

**KORPUSAVUSTEINEN DISKURSSINTUTKIMUS VAM-
MAINEN-SANASTA SUOMI24-KESKUSTELUFOORU-
MILLA**

Eero Kallio
Kandidaatintutkielma
Suomen kieli
Kieli- ja viestintätieteiden lai-
tos
Jyväskylän yliopisto
Kevät 2021

JYVÄSKYLÄN YLIOPISTO

Tiedekunta Humanistis-yhteiskuntatieteellinen tiedekunta	Laitos Kieli- ja viestintätieteiden laitos
Tekijä Eero Kallio	
Työn nimi Korpusavusteinen diskurssintutkimus <i>vammainen</i> -sanasta Suomi24-keskustelufoorumilla	
Oppiaine Suomen kieli	Työn laji Kandidaatintutkielma
Aika 04/2021	Sivumäärä 21
<p>Tiivistelmä</p> <p>Tämä tutkimus tarkastelee <i>vammainen</i>-sanaa Suomi24-keskustelufoorumilla. Tutkimuksessa keskityn siihen, millaisissa yhteyksissä <i>vammainen</i>-sanaa käytetään ja millaisia asenteita ja tunteita sanan käyttöön liittyy. Tutkimus hyödyntää korpusavusteista diskurssintutkimusta eli CADS-tutkimusta. Tutkimuksessa yhdistyy kvantitatiivinen sekä kvalitatiivinen tutkimus. Tutkimuskysymykset ovat: Mitkä ovat <i>vammainen</i>-sanon kollokaatit Suomi24-korpuksessa? Millaisia diskurssiprosodioita kollokaateista voidaan päätellä?</p> <p>Tutkimuksen aineistona toimii Suomi24-keskustelufoorumista luotu kirpus, joka kattaa vuodet 2001-2017. Korpus löytyy Kielipankin Korp-palvelusta. Aineisto on kerätty hakusanalla <i>vammainen</i>. Aineisto on tallennettu Exceliin, josta se on siirretty tekstitiedostona AntConc-ohjelmaan, joka on laskenut MI-menetelmällä merkittävimmät kollokaatit. Osa kollokaateista karsittiin niin, että jäljelle jäi 100 kollokaattia, joista diskurssiprosodioita kyettiin päättämään.</p> <p>Tutkimuksessa löydettiin kahdeksan diskurssiprosodiaa: <i>vammaisuuden aste</i>, <i>vammaisten oikeudet</i>, <i>vähemmistöpuhe</i>, <i>vammainen haukkumasanana</i>, <i>synnytys- ja aborttipuhe</i>, <i>syrijintäpuhe</i>, <i>henkilökohtainen suhde vammaisiin</i> sekä <i>vammaisten toimijuus</i>. Lisäksi sekalaisille sanoille oli oma ryhmänsä. Ensimmäiset kolme diskurssiprosodiaa kattoivat suurimman osan kollokaateista, ja ne olivat sävyiltään yleisesti neutraaleja tai positiivisia. Yksi diskurssiprosodia, <i>vammainen haukkumasanana</i>, oli selkeästi negatiivinen. Tutkimustuloksia voidaan käyttää tulevaisuudessa esimerkiksi vertailemalla niitä muunlaisiin korpuksiin.</p>	
Asiasanat korpusavusteinen diskurssintutkimus, diskurssiprosodia, kollokaatio, verkkokeskustelu	
Säilytyspaikka Jyväskylän yliopisto	
Muita tietoja	

TAULUKOT

Taulukko 1. Diskurssiprosodiat ja niissä olevien kollokaattien lukumäärä.	11
Taulukko 2. Diskurssiprosodia <i>vammaisuuden aste</i>	12
Taulukko 3. Diskurssiprosodia <i>vammaisten oikeudet</i>	13
Taulukko 4. Diskurssiprosodia <i>vähemmistöpuhe</i>	15
Taulukko 5. Diskurssiprosodia <i>vammainen haukkumasanana</i>	16
Taulukko 6. Diskurssiprosodia <i>synnytys- ja aborttipuhe</i>	16
Taulukko 7. Diskurssiprosodia <i>syrjintäpuhe</i>	17
Taulukko 8. Diskurssiprosodia <i>henkilökohtainen suhde vammaisiin</i>	17
Taulukko 9. Diskurssiprosodia <i>vammaisten toimijuus</i>	18
Taulukko 10. <i>Sekalaisten ryhmä</i>	18

SISÄLLYS

1	JOHDANTO	1
2	TEORIA.....	3
	2.1 Diskurssintutkimus.....	3
	2.2 Korpuslingvistiikka	4
	2.3 Korpusavusteinen diskurssintutkimus.....	4
3	AINEISTO JA MENETELMÄT	7
	3.1 Suomi24-keskustelufoorumi ja -korpus.....	7
	3.2 Korp-konkordanssityökalu aineistonkeruumenetelmänä	8
	3.3 AntConc-ohjelma apuna aineiston analysoimisessa.....	9
	3.4 Kollokaattien rajausta ja luokittelu	10
4	ANALYYSI JA TULOKSET	11
	4.1 <i>Vammainen</i> -sanalla kollokaattien diskurssiprosodiat	11
	4.1.1 Vammaisuuden aste.....	12
	4.1.2 Vammaisten oikeudet	13
	4.1.3 Vähemmistöpuhe	15
	4.1.4 <i>Vammainen</i> haukkumasanana.....	15
	4.1.5 Synnytys- ja aborttipuhe	16
	4.1.6 Syrjäntäpuhe	17
	4.1.7 Henkilökohtainen suhde vammaisiin	17
	4.1.8 Vammaisten toimijuus.....	18
	4.1.9 Sekalainen.....	18
5	PÄÄTÄNTÖ.....	19
	5.1 Yhteenveto ja johtopäätökset.....	19
	5.2 Arviointi ja mahdollinen jatkotutkimus	20
	LÄHTEET	22

LIITTEET

Liite 1: MI-arvon mukaan järjestetty lista kollokaateista sekä niiden frekvensseistä.

1 JOHDANTO

Suomella on tarve saada tietoa vammaisuudesta. YK on laatinut vammaisten henkilöiden oikeuksia koskevan yleissopimuksen, jonka noudattaminen vaatii tietoa vammaisten ihmisten asemasta yhteiskunnassa. Esimerkiksi THL kerää tietoa vammaisuudesta, jotta vammaisten yhdenvertaisuus yhteiskunnassa toteutuu. Lisäksi vammaisuutta koskeva tieto voi lisätä itsetuntemusta vammaisilla yksilöillä tai ryhmillä. Vammaisuuteen liittyvät asenteet eivät aina ole ilmeisiä. Tutkimuksissa kansalaiset tukevat vammaisuuden vastuun jakamista, mutta tarkemmin tarkasteltuna tuki kohdistuu usein vain samankaltaisiin ihmisiin tai läheisiin. Myös yhteiskunnalliset muutokset ja esimerkiksi ihmisoikeussopimukset luovat tiedontarvetta. (Nurmi-Koikkalainen ym. 2017: 4, 10, 12–13.)

Vammaisiin liittyviä asenteita kannattaa tutkia suomalaisella keskustelufoorumilla, sillä tällainen tutkimus voi antaa laajemmin vastauksia siihen, miten vammaisiin ja esimerkiksi heidän oikeuksiinsa suhtaudutaan Suomessa. Vastaukset voivat olla relevantteja esimerkiksi vammaisten oikeuksien parantamisessa ja tasa-arvon lisäämisessä. Tutkimus voi myös kartoittaa sitä, miten vammaisista puhutaan Suomessa tai ainakin suomenkielisessä sosiaalisessa mediassa. Yksi mahdollinen näkökulma on syrjintään liittyvä. Kotikielen seuran aikakauslehdessä *Virittäjässä* (2/2020) julkaistiin artikkeli, jonka mukaan vastakkainasettelujen hillitsemisen ja yhteiskunnallisen keskusteluilmapiirin kannalta on tärkeää ymmärtää syrjivän kielenkäytön mekanismeja ja riskejä. Näiden ymmärtämisestä voi hyötyä myös keskustelupalstojen ylläpito. (Määttä, Suomalainen & Tuomarla 2020: 210.)

Tarkoitukseni on siis tutkia niitä asenteita ja merkityksiä, joita *vammainen*-sanaan liitetään. Käytän Suomi24:ää kansalaiskeskustelujen aineistona saadakseni aiheesta kuvan, jonka voi toivottavasti yleistää suomalaisiin myös Suomi24-keskustelufoorumin ulkopuolella. Käytännössä Suomi24-keskustelufoorumin aineisto on sen käyttäjien kommentteista koottu korpus. Korpusavusteinen diskurssitutkimus toimi tämän tutkimuksen aineiston keruu- ja analyysimenetelmänä. Tutkimuskysymykseni ovat:

1. Mitkä ovat *vammainen*-sanan kollokaatit Suomi24-korpuksessa?
2. Millaisia diskurssiprosodioita kollokaateista voidaan päätellä?

Korpusavusteista diskurssintutkimusta vammaisiin liittyvästä aiheesta ei ole Suomessa tehty, ja ala on Suomessa vielä suhteellisen nuori. Melko tuore suomalainen esimerkki korpusavusteisesta diskurssintutkimuksesta on Jantunen vuoden 2018 artikkeli Homot ja heterot Suomi24:ssä. Analyysi digitaalisista diskursseista. Artikkelissa Jantunen tutkii homouden ja heterouden diskursseja käyttäen aineistonaan korpusta Suomen suurimmasta keskustelupalstasta (Jantunen 2018a: 3). Jantunen mukaan ensimmäinen suomalainen laaja korpusavusteinen diskurssintutkimus on vuonna 2018 julkaistu japaninsuomalaisten kielipuhetta koskeva analyysi (Jantunen 2018b: 26). Toinen samana vuonna julkaistu tutkimus on Suomi24:n kansa-sanasta tehty korpusavusteinen diskurssianalyysi (Johansson, Jantunen, Heimo, Ahonen & Laippala 2018). Muilla aloilla asenteita vammaisuutta kohtaan on tutkittu ainakin erityispedagogiikassa (esim. Karjalainen 1993). Vammaisten representaatioista mediassa taas on tehty kandidaatintutkielma (Hakala 2018). Tutkimukseni täydentää suomalaisen korpusavusteisen diskurssintutkimuksen kenttää aiheella, jota ei olla kentällä aiemmin tutkittu, sekä vammaisiin liittyvää tutkimusta hyödyntämällä uudenlaista lähestymistapaa.

Tutkimus etenee niin, että aluksi esittelen tutkimuksen teoreettisen viitekehyyksen teoriaosuudessa. Korpusavusteisen diskurssintutkimuksen esitelläkseni minun on ensin esiteltävä korpuslingvistiikka sekä diskurssintutkimus. Tämän jälkeen esittelen aineiston sekä käyttämäni analyysimenetelmät tarkemmin. Sitten esittelen analyysiani ja sen tuloksia. Viimeisessä luvussa annan yhteenvedon tutkimuksesta sekä pohdin jatkotutkimuksen mahdollisuuksia.

2 TEORIA

2.1 Diskurssintutkimus

Diskurssintutkimuksessa on kyse diskurssin tai diskurssien tutkimuksesta. Diskurssi itsessään on laaja käsite, jota käytetään eri aloilla eri tavoilla. Kielitieteissä sillä on esimerkiksi saatettu joskus viitata pelkästään puhuttuun dialogiin kontrastina kirjoitettuun kieleen. Tällöin kyetään keskittymään puhutulle dialogille tyypillisiin piirteisiin kuten puheenvuoroihin. (Fairclough 1992: 3.) Diskurssia voidaan tutkia myös tekstin rakenteellisella tasolla: esimerkiksi ruokareseptit ovat usein jo rakenteensa puolesta tunnistettavia (Baker 2006: 3). Tyypillisemmin diskurssilla kuitenkin viitataan laajempiin puhutun tai kirjoitetun kielen aineistoihin. Tällainen näkökulma korostaa kielen tuottamista ja vastaanottamista eli puhumista tai kirjoittamista ja kuulemistä tai lukemista. Tällöin teksti itsessään on ainoastaan yksi osa diskurssia: diskurssiin kuuluu myös tekstin tulkitseminen ja kielenkäytön tilanteen konteksti. (Fairclough 1992: 3.) Diskurssin voi siis määritellä käytänteiksi, jotka muodostavat merkityksiä tietyn aiheen ympärillä. Koska käytänteitä ja merkityksiä on useita, myös diskursseja voi olla useita. (Baker 2006: 4.) On siis huomioitava, että diskurssi ja yksittäiset diskurssit ovat eri asioita. Myös eri sosiaaliset tilanteet tuottavat erilaisia diskursseja: voimme puhua esimerkiksi luokkahuoneen diskurssista tai sanomalehtidiskurssista (Fairclough 1992: 3).

Diskurssit eivät ainoastaan heijasta tai kuvaa sosiaalisia suhteita tai asenteita vaan myös luovat ja vahvistavat niitä (Fairclough 1992: 3–4.) Tämän vuoksi diskursseilla voi nähdä olevan voimaa, ja ne ovat mielekkäitä tutkimuskohteita. Tämä on tutkimukseni näkökulma diskurssintutkimuksesta: tutkimalla diskurssia *vammainen*-sanan ympärillä pyrin selvittämään, mitä asenteita ja arvoja vammaisiin mahdollisesti liittyy.

Mitä tahansa puheenaihetta ympäröi yleensä useampi diskurssi. Eri diskurssit voivat olla ristiriitaisia, eivätkä ne välttämättä kuvaa tarkasti ainakaan yksittäisen ihmisen persoonallisuutta tai ajatusmaailmaa. Lisäksi diskurssit muuttuvat ja kehittyvät jatkuvasti sekä ovat vuorovaikutuksissa toisiinsa. Näistä syistä diskursseja voi olla vaikea kuvailla. On myös otettava huomioon, että diskurssin nimeäminen tai kuvaileminen tulee aina olemaan jokseenkin subjektiivista. (Baker 2006: 4.) Toinen henkilö voisi tarkastella esimerkiksi minun tutkimukseni aineistoa ja löytää aivan erilaisia diskursseja. On siis tärkeää perustella löytämänsä diskurssit huolellisesti ja avoimesti. Tutkijan kannattaa myös pitää mielessä, että diskurssien ulkopuolelle astuminen voi olla mahdotonta: tiedostetut tai tiedostamattomat diskurssit vaikuttavat diskurssien havaitsemiseen ja nimeämiseen (Baker 2006: 4).

2.2 Korpuslingvistiikka

Kielentutkimuksella oli taipumus suuntautua ennen lähinnä sisäänpäin ja nojata kielentutkijoiden omaan intuitioon ja mieleen. Nykyään tietokoneiden ansiosta kielentutkimuksessa voidaan kuitenkin käyttää valtavia määriä dataa. Korpuslingvistiikan ytimessä onkin suurien tekstimäärien hyödyntäminen, jolloin tutkimusta voidaan perustella objektiivisella aineistolla tutkijan oman intuition sijaan. (Sinclair 1991: 1.) Sana korpus viittaa laajaan sähköiseen aineistoon luonnollista käytettyä kieltä – laajalla voidaan viitata tuhansiin, miljooniin tai jopa miljardeihin sanoihin (Baker 2006: 2).

Korpuslingvistiikan vahvuutena voidaan pitää sitä, että laajojen aineistojen avulla tutkija voi vahvistaa omia tulkintojaan objektiivisin tuloksin. Korpuslingvistinen tutkimus siis hyödyntää kvantitatiivisia menetelmiä. On tosin huomioitava, että myös kvalitatiivinen analyysi on tärkeä osa korpuslingvistisiä menetelmiä hyödyntäviä tutkimuksia. (Baker 2006: 2.)

2.3 Korpusavusteinen diskurssintutkimus

Korpusavusteisessa diskurssintutkimuksessa on kyse diskurssintutkimuksesta, joka hyödyntää sähköisiä korpuksia. Tämän lisäksi se hyödyntää niitä korpuslingvistisiä kvantitatiivisia ja tilastollisia menetelmiä, jotka koetaan tutkimuksen kannalta tarpeellisiksi ja hyödyllisiksi. Tästä juontuu nimitys korpusavusteinen. Tutkimuksessa siis yhdistyy kvantitatiivinen sekä kvalitatiivinen lähestymistapa. (Partington, Duguid & Taylor 2013: 10-11.)

Korpusavusteista diskurssintutkimusta kutsutaan myös nimellä CADS eli *corpus assisted discourse studies*. Termi keksittiin vuonna 2004, mutta CADS-tutkimusta on tehty jo aiemmin, luultavasti jo hieman korpuslingvistiikan syntymän jälkeen. Melko varhainen esimerkki CADS-tutkimuksesta on vuoden 1996 tutkimus, jossa Krishnamurthy tutkii sanoja *ethnic*, *racial* sekä *tribal*. Hän vertailee sanojen käyttöä ja määritelmiä hyödyntäen korpusta, uutislehdistöä sekä sanakirjoja. (Partington, Duguid & Taylor 2013: 10.)

Korpusavusteista diskurssintutkimusta voi hyödyntää monenlaisilla eri aihealueilla. Se ei sido itseään mihinkään tiettyyn diskurssintutkimuksen suuntaukseen, kuten esimerkiksi kriittiseen diskurssintutkimukseen (Partington, Duguid & Taylor 2013: 10). CADS-tutkimuksen tavoitteena voi pitää sellaisten merkitysten paljastamista, jotka eivät ole itsestäänselviä. Kieltä käytetään ainakin jossain määrin automaattisesti, jolloin jotkin kielivalinnat tehdään jokseenkin tiedostamattomasti. Edes yksittäiset kirjailijat eivät aina tiedosta kaikkia tekstiensä välittämiä merkityksiä. (Partington, Duguid & Taylor 2013: 11.)

Diskurssiprosodia tai sitä hyvin lähellä olevat ilmiöt tunnetaan myös nimillä evaluatiivinen tai semanttinen prosodia (Partington, Duguid & Taylor 2013: 58), mutta käytän tutkimuksessani ainoastaan nimitystä diskurssiprosodia selkeyden vuoksi. Diskurssiprosodiassa on kyse siitä, että sanat ilmenevät erilaisissa diskursiivisissa konteksteissa, joista sanaan liitettyjä asenteita ja arvoja voidaan päätellä. Esimerkiksi englannin kielen lemmaa *cause* käytetään usein kuvaamaan onnettomuuksia tai muita negatiivisia tapahtumia, kun taas lemmaa *provide* käytetään tyypillisemmin haluttujen tai muuten positiivisten tapahtumien kuvaamiseen. (Stubbs 2001: 65-66.) Asenteiden ja arvojen lisäksi voidaan tarkastella myös sanan pragmaattisia funktioita (Jantunen 2018b: 29).

Kenties helpoin tapa tutkia diskurssiprosodioita on kollokaatioiden kautta (Partington, Duguid & Taylor 2013: 58). Kollokaatilla viitataan siihen, kun yksi sana esiintyy toisen sanan kanssa keskimääräistä useammin (Partington & Marchi 2015: 217). Sanoja käytetään usein pareissa tai ryhmissä tiettyjen toisten sanojen kanssa, ja tästä on kollokaatiossa kyse (Sinclair 1991: 115). Kollokaatioiden, diskurssiprosodioiden ja muiden leksikaalisten myötäesiintymärakenteiden hyödyntäminen kertoo CADS-tutkimuksen fraseologisesta lähestymistavasta. Fraseologisella kielikäsitteellä viitataan psykolingvistiseen näkökulmaan, jonka mukaan kielenkäyttö perustuu tietoihin ja tiedostamattomiin valintoihin, jotka tuottavat toistuvia rakenteita. (Jantunen 2018b: 27.)

Esimerkiksi jos tietyn sanan yleisimmät kollokaatit korpuksessa ovat kaikki vahvasti negatiivisia, voidaan päätellä, että sanan diskurssiprosodia on luultavasti negatiivinen. Tämän vuoksi tutkimukseni aineisto hyödyntää listaa *vammainen*-sanan yleisimmistä kollokaateista. Kollokaattien käyttäminen lisää tutkimukseen

kvantitatiivisen piirteen, joka voi auttaa ainakin joidenkin diskurssiprosodioiden havaitsemisessa. Se myös toivottavasti lisää tutkimuksen objektiivisuutta. On tosin otettava huomioon, että itse analyysi ja diskurssiprosodioiden havaitseminen on kvalitatiivinen prosessi, johon tutkijan oma tietoinen ja tiedostamaton puolueellisuus vaikuttaa.

3 AINEISTO JA MENETELMÄT

Kerron tässä luvussa käyttämästäni aineistosta ja menetelmistä. Esittelen Suomi24-keskustelufoorumia ja kerron sen laajuudesta. Kerron myös Korp-työkälystä, josta löytyy hyödyntämäni Suomi24-korpus. Lopuksi puhun AntConc-ohjelmasta, jota käytin Suomi24-korpuksen käsittelyssä.

3.1 Suomi24-keskustelufoorumi ja -korpus

Keskustelufoorumi Suomi24 on suomalainen sosiaalisen median sivusto, joka sisältää arkisia kansalaiskeskusteluja. Keskustelufoorumilla voi kuitenkin esiintyä anonyymisti, mikä sallii keskustelun intiimeistäkin aiheista. Tutkijat voivat hyödyntää sivustoa aineistona esimerkiksi vihapuheeseen tai hyvinvointiin liittyvissä tutkimuksissa. Otollisen tutkimusaineiston keskustelufoorumista tekee se, että kyseessä on poikkeuksellisen laaja aineisto sosiaalisen median mittakaavassa kansainvälisestäkin. Esimerkiksi keväällä 2016 keskustelufoorumille kirjoitettiin 15 000-20 000 viestiä päivässä. Se sisältää aitoa ja luonnollista suomalaisten käyttämää kieltä eri puolilta maata. Eettisesti käyttäjien yksityisyys on turvattu, sillä keskustelufoorumista koottu korpus ei sisällä viestin lähettäjien IP-osoitteita tai muita tietoja, joilla lähettäjät voitaisiin helposti yksilöidä. (Lagus, Pantzar, Ruckenstein & Ylisiurua 2016: 5-6, 9-10, 40.)

Aineistonani käyttämä Suomi24-korpus sisältää keskustelufoorumille kirjoitetut viestit vuodesta 2001 vuoteen 2017. Keskustelufoorumi sisältää yli 70 miljoonaa viestiä sekä erinäisiä keskustelualeita (Lagus, Pantzar, Ruckenstein & Ylisiurua 2016: 5-6), joista en ole rajannut mitään pois. Korpus on tallennettu Kielipankkiin, jossa se on kaikille saatavilla www-selaimen avulla (Lagus, Pantzar, Ruckenstein & Ylisiurua 2016: 11).

3.2 Korp-konkordanssityökalu aineistonkeruumenetelmänä

Tutkimukseni aineiston keräämiseksi hyödynsin Kielipankin Korp-työkalua. Korp on verkkopohjainen konkordanssityökalu, jolla voidaan hakea ja tallentaa konkordanssihakuja tekstikorpuksista. Korpissa on saatavilla laajoja korpuksia suomeksi sekä suomenruotsiksi. Suomi24-korpuksen lisäksi se sisältää esimerkiksi kirja- ja sanomalehtiaineistoja sekä muita internet-aineistoja. (Kielipankki 2021; Lagus, Pantzar, Ruckenstein & Ylisiurua 2016: 16.)

Korpin Suomi24-korpusta voi hyödyntää kuka tahansa www-selaimen kautta. Haku voi sisältää yksittäisen sanan, sanan osan tai sanan perusmuodon tai sanaluokan, sillä Korpin aineistoissa on tiedot kielen eri lauseenjäsennyksistä sekä sanaluokista. Myös usean sanan haut onnistuvat. Hakutulokset näkyvät konkordanssina, joka on lyhyt lausepätkä haetun sanan ympärillä, josta voi nähdä, missä kontekstissa sanaa on käytetty. Konkordanssit ovatkin kielitieteissä yleinen tapa tarkastella sitä, missä kontekstissa jotain tiettyä sanaa käytetään. (Lagus, Pantzar, Ruckenstein & Ylisiurua 2016: 15; Sinclair 1991: 32.)

Korpuslingvistiikassa hyödynnetään kvantitatiivista tilastointia, johon liittyy useita eri termejä, joista tutkimukseni kannalta oleellisimpia selitän tässä. Frekvenssillä viitataan siihen, kuinka monta kertaa jokin sana esiintyy tekstissä (Sinclair 1991: 30). Lemmalla viitataan sanamuotoja yhdistävään perusmuotoon (Sinclair 1991: 41). Esimerkiksi *vammainen*-sanana lemmalla on *vammainen*, joka pitää sisällään kaikki sanan eri taivutusmuodot. Tämä on tutkimukseni kannalta relevanttia, sillä käytän aineistona lemmatisoitua tekstiä.

Tutkimukseni aineisto on koottu hakemalla sanoja, joiden perusmuoto on *vammainen*. Tällöin hakutulokseen siis sisältyvät myös kaikki *vammainen*-sanana taivutusmuodot. Myös isot ja pienet alkukirjaimet on samaistettu, jotta molemmat vaihtoehdot sisältyvät hakutulokseen. Aikaväliksi on asetettu koko Suomi24-korpus eli päivämäärästä 1.1.2001 päivämäärään 31.12.2017. Tällöin aineisto on mahdollisimman laaja, mikä parantaa tutkimuksen reliabiliteettia. Toisaalta on otettava huomioon, että aineisto ei kata pelkästään lähimenneisyyttä vaan myös 20 vuotta vanhoja keskusteluja. *Vammainen*-sanaa ja sen eri taivutusmuotoja löytyy tällä tavalla Korp-työkalulla haettuna 70 452 kappaletta – tulosten järjestys on satunnainen. Konkordansseja saa Korp-työkalulla enintään 1000 per sivu, joten sivujen määrä oli vajaa 71. Keräsin tästä 71 sivusta yhteensä 12 sivua siten, että poimin joka kuudennen sivun. Tämä sekä hakutuloksen satunnainen järjestys varmistavat, että aineisto todella on satunnainen. Aineiston koko eli 12 sivua, mikä merkitsee 12 000 konkordanssia, varmistaa, että aineisto on mahdollisimman laaja muttei esimerkiksi niin laaja, että AntConc-ohjelmalla olisi vaikeuksia käsitellä sitä. Nämä 12 000 konkordanssia on ladattu Excel-ohjelmaan, jossa niistä poistettiin kaksoiskappaleet. Toimenpiteen

jälkeen aineistoon jäi 11 107 konkordanssia. Excel-tiedosto sisältää sekä lemmatisoimattoman että lemmatisoidun version konkordansseista. Lemmatisoitu konkordanssi tarkoittaa konkordanssia, jonka kaikki sanat on muutettu lemmoiksi eli perusmuotoon. Lemmatisoitu aineisto on siirretty Wordiin, jossa sitä on siivottu poistamalla turhia välilyöntejä. Tällöin aineistoa on miellyttävämpi ja helpompi tarkistella.

3.3 AntConc-ohjelma apuna aineiston analysoimisessa

AntConc on ohjelma, jolla voidaan tutkia korpuksia. Ohjelma on vapaasti saatavilla internetissä. Käytännössä AntConc-ohjelmassa avataan tekstitiedosto, jolloin ohjelman työkaluja voidaan hyödyntää tiedoston sisältämän tekstin tarkisteluun. Tässä tutkimuksessa tekstitiedosto siis sisältää 11 107 lemmatisoitua konkordanssiriviä. AntConc-ohjelmassa voidaan muun muassa hakea tietyn sanan konkordansseja, tehdä avainsanahakuja sekä laskea ja tilastoida frekvenssejä. Tutkimuksessani hyödynnän AntConc-ohjelmaa kollokaatioiden tarkasteluun. Käyttämäni AntConcin versio on vuonna 2019 julkaistu 3.5.8. (Anthony 2019.)

Kollokaatiolla viitataan siihen, kun tietty sana myötesiintyy toisen sanan tai sanojen kanssa niin usein, että kyseessä on tilastollisesti merkittävä ilmiö. Ilmiön nimi on kollokaatio, kun taas kollokaatit ovat ne sanat, jotka myötäesiintyvät. (Baker 2006: 95-96.) Kollokaatti ei välttämättä esiinny tutkittavan sanan vieressä, vaan se voi sijaita useamman sanan edellä tai takana. Tässä tutkimuksessa käytetty tarkasteluväli on 4L-4R, eli kollokaatin suurin mahdollinen etäisyys on neljän sanan päässä haetusta sanasta. Tämä johtuu siitä, että tätä tarkasteluväliä on pidetty optimaalisena, ja se sopii suomenkielisten lauseiden tutkimiseen (Jantunen 2018b: 29). Toinen huomioitava asia on vähimmäisfrekvenssi. Vähimmäisfrekvenssi säätelee sitä, kuinka monta kertaa kollokaatin on esiinnyttävä aineistossa, jotta AntConc ottaa sen huomioon. Tässä tutkimuksessa vähimmäisfrekvenssi on asetettu 50:een. Tämä takaa sen, että ainoastaan yhden käyttäjän keskustelufoorumien kirjoitukset eivät sotke tuloksia. Tutkimuksessa on käytetty niin kutsuttua MI-testiä (*Mutual Information*). Tämä johtuu siitä, että MI-testi korostaa kollokaatiohaussa sisältösanoja kieliopillisten sanojen sijaan. Kieliopillisilla sanoilla on tyypillisesti suuri frekvenssi, joten ne saattavat esiintyä usein minkä tahansa sanan lähellä. MI-testi ottaa tämän huomioon ja vertaa, kuinka usein sana esiintyy korpuksessa yleensä suhteessa siihen, kuinka usein se esiintyy haetun sanan lähellä. (Baker 2006: 100-101.)

3.4 Kollokaattien rajausta ja luokittelu

Kaiken kaikkiaan AntConc löysi MI-menetelmällä tilastollisesti merkittäviä kollokaatteja 312 kappaletta. Valitsin näistä tarkasteltavakseni 100 merkittävintä. Karsin yhteensä 15 kollokaattia pois. Suurin osa rajatuista kollokaateista oli kielioppisanoja kuten esimerkiksi *tai*, *vai* ja *sekä*. Diskurssiprosodioiden etsiminen tällaisista sanoista on vaikeaa, ja MI-menetelmällä pyritään karsimaan niitä pois. Myös lyhenteet (*jne*, *ym*, *mm*) karsittiin, sillä ne edustavat useampaa sanaa.

Varsinaiset tutkimuksessa tarkastellut kollokaatit on jaoteltu semanttisiin merkitysryhmiin ja diskurssiprosodioiden. Jaottelu on tehty siten, että jokaisen kollokaatin kohdalla on tarkasteltu sen konkordansseja. Konkordansseja on tarkasteltu siten, että kollokaatille on pyritty löytämään jokin yleinen merkitys ja käyttöyhteys.

Kollokaattien luokittelu voi olla haasteellista. *Vammainen*-sana itsessään voi olla joko adjektiivi tai substantiivi, ja myös useat kollokaatit saattavat olla monitulkintaisia. Tekstin lemmatisoitu muoto on otettava huomioon kuin myös se, ettei AntConc ole täydellinen ohjelma. Näiden seikkojen takia konkordanssien tarkistelu on niin tärkeää – kollokaattia saatetaan käyttää tavalla, jota ei ensisilmäyksellä välttämättä olettaisi. Esimerkiksi kollokaattilistalta löytyvällä eläke-sanalla viitataan lähes aina eläkeläiseen. Syrjäytyä-verbi taas viittaa usein syrjäytyneeseen henkilöön. Adjektiivi vanha taas viittaa konkordansseissa tavallisesti vanhempaan eli isään tai äitiin. Sanoja voidaan myös käyttää vertauskuvallisesti tai kirjaimellisesti. Koskea- ja tuntea-verbit löytyvät listalta, mutta konkordansseissa niillä ei yleensä viitattu fyysiseen koskemiseen tai tuntemiseen.

4 ANALYYSI JA TULOKSET

Tarkastelen tässä luvussa 100 merkittävintä kollokaattia, joita *vammainen*-sana aineistosta sai. On huomioitava, että merkittävintä ei tässä tapauksessa välttämättä tarkoita MI-mittauksella tilastollisesti merkittävintä – osa kollokaateista karsiutui pois. Esittelen diskurssiprosodiat, joihin olen jakanut kollokaatit. Pohdin luokittelun pätevyyttä ja mahdollisia ongelmia ja kysymyksiä sekä perustelen esimerkein valintojani liittyen sekä diskurssiprosodioihin että niihin sijoitettuihin kollokaatteihin. Koska kollokaatteja on 100, minun on keskityttävä joihinkin kollokaatteihin enemmän kuin muihin: keskityn erityisesti tyypillisiin kollokaatteihin sekä kollokaatteihin, joiden merkitys ei välttämättä vaikuta ilmeiseltä.

4.1 *Vammainen*-sanan kollokaattien diskurssiprosodiat

Kaiken kaikkiaan ryhmiä löytyi yhdeksän: kahdeksan diskurssiprosodiaa sekä yksi sekalaisten ryhmä. Yleisimmät diskurssiprosodiat olivat *vammaisuuden aste*, *vammaisten oikeudet* sekä *vähemmistöpuhe*. Loput diskurssiprosodiat olivat nimiltään *vammainen haukkumasanana*, *synnytys- ja aborttipuhe*, *syrjintäpuhe*, *henkilökohtainen suhde vammaisiin* sekä *vammaisten toimijuus*. Diskurssiprosodiat sekä niissä olevien kollokaattien lukumäärän näkee taulukosta 1. Koska kollokaatteja on 100, lukumäärä on myös prosentuaalinen määrä.

Taulukko 1. Diskurssiprosodiat ja niissä olevien kollokaattien lukumäärä.

Diskurssiprosodia	Kollokaattien lukumäärä
Vammaisuuden aste	23
Vammaisten oikeudet	23

Vähemmistöpuhe	20
<i>Vammainen</i> haukkumasanana	8
Synnytys- ja aborttipuhe	7
Syrjintäpuhe	4
Henkilökohtainen suhde vammaisiin	4
Vammaisten toimijuus	2
Sekalainen	9

4.1.1 Vammaisuuden aste

Vammaisuuden aste oli yleisin ryhmä *vammaisten oikeuksien* kanssa. Tämä diskurssiprosodia sisälsi luokittelua siitä, millä tavalla joku on vammainen. Kollokaatteihin kuului substantiiveja, verbejä ja yksi adjektiivi, mutta ylivoimaisesti suurin osa sanoista oli adverbejä. Taulukko 2 havainnollistaa tätä.

Taulukko 2. Diskurssiprosodia *vammaisuuden aste*.

Adverbit vaikeasti, lievästi, fyysisesti, pahasti, vakavasti, henkisesti, jotenkin, muuten, esimerkiksi, oikeasti, erittäin, todella, täysin	Substantiivit cp, invalidi, lapsi , kehitys, pyörätuoli, aikuinen, poika
	Verbit tarkoittaa, vammautua Adjektiivit vaikea

Vammaisuuden aste -diskurssiprosodian 23 kollokaatista 13 ovat adverbejä. Suurin osa niistä edustaa tätä prosodiaa varsin selvästi. Voidaan puhua esimerkiksi vaikeasti, lievästi, fyysisesti tai henkisesti vammaisista henkilöistä. Keskustelut liittyvät usein juuri siihen, miten ihmiset voivat olla vammaisia eri tavoilla (esimerkki 1). Keskustelua käydään siitä, miten tietynlainen vammaisuus voi näkyä tai olla näkymättä elämässä ulkopuolisille (esimerkki 2) sekä siitä, miten se voi rajoittaa omaa elämää tietyllä tavalla (esimerkit 3 ja 4).

- 1) On aivan eri asia olla **lievästi** kuulovammainen kuin rampa afaattinen kuurosokea (jälkimmäinen on sitten paljolti muiden avun varassa, että sitä sitten voi miettiä ...)
- 2) eikä kaikki vammat ole downin syndoomia, vaan **vaikeasti** kehitysvammainen voi olla päälle päin ihan normaalin näköinen.

- 3) Jumppa, lenkkeily ja kaikkalainen kunnon kohotus paikallaan, jollet ole **fyysisesti** niin sairas tai *vammainen*, että pystyt liikkumaan.
- 4) Jos olet millään tavalla *vammainen* - etenkin **henkisesti** tai sosiaalisesti - niin oikeastaan on parempi vetää itsensä hirteen koska et tule koskaan olemaan onnellinen toisen ihmisen kanssa.

Tässä diskurssiprosodiassa siis saatettiin esimerkkien 1 ja 2 tapaan kommentoida itse vammaisuuden astetta eli sitä, mitä vammaisuus voi olla tai mitä ihmiset olettavat sen olevan. Tarkoittaa-verbiä käytetään vastaavalla tavalla kommentoimaan sitä, ettei vammaisuus välttämättä *tarkoita* sitä, mitä oletetaan (esimerkki 5).

- 5) En siis **tarkoittanut** aiemmin että *vammaiset* olisivat automaattisesti täys-vihanneksia...

Toisenlainen tapa taas näkyy esimerkeissä 3 ja 4, joissa spesifioitiin tietynlainen vammaisuus, jota kommentoitiin. Tämä näkökulma oli myös substantiivikollokaateissa *lapsi*, *aikuinen* ja *poika*. Lapsi- ja poika-kollokaateissa puhuttiin lapsen vammaisuudesta, eli sillä viitattiin tavallisesti synnyntäiseen vammaisuuteen (esimerkit 6 ja 7).

- 6) Alkoholivammaisten **lasten** määrä lisääntyy jatkuvasti ja kumuloituu tulevaisuudessa.
- 7) Minulla oikein sattuu sydämeen kun ajattelen mitä tämä *vammainen poika* kuulee äidiltään, tietenkin vammaisen lapsen vanhempana on todella rankkaa olen yksi niistä itsekin mutta ei se ole lapsen vika

Aikuinen olisi sopinut varmasti muihinkin ryhmiin. Kollokaatti kuitenkin viesti vammaisen henkilön tarpeista ja ympäristöstä, jotka taas ovat kytköksissä siihen, millä tavalla aikuinen henkilö on vammainen - vammaisesta aikuisesta puhuttaessa puhuttiin usein ammattihoidosta ja palvelukodeista (esimerkit 8 ja 9). Toisaalta tällainen puhe voisi kuulua myös ainakin *vammaisten oikeudet* -diskurssiprosodiaan.

- 8) olen toiminut kehitysvammaisten parissa ja erästä *vammaista aikuista* hoitaessani hän kertoi uskostaan ja tiedän sen olevan vilpittöntä, vaikkei hän voikaa ymmärtää paljonkaan uskon asioista.
- 9) Kehitysvammaiselle **aikuiselle** ei riitä palvelukodin hyvä asunto ja ammattilaisten antama hyvä hoito.

4.1.2 Vammaisten oikeudet

Vammaisten oikeudet -diskurssiprosodia sai yhtä paljon kollokaatteja kuin vammaisuuden aste. Esittelen kollokaatteja taulukossa 3.

Taulukko 3. Diskurssiprosodia *vammaisten oikeudet*

Palvelut: avustaja, laitos, palvelu, hoitaa, tukea, huolehtia	Ihmisoikeudet : oikeus, tappa, kohtaan, kohdella, myöntää, asema, seksi, erilainen, suhtautua, terve, ihminen, usein, miksi, kieltää, miten
Lakisäädökset: pysäköintilupa, koskea	

Kyseessä on melko laaja ryhmä. Sanaluokista se sisälsi eniten verbejä (yhdeksän) ja substantiiveja (kahdeksan). Se sisältää puhetta vammaisten asemasta, eri laeista ja oikeuksista sekä palveluista ja niitä tarjoavista instituutioista. Ryhmä tuntui hankalalta ehkä sen takia, että puhe on usein filosofista tai muuten abstraktia.

Suurimmassa semanttisessa ryhmässä, jolle annoin nimeksi *ihmisoikeudet*, on kyse vammaisten oikeuksista (esimerkki 10). Se sisältää myös viittauksia ihmisoikeuksien rikkomisesta (esimerkki 11). Kysymyssanoilla pohditaan vammaisten elämänlaatua ja oikeuksia (esimerkit 12, 13 ja 14).

- 10) Täytyy kyl sanoo et kuhan YK:n *vammaisten oikeusdet* nappaa voimaan niin me ollaan kaikki paremmassa maailmassa.
- 11) Tämä laki mahdollistaisi vanhusten ja *vammaisten* keskitysleirien aukottoman ja **TAPPAVAN** ylivallan.
- 12) milloin pääsemme näkemään sitä, **miten** *vammainen* elää sinut itsensä ja vammaisuutensa kanssa?
- 13) **Miksi** *vammainen* ei saisi osallistua politiikkaan?
- 14) **Miksi** *vammainen* ei saisi omistaa kunnon duunipaikkaa?

Palvelut-ryhmässä on kyse vammaisten palveluista ja niitä tarjoavista instituutioista (esimerkit 15 ja 16).

- 15) Olen samaa mieltä siitä että veroja ikänsä maksaneet vanhuksset ja *vammaiset* ansaitsevat ammattitaitoiset **avustajat** avustajikseen.
- 16) Vanhusten ja *vammaisten* hoitolaitosten huolenpito ei ole kovin korkea.

Lakisäädöksissä on kyse laeista. *Pysäköintilupa* viittaa vammaisia koskevaan pysäköintilakiin (esimerkki 17) ja koskea-verbillä viitataan usein vammaisia *koskeviin* lakeihin (esimerkki 18).

- 17) Poikkeuksena voi kuitenkin olla vammaisen **pysäköintilupa**: poliisi voi myöntää vaikeasti *vammaiselle* henkilölle tai vaikeasti *vammaisen* henkilön kuljettamista varten **pysäköintiluvan**.

- 18) Ohjelmassa kuvattiin niin hyvin myös erään *vammaisia koskevan* lain käsittelyä alkaen 90-luvun puolelta tähän päivään asti, että oikeastaan oli apea mieli eilen koko lopun päivää sen takia

4.1.3 Vähemmistöpuhe

Tässä diskurssiprosodiassa on kyse vammaisten liittämistä muihin ryhmiin, joita pidetään keskustelussa vähemmistöryhminä. Vammaiset eivät välttämättä ole keskustelun aihe tässä diskurssiprosodiassa vaan yksi esimerkkiryhmä muiden ryhmien joukossa.

Taulukko 4. Diskurssiprosodia *vähemmistöpuhe*.

<p>Substantiivit: vanhus, eläkeläinen, maahanmuuttaja, opiskelija, eläke, hoito, nainen, <i>nuori, sokea, sairas, työtön, köyhä, homo, juutalainen</i></p>	<p>Verbit: syrjäytyä, kärsiä Adverbit: yhtä, samoin, kuten Adjektiivit: heikko</p>
---	---

Tämän diskurssiprosodian 20 kollokaatista 14 on substantiiveja. Kursiivilla kirjoitetut substantiivit voitaisiin tulkita myös adjektiiveiksi, mutta konkordansseissa niitä käytettiin substantiiveina. Tässä diskurssiprosodiassa esiintyi paljon luettelointia, ja luettelot sisälsivät usein myös muita tämän diskurssiprosodian kollokaatteja (esimerkit 19, 20 ja 21). Myös adverbit liittyivät luettelointiin (esimerkki 22).

- 19) Viranomaisen silmissä korkean riskin lapsia ovat yksinhuoltajan lapsi, **maahanmuuttaja**, adoptiolapsi, *vammainen*, **sairas** eli ne, joiden tulisi saada erityistä tukea ja sosiaalihuollon hyvää palvelua
- 20) He julistivat vuosikausia kommunistit, sosiaalidemokraatit, **juutalaiset**, bolševikit, romanit, **homot**, *vammaiset* täiksi epäihmismäisiksi hirviöiksi ja loisiksi.
- 21) Seuraavaksi tulee liittää jäseneksi Veikkaus Oy, sitten sotaveteraanit, vammaiset, **eläkeläiset, työttömät**, yksinhuoltajat ja muut.
- 22) toivotatko siellä, että vähemmistöt **kuten** esim. *vammaiset*, mielenterveyspotilaat ja **homot** ja lesbot ovat kakkosluokan ihmisiä?

Eläke-substantiivilla viitattiin eläkeläiseen ja syrjäytyä-verbillä syrjäytyneeseen, mutta AntConc tulkitsi niiden lemmat tällä tavalla.

4.1.4 *Vammainen* haukkumasanana

Tässä diskurssiprosodiassa *vammainen*-sanaa käytetään pejoratiivisesti toista haukkuessa, tai se mielletään keskustelussa haukkumasanaksi. Muitakin kollokaatteja

on käytetty haukkumiseen – esimerkiksi jotakuta voidaan haukkua cp-vammaiseksi – mutta tämän ryhmän tapaukset ovat selviä. Ne eivät myöskään viittaa todelliseen vammaisuuteen.

Taulukko 5. Diskurssiprosodia *vammainen haukkumasanana*.

Spesifioitu mutta ei-todellinen vammaisuus: sosiaalisesti, asenne, tunne	Verbit: taitaa, kirjoittaa
Muut: vittu, ruma, tyhmä	

Ensimmäisessä ryhmässä puhutaan *vammaisuuden asteen* diskurssiprosodiaa muistuttavasti tietynlaisesta vammaisuudesta. Kyse ei kuitenkaan ole todellisesta vammaisuudesta, vaan sanaa *vammainen* käytetään haukkumasanana. Tein päätöksen erottaa tällaiset ensisijaisesti pejoratiivisesti käytetyt tapaukset omaksi ryhmäkseen. Pejoratiivisessa käytössä keskustelu itsessään ei tyypillisesti liity vammaisiin. Asenne- ja tunne-sanat ovat etuliitteitä (esimerkit 23 ja 24). Kirjoittaa- ja taitaa-verbit liittyvät usein tilanteisiin, joissa haukutaan toista käyttäjää (esimerkit 25 ja 26).

- 23) Niin kauan kun **asennevammaisia** on olemassa niin kauan on vammaisia asenteita
- 24) Kyllä tuollainen **tunne vammainen** ei ole rakastanut sinua hetkeäkään vaan tuo on ollut puhdas hyväksikäytösuhde.
- 25) Taisi kolahtaa kun itse **kirjoitat** kuin *vammainen*.
- 26) Suurin *vammainen* **taidat** olla ihan itse.

4.1.5 Synnytys- ja aborttipuhe

Tämä diskurssiprosodia liittyi keskusteluun abortista ja synnytyksestä.

Taulukko 6. Diskurssiprosodia *synnytys- ja aborttipuhe*.

Vanhempi: synnyttää, abortoida, vanha, riski	Sikiö: sikiö, syntyä, vauva
---	------------------------------------

Keskustelu liittyi kehitysvammaisen sikiön aborttiin ja synnyttämiseen. Sana vanha viittaa vanhempaan. Kommenteissa keskustellaan abortin hyväksyttävyydestä (esimerkit 27 ja 28). Esimerkistä 28 myös näkee uskontoon liittyvän puheen. Suomi24:n uskontopuhe on havaittu myös heteroutta ja homoutta käsittelevässä CADS-tutkimuksessa (Jantunen 2018a: 8-9).

- 27) Te tyrkytätte meille jotka eivät halua **synnyttää** epämuodostuneita/*vammaisia*/vakavasti sairaita!
- 28) ehdotettiin aborttia vaikeasti *vammaisen vauvan* pelossa, mutta sieltä syntyikin aivan terve poika vauva. Jos tässä kohden ei luotettaisi jumalan johdatukseen niin olisiko sinun mielestäsi käynyt oikein, jos tämän vauvan elämänlanka olisi katkaistu abortilla?

4.1.6 Syrjintäpuhe

Syrjintäpuheessa on kyse puheesta, jossa käsitellään syrjintää *vammaisia* kohtaan. Puhe ei siis itsessään välttämättä ole syrjivää.

Taulukko 7. Diskurssiprosodia *syrjintäpuhe*

Syrjintäpuhe: pilkata, syrjiä, haukkua, kiusata

Kommentit käsittelevät vammaisten kokemaa syrjintää usein kriittiseen sävyyn (esimerkit 29 ja 30). Puhe ei aina kuitenkaan ole kritisoivaa tai tuomitsevaa (esimerkki 31).

- 29) **Kiusaavat** *vammaisia* ja vähempiosaisia niin kauan kuin vain voivat ja valehtelevat niin pitkään kuin henki pihisee
- 30) On surullista ja jopa oksettavaa että mira **haukkuu** avoimesti *vammaisia* ihmisiä.
- 31) Ymmärrän ihmisten mielipiteet siinä, että *vammaisia* ei saisi **pilkata**, mutta come on!

4.1.7 Henkilökohtainen suhde vammaisiin

Tässä diskurssiprosodiassa puhutaan omista kokemuksista ja toiminnasta vammaisten kanssa.

Taulukko 8. Diskurssiprosodia *henkilökohtainen suhde vammaisiin*.

Henkilökohtainen suhde vammaisiin: parissa, tuntea, kanssa, kokemus

Tuntea-verbi viittaa tässä yhteydessä vammaisen henkilön tuntemiseen (esimerkki 32). Joissain kommentteissa puhutaan työskentelystä vammaisten parissa (esimerkki 33).

- 32) En **tunne** ainoatakaan "*vammaista*" jolla olisi korvien välissä mitään muuta kuin täysi pakka
- 33) olen ollut töissä kehitysvamma laitoksessa sekä **kokemuksia** minulla on *kehitysvammaisesta* serkustani sekä siskoni miehen tyttärestä jolla on mosaiikki down .

4.1.8 Vammaisten toimijuus

Tässä diskurssiprosodiassa on vain kaksi kollokaattia, ja harkitsin sen kollokaattien siirtämistä sekalaisiin. Päätin pitää diskurssiprosodian, sillä siitä kuitenkin erottui jokseenkin uniikki näkökulma.

Taulukko 9. Diskurssiprosodia *vammaisten toimijuus*.

Vammaisten toimijuus: henkilö, joukko
--

Diskurssiprosodiassa puhutaan vammaisista toimijoina. Henkilöt toimivat yksilöinä (esimerkit 34 ja 35). Joukosta puhuttaessa näkökulma on selvästi ulkopuolinen (esimerkki 36).

- 34) *Vammaiset* **henkilöt** ei todellakaan hallitse mitään vastaavaa kaverisuhteiden rakentamista.
- 35) Kunta ei kuitenkaan toimi työnantajana, vaan *vammainen* **henkilö**
- 36) Joskus jossain kylpylässä oli **joukko** kehitys *vammaisia* saattajien kanssa.

4.1.9 Sekalainen

Sekalaisten ryhmä oli joukko sanoja, joita oli vaikea sijoittaa mihinkään tiettyyn ryhmään.

Taulukko 10. *Sekalaisten* ryhmä.

Adverbit: tuskin, myös, kohta, toki, jopa, täällä	Substantiivit: tapa, sana Verbit: liikkua	
--	--	--

Suurin osa ryhmästä muodostuu adverbeista. Tämä on oletettavaa, sillä partikkeleita käytetään tuomaan lauseeseen lisämerkityksiä (VISK § 794), joten niistä itsestään voi olla vaikeaa löytää yleistettäviä merkityksiä. Ryhmän substantiivit ja verbit taas ovat sen verran abstrakteja ja yleisluontoisia, ettei niiden käyttöä voinut yleistää mihinkään tiettyyn diskurssiprosodiaan.

5 PÄÄTÄNTÖ

5.1 Yhteenveto ja johtopäätökset

Tutkimukseni tarkoituksena on ollut selvittää sekä tarkastella sitä, millaisia diskurssiprosodioita *vammainen*-sana saa Suomi24-keskustelufoorumilla. Ensimmäinen tutkimuskysymykseni oli, että mitkä ovat *vammainen*-sanan kollokaatit Suomi24-korpuksessa. Kollokaatteja oli kaiken kaikkiaan 312 kappaletta, mutta ainoastaan 100 MI-arvoltaan suurinta sanaa, jotka eivät olleet kieliopillisia sanoja tai lyhenteitä, tarkasteltiin. Kollokaatit olivat sanaluokaltaan adverbeja, substantiiveja, adjektiiveja, verbejä sekä adpositioita. Useat kollokaatit olivat sellaisia, että niiden todellinen merkitys avautui vasta konkordansseja tarkastelemalla.

Tutkimuskysymyksistä toinen oli, että millaisia diskurssiprosodioita kollokaateista voidaan päätellä. Diskurssiprosodioita oli kahdeksan: *vammaisuuden aste*, *vammaisten oikeudet*, *vähemmistöpuhe*, *vammainen haukkumasanana*, *synnytys- ja aborttipuhe*, *syrjintäpuhe*, *henkilökohtainen suhde vammaisiin* sekä *vammaisten toimijuus*. Tämän lisäksi yhdeksän kollokaattia jäi omaan ryhmään nimeltä *sekalainen*. *Vammaisuuden aste*, *vammaisten oikeudet* ja *vähemmistöpuhe* olivat suurimpia ryhmiä ja sisälsivät yhteensä 66 kollokaattia 100:sta. Kokonaisuudessaan diskurssiprosodiat olivat sävyiltään positiivisia, neutraaleja tai negatiivisia. Diskurssiprosodioista nousee kuitenkin enemmän positiivisuus tai neutraalius esille kuin negatiivisuus. *Vammaisuuden aste* -diskurssiprosodiassa keskusteltiin esimerkiksi siitä, mitä eri asioita vammaisuus voi tarkoittaa. *Vammaisten oikeudet* sekä *vähemmistöpuhe* -diskurssiprosodiat olivat usein vammaisten oikeuksia puolustelevia ja vähemmistöjen syrjintää vastustavia. Ainoa selkeästi negatiivinen ryhmä oli *vammainen haukkumasanana*. Tämä ryhmä sisälsi kahdeksan kollokaattia. Vaikuttaa siis siltä, että puhe vammaisista Suomi24-keskustelufoorumilla on melko monipuolista eikä ainakaan yksipuolisen negatiivista. On toisaalta

huomioitava, että puhekielisempi vammaisen-sana ei näy tuloksissa. Sitä mahdollisesti käytetään negatiivisempaan sävyyn. Lisäksi tutkimus sisälsi kommentteja sekä vammaisilta että ei-vammaisilta henkilöiltä. Tutkimuksesta ei siis voi päätellä esimerkiksi ei-vammaisten asenteita vammaisia kohtaan tai vammaisten käsityksiä itsestään tai yhteiskunnasta.

5.2 Arviointi ja mahdollinen jatkotutkimus

Vammainen-sana on monitulkintainen, sillä se voi olla substantiivi tai adjektiivi. Tämä voi tehdä siitä jokseenkin haastavan sanan tutkia. Toisaalta se teki tutkimuksesta myös mielenkiintoisemman.

On huomioitava, että kollokaattien analyysin prosessi on subjektiivinen – kollokaatin frekvenssi saattaa olla useita satoja, ja konkordansseja on tarkasteltu silmämääräisesti ja manuaalisesti eikä empiirisesti esimerkiksi tilastollisia työkaluja hyödyntäen. Toiseksi kollokaateilla on usein monia eri käyttöyhteyksiä, jolloin tein päätöksen sen mukaan, mikä käyttöyhteys ja merkitys vaikutti olevan yleisin. Kollokaatteja analysoidessa minun on tiedostettava ainakin se, että tutkijan oma ennakkokäsitys siitä, mihin diskurssiprosodiaan kollokaatin tulisi intuitiivisesti kuulua, on saattanut vaikuttaa erityisesti sellaisiin kollokaatteihin, joilla on useampia käyttöyhteyksiä tai merkityksiä. Tässä tutkijan omaa subjektiivisuutta olisi voinut vähentää ainakin se, että kollokaatteja ei olisi rajattu ainoastaan yhteen diskurssiprosodiaan. Toisaalta se olisi saattanut lisätä monitulkintaisuutta joidenkin kollokaattien kohdalla niin paljon, ettei luokittelusta enää olisi hyötyä. Pyrin tarkistelemaan jokaisen kollokaatin konkordansseja useasti ja tarkkaan, jotta löytäisin yleisimmän käytötavan.

Korpusavusteinen diskurssintutkimus sopi tutkimukseeni hyvin, sillä se pyrki selvittämään kansalaiskeskustelusta merkityksiä, jotka eivät välttämättä ole selviä. CADS-tutkimuksen yksi tavoitteista on tällainen epäselvien tai piilossa olevien merkitysten paljastaminen (Partington, Duguid & Taylor 2013: 11). Suomi24-keskustelufoorumi on niin laaja, että siitä vapaasti saatavilla olevan korpuksen ja tilastollisten menetelmien hyödyntäminen on otollista. Lisäksi Suomi24 sisältää arkista luonnollista kieltä (Lagus, Pantzar, Ruckenstein & Ylisiurua 2016: 40), jonka tutkimista diskurssintutkimuksessa painotetaan (Partington, Duguid & Taylor 2013: 4).

Suomi24-keskustelufoorumia ei voi välttämättä yleistää koko Suomen väestöön. Kirjoittajien ei tarvitse rekisteröityä sivulle eikä aineistoon kuulu esimerkiksi IP-osoitteet tai muut tiedot, joilla käyttäjiä voitaisiin tunnistaa. Käyttäjiä ei voida siis demografisesti yksilöidä. (Lagus, Pantzar, Ruckenstein & Ylisiurua 2016: 9.) Voi siis olla mahdollista, että esimerkiksi tietty ikä tai sukupuoli on keskustelufoorumilla ylliedustettuna eikä tarkasti kuvaa koko Suomea. Anonymiteetin vuoksi on myös

mahdollista, että viestejä kirjoitetaan tahallisessa häiriköinti- tai provosointimielessä tavalla, joka ei kuvaa yhteiskuntaa keskustelufoorumien ulkopuolella.

Useamman kollokaatin tarkastelu olisi voinut synnyttää lisää diskurssiprosodioita tai tuoda uusia näkökulmia jo olemassa oleviin prosodioihin – nyt jotkin diskurssiprosodiat jäivät hyvin pieniksi. Myös kollokaattien konkordanssien systemaattisempi tarkastelu tilastollisia menetelmiä hyödyntäen olisi voinut tuottaa tarkempia tuloksia ja parantaa tutkimuksen reliabiliteettia.

Jatkotutkimusmahdollisuuksia on ainakin tutkimuksen laajentaminen muihin korpuksiin. Olisi mielenkiintoista verrata tutkimustuloksia vastaavaan tutkimukseen, joka käyttäisi esimerkiksi sanomalehtikorpuista ja nähdä, löytyykö niistä samoja vai aivan erilaisia diskurssiprosodioita ja missä suhteessa.

LÄHTEET

- Anthony, L. (2019): *AntConc. (Version 3.5.8). Computer Software*. Tokyo, Japan: Waseda University. <https://www.laurenceanthony.net/software> 16.4.2021
- Baker, Paul 2006: *Using corpora in discourse analysis*. London; New York. Continuum
- Fairclough, Norman 1992: *Discourse and Social Change*. Cambridge, UK. Polity Press
- Hakala, Marjaana 2018: *VAMMAISUUDEN MEDIAREPRESENTAATIO. Toiseuttavia stereotyyppioita vai progressiivisia lähestymistapoja?* Kandidaatintutkielma, Helsingin yliopisto
- Jantunen, J. H. 2018a: *Homot ja heterot Suomi 24:ssä. Analyysi digitaalisista diskursseista. Puhe ja kieli*, 38 (1), 3-22
- Jantunen, J. H. 2018b: *Korpusavusteinen diskurssintutkimus (CADS): analyysiesimerkki homouden ja heterouden digitaalisista diskursseista. AFinLA-e*, 11
- Johansson, M.; Jantunen, J. H.; Heimo, A.; Ahonen, M. & Laippala, V. 2018: *Verkkokeskustelujen kansa: korpusavusteinen diskurssianalyysi Suomi24-keskustelupalstasta. Sananjalka*, 60 (60), 96-117. doi: 10.30673/sja.69963
- Karjalainen, Marja 1993: *Opiskelijoiden asenteet vammaisia oppilaita ja heidän liikunnan opettamista kohtaan. Pro gradu -tutkielma, Jyväskylän yliopisto*
- Kielipankki 2021: Korp-käyttöliittymän käyttöohjeet <https://www.kielipankki.fi/tuki/korp/>
- Lagus, K.; Pantzar, M.; Ruckenstein, M. & Ylisiurua, M. 2016: *Suomi24. Muodonantaa aineistolle. Helsingin yliopiston valtiotieteellisen tiedekunnan kuluttajatutkimuskeskuksen julkaisuja 10. Helsinki: Helsingin yliopisto*
- Määttä, S.; Suomalainen, K. & Tuomarila, U.: 2020: *Maahanmuuttovastaisen ideologian ja ryhmäidentiteetin rakentuminen Suomi24-keskustelussa. Virittäjä*, 124 (2)
- Nurmi-Koikkalainen, P.; Ahola, S.; Gissler, M.; Halme, N.; Koskinen, S.; Luoma, M.; Malmivaara, A.; Muuri, A.; Sainio, P.; Sääksjärvi, K.; Väyrynen, R. (2017): *Tietoa ja tietotarpeita vammaisuudesta : Analyysia THL:n tietotuotannosta. THL Työpaperi 38/2017*
- Partington, A. & Marchi, A. 2015: *Using corpora in discourse analysis. The Cambridge Handbook of English Corpus Linguistics*, 216-234. Cambridge. Cambridge University Press

Partington, A.; Duguid, A. & Taylor, C. 2013: *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies (CADS)*. Amsterdam: John Benjamins Publishing Company

Sinclair, John 1991: *Corpus, concordance, collocation*. Oxford. Oxford University Press

Stubbs, Michael 2001: *Words and Phrases. Corpus studies of Lexical Semantics*. Blackwell, Oxford

VISK = Auli Hakulinen, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen ja Irja Alho 2004: *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura. <http://scripta.kotus.fi/visk> 15.4.2021.

LIITTEET

LIITE 1:

MI-arvon mukaan järjestetty lista kollokaateista sekä niiden frekvensseistä.

kollokaatti	frekvenssi	MI-arvo
pysäköintilupa	87	5.26348
vaikeasti	319	5.26053
vanhus	1009	5.02349
lievästi	67	4.90292
sosiaalisesti	50	4.89573
sairas	750	4.81635
cp	69	4.73236
fyysisesti	66	4.72595
parissa	60	4.71130
invalidi	68	4.69723
pahasti	94	4.64140
vakavasti	70	4.59266
pilkata	90	4.57867
eläkeläinen	88	4.57579
syрjäytyä	56	4.56446
syрjä	68	4.51336
henkisesti	99	4.50777
syntyä	497	4.43159
työtön	183	4.42395
köyhä	181	4.40809

homo	168	4.32242
maahanmuuttaja	52	4.25524
henkilö	293	4.25262
sokea	52	4.21535
abortoida	66	4.19091
haukkua	129	4.17821
lapsi	2726	4.17283
vittu	91	4.16020
opiskelija	56	4.10362
synnyttää	90	4.06020
ruma	90	3.98722
jotenkin	106	3.97434
juutalainen	51	3.89188
asenne	82	3.85940
oikeus	312	3.82845
tappaa	115	3.79303
kohtaan	87	3.76767
kohdella	54	3.76575
muuten	216	3.75653
kiusata	73	3.75138
riski	53	3.72957
myöntää	57	3.72828
kehitys	58	3.71130
asema	64	3.70381
avustaja	113	3.69859
esimerkiksi	91	3.69032
sikiö	60	3.67959
pyörätuoli	69	3.64324
seksi	119	3.64034
tuskin	57	3.62136
eläke	60	3.60324
hoito	146	3.58824
laitos	53	3.58124
heikko	78	3.57044
palvelu	92	3.55304
hoitaa	202	3.54721
tyhmä	68	3.54452
koskea	83	3.52665
erilainen	95	3.52183
tarkoittaa	148	3.49729
suhtautua	56	3.49629
joukko	51	3.48488
yhtä	80	3.48377

liikkua	51	3.46885
vaikea	119	3.46316
tukea	58	3.46076
terve	314	3.45646
tapa	180	3.45273
oikeasti	80	3.44327
aikuinen	78	3.43701
samoin	52	3.43446
nainen	396	3.43177
kärsiä	75	3.42705
vammautua	65	3.42382
tuntea	117	3.42180
myös	567	3.41902
nuori	148	3.41561
huolehtia	63	3.41556
sana	90	3.37449
vanha	267	3.37254
kanssa	373	3.36966
kohta	116	3.36634
kokemus	58	3.35983
erittäin	56	3.34873
ihminen	987	3.34155
toki	56	3.33544
kuten	159	3.30309
poika	86	3.28790
jopa	123	3.27878
usein	85	3.25445
todella	188	3.23983
täysin	92	3.23487
taitaa	62	3.23365
miksi	180	3.23353
kirjoittaa	96	3.22659
kieltää	54	3.22460
vauva	60	3.22016
miten	251	3.21796
tunne	53	3.21683
täällä	123	3.21087

