамматические конструкции различны, что может вызвать проблемы в изучении языка и переводе.

В данных ниже таблицах показаны частичные соответствия между временами финского и русского языков и выражением вида в этих двух языках.

Таблица 16. Частичные соответствия между выбором финского объекта и видом русского глагола.

	в русском несовершенный вид	в русском совершенный вид
в финском разделительный объект в утвердительном предложении	чаще	
в финском разделительный объект в отрицательном предложении		чаще
в финском разделительном объект, выражающем неопределенность	возможны оба	
в финском винительный объект		чаще
в финском определенные лексические значения падежа объекта	возможны оба	

Таблица 17. Частичные соответствия финских времен глагола русским временам и видам.

	настоящее время финского	имперфект финского	перфект финского
будущее простое русского	совершенный вид		совершенный вид
будущее сложное русского	несовершенный вид		
настоящее время русского	несовершенный вид	несовершенный вид	несовершенный вид
прошедшее время русского		несовершенный / совершенный вид	несовершенный / совершенный вид

Метод и материал исследования

В настоящем исследовании мы используем корпусный метод. Корпус – это коллекция текстов, собранных для лингвистических исследований (Sinclair 2004, 23). Исследователи могут либо составить свои собственные корпуса, либо использовать существующие корпуса (Lairpala & Palander-Collin 2020, 461). Параллельные корпуса часто используются в переводоведении. Это многоязычные корпуса, состоящие из

оригинальных текстов и переводов этих текстов на один или несколько языков. (Tiittula & Kolehmainen 2020, 955–956.) В данном исследовании мы используем составленный нами двуязычный параллельный корпус, где все оригинальные тексты представлены на русском языке, а переводы – на финском языке.

Исследовательским материалом для данного исследования является ограниченная подборка статей из русскоязычной Википедии и машинных переводов этих статей на финский язык. Мы выбирали статьи из Википедии с помощью функции «Случайная статья», однако таким образом, чтобы статьи были достаточной длины и содержали как можно больше подходящего для исследования материала. Материал был собран в январе 2023 года. Список статей, включенных в материал исследования, находится в приложении 1 к настоящей работе. Кроме того, в приложении 2 приведен пример одной из статей, включенных в материал исследования, и машинный перевод этой статьи. Мы выполнили машинный перевод статей на финский язык с помощью Google Переводчика. Мы собрали параллельный корпус статей и машинных переводов с помощью корпусной программы Sketch Engine. Собранный материал также был проанализирован и обработан с использованием этой программы. Всего материал содержит 33 русскоязычные статьи Википедии и 33 машинных перевода этих статей на финский язык. Размер корпуса, содержащего оригинальные русские статьи, составляет 54 519 словоупотреблений, а размер корпуса, содержащего финские машинные переводы, — 45 298 словоупотреблений. Размеры корпусов достаточны для данного исследования.

Наше исследование продвигалось следующим образом: с помощью корпусной программы мы искали наиболее часто встречающиеся в текстах глаголы. Программа позволяет получить частотные списки, показывающие эту информацию. Более подробно мы рассматриваем тексты с помощью функции конкорданса, которая позволяет нам изучить контексты, в которых появляются глаголы. С помощью функции параллельного конкорданса мы могли одновременно рассматривать и русский текст, и финский перевод.

Используя эти функции, мы выбрали глаголы, на которых остановимся в данном исследовании. С помощью частотного списка корпуса русского языка мы нашли пятнадцать видовых пар глагола, переводные соответствия глаголов которых на финском языке мы начали рассматривать (таблица 6). Все найденные переводные соответствия мы собрали в таблице 8. Из этих пятнадцати видовых пар мы выбрали пять

пар для исследований. Мы выбрали глаголы, переводные соответствия которых были ясны и не слишком двусмысленны. Например, такими глаголами были глаголы движения, для которых часто существовало множество различных переводных соответствий. Для исследования были выбраны следующие видовые пары:

- выигрывать / выиграть
- выпускать / выпустить
- играть / сыграть
- начинать / начать
- писать / написать

Финские соответствия глаголов, содержащихся в этих видовых парах, приведены в таблице 10. Исследованные глаголы встречаются в корпусе в общей сложности 331 раз. Из них в 164 случаях вид несовершенный и в 167 случаях совершенный. Таким образом, глаголы обоих видов встречаются почти одинаково часто.

Основные результаты исследования

Для всех пяти видовых пар мы рассмотрели в корпусе все случаи, в которых встречаются глаголы, содержащиеся в этих видовых парах. Мы классифицировали случаи на четыре группы в зависимости от передачи значения вида. К первой группе относятся случаи, когда значение вида хорошо передано в переводе. Вторую группу составляют случаи, в которых значение вида передается лишь частично. Третью группу составляют случаи, когда значение вида вообще не передается. Четвертую группу составляют неоднозначные случаи, по которым нельзя с уверенностью сказать, передано значение вида или нет. Мы также классифицировали случаи, основанные на качестве машинного перевода, на три группы: хорошие, слабые и непригодные для использования.

Как мы заявляли ранее, нам предстояло исследовать в общей сложности 331 случай. Всего в 326 случаях значение вида глагола было хорошо передано в переводе. Таким образом, всего было всего пять случаев неудачного перевода. Значение было частично передано в двух случаях. В одном случае значение вообще не было передано. Мы отнесли два случая к категории неясных. Значение вида было передано хорошо в 98,5% случаев. Глаголы несовершенного вида показали себя в исследовании немного лучше, и в 98,8% их значения были переданы хорошо. Для глаголов совершенного вида

соответствующий показатель составил 98,2%. С другой стороны, единственный случай, когда значение вида не было передано, был среди глаголов несовершенного вида.

Анализируя качество машинных переводов, мы заметили, что 57,1% переводов были классифицированы нами как хорошие. Однако доля слабых переводов (40,8%) также значительна. Доля непригодных для использования переводов составила всего 2,1%. Для восьми глаголов из десяти доля хороших переводов была выше доли слабых. Таким образом, только для двух глаголов доля слабых переводов была выше. Для глаголов несовершенного вида качество перевода оказалось лучше, чем для глаголов совершенного вида. В случаях, содержащих глагол несовершенного вида, качество перевода было хорошим в 62,8% случаев, тогда как для глаголов совершенного вида соответствующая доля составила 51,5%. Все переводы, отнесенные к непригодным для использования, относились к числу случаев, содержащих глаголы совершенного вида. Однако здесь стоит напомнить, что слабое качество машинных переводов не обязательно было связано с изучаемым глаголом, но слабое качество могло быть вызвано и другими причинами. Однако похоже, что глаголы несовершенного вида показали себя лучше в исследовании как с точки зрения передачи значения вида, так и с точки зрения качества машинного перевода.

Далее мы представляем отдельные случаи, которые представляют случаи, отнесенные к различным группам как можно шире. Естественно, мы больше обращаем внимание на те случаи, когда были неясности или другие проблемы. Однако мы начнем с самых простых случаев.

(79) <<Я пишу песни, и мой голос -- всего лишь способ донести эти тексты>>

(80) "Kirjoitan kappaleita, ja ääneni on vain tapa saada nuo sanat perille."

Вот пример случая, когда значение вида передано хорошо и качество перевода хорошее. В финском предложении объект находится в партитиве, поэтому это нерезультативный аспект, соответствующий русскому несовершенному виду.

(81) В дальнейшем, уже как актёрский дуэт, Михаил Вавич и Татьяна Павлова сыграли в целом ряде отечественных и зарубежных кинофильмов.

(82) Tulevaisuudessa Mikhail Vavich ja Tatjana Pavlova soittivat jo näyttölevänä duetona useissa kotimaisissa ja ulkomaisissa elokuvissa.

В данном случае налицо ошибка выбора финского соответствия. Глагол *soittivat* следует заменить глаголом *näyttelevät*. Однако слабое качество перевода не влияет на передачу значения вида.

- (83) Во время грунтового сезона Шафаржова выиграла домашний 100-тысячник из цикла ITF в Праге.
- (84) Savikauden aikana Shafarova voitti Prahassa ITF -kierroksen 100 000. kotikiirroksen.

Вот пример случая, когда перевод непригоден для использования. Первоначальный смысл предложения не передан. Однако значение вида передано. Аспект финского предложения результативный, поскольку объект предложения стоит в винительном падеже.

- (85) В конце ноября Big Machine Records выпустили плейлист Reputation Stadium Tour для стриминговых сервисов.
- (86) Marraskuun lopun suuri kone Records julkaisee Maine- soittolista Stadion Kierros suoratoistopalveluihin.

На наш взгляд, значение вида в данном случае передается лишь частично. В переводе объект стоит в именительном падеже, что в данном контексте грамматически неверно. Качество перевода непригодное, поскольку английские названия частично переведены на финский. Интересно, что в финском переводе присутствует настоящее время, хотя в русском тексте есть прошедшее время.

- (87) Во второй половине сезона чешка выигрывает свой второй турнир WTA -- в Форест-Хилсе и дошла до полуфинала на 75-тысячнике в Пуатье.
- (88) Kauden toisella puoliskolla tšekki voitti toisen WTA-turnauksensa Forest Hillsissä ja pääsi semifinaaliin 75 000:lla Poitiersissa.

В русском тексте глагол стоит в настоящем времени, а в финском — в прошедшем времени. В русском предложении вид глагола несовершенный. В финском предложении объект стоит в винительном падеже, поэтому аспект предложения результативный. Поэтому в данном примере исходное значение вида вообще не передается.

- (89) Выступала в 2005 -- 2007 годах в Германии за команду <<Анрохт> >, позже играла за французский <<Мондевилль> >.
- (90) Hän pelasi Saksassa vuosina 2005-2007 Anrochtin joukkueessa, myöhemmin ranskalaisessa Mondevillessä.

В русском примере есть два глагола, а в финском только один. Машинный переводчик перевел глагол *выступала* как *pelasi*. Последний глагол *играла* практически исчез из финского перевода. С другой стороны, это также можно интерпретировать как означающее, что первый глагол был опущен и в перевод включен только последний глагол. В данном случае глагол *выступать* можно было бы перевести, например, глаголом *edustaa (joukkuetta)*. Однако перевод хорош и без этого глагола. Поскольку мы не можем знать логику переводчика, лежащую в основе перевода, мы не можем сказать,

передается ли значение вида для изучаемого глагола. Поэтому данный случай сейчас остается неясным.

Заключение

В настоящей работе мы рассмотрели передачу значения вида русского глагола в финских машинных переводах. Целью было выяснить, насколько хорошо машинный переводчик справляется с передачей значения. Мы представили понятие вида и способы его выражения в русском и финском языках. Мы также сравнили выражение вида в этих языках и представили сходства и различия рассматриваемых средств выражения.

Исследование мы провели корпусным методом. Мы собрали параллельный корпус, содержащий тексты на русском языке и их машинные переводы на финский язык. Сначала мы изучали часто встречающиеся в текстах глаголы и их переводные соответствия с помощью корпусной программы. В итоге из материала мы выбрали глаголы для изучения. Всего нужно было изучить десять глаголов, и они образовали в общей сложности пять видовых пар. На примерах мы представили, как машинный переводчик передает значение вида. В то же время мы также обратили внимание на качество машинного перевода.

Исследовательским вопросом в данном исследовании было выяснить, как машинный переводчик (Google Переводчик) справляется с передачей значения вида русских глаголов на финский язык. Машинный переводчик прекрасно справился с этой задачей, поскольку не менее чем в 98 процентах случаев значение вида было передано хорошо. Среди более чем трехсот случаев было лишь несколько неясных случаев и случаев, в которых значение было передано лишь частично или не было передано вовсе.

Гипотеза в данном исследовании заключалась в том, что машинный переводчик хорошо справляется лишь с некоторыми переводами. Эту гипотезу можно проверить с помощью двух разных показателей, основанных на передаче значения вида и качестве машинного перевода. Что касается передачи значения вида, то гипотеза оказалась не совсем верной, ведь такое подавляющее большинство случаев были успешными. Предполагалось, что передача значения вида создаст некоторые трудности для машинного переводчика. Количество вызовов и трудностей оказалось гораздо меньшим, чем мы ожидали. С

другой стороны, по качеству машинных переводов ситуация более ровная. Доля хороших переводов (57,1%) была выше доли слабых (40,8%) и непригодных (2,1%).

Хотя в данном исследовании машинный переводчик достаточно хорошо справился с передачей значения вида, результаты нельзя обобщить на явление более широко. Однако в итоге эти результаты применимы только к выбранным глаголам, выбранному материалу и выбранному машинному переводчику. Только десять глаголов были включены в окончательные исследования. Если бы для исследования были выбраны разные глаголы, результаты могли бы быть совершенно другими. Однако на основании данного исследования может показаться, что использование машинного переводчика для исследуемого объекта кажется многообещающим.

Материал исследования состоял из ограниченной выборки русскоязычных статей Википедии и их машинных переводов на финский язык. Мнений по поводу этого выбора может быть много. В Википедии было легко искать тексты для исследования, потому что там огромное количество статей. С другой стороны, Википедия — открытая энциклопедия и писать в ней может практически каждый. Поэтому качество текста статей может значительно различаться между статьями или даже внутри статей, поскольку на протяжении многих лет статья могла быть написана большим количеством людей с разными навыками и стилями письма. Разные тексты могли бы привести к совершенно разным результатам.

Однако существует множество возможностей для дальнейшего исследования темы данной работы. Основываясь только на данной работе, можно было бы продолжать использовать тот же метод и тот же материал. Однако в данной работе были рассмотрены только пять видовых пар, то есть десять глаголов, поэтому предстоит изучить большое количество пар. Однако, когда исследовательский материал очень ограничен, может потребоваться лишь небольшое количество случаев для изучения большей части глаголов. Помимо видов, больше внимания можно было бы уделить изучению времен глаголов, что может быть важно с точки зрения вида или продолжительности ситуации. В финском языке прошедшее время может выражаться тремя временами, а в русском — только одним, поэтому машинный переводчик должен уметь выбирать подходящее время для переводного соответствия.

В данном исследовании мы использовали Google Переводчик в качестве машинного переводчика. Мы могли бы выбрать другой машинный переводчик вместо Google Переводчик и посмотреть, какие результаты он даст. Также можно было бы использовать два или даже более машинных переводчика параллельно и сравнивать, насколько результаты будут различаться для разных МП. Конечно, можно было бы также изучить развитие МП, проведя то же исследование еще раз позднее. Машинные переводы данного исследования были выполнены в январе 2023 года. Вполне возможно, что машинный переводчик появился после этого времени.

Конечно, можно было бы также увеличить объем исследовательского материала. В корпус можно было бы включить значительно большее количество статей. Материал исследования также мог быть полностью заменен чем-то другим и изучаться какие-то совершенно другие и разные типы текстов. В более масштабном исследовании можно было бы использовать, например, большие параллельные корпуса, которые могли бы содержать самые разные типы текстов и их переводы. Если бы тексты были широко взяты из разных источников, на их основе можно было бы получить результаты, которые было бы легче обобщить. Было бы интересно изучить, например, как машинный переводчик справится с переводом художественной литературы и как значение вида будет передано в таких текстах.