

**KORPUSAVUSTEINEN KRIITTINEN DISKURSSINTUTKI-  
MUS NIMITYKSISTÄ *MUSTALAINEN* JA *ROMANI*  
SUOMI24-KESKUSTELUPALSTALLA**

Erika Kilpeläinen  
Kandidaatintutkielma  
Suomen kieli  
Kieli- ja viestintätieteiden lai-  
tos  
Jyväskylän yliopisto  
Kevät 2023

# JYVÄSKYLÄN YLIOPISTO

Tiedekunta Humanistis-yhteiskuntatieteellinen	Laitos Kieli- ja viestintätieteiden laitos
Tekijä Erika Inkeri Kilpeläinen	
Työn nimi Korpusavusteinen kriittinen diskurssintutkimus nimityksistä <i>mustalainen</i> ja <i>romani</i> Suomi24-keskustelupalstalla	
Oppiaine Suomen kieli	Työn laji Kandidaatintutkielma
Aika Kevät 2023	Sivumäärä 22 + liitteet
<p>Tiivistelmä</p> <p>Tutkielmassa selvitetään, minkälaisia diskursseja nimitysten <i>mustalainen</i> ja <i>romani</i> ympärille rakentuu, ja kuinka nämä diskurssit eroavat toisistaan. Tutkielmassa on mukana myös kriittinen näkökulma: siinä pohditaan vihapuheen roolia vähemmistöihin kohdistuvassa syrjinnässä. Tutkimuksen teoriana toimivat korpusavusteinen diskurssintutkimus sekä kriittinen diskurssintutkimus. Tutkimusaineisto koostuu Suomi24:n keskustelupalstan kommentteista vuosilta 2014–2020, ja se on kerätty Korp-konkordanssityökalun avulla. Tutkimusmenetelmänä on hyödynnetty avainsana-analyysiä. Tutkimuskysymykset ovat 1) Minkälaisia diskursseja nimityksistä <i>mustalainen</i> ja <i>romani</i> käytetään? 2) Miten diskurssit eroavat toisistaan? 3) Millainen rooli vihapuheella on vähemmistöihin kohdistuvassa syrjinnässä?</p> <p>Nimityksen <i>mustalainen</i> korpuksesta nousi esiin neljä diskurssia, ja nimityksen <i>romani</i> korpuksesta viisi. Tutkielma keskittyy molempien aineistojen kolmeen laajimpaan diskurssiin. M-korpuksen laajimmiksi diskursseiksi muodostuivat vihan diskurssi, stereotyyppisten elämäntapojen diskurssi sekä elättämisen diskurssi. R-korpuksen laajimmat diskurssit olivat yhteiskunnallisen aseman diskurssi, romanikielen ja -kulttuurin diskurssi sekä vähemmistöjen oikeuksien diskurssi.</p> <p>Tutkimuksen tuloksena muodostuneet diskurssit eroavat toisistaan merkittävästi: nimityksen <i>mustalainen</i> ympärillä käyty keskustelu Suomi24:ssä on selvästi vihamielisempää ja tuomitsevampaa kuin keskustelu, jota käydään <i>romani</i>-nimityksen yhteydessä. Vihapuheen rooli vähemmistöihin kohdistuvassa syrjinnässä nähdään toiseuttavana, leimaavana ja eriarvoistavana.</p>	
Asiasanat Korpusavusteinen diskurssintutkimus, kriittinen diskurssintutkimus, romanit, Suomi24, vihapuhe	
Säilytyspaikka Jyväskylän yliopisto	
Muuta tietoa	



## KUVIOT

KUVIO 1	M-korpukset diskurssiprosodiat.....	9
KUVIO 2	R-korpuksen diskurssiprosodiat.....	14

## TAULUKOT

KUVIO 1	M-korpukset diskurssiprosodiat.....	9
TAULUKKO 1	M-korpuksen diskurssit.....	9
TAULUKKO 2	Vihan diskurssi.....	10
TAULUKKO 3	Stereotyyppisten elämäntapojen diskurssi .....	11
TAULUKKO 4	Elättämisen diskurssi.....	12
KUVIO 2	R-korpuksen diskurssiprosodiat.....	14
TAULUKKO 5	R-korpuksen diskurssit.....	14
TAULUKKO 6	Yhteiskunnallisen aseman diskurssi.....	15
TAULUKKO 7	Vähemmistöjen oikeuksien diskurssi.....	16
TAULUKKO 8	Romanikulttuurin ja -historian diskurssi.....	17

## SISÄLLYS

1	JOHDANTO .....	1
2	KORPUSAVUSTEINEN DISKURSSINTUTKIMUS.....	3
3	MENETELMÄ JA AINEISTO .....	6
3.1	Suomi24-korpus .....	6
3.2	Avainsana-analyysi.....	7
4	ANALYYSI JA TULOKSET .....	8
4.1	Mustalainen-nimityksen diskurssiprosodiat ja diskurssit .....	8
4.1.1	Vihan diskurssi .....	9
4.1.2	Stereotyyppisten elämäntapojen diskurssi.....	11
4.1.3	Elättämisen diskurssi.....	12
4.2	Romani-nimityksen diskurssiprosodiat ja diskurssit .....	13
4.2.1	Yhteiskunnallisen aseman diskurssi .....	15
4.2.2	Vähemmistöjen oikeuksien diskurssi.....	16
4.2.3	Romanikulttuurin ja -historian diskurssi.....	17
4.3	Tulosten vertailu ja pohdinta .....	18
	LÄHTEET.....	20

## LIITTEET

# 1 JOHDANTO

Kandidaatintutkielmassani tutkin, kuinka romaneista keskustellaan Suomi24-keskustelupalstalla. Tarkoitukseni on ensisijaisesti selvittää, minkälaisia sanoja romanien nimityksiä *mustalainen* ja *romani* käytettäessä esiintyy, minkälaisia diskursseja niiden ympärille rakentuu ja kuinka ne eroavat toisistaan. Lisäksi pohdin vihapuheen roolia vähemmistöihin kohdistuvassa syrjinnässä.

Vihapuhe uhkaa sananvapautta, rakentavaa yhteiskunnallista keskustelua ja näin ollen koko demokraattista järjestelmää. (Saresma, Pöyhtäri, Knuutila, Kosonen, Juutinen, Haara, Tulonen, Nikunen & Rauta 2022: 12). Vihapuheen voi määritellä monin eri tavoin. Saresma ym. (2022: 12) ovat määritelleet sen tarkoittavan halventavia, uhkaavia tai leimaavia ilmaisuja, jotka liittyvät yksilön tai ihmisryhmän oletettuihin tai todellisiin ominaisuuksiin, ja joiden taustalla on suvaitsemattomuus tai syrjintä. Näitä ominaisuuksia voivat olla esimerkiksi sukupuoli, ihonväri tai etninen tausta. Jauholan, Siltalan ja Niemisen (2022: 9) selvityksen mukaan yksi yleisimmistä vähemmistöihin kohdistuvista vihapuheen tai häirinnän ilmenemismuodoista on sanallinen loukkaus. Suomi24:n keskustelut koostuvat kommentteista, joita käyttäjät voivat lähettää alustalle anonyymisti, nimimerkin turvin (Lagus, Pantzar, Ruckenstein & Ylisiurua 2016: 6). Anonymiteetti mahdollistaa sananvapauden, mutta toisaalta myös asiattoman ja häiritsevän käytöksen: sivustolla on näkyvillä erilaista vihamielistä ja asiatonta kielenkäyttöä, väkivaltaa ja rasismia (Harju 2018: 63; Pöyhtäri, Haara & Raittila 2013: 48). Suomessa vihan ja halveksunnan kohteita ovat olleet esimerkiksi ruotsalaiset, venäläiset, saamelaiset sekä romanit (Pöyhtäri ym. 2013: 33).

Romanit ovat kielellinen, etninen ja kulttuurinen vähemmistö. Suomessa on arviolta 13 000–14 000 romanialaista. (Romanit 2020a.) Romaniväestön omien kokemusten sekä monien selvitysten mukaan heihin kohdistuva syrjintä ja vihapuhe ovat lisääntyneet (Sosiaali- ja terveystieteiden tutkimuskeskus 2019: 10). Syrjintä näkyy esimerkiksi romanilasten koulukiusaamisena ja epäyhdenvertaisina työllistymismahdollisuuksina. Romanit syytetään muuta helpommin varkauksista – olivat he sitten työntekijän tai asiakkaan roolissa. Heille myös asetetaan usein erityisehtoja asunnon vuokraamisen yhteydessä. (Yhdenvertaisuus 2023a.) Ennakkoluuloihin pohjautuva syrjintä huojuttaa turvallisuudentunnetta sekä viranomaisluottamusta ja aiheuttaa tunnetta

ulkopuolisuudesta sekä ongelmia elämän eri osa-alueilla. (Korhonen, Jauhola, Oosi & Huttunen 2016: 58.) Romaneita on ajan saatossa syytetty kiertolaisuudesta, mutta samaan aikaan heitä on kuitenkin usein kielletty asettumasta aloilleen. Lisäksi heitä on rankaistu pelkästään romanina olemisesta tai siitä, että he nimittävät itseään romaneiksi. (Heikinheimo 2015: 46–47.) Romaneihin kohdistuva syrjintä on vakava ihmisoikeusongelma, josta Suomea on moitittu kansainvälisten ihmisoikeuksien valvovien elinten toimesta (Suomen romaniyhdistys 2023).

Vuonna 1992 kielitoimisto antoi suosituksen käyttää virallisissa yhteyksissä nimitystä *romani* nimityksen *mustalainen* sijaan (Friman-Korpela 2012: 229). Nimitys *romani* tulee kansainvälisen kielen ihmistä tarkoittavasta sanasta *rom*, ja se on suositeltu ja neutraali nimitys (Romano Missio 2022). Nimitys *mustalainen* sen sijaan voidaan assosoida negatiivisiin stereotypioihin, jotka ovat juurtuneet eurooppalaiseen kulttuuriin. Sen vuoksi termin käyttö julkisessa keskustelussa on lopetettu, mutta sitä käytetään yhä paljon arkikeskustelussa. (Stauber & Vago 2007.) Nimitys *mustalainen* koetaan halventavana ja leimaavana, kun sitä käytetään yhteisön ulkopuolella (Friman-Korpela 2009: 95). Valitsin tutkimukseeni nimitykset *mustalainen* ja *romani* juuri niihin liittyvien yhteiskunnallisten ja asenteellisten merkityserojen vuoksi sekä selvittääkseni, kuinka niiden käyttöerot näkyvät keskustelupalstalla. Aiheen valintaa ohjasi myös oma toiveeni tasa-arvoisemmasta ja vähemmistöille oikeudenmukaisemmasta yhteiskunnasta. Näiden asioiden myötä tutkimuskysymykseni muotoutuivat seuraavanlaisiksi: 1) Minkälaisia diskursseja nimityksistä *mustalainen* ja *romani* käytetään? 2) Miten diskurssit eroavat toisistaan? 3) Millainen rooli vihapuheella on vähemmistöihin kohdistuvassa syrjinnässä?

Tutkimusta suomalaisiin sekä ulkomaalaisiin romaneihin liitetystä diskursseista Helsingin Sanomissa ovat tehneet Toiviainen (2012) sekä Koivuranta (2013). Tuomainen (2014) on tutkinut romanikerjäläisilmiön diskursseja viranomaisraporteissa ja kansanedustajien kysymyksissä. Pietikäinen (2000) on tutkinut, kuinka etnisten vähemmistöjen ja enemmistön välistä eroa rakennetaan diskurssiivisesti uutisissa ja pääkirjoituksissa. Kriittistä korpusavusteista diskurssintutkimusta on tehty muiden vähemmistöjen representaatiosta Yhdysvaltain (Mahmoud & Bahareh 2017) sekä Iso-Britannian uutismediassa (Baker, Gabrielatos & McEnery 2013). Jantunen (2018a) on tutkinut muun muassa seksuaalivähemmistöihin liitettyjä diskursseja Suomi24-keskustelupalstalla hyödyntäen korpusavusteista diskurssintutkimusta sekä avainsana-analyysiä. Romaneihin liitettyjä diskursseja Suomi24:ssä ei ole nähdäkseni vielä tutkittu, ja edellinen romanidiskursseihin liittyvä tutkimus painottuu uutismediaan. Oma tutkimukseni antaa aiheesta erilaisen näkökulman, sillä Suomi24:ään voi kirjoittaa kuka tahansa, toisin kuin uutismediaan.

## 2 KORPUSAVUSTEINEN DISKURSSINTUTKIMUS

Tutkielmani teoreettinen viitekehys on korpusavusteisen diskurssintutkimus eli CADS (corpus-assisted discourse studies). CADS on kielen muotoa ja/tai toimintaa kommunikatiivisena diskurssina käsittelevä tutkimussuunta, jonka analyyseissä käytetään tietokonepohjaisia korpuksia (Partington, Duguid & Taylor 2013: 10). CADS:ssa yhdistyvät luonteeltaan määrällisen **korpuslingvistiikan** (corpus linguistics, CL) sekä laadullisen **diskurssintutkimuksen** (discourse analysis, DA) teoreettiset lähtökohdat (Ancarno 2020: 165–166). CADS pyrkii paljastamaan tekstistä merkityksiä, jotka eivät ole ilmiselvästi huomattavissa tutkittavassa diskurssityypissä. Yhdistämällä kvantitatiivinen lähestymistapa, eli tilastollinen yleiskatsaus suuresta diskurssimäärästä, sekä kvalitatiivinen lähestymistapa, eli tiettyjen diskurssin osien läheinen ja yksityiskohdainen analyysi, mahdollistetaan diskurssityypin prosessien ymmärtäminen eli piilomerkitysten löytäminen. (Partington ym. 2013: 11.)

CL:llä tarkoitetaan kielenkäytön, sen vaihtelun sekä kielen rakenteen tutkimista laajojen aineistojen avulla. Koska tutkimus pohjautuu suuriin aineistoihin ja niistä laskennallisesti saatuihin tuloksiin, tutkimustulokset ovat yleistettävissä ja kielestä tai ilmiöstä on mahdollista saada laajempi kuva. (Kyröläinen & Laippala 2021.) DA sen sijaan keskittyy siihen, mitä kielellä tehdään ja kuinka kieli ja sen käyttö yhdistyvät muihin yhteiskunnallisiin sekä sosiaalisiin rakenteisiin. Sen kielikäsitteellisyys korostaa vuorovaikutusta ja on luonteeltaan dialogista ja sosiaalista. (Pälli & Lillqvist 2021.)

Diskurssit voidaan määritellä erilaisiksi tavoiksi käyttää kieltä. Ne kietoutuvat yhteiskunnan käytäntöihin ja rakenteisiin, ja ne elävät jatkuvasti vuorovaikutuksessa muotoutuen, toisiinsa sulautuen tai toisistaan irtaantuen. (Baker 2006: 3–4.) Tässä tutkimuksessa tarkastelen diskursseja vakiintuneina tapoina puhua romaneista eri nimitysten yhteydessä.

CADS-tutkimus pohjautuu **fraseologiseen kielikäsitteeseen**, jonka mukaan kieli on toistuvaa, jaettua ja kontekstuaalista. Siinä korostuu näkemys, jonka mukaan kieltä tuotetaan fraseologisesti eli muistissa olevien ja kielenkäyttäjien jakamien



monisanaisten yksiköiden, leksikaalis-kieliopillisten sekä leksikaalis-semanttisten assosiaatioiden kautta. Se näkyy CADS:ssa yleisimmin myötäesiintymärakenteisiin, kuten diskurssiprosodioihin keskittymisenä. (Jantunen 2018b: 26–27.) Fraseologinen kielikäsite havainnollistuu esimerkiksi Hoeyn (2005) leksikaalisen primingin periaatteessa (lexical priming) sekä Sinclairin (1991) idiomiperiaatteessa (idiom principle). Omassa tutkimuksessa fraseologinen kielikäsite näkyy fraseologisiksi ilmaisuiksi käsiteltävien avainsanojen sekä diskurssiprosodioiden tarkastelussa.

CADS seuraa usein kolmivaiheista etikettiä: sana- tai avainsanaluetteloiden muodostaminen keskeisten semanttisten osa-alueiden tunnistamiseksi, kollokaatio-luettelon muodostaminen mallien tunnistamiseksi tai termien luokittelun aloittamiseksi sekä konkordanssiluettelon muodostaminen muiden hallitsevien mallien tutkimiseksi. Etikettiin voidaan lisätä myös neljäs vaihe, joka on läheinen, kielellinen ja diskurssianalyttisiin menetelmiin perustuva analyysi. Tutkijat eivät käytä vaiheita systemaattisesti, vaan soveltavat näitä omaan tutkimukseensa. (Ancarno 2020: 177.) Tässä tutkimuksessa vaiheet näkyvät avainsanaluetteloiden muodostamisena, niiden luokitteluksena diskurssiprosodioihin ja diskursseihin sekä niiden analysoimisena.

**Avainsanoilla** tarkoitetaan korpuslingvistiikassa tavallisimmin sanoja, jotka kuvaavat tekstin merkittäviä ominaisuuksia. Avainsanat toimivat siis tekstin ”avaimena”. (Kyröläinen & Laippala 2021.) Avainsanaluettelot ovat vertailevia tutkimustuloksia, joiden avulla voidaan todeta, kuinka tyypillinen tietty korpus on sanojen esiintymistiheyden suhteen. (Ancarno 2021: 169.) Avainsanaisuus sen sijaan on ominaisuus, joka sanoilla voi olla tietyssä tekstissä tai tekstikokonaisuudessa ja joka viittaa sanojen tärkeyteen. Se heijastaa sitä, mistä tekstissä on todella kyse, pikkuseikkoja sekä merkityksettömiä yksityiskohtia välttämällä. (Scott & Tribble 2006: 55–56.) **Diskurssiprosodioilla** eli merkitysryhmillä tarkoitetaan, että sanan, lauseen tai lemman ja niitä ympäröivien sanojen väliltä voi löytää johonkin diskurssiin viittaavia malleja (Baker 2006: 87). Ryhmittelyn avulla voidaan tunnistaa sosiaalistuneita merkityksiä ja puhetapoja, sillä avainsanat yksin eivät luo diskursseja (Jantunen 2018a: 8).

CADS:n käytössä on otettava huomioon myös sen rajoitteet. Analyysi ei esimerkiksi ota huomioon monia aineiston kontekstiin liittyviä ominaisuuksia (Flowerdew 2008: 115). Siinä on myös aina mahdollisuus tutkijan näkemyksen subjektiivisuuteen (Baker 2006: 10–11). Tutkija nimittäin valitsee tekstit, niiden analysointitavan sekä tutkimusmenetelmät (Baker, Gabrielatos, KhosraviNik, Krzyżanowski, McEnery & Wodak 2008: 277).

CADS ei ole sidottu mihinkään tiettyyn diskurssianalyysin koulukuntaan, ei esimerkiksi **kriittiseen diskurssianalyysiin** (CDA). Toisin kuin CDA:lla (critical discourse analysis), sillä ei ole yleistä poliittista agenda, ja sillä on hyvin erilaiset asenteet ja perinteet siitä, miten kielitietoa tulisi hallita. (Partington ym. 2013: 10.) CDA:n mukaan diskurssit ovat sekä sosiaalisesti konstitutiivisia että sosiaalisesti mukautuvia

(Blommaert 2005: 25). Se tutkii vallan, kielen ja ideologian välisiä suhteita (Fairclough 1995: 23). Se lähtee usein liikkeelle yhteiskunnallisesta epäkohdasta ja sille on ominaista ottaa avoimesti kantaa: määrite kriittinen voidaan ymmärtää tutkijan tehtävänä tuoda valtasuhteita näkyviksi, purkaa konventionaalistuneita ideologioita ja tarjota vaihtoehtoisia toimintatapoja (Solin 2012: 558).

Tutkimukseni käsittelee syvälle historiaan juurtunutta, yhteiskunnallista ongelmaa, joka tässä tapauksessa on etninen syrjintä: valtaväestön epäluulolla ja vihamielisyydellä on ollut ja on yhä suuria vaikutuksia romanien sosiaaliseen asemaan, kulttuuriseen itseymmärrykseen sekä suhteeseen ympäröivään yhteiskuntaan (Tervonen 2009: 82). Tämän vuoksi tarkastelen tuloksia myös kriittisestä näkökulmasta. CL ja CDA täydentävät toisiaan: objektiiviset, kvantitatiiviset lähestymistavat paljastavat tutkimuksen, havaintojen ja johtopäätösten yleisyyden tai luotettavuuden, mikä suojaa aineiston yli- tai alitulkitsemiselta (Baker ym. 2008: 297). Lisäksi ne antavat tilastollisen validiteetin rinnastussuhteille ja osoittavat todellisuuden tekstuaalisen rakentamisen suuntauksia, joita sitten voidaan analysoida ideologisesti (Hunt 2015: 269).

## 3 MENETELMÄ JA AINEISTO

### 3.1 Suomi24-korpus

Tutkimusaineistoni on Suomi24-korpus, joka koostuu Suomi24-keskustelufoorumin keskusteluista vuosilta 2001–2020. Se on tarkoitettu ei-kaupalliseen opetus- ja tutkimustarkoitukseen, eikä sen käyttöön tarvita erityistä lupaa. (Lagus ym. 2016: 5.) Korpuksen koko on 4,5 miljardia sanetta ja 391 miljoonaa virkettä (City Digital Group 2021). Kielitieteellisen tutkimuksen kannalta Suomi24 moniulotteinen tutkimuskohde, joka tarjoaa laaja-alaisen kokoelman arkityylistä keskustelua sekä puheenomaista kieltä (Lagus ym. 2016: 40).

Tämän tutkimuksen Suomi24-korpus on kerätty Kielipankin tarjoaman Korpuhakukäyttöliittymän avulla. Aineisto koostuu vuosilta 2014–2020 kerätyistä keskustelufoorumin keskusteluista. Valitsin kyseisen aikavälin saadakseni mahdollisimman tuoretta aineistoa, sillä keskustelu vähemmistöjen asemasta ja oikeuksista yhteiskunnassa on lisääntynyt viime vuosina. Aineistoa kerätessäni käytin Korpin laajennettua hakua, hakusanoistani sanoja *mustalainen* ja *romani*. Näiden hakusanojen käyttäminen jätti tutkimukseni ulkopuolelle muut sanat, joilla romaneita mahdollisesti nimitetään. Hakusanat olivat perusmuodossa, jotta tuloksissa näkyivät sanojen kaikki taivutusmuodot. Aineistoni on lemmatisoitu. Kaikki yhden sanan sanamuodot on siis muutettu sanan perusmuotoon (Sinclair 1991: 41). Saamani hakutulokset näkyivät kappalekontekstissaan. Nimitän sanan *mustalainen* tuloksia M-korpuksiksi ja sanan *romani* R-korpuksiksi. M-korpuksen koko oli 558 727 sanetta ja R-korpuksen 595 962 sanetta. Käsitelin aineistot Excelissä, jossa poistin kommenttien joukosta kaksoiskappaleet sekä turhat välilyönnit. Poistin R-korpuksesta viestit, joissa sana *romania* esiintyi maan nimenä eikä nimityksen *romani* partitiivimuotona. M-korpuksesta lopulliseen käyttöön jäi 16 322 konkordanssia ja R-korpuksesta 20 323.

## 3.2 Avainsana-analyysi

Hyödynsin tutkielmassani **avainsana-analyysia**, jonka avulla tunnistetaan sanoja, joita esiintyy tilastollisesti useammin yhdessä tekstiaineistossa verrattuna toiseen. Menetelmää voidaan käyttää tekstilajille ominaisten piirteiden määrittämiseen sekä analysoitujen tekstien välittämien diskurssien kuvaamiseen (Kyröläinen & Laippala 2021). Toteutin tutkielmani analyysin AntConc-korpusohjelmiston avulla. Ohjelmistoa voidaan hyödyntää avainsana-analyysissä, jota olen käyttänyt tässä tutkimuksessa menetelmänä. M- ja R-korpus toimivat tutkielmassani toistensa referenssikorpuksina, eli vertasin niitä toisiinsa. Ohjelmisto vertaa korpusten sanoja ja frekvenssejä, ja jättää listasta pois ne sanat, jotka esiintyvät molemmissa korpuksissa suurin piirtein yhtä usein (Scott & Tribble 2006: 59). Frekvensseillä tarkoitetaan sanojen alttiutta esiintyä yhdessä (Baker 2006: 2). Vertailun tuloksena sain kaksi avainsanalistaa, joiden avainsanaisuusarvo on laskettu log-likelihood-testin avulla. Testissä saadut avainsanat ovat tilastollisesti merkitseviä, kun niiden avainsanaisuusarvo on 3.8 tai korkeampi (Anthony 2022). Kaikki tutkimukseni avainsanat ylittävät tämän arvon.

Alun perin M-korpuksen keskeisimpiä avainsanoja oli 96 ja R-korpuksen avainsanoja 131. Poistin avainsanojen joukosta vieraskieliset sanat, funktiosanat sekä kaksi kertaa, vaikkakin eri muodossa esiintyvät sanat. Tällaiset sanat poistettaessa jäljelle jäivät ainoastaan kulttuurillisesti merkitsevät avainsanat (Baker 2006: 147). Sanalla *romani* tarkoitetaan henkilön lisäksi myös romanikieltä, mistä syystä osa R-korpuksen avainsanoista voi esiintyä merkitsevämpänä kuin mitä se todellisuudessa on. Ohjelma ei myöskään tunnistanut kaikkia avainsanoja oikein: se listasi avainsanojen sanoja, jotka esiintyivät pääasiassa muiden sanojen osana. Esimerkiksi avainsanaksi luokiteltu *ukko* esiintyi todellisuudessa muissa sanoissa, kuten *joukko* ja *lukko*. Päädyin poistamaan myös tällaiset sanat. Poistin myös M-korpuksesta avainsanan *mustalainen* ja R-korpuksesta avainsanan *romani*, sillä keskityn tutkielmassani juuri näiden sanojen kontekstiin. Sanojen poistamisen jälkeen M-korpuksesta tutkimukseeni päätyi lopulta 69 avainsanaa (liite 1) ja R-korpuksesta 95 (liite 2). Avainsanat on järjestetty liitteisiin niiden avainsanaisuuden mukaan.

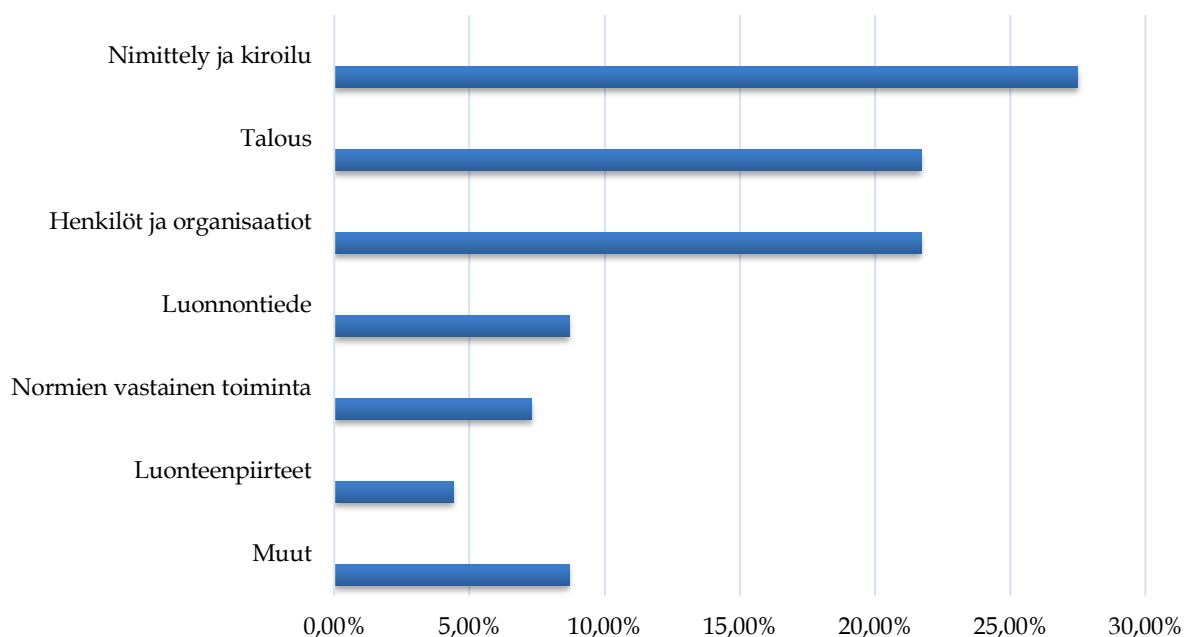
Jaottelin avainsanat diskurssiprosodioihin niiden denotatiivisen merkityksen mukaan, ottaen huomioon myös niiden käyttökontekstin. Sanat, jotka eivät kuuluneet selkeästi mihinkään diskurssiprosodiaan, sijoitin ryhmään Muut. Luokittelin avainsanat edelleen laajempiin diskurssisiin, jotka sisältävät sanoja useasta eri diskurssiprosodiasta. Perusteena luokittelulle toimi sanojen konteksti sekä selvärajaiset aihepiirit, joihin sanat kytkeytyivät. Sanat, jotka eivät kuuluneet selkeästi mihinkään aihepiiriin, jäivät luokittelun ulkopuolelle.

## 4 ANALYYSI JA TULOKSET

Tässä luvussa tarkastelen nimitysten *mustalainen* ja *romani* avainsanoja sekä niistä muodostuneita diskurssiprosodioita ja diskursseja. Tarkastelen ensimmäisenä M-korpusta. Sen jälkeen siirryn R-korpuksen tarkasteluun.

### 4.1 Mustalainen-nimityksen diskurssiprosodiat ja diskurssit

Avainsanoista muodostui yhteensä kuusi eri diskurssiprosodiaa (kuvio 1). Suurimmiksi diskurssiprosodioiksi muodostuivat nimittelyn ja kiroilun sekä henkilöiden ja organisaatioiden diskurssiprosodiat. Nimittely ja kiroilu -prosodia sisältää haukkuma- ja kirosanoja (19 kpl), kuten *äpäpä* ja *vittu*. Henkilöt ja organisaatiot -prosodia koostuu erisnimistä (esim. *Mäenpää*), yleisnimistä (esim. *yhteiskunta*, *maanviljelijä*) sekä karitiivijadjektiivista *vastikkeeton* (15 kpl). Talous-diskurssiprosodiaan sisältyy kulutushyödykkeisiin ja rahaan liittyviä substantiiveja (15 kpl), kuten *sosiaaliaavustus*, sekä verbejä kuten *elättää*. Luonnontiede-prosodia koostuu eläinten, maiden, kaupunkien sekä luonnonilmiöiden nimistä (6 kpl), kuten *hevonen*, *Lieksa* ja *kaamos*. Normien vastaiseen toimintaan sisältyy rangaistavaa tai yleisesti paheksuttua toimintaa kuvaavia verbejä (5 kpl), kuten *tappaa* ja *varastaa*. Luonteenpiirteet ja ominaisuudet -prosodia suhteellisia ominaisuuksia ja luonteenpiirteitä kuvaavia adjektiiveja (3 kpl), kuten *rehellinen* ja *hullu*.



KUVIO 1 M-korpuksen diskurssiprosodiat

Diskurssiprosodioista muodostui yhteensä neljä diskurssia (taulukko 1). Diskurssien ulkopuolelle jäi viisi avainsanaa. Käsittelen seuraavaksi tarkemmin kolmea suurinta diskurssia, joita ovat vihan diskurssi, stereotyyppisten elämäntapojen diskurssi sekä elättämisen diskurssi.

TAULUKKO 1 M-korpuksen diskurssit

Diskurssit	Avainsanojen määrä	Prosenttiosuus (%)
Vihan diskurssi	22	34,4 %
Stereotyyppisten elämäntapojen diskurssi	22	34,4 %
Elättämisen diskurssi	13	20,3 %
Historian ja politiikan diskurssi	7	10,9 %
<b>Yhteensä:</b>	<b>64</b>	<b>100 %</b>

#### 4.1.1 Vihan diskurssi

M-korpuksen laajimpia diskursseja on vihan diskurssi (taulukko 2). Diskurssi koostuu kolmen eri diskurssiprosodian sanoista. Avainsanat kuuluvat muun muassa nimittelyyn ja kiroiluun ja normien vastaiseen toimintaan.

TAULUKKO 2 Vihan diskurssi

Diskurssiprosodiat	Avainsanat
Nimittely ja kiroilu	<i>juoppo, loinen, manne, mamut, matut, mutakuono, neekeri, paskasakki, pummi, riesa, ryssä, äpäpä, helvetti, paska, perkele, saatana, vittu, vituttaa</i>
Normien vastainen toiminta	<i>ampua</i>
Luonteenpiirteet ja ominaisuudet	<i>hullu</i>
Muut	<i>perse, hopi</i>

Haukkuma- ja kirosanat ovat affektisia ilmauksia, joilla osoitetaan suhtautumista puheenalaisen asiaan tai puhekumppaniin (VISK § 1707). Haukkumasanoja ei käytetä keskustelussa ainoastaan romaniin halventamiseen: esimerkki 1 osoittaa, kuinka romaneita verrataan muiden vähemmistöjen tai kansojen kanssa. Vertailua voidaan ilmaista esimerkiksi komparatiivisanan ja *kuin*-ilmauksen avulla (VISK § 1172), mikä näkyy esimerkissä 1. Esimerkissä 2 *saatana* toimii lausekkeessa haukkumanimen genetiivimäärityksenä, nimittelytarkoituksessa (nimittelystä ks. VISK § 1726). Kirosanat käyttö lausekkeessa vahvistaa puhuttelun kielteistä sävyä. Vihaa ilmaistaan myös tunnetta ja toimintaa kuvaavilla verbeillä kuten *vituttaa* ja *ampua* (esimerkit 3 ja 4). *Vituttaa* on johdos affektista sävyä ilmaisevasta kirosanasta *vittu* (VISK § 1725). Sen avulla ilmaistaan negatiivista suhtautumista romaneihin (esimerkki 3). Verbi *ampua* viittaa esimerkissä 4 romaniin tappamiseen, ja se tuodaan kommentissa esiin positiivisena, palkitsemisen arvoisena asiana.

- 1) Jätejenkit ja roskaryssät ovat molemmat maapallon pohjasakkaa, melkein pahempia kuin somalit ja *mustalaiset*. Melkein.
- 2) *Mustalaiset* huijasivat eläkeläismiehiä 200000 euron edestä, eivätkö nuo **saatanan** työtävieroivat elikot lopeta koskaan älyllisesti heikompimielisten ihmisten huijaamista ja ryöväämistä. Kuvitelkaa kuinka paljon parempi paikka maamme ja koko maailmakin olisi asua sekä elää ilman *mustalaisia*.
- 3) Enempi **vituttaa** nuo *mustalaiset* ne ne vasta ihme sakkia ja ne vois lähettää takas kotiinsa.
- 4) Ai **ampui** *mustalaisia*? No siitähän pitäisi antaa mitali! 😊

Suhteellista ominaisuutta kuvaavalla adjektiivilla *hullu* (ks. VISK § 605) viitataan romaniin irrationaaliseen käytökseen – joskin epäsuorasti (esimerkki 5). Esimerkissä 5 rinnastuvat lauseenjäsenet *hullu* sekä *mustalainen*. *Joko-tai* on rinnastava parikonjunktio (VISK § 815). Kommentteissa myös kehoitetaan romaneita lähtemään Suomesta, ja kehoitusta vahvistamassa on nopeampaan liikkeeseen kannustava avainsana *hopi* (esimerkki 6).

- 5) Mikävitun lääkitys tällä keharikakaralla oikein on? se on joko täys **hullu** tai sitten *mustalainen*.
- 6) Menehän mustalainen romaniaan. **hopihopi!**

Vihan diskurssi koostuu romaneita sekä muita vähemmistöjä halventavista, uhkaavista ja leimaavista viesteistä. Romaneita usein myös verrataan tai rinnastetaan muihin vähemmistöihin. Heitä myös kehoitetaan lähtemään pois Suomesta.

#### 4.1.2 Stereotyyppisten elämäntapojen diskurssi

Toinen M-korpuksen laajimmista diskursseista on stereotyyppisten elämäntapojen diskurssi (taulukko 3), joka koostuu kuuden eri diskurssiprosodian sanoista. Avainsanat kuuluvat muun muassa henkilöihin ja organisaatioihin sekä normien vastaiseen toimintaan.

TAULUKKO 3 Stereotyyppisten elämäntapojen diskurssi

Diskurssiprosodiat	Avainsanat
Henkilöt ja organisaatiot	<i>mustalaisnainen, somali, puolimustalainen, kaveri, yhteiskunta</i>
Normien vastainen toiminta	<i>varastaa, valehdella, tappaa</i>
Luonnontiede	<i>kärpänen, hevonen</i>
Talous	<i>auto, viina, kartonki, työ, Bmw, syödä</i>
Luonteenpiirteet ja ominaisuudet	<i>rehellinen, laiska</i>
Muut	<i>vaihtaa, kuolla</i>

Romanien oletettua rikollisuutta ja epärehellisyyttä tuodaan kommenteissa esiin muun muassa vitsien ja sanontojen muodossa (esimerkit 7 ja 8) sekä normienvastaista toimintaa kuvaavilla verbeillä kuten *varastaa* ja *valehdella* (esimerkki 9). Luonteenpiirrettä kuvaava adjektiivi *rehellinen* esiintyy viesteissä, joiden tarkoitus onkin korostaa romanien oletettua epärehellisyyttä (esimerkki 10). Romanien rikollisuuden korostaminen näkyy myös esimerkissä 11, jossa sanotaan vankiloiden olevan täynnä romaneita. Viesteissä viitataan usein romaneihin yhtenä, samalla tavalla ajattelevana joukkona ilman erilaisia elämäntapoja ja -katsomuksia. Tämä näkyy muun muassa kvanttoripronominin *kaikki* käytössä esimerkissä 9.

- 7) Jos samassa tilassa on **somali**, mustalainen, muslimi ja venäläinen, miksi suomalaista kutsutaan? -Vartijaksi.
- 8) Tuttua kauraa, sanoi *mustalaisen* **hevonen** kun naapurin pellostä söi.
- 9) *Mustalaiset* kaikki **varastaa, valehtelee** ja huijaa. Opettaa pentunsakin tekemään niin. Ne lietsoo ongelmia läsnäolollaan. *Mustalaiset* mannet ovat kaikki paskoja! Painukoon helvettiin! Mä injoan niitä. Enkä voi sietää niitä. Pysykköön poissa mun läheltä. Vtun *mustalaiset*.



- 10) en oo kyllä kertaakaan nähnyt saatika kuullut että olisi joku mustalainen ottanut **rehellistä** yks ykköstä :D aina pitää olla soittelemassa suvulle tai muutenvaan kaivetaan puukko esiin. Säälittävät limasiat
- 11) Vankilat ovat täynnä muslimeja ja *mustalaisia*. Uskotko aikuisten oikeasti, että **yhteiskunta** heittäisi minut niiden sekaan raiskattavaksi? Varsinkin, kun on yleisessä tiedossa se fakta, että olen rasisti.

Talous-diskurssiprosodian sanat esiintyivät viesteissä, joissa puhutaan romanien väitettuihin tapoihin ja tottumuksiin liittyvistä hyödykkeistä, kuten alkoholista ja autoista (esimerkki 12). Tällaiset substantiivit, kuten *työ* liittyvät usein keskusteluun romanien väitetyistä työntekotottumuksista (esimerkki 13). Esimerkissä 14 verbi *kuolla* on osa vanhaa sanontaa, jonka voi tulkita viittaavan romanien valmiuteen tehdä kyseenalaisia asioita rahan säästämisen eteen.

- 12) *Mustalaiselle* työ merkitsee **autoon** hyppäämistä ja puukko kädessä putsamaan yöllisiä kulkijoita.
- 13) Helmikuussa päässyt vapaaksi, onkohan tehnyt päivääkään levyseppähitsarin **töitä**? Meinaan kun on *mustalainen*, eikös romanikulttuuri kiellä rehellisen työntöön...
- 14) Se on sitten niinkin *mustalaisella*, " juuri kun sai oppimaan hevosensa olemaan syömättä niin se otti ja **kuoli** ".

Stereotyyppisten elämäntapojen diskurssin muodostava keskustelu liittyy nimensä mukaisesti romanien elämäntapoihin ja on täynnä stereotyyppisiä: romaneihin yhdistetään tietynlainen käyttäytyminen ja hyödykkeet. Keskustelu sisältää oletuksia, yleistyksiä, kokemuksia ja pohdintoja muun muassa romanien työntekotottumuksista, epärehellisyydestä ja rikollisuudesta.

#### 4.1.3 Elättämisen diskurssi

Seuraavaksi laajin diskurssi M-korpuksessa on elättämisdiskurssi (taulukko 4). Se sisältää avainsanoja viidestä eri diskurssiprosodiasta, ja avainsanat liittyvät muun muassa henkilöihin ja organisaatioihin sekä talouteen.

TAULUKKO 4 Elättämisen diskurssi

Diskurssiprosodiat	Avainsanat
Henkilöt ja organisaatiot	<i>vastikkeeton, veronmaksaja, duunari, maanviljelijä, sossu</i>
Talous	<i>sosiaalivastuu, maatalous, ruoka, maksaa, elättää</i>
Normien vastainen toiminta	<i>loisia</i>
Nimittely ja kiroilu	<i>maatalousloinen</i>
Luonnontiede	<i>afrikka</i>

Henkilöt, organisaatiot ja ryhmät -prosodiaan kuuluvilla sanoilla viitattiin kommentteissa erilaisten tukien varassa eläviin ihmisiin, sosiaalivirastoon tai niihin, jotka tekevät töitä ja maksavat veroja (esimerkki 15). Kommenteissa sanotaan usein maatalouden harjoittajia ja romaneita loisimisesta ja heidät usein rinnastetaan keskenään. Näissä kommentteissa on havaittavissa talouden, normien vastaisen toiminnan sekä nimittelyn piiriin sisältyviä sanoja, ja nämä näkyvät esimerkeissä 16 ja 17.

- 15) **Veronmaksajilta** huijatus ja rikoksista saadut rahat *mustalaiset* pistävät viinaan ja huumeisiin.
- 16) Talikkokommarin palkka on oikeasti 4euroa tunti, loput ne saa tukiaisina keskimäärin 90 000 per tila, vain 4 euroa ne pystyy tekemään tulosta eli **maatalous** ei ole edes elinkeino se on verrattavissa maahanmuuttajien ja *mustalaisten* laiskuuteen ja työtä karttavaksi loisiksi joita veronmaksajat joutuu elättämään vasten tahtoaan.
- 17) **Maatalousloiset** on mustalaisiin ja mamuihin verrattavia sosiaaliavustuksilla elätettäviä pummeja!!!! tosin tulevat paljon kalliimmiksi.

Elättämiskurssiin sisältyy myös luonnontiede-diskurssiprosodian sana *Afrikka*, jota käytetään usein puhuttaessa "Afrikan mustalaisista". Tällä keskustelussa toistuvalla metaforalla viitataan somaleihin, joihin romaneita ja heidän työntekotottumuksiaan verrataan.

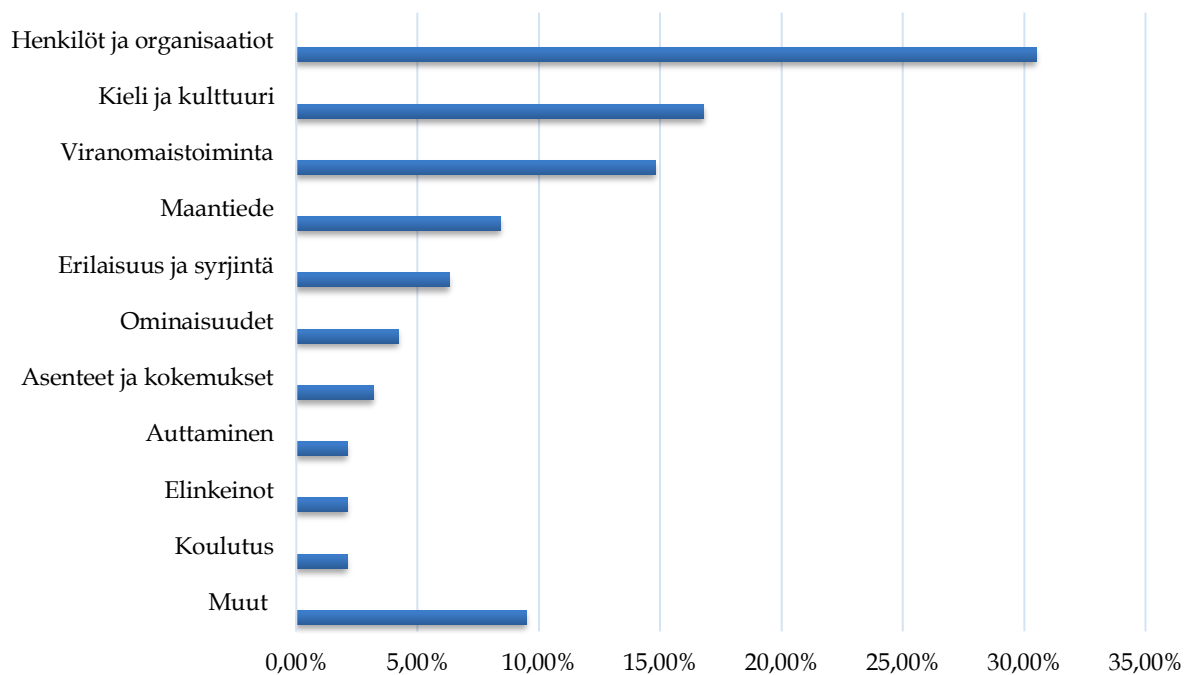
- 18) Taitaa olla niin, että **Afrikan mustalaiset**, somalit, käyvät enemmän duunissa kuin täällä sa-toja vuosia maleskelleet Intiasta tulleet.

Elättämisen diskurssi sisältää siis keskustelua, joka käsittelee romanien väitettyä elämistä sosiaaliturvan varassa. Keskustelussa pohditaan myös, ovatko romanit ansainneet yhteiskunnan heille maksamia rahoja. Myös tässä diskurssissa romaneita verrataan muihin vähemmistöihin.

## 4.2 Romani-nimityksen diskurssiprosodiat ja diskurssit

Nimityksen romani avainsanoista muodostui yhteensä kymmenen diskurssiprosodiaa (kuvio 3). Suurin niistä oli henkilöt ja organisaatiot -prosodia, joka koostuu yksilöitä ja ryhmiä kuvaavista yleisnimistä (esim. *romaninuori*, *yhteisö*), erisnimestä *Kinnunen* sekä karitiiviadjektiivista *paperiton* (29 kpl). Toiseksi laajimman diskurssiprosodian, kielen ja kulttuurin, muodostivat substantiivit (16 kpl) kuten *romanikulttuuri* sekä *suomenkielinen*. Viranomaistoiminta-prosodia sisältää viranomaisten toimintaan viittaavia substantiiveja, kuten *hanke*, verbejä, kuten *kehittää* ja adjektiiveja, kuten *virallinen* (14 kpl). Maantieteelliset alueet, kuten *Eurooppa* ja *Israel*, muodostavat maantiede-prosodian (8 kpl). Erilaisuus ja syrjintä -prosodia koostuu verbeistä (*syrjiä*,

*kohdistua*), substantiiveista (esim. *ennakkoluulo*) sekä adjektiivista *etninen* (6 kpl). Asenteisiin ja kokemuksiin sisältyy mentaalista toimintaa kuvaavia verbejä kuten *hyväksyä* ja *kunnioittaa* (3 kpl). Auttaminen-diskurssiprosodia sisältää avunantoon ja sen pyytämiseen liittyvät verbit *auttaa* ja *pyytää*. Elinkeinot-diskurssiprosodiaan sisältyy verbi *kerjätä* sekä substantiivi *rikollisuus*. Koulutus-diskurssiprosodia taas koostuu verbistä *opiskella* ja substantiivista *koulutus*.



KUVIO 2 R-korpuksen diskurssiprosodiat

R-korpuksen avainsanoista muodostui viisi eri diskurssia (taulukko 5). Diskurssien ulkopuolelle jäi neljä avainsanaa. Käsittelen seuraavaksi tarkemmin kolmea laajinta diskurssia. Pyöristämisen vuoksi diskurssien yhteenlaskettu prosenttiosuus ei ole tasan 100 %.

TAULUKKO 5 R-korpuksen diskurssit

Diskurssi	Avainsanojen määrä	Prosenttiosuus (%)
Yhteiskunnallisen aseman diskurssi	36	39,6 %
Vähemmistöjen oikeuksien diskurssi	28	30,7 %
Romanikulttuurin ja -historian diskurssi	15	16,5 %
Toiseuttamisen diskurssi	7	7,7 %
Auttamisen diskurssi	5	5,5 %
<b>Yhteensä:</b>	<b>91</b>	<b>100,1 %</b>

#### 4.2.1 Yhteiskunnallisen aseman diskurssi

R-korpuksen laajin diskurssi on yhteiskunnallisen aseman diskurssi (taulukko 6). Diskurssi sisältää kahdeksan eri merkitysryhmän sanoja, jotka liittyvät muun muassa henkilöihin ja organisaatioihin sekä viranomaistoimintaan.

TAULUKKO 6 Yhteiskunnallisen aseman diskurssi

Diskurssiprosodiat	Avainsanat
Henkilöt ja organisaatiot	<i>romaninuori, kerjäläinen, johtaja, kansanryhmä, kieliryhmä, yhteisö, ryhmä, romaniperhe, romaniväestö, romaniyhteisö, romanisuku, valtaväestö, kinnunen</i>
Koulutus	<i>opiskella, koulutus</i>
Viranomaistoiminta	<i>selvitys, hanke, palvelu, ohjelmatarjonta, määräaika, romaniasia, neuvottelukunta, tarjota, virallinen</i>
Elinkeinot	<i>kerjätä, rikollisuus</i>
Ominaisuudet	<i>mahdollisuus, asema</i>
Asenteet ja kokemukset	<i>hyväksyä</i>
Maantiede	<i>eurooppa, romania, bulgaria</i>
Muut	<i>tausta, romaniveri, tieto, liittyä</i>

Henkilöt ja organisaatiot -prosodian yleisnimet nostavat esiin muun muassa erilaisten romanien muodostamien ryhmien (esim. *romaniperhe, romaniväestö*) asemaa suhteessa muuhun yhteiskuntaan, kuten *valtaväestöön* (esimerkki 19). Vertailu näkyy *se*-sanana inessiivin *siinä* sekä vertailukonjunktion *kuin* käytössä. Esimerkki 19 havainnollistaa myös keskustelua romanien koulutuksesta, mikä näkyy toimintaa kuvaavan verbin *opiskella* käytössä. Substantiivi selvitys esiintyy keskustelussa viranomaisten toiminnasta romanien aseman suhteen (esimerkki 20). Esimerkki 21, jossa esiintyy toimivaa kuvaava verbi *kerjätä*, havainnollistaa keskustelua väitetyin aseman ja väitettyjen tapojen muutoksesta, mistä kielii kieltosana *ei* sekä negatiivinen, temporaalinen partikkeli *enää*. *Mahdollisuus*, joka on adjektiivikantainen ominaisuudennimi (ks. VISK § 176), esiintyy keskustelussa monenlaisissa romanien asemaan liittyvissä konteksteissa, kuten asumisesta (esimerkki 22) tai elinkeinoista puhuttaessa.

19) *Romanit opiskelevat* ja suorittavat ammatteihin vaadittavia tutkintoja siinä kuin *valtaväestökin*.

20) 27.3.2014 11:17 Ruotsin hallitus on kutsunut hotelli- ja ravintola-alan edustajat kriisikokoukseen *romanien* syrjinnän takia, kertoo uutistoimisto TT. Taustalla on tapaus joka sattui tällä viikolla Tukholman Sheraton-hotellissa romanien syrjintää käsittelevän *selvityksen* julkaisun yhteydessä.

21) *Romanit* eivät enää *kerjää* ja varasta eivätkä petä ihmisiä, vaan pyrkivät kouluttautumaan ja pääsemään mukaan normaaliin työelämään.

- 22) Tällä hetkellä Suomessa ihmiset saa vuokrata tai myydä asuntonsa vapaasti kenelle tahansa. *Romanit* ovat itse omalla käytöksellään pilanneet **mahdollisuudet** saada asunto yksityiseltä markkinoilta. Joten herättää *romanit* nyt on 2015 vuosi.

R-korpuksessa keskustellaan myös muualta Suomeen tulevista romaneista sekä romanien asemasta kansainvälisellä tasolla. Keskustelussa on tyypillisesti paljon maantiede-prosodian erisnimiä, kuten *Eurooppa* (esimerkki 23):

- 23) Olen usein miettinyt, miksi *romanit* ovat syrjittyjä jokaisessa **Euroopan** maassa.

Yhteiskunnallisen aseman diskurssi koostuu keskustelusta, jossa käsitellään romanien yhteiskunnallista asemaa muun muassa koulutuksen, sopeutumisen ja asumisen kautta. Näistä näkökulmista keskustellaan myös kansainvälisellä tasolla. Lisäksi keskustelussa pohditaan aseman muutosta.

#### 4.2.2 Vähemmistöjen oikeuksien diskurssi

Vähemmistöjen oikeuksien diskurssi (taulukko 7) muodostuu viiden eri diskurssiprosodian sanoista. Sanat kuuluvat muun muassa kieleen ja kulttuuriin sekä henkilöihin ja organisaatioihin.

TAULUKKO 7 Vähemmistöjen oikeuksien diskurssi

Diskurssiprosodiat	Avainsanat
Kieli ja kulttuuri	<i>kieli, vähemmistökieli, kansalliskieli, pakoruotsi, viittomakieli, romanikieli, äidinkieli, meänkieli, kielilaki, ruotsinkielinen, suomenkielinen, viittomakielinen</i>
Henkilöt ja organisaatiot	<i>turvapaikanhakija, maahanmuuttaja, paperiton, saamelainen, romanialainen, homoseksuaali, vähemmistö</i>
Viranomaistoiminta	<i>perustuslaki, oikeus, säätää, kehittää</i>
Ominaisuudet	<i>vammaisuus</i>
Muut	<i>asia, yhtäläinen, saapua, anteeksi</i>

Keskustelu romanien sekä muiden vähemmistöjen kieleen ja kulttuuriin oikeuksista näkyy esimerkissä 24, joka sisältää kolmen eri diskurssiprosodian sanoja: henkilöt ja organisaatiot (*saamelainen*), viranomaistoiminta (*oikeus, kehittää*) ja kieli ja kulttuuri (*kieli*). Esimerkki 24 on suora lainaus Suomen perustuslain 17 §:n kolmannesta momentista. Esimerkissä 25 viitataan myös perustuslakiin, ja siinä nostetaan esiin myös viittomakielisten sekä vammaisten muodostamien vähemmistöjen oikeudet. Samanlaisuutta kuvaava adjektiivi *yhtäläinen* esiintyy R-korpuksessa pääasiassa sanan *oikeus* kanssa (26).

- 24) **Saamelaisilla** alkuperäiskansana sekä *romaneilla* ja muilla ryhmillä on **oikeus** ylläpitää ja **kehittää** omaa **kieltään** ja kulttuuriaan.
- 25) Samassa perustuslain pykälässä taataan myös saamelaisilla alkuperäiskansana sekä *romaneille* ja muille ryhmille oikeus ylläpitää ja kehittää omaa kieltään ja kulttuuriaan. Saamelaisien oikeudesta käyttää saamen kieltä viranomaisessa säädetään lailla. Viittomakieltä käyttävien sekä **vammaisuuden** vuoksi tulkitsemis- ja käännösapua tarvitsevien oikeudet turvataan lailla.
- 26) Jokaisella Suomen *romanilla* on **yhtäläiset** oikeudet käydä koulua ja työllistyä.

Diskurssi koostuu vähemmistöjen oikeuksia käsittelevistä viesteistä, joissa viitataan usein Suomen perustuslakiin. Keskustelu ei rajoitu ainoastaan romanien oikeuksiin. Siinä pohditaan myös muiden vähemmistöjen kielellisiä ja kulttuurisia oikeuksia.

#### 4.2.3 Romanikulttuurin ja -historian diskurssi

Romanikulttuurin ja -historian diskurssi (taulukko 8) koostuu kuuden eri merkitysryhmän sanoista.

TAULUKKO 8 Romanikulttuurin ja -historian diskurssi

Diskurssiprosodiat	Avainsanat
Kieli ja kulttuuri	<i>kulttuuri, lippu, romanikulttuuri, kansallispäivä</i>
Henkilöt ja organisaatiot	<i>romaninainen, romanimies, serbi</i>
Asenteet ja kokemukset	<i>kunnioittaa</i>
Maantiede	<i>helsinki, israel, yhdysvallat, suomi, ruotsi</i>
Erilaisuus ja syrjintä	<i>etninen</i>
Ominaisuudet	<i>liikkuvuus</i>

Romanien kulttuuriin liittyvä keskustelu ilmenee tyypillisesti kulttuuriin liittyvistä substantiiveista, kuten *kansallispäivä* (esimerkki 27). Substantiiveja käytettiin usein, kun keskusteltiin tavoista, joita romanikulttuuriin kuuluu tai ei kuulu. Yleisnimet *romaninainen* (esimerkki 28) ja *romanimies* esiintyivät tyypillisesti juuri jommankumman mainitun sukupuolen tapoihin tai kulttuuriin viittaavassa keskustelussa. Verbi *kunnioittaa* esiintyy myös tyypillisesti perinteisiin ja tapoihin liittyvässä keskustelussa: esimerkiksi kunnioitukseen toisia romaneita tai romanikulttuurin tapoja kohtaan (esimerkki 29).

- 27) Tänäpäin vietetään *romanien kansallispäivää* ja luin lehdestä, että he juhliivat asiaa eri tavoin niin tiedättekö milläläilla tätä juhlapäivää vietetään Kokkolassa ja voiko tuota juhlintaa nähdä jossain käytännössä olisin ehkä kiinnostunut näkemään?
- 28) Ihan suomen *romani* olen ☺ Ja joo ihan oikeassa olet, varmaan aika harvoin näkyy kaupungilla raskaana olevia **romaninaisia**. Se johtuu siitä että mustalaisyksikulttuurissa raskaus yms. on häpeää. Vanhempia ihmisiä pitää kunnioittaa, ja ei ole soveliasta näyttäytyä heidän edessään maha pystyssä :D Raskaus salataan siis ihan loppuun asti vanhemmilta.

- 29) Mikäli kohtaat muunnimisiä *romaneja*, heidän nimensä on väärennetty, he eivät ole romaneja, heidän vanhempansa eivät **kunnioittaneet** romanien perinteitä tai he valehtelevat.

Maantiede-prosodian avainsanat esiintyivät tyypillisimmin historiaan liittyvässä keskustelussa. Esimerkissä 30 keskustellaan siitä, kuinka romanit ovat saapuneet Suomeen ja milloin. Myös avainsanat *liikkuvuus* ja *etninen* esiintyivät R-korpuksessa pääasiassa historiaan liittyvässä keskustelussa (esimerkit 31 ja 32).

- 30) *Romanit* ovat tulleet **Suomeen** noin viisisataa vuotta sitten pääasiassa lännestä **Ruotsin** kautta.
- 31) Se oli niitäalkuaikojia kun vapaa **liikkuvuus** toi ulkomaan *romanit* **Helsinkiin**.
- 32) Koko Saksan armeija syyllistyi itärintamalla **etnisiin** puhdistuksiin toisen maailmansodan aikana ja myös heidän juutalaisiin, *romaneihin* ja vammaisiin kohdistuneet toimensa ovat tiedossa.

Diskurssi muodostuu romanien kulttuuria, siihen kuuluvia tapoja, perinteitä ja historiaa käsittelevästä keskustelusta. Historiakeskustelussa nousee usein esiin se, kuinka romanit ovat päätyneet jostakin johonkin, ja mitä he ovat historian saatossa joutuneet kokemaan.

### 4.3 Tulosten vertailu ja pohdinta

Tutkimukseni tarkoitus oli ensisijaisesti selvittää, millaisia diskursseja nimitysten *mustalainen* ja *romani* ympärille rakentuu ja kuinka kyseiset diskurssit eroavat toisistaan. Jaottelin M-korpuksen avainsanat neljään diskurssiin: vihan diskurssiin, stereotyyppisten elämäntapojen diskurssiin, elättämisen diskurssiin sekä historian ja politiikan diskurssiin. R-korpuksen diskursseja muodostin yhteensä viisi: yhteiskunnallisen aseman diskurssi, vähemmistöjen oikeuksien diskurssi, romanikulttuurin ja -historian diskurssi, toiseuttamisen diskurssi sekä auttamisen diskurssin. Luokittelin avainsanat niiden denotatiivisen merkityksen sekä käyttökontekstin perusteella. Osa avainsanoista jäi diskurssien ulkopuolelle. Keskityin tarkastelemaan kolmea suurinta diskurssia molemmista korpuksista.

Diskursseissa on hieman samankaltaisuuksia: molemmat korpuukset sisältävät keskustelua romanien historiasta ja tavoista. Korpuksen diskurssit ovat kuitenkin luonteiltaan selkeästi erilaisia. Huomattavin ero tuloksissa on, että keskustelu nimityksen *mustalainen* ympärillä Suomi24:ssä on selkeästi vihamielisempää, tuomitsevämpää ja negatiivisävytteisempää kuin keskustelu, jota käydään nimityksen *romani* ympärillä. *Romani*-nimitystä ympäröivän keskustelun luonne on huomattavasti yhteiskunnallisempaa ja asiallisempaa. Tämä oli nähtävissä jo vihapuheeksi

tunnistettavien avainsanojen määrien erossa, mutta se nousi esiin myös diskurssien kontrastissa. Tutkimuksessani paljastuneet diskurssit ovat osin samankaltaisia diskursseja kuin Toiviaisen (2012) sekä Koivurannan (2013) tutkimuksissa. Tutkimustani aiemmasta tutkimuksesta erottaa eniten juuri vihan diskurssi, joka ei ole Helsingin Sanomissa yhtä vahvasti nähtävissä kuin Suomi24:ssä. Tästä huolimatta esimerkiksi ulkomaalaisromaneista kirjoitettiin Koivurannan (2013: 47) mukaan lähes poikkeuksetta vain negatiivisten uutisten yhteydessä.

Kolmas tutkimuskysymykseni kuului: Millainen rooli vihapuheella on vähemmistöihin kohdistuvassa syrjinnässä? Käyttämämme sanat eivät ole vain sanoja. Ne kantavat mukanaan historiaa ja valta-asetelmia (Jäntti & Martin 2007). Ryhmien tai yksittäisten henkilöiden yksilöiminen nimeämällä heidät pilkkaavasti, halventavasti tai leimaavasti on yksinkertaisin kielellisen ja retorisen syrjinnän muoto (Reisigl & Wodak 2001: 45). Nimitys *mustalainen* perustuu ulkoisiin tunnuspiirteisiin, minkä vuoksi sitä voidaan pitää yhteisön ulkopuolelta tulevana syrjivänä, rasistisena ja leimaavana. Etniset nimitykset voivat toimia diskursiivisen vallan muotona, jota käytetään muun muassa ryhmärajojen osoittamiseen, marginalisointiin tai oikeuksien tavoittelemiseen. Nimitykset voivat olla sävyltään loukkaavia – tätä loukkaavuuden tunnetta ei voi määritellä tai arvioida kukaan ulkopuolinen. (Pietikäinen 2002: 250.) Monet nimitykset, joita etnisistä ryhmistä käytetään, korostavat valkoisen kulttuurin valtaa määritellä ja nimetä kansallisuus sekä etnisuus (Friman-Korpela 2009: 91). Vihapuheella voi siis nähdä olevan vähemmistöjen syrjinnässä toiseuttava, leimaava ja eriarvoistava rooli.

Romanit kategorisoidaan usein yhdeksi ryhmäksi ilman, että heidät kohdattaisiin yksilöinä elämän eri osa-alueilla (Romanit 2020b). Varsinkin sellaiset suomalaiset, joilla ei ole kontakteja romaneihin, näkevät heidät ryhmänä, eivät yksilöinä (Stenroos 2014: 434). Ihmisiä tulee kutsua sillä nimityksellä, jolla he itse toivovat itseään kutsutavan. Tutkimukseni keskittyy kieleen ja vihapuheeseen, mutta vihapuhe on vain pieni osa laajempaa ilmiötä. Romanivastaisuuteen (*antigypsyism*) kuuluu laaja kirjo syrjiviä ilmauksia ja käytäntöjä, joihin lukeutuu monia implisiittisiä ja piilotettuja ilmenemismuotoja. Siinä ei ole kyse vain siitä, mitä sanotaan, vaan myös siitä, mitä tehdään ja jätetään tekemättä (Antigypsyism 2017: 2). Koska tutkielmani aihe on osa laajaa ja syvälle juurtunutta ongelmaa, se vaatii vielä paljon syvempää käsittelyä. Kandidaatintutkielmani antaa mahdollisuuden vain aiheen pinnalliseen tarkasteluun. On myös tärkeää huomioda, että enemmistökulttuurin edustajana en ole täysin pätevä kirjoittamaan aiheesta ymmärtäen sitä kokonaisuudessaan. Se ei siis sovellu aiheen yleispäteväksi tiedoksi, vaan toimii enemmänkin osana tärkeää keskustelua ja tietoisuuden levittämistä. Citizen Mindscapes -hankkeen yhteydessä toteutetun kyselyn mukaan Suomi24:n keskustelijajoukko koostuu valtaosin keski-ikäisistä miehistä (Harju 2018: 54–55). Tutkimusaineistoni ei siis voi sanoa edustavan yhteiskunnan



jäsenten ajatuksia monipuolisesti. Lisäksi tavat luokitella avainsanoja ovat erilaisia, joten eri tutkijalla voisi olla erilainen näkemys diskurssiprosodioista sekä diskursseista. Osan sanoista olisi voinut sijoittaa myös toiseen diskurssiin, sillä niiden esiintymiskonteksti oli vaihteleva.

Saresma ym. (2022: 181–186) esittävät toimenpide-ehdotuksia, joilla voi vaikuttaa vaikuttaa verkkovihan tuottajiin sekä verkkovihan ennaltaehkäisyyn ja sen vähentämiseen. Ehdotuksiin lukeutui muun muassa tutkimustiedon tuottaminen. Tutkielmani antaa tähän kategoriaan kattavan kuvauksen siitä, millaista keskustelua Suomi24:ssä on käyty tietyinä ajankohtana, ja sen tarkoitus on tuoda yhä näkyvämmäksi yhteiskunnassa piilevää, rakenteellista ongelmaa. Jatkotutkimusta ajatellen samankaltaista tutkimusta olisi hyvä toteuttaa suuremmalla ja systemaattisemmalla mittakaavalla, erilaisilla internetin ja sosiaalisen median alustoilla.

## LÄHTEET

- Ancarno, Clyde 2020: Corpus-assisted discourse studies. – Anna De Fina & Alexandra Georgakopoulou (toim.), *The Cambridge Handbook Discourse Studies*. S. 165–185. Cambridge University Press.
- Anthony, Laurence. (2022). AntConc (Version 4.2.0) [Computer Software]. Tokyo, Japan: Waseda University. – <https://www.laurenceanthony.net/software/antconc/releases/AntConc341/help.pdf> 2.2.2023.
- Antigypsyism 2017: Antigypsyism – A Reference Paper. – [https://ccme.eu/wp-content/uploads/2018/12/2016-07-06-Antigypsyism\\_a\\_reference\\_paper.pdf](https://ccme.eu/wp-content/uploads/2018/12/2016-07-06-Antigypsyism_a_reference_paper.pdf) 10.4.2023.
- Baker, Paul 2006: *Using Corpora in discourse analysis*. London: Continuum.
- Baker, Paul, Gabrielatos, Costas, KhosraviNik, Majid, Krzyżanowski, Michał, McEnery, Tony & Wodak, Ruth 2008: A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse and society* 19 (3) s. 273–206.
- Baker, Paul, Gabrielatos, Costas & McEnery, Tony 2013: *Discourse Analysis and Media Attitudes: The Representation of Islam in the British Press*. Cambridge University Press.
- Blommaert, Jan 2005: *Discourse: A Critical Introduction*. Cambridge University Press.
- City Digital Group 2021: Suomi24 virkkeet -korpus 2001-2020, Korp-versio [tekstikorpus]. Kielipankki. – <http://urn.fi/urn:nbn:fi:lb-2021101525> 10.12.2022.
- Fairclough, Norman 1995: *Critical discourse analysis: the critical study of language*. London: Longman.
- Flowerdew, Lynne 2008: Corpora and context in professional writing. – Vijay K. Bhatia, John Flowerdew & Rodney H. Jones (toim.), *Advances in discourse studies* s. 115–126. London: Routledge.
- Friman-Korpela, Sarita 2009: Mustalaisesta romaniksi. Romanien itsenimeämisen politiikkaa. – Jere Jäppinen (toim.), *Varokaa, mustalaisia! Väärinymmärryksen historiaa. Suomen romanien vaiheita ja kulttuuria*. Helsinki: Helsingin kaupungin museo.
- 2012: Suomen romanipolitiikan kansainväliset yhteydet. – Panu Pulma (toim.), *Suomen Romanien historia* s. 211–235. Helsinki: suomalaisen kirjallisuuden seura.
- Harju, Auli 2018: Suomi24-keskustelut kohtaamisten ja törmäysten tilana. *Media ja viestintä* (41) 1 s. 51–74. – <https://doi.org/10.23983/mv.69952> 27.3.2023
- Heikinheimo, Mikko 2015: *Romanit*. Tampere: Mediapinta.
- Hoey, Michael 2005: *Lexical Priming : A new theory of words and language*. Lontoo: Routledge.

- Hunt, Sally 2015: Representations of Gender and Agency in the Harry Potter Series. – Paul Baker & Tony McEnery (toim.), *Corpora and Discourse studies: Integrating Discourse and Corpora* s. 266–284.
- Jantunen, Jarmo Harri 2018a: Homot ja heterot Suomi24:ssä analyysi digitaalisista diskursseista – *Puhe ja Kieli* 38(1) s. 3–22, <https://doi.org/10.23997/pk.65488> 23.2.2023.
- 2018b: Korpusavusteinen diskurssintutkimus (CADS): analyysiesimerkki homouden ja heterouden digitaalisista diskursseista. – Lauri Haapanen, Liisa Kääntä & Lotta Lehti (toim.), *Diskurssintutkimuksen menetelmistä*. AFinLA-E: Soveltavan Kielitieteen Tutkimuksia 11 (2018) s. 20–44. – <https://doi.org/10.30660/afinla.69259> 15.11.2022.
- Jauhola, Laura, Siltala, Jenna & Nieminen, Kati 2022: "Että puututtaisiin konkreettisesti" Seurantaselvitys vihapuheesta ja häirinnästä ja niiden vaikutuksista eri vähemmistöryhmiin. Oikeusministeriön julkaisuja 10/2022. – <http://urn.fi/URN:ISBN:978-952-324-264-7> 1.4.2023
- Jääntti, Saara & Martin, Maisa 2007: Sanat kantavat valtaa – *Tiedonjyvä* 42 (4) s. 25.
- Koivuranta, Sanna 2013: "Sinnikkäät leirytyjät" *Ulkomaalaisromanien representointi Helsingin Sanomien uutisissa*. Pro gradu -tutkielma. Jyväskylän yliopiston kielten laitos. – <http://urn.fi/URN:NBN:fi:ju-201312032720> 16.11.2022.
- Korhonen, Nita, Jauhola, Laura, Oosi, Olli, Huttunen, Hannu-Pekka, Owl Group Oy 2016: "Usein joutuu miettimään, miten pitäisi olla ja minne menemättä" *Selvitys vihapuheesta ja häirinnästä ja niiden vaikutuksista eri vähemmistöryhmiin*. Oikeusministeriön julkaisu 7/2016. – <http://urn.fi/URN:ISBN:978-952-259-496-9> 20.2.2023.
- Kyröläinen, Aki-Juhani & Laippala, Veronika 2021: Määrällinen korpuslingvistiikka. Milla Luodonpää-Manni, Markus Hamunen, Reetta Konstenius, Matti Miestamo, Urpo Nikanne & Kaius Sinnemäki (toim.), *Kielentutkimuksen menetelmiä I–V*. Helsinki: Suomalaisen kirjallisuuden seura – <https://doi.org/10.21435/skst.1457> 20.2.2023.
- Lagus, Krista, Pantzar, Mika, Ruckenstein, Minna & Ylisiurua, Marjoriikka 2016: Suomi24. Muodonantoa aineistolle. Helsinki: Unigrafia. – <http://hdl.handle.net/10138/163190> 1.3.2023
- Mahmoud, Samaie & Bahareh, Malmir 2017: US news media portrayal of Islam and Muslims: a corpus-assisted Critical Discourse Analysis – *Educational Philosophy and theory* 49 (14) s. 1351–1366. – <https://doi.org/10.1080/00131857.2017.1281789> 21.2.2023.
- Partington, Alan, Duguid, Alison & Taylor, Charlotte 2013: *Patterns and Meanings in Discourse: Theory and Practice in Corpus-assisted Discourse Studies (CADS)*. John Benjamins Publishing Company.
- Pietikäinen, Sari 2000: *Discourses in differentiation: ethnic representations in newspaper texts*. Väitöskirja. Jyväskylän yliopisto.
- 2002: Media, kieli ja etniset identiteetit. – Sirkka Laihiala-Kankainen, Sari Pietikäinen & Hannele Dufva (toim.), *Moniääninen Suomi: kieli, kulttuuri ja identiteetti*. Jyväskylän yliopisto: soveltavan kielentutkimuksen keskus.
- Pälli, Pekka & Lillqvist, Ella 2021: Diskurssianalyysi. Milla Luodonpää-Manni, Markus Hamunen, Reetta Konstenius, Matti Miestamo, Urpo Nikanne & Kaius Sinnemäki (toim.), *Kielentutkimuksen menetelmiä I–V*. Helsinki: Suomalaisen kirjallisuuden seura – <https://doi.org/10.21435/skst.1457> 20.2.2023.
- Pöyhtäri, Reetta, Haara, Paula & Raittila, Pentti 2013: *Vihapuhe sananvapautta kaventamassa*. Tampere University press. – <https://urn.fi/URN:ISBN:978-951-44-9249-5> 27.3.2023
- Reisigl, Martin & Wodak, Ruth 2001: *Discourse and discrimination: rhetorics of racism and antisemitism*. London: Routledge.
- Romano Missio. 2023. – <https://romanomissio.fi/tietoa-romaneiden-kulttuurista/> 20.2.2023.
- Romanit 2020a. – <https://romanit.fi> 20.2.2023.
- 2020b. Romanian asema Suomessa. – <https://romanit.fi/romanien-asema-suomessa/> 5.4.2023.
- Saresma, Tuija, Pöyhtäri, Reetta, Knuutila, Aleks, Kosonen, Heidi, Juutinen, Marko, Haara, Paula, Tulonen, Urho, Nikunen, Kaarina & Rauta, Jenita 2022: *Verkkoviha: Vihapuheen tuottajien ja levittäjien verkostot, toimintamuodot ja motiivit*. Valtioneuvoston kanslia. – <https://urn.fi/URN:ISBN:978-952-383-298-5> 25.3.2023.
- Scott, Mike & Tribble, Christopher 2006: *Textual Patterns: Key Words and corpus analysis in language education*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Sinclair, John 1991: *Corpus, Concordance, Collocation*. Oxford University Press.

- Solin, Anna 2012: Kriittinen diskurssintutkimus. – Vesa Heikkinen, Petri Lauerma, Mikko Lounela, Ulla Tiirilä ja Eero Voutilainen (toim.), *Genreanalyysi – tekstilajitutkimuksen käsikirja*. s. 558–563. Helsinki: Gaudeamus.
- Sosiaali- ja terveysministeriö. 2019. *Finlands national Roma policy (ROMPO) 2018–2022*. Sosiaali- ja terveysministeriön julkaisuja 2019:25 – <http://urn.fi/URN:ISBN:978-952-00-3916-5> 20.2.2023.
- Stauber, Roni & Vago, Raphael 2007: *The Roma – a Minority in Europe: Historical, Political and Social Perspectives*. Central European University Press.
- Stenroos, Marko 2014: Meitä on moneksi – romanien monet kasvot. – Panu Pulma (toim.), *Suomen Romanien historia* s. 430–435. Helsinki: suomalaisen kirjallisuuden seura.
- Suomen romaniyhdistys ry 2023: Syrjinnän vastainen työ. – <https://www.suomenromaniyhdistys.fi/tietoa-meista/syrjinnan-vastainen-tyo/#> 8.4.2023.
- Tervonen, Miika 2009: Mustalaisia vai ihmisiä? Väärinymmärryksen historiaa – ja nykypäivää. – Jere Jäppinen (toim.), *Varokaa, mustalaisia! Väärinymmärryksen historiaa. Suomen romanien vaihteita ja kulttuuria*. Helsinki: Helsingin kaupungin museo.
- Toiviainen, Minna Marjaana 2012: *Eksoottisia uhreja, toivottomia sankareita? Romanidiskurssit Helsingin sanomissa 2000–2010*. Pro gradu -tutkielma. Helsingin yliopisto, Valtiotieteellinen tiedekunta, sosiaalitieteiden laitos.
- Tuomainen, Leena 2014: *Romanikerjäläisilmiö viranomaispuheen rakentamana: diskurssianalyysi kerjäläisilmiöstä viranomaisraporteissa ja kansanedustajien kysymyksissä*. Pro gradu -tutkielma. Tampereen yliopisto, yhteiskunta- ja kulttuuritieteiden yksikkö. – <https://urn.fi/URN:NBN:fi:uta-201406061633>
- 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. – <http://scripta.kotus.fi/visk> 26.3.2023.
- Yhdenvertaisuus 2023a: *Syrjintä eri väestöryhmissä*. – <https://yhdenvertaisuus.fi/vahemmistot-suomessa> 20.3.2023.
- 2023b: *Vihapuhe*. – <https://yhdenvertaisuus.fi/vihapuhe> 28.3.2023.

# LIITTEET

## LIITE 1 M-KORPUKSEN AVAINSANAT

Avainsana	Sija	Frekvenssi	Avainsanaisuus
manne	1	573	156.319
maanviljelijä	2	158	147.174
hevonen	3	358	142.373
elättää	4	412	135.106
lieksa	5	197	103.267
loinen	6	200	91.756
paska	7	331	87.523
vittu	8	202	83.858
loisia	9	159	82.851
veronmaksaja	10	271	82.631
auto	11	820	76.482
maatalousloinen	12	53	75.598
syödä	13	265	72.344
maksaa	14	894	68.205
saatana	15	161	67.482
perkele	16	92	64.699
sosiaaliavustus	17	62	63.644
kuningas	18	149	61.439
somali	19	939	54.953
rehellinen	20	304	54.545
neekeri	21	63	53.780
ryssä	22	123	52.236
kuolla	23	279	51.851
juoppo	24	142	51.037
varastaa	25	502	50.955
mamut	26	192	47.813
vastikkeeton	27	43	47.664
helvetti	28	188	44.041
duunari	29	68	43.146
pummi	30	82	42.940
afrikka	31	175	42.113
tappaa	32	373	41.932
perse	33	106	41.777
työ	34	1923	40.353
viina	35	125	39.863
hurtta	36	37	39.692
kartonki	37	27	38.512
hullu	38	140	38.339
paskasakki	39	55	39.109
vaihtaa	40	216	35.460
vituttaa	41	64	34.699
valehdella	42	160	33.994
yhteiskunta	43	543	33.277

tehdä	44	2669	32.544
laiska	45	104	32.062
bmw	46	34	31.714
tolkku	47	33	30.459
mennä	48	1375	30.282
maalaisliitto	49	21	29.953
äpäpä	50	26	29.878
mäenpää	50	26	29.878
ruoka	52	240	28.842
siitoin	53	20	28.527
hopi	53	20	28.527
viipuri	55	39	28.370
riisa	56	61	28.117
raha	57	915	26.695
mutakuono	58	27	26.650
kärpänen	59	23	25.839
ampua	60	172	25.688
mustalaisnainen	61	67	25.959
maatalous	62	41	25.643
sossi	63	237	25.019
kaveri	64	281	24.933
matut	65	100	24.643
kaamos	66	17	24.248
puolimustalainen	67	33	24.212
tuotanto	68	25	24.094
talo	69	348	24.032

## LIITE 2 R-KORPUKSEN AVAINSANAT

Avainsana	Sija	Frekvenssi	Avainsanaisuus
kieli	1	922	394.114
kulttuuri	2	959	174.755
romaninainen	3	194	156.999
romania	4	688	148.501
suomi	5	3978	143.156
romanimies	6	187	136.648
ruotsi	7	1035	135.176
saamelainen	8	508	128.935
viittomakieli	9	94	126.630
syrjintä	10	336	124.203
valtaväestö	11	621	121.091
ruotsinkielinen	12	297	117.962
vähemmistökieli	13	87	117.199
helsinki	14	352	107.567
etninen	15	305	102.170
ryhmä	16	472	101.863
maahanmuuttaja	17	529	92.335
romanikulttuuri	18	174	91.250
vähemmistö	19	505	83.706
bulgaria	20	247	78.891
asema	21	278	74.361
eurooppa	22	346	74.011
romaniväestö	23	80	62.302
kerjäläinen	24	204	61.950
romanikieli	25	86	60.098
kansallispäivä	26	65	58.668
romanilapsi	27	61	57.824
kerjätä	28	298	53.089
virtanen	29	76	52.026
romanialainen	30	171	51.308
syrjiä	31	264	48.024
kehittää	32	101	48.011
oikeus	33	433	46.889
virallinen	34	110	45.488
neuvottelukunta	35	39	44.610
suomenkielinen	36	131	44.602
kansalliskieli	37	32	43.106
homoseksuaali	38	97	42.582
perustuslaki	39	57	42.581
romaniperhe	40	59	41.880
kohdistua	41	140	40.902
palvelu	42	139	40.168
tieto	43	401	39.827
hanke	44	44	37.440
romanikerjäläinen	44	44	37.440
asia	46	1629	37.214

tarjota	47	190	37.081
opiskella	48	142	36.414
ennakkoluulo	49	111	34.523
patmos	50	24	32.330
diakonissalaitos	51	33	31.974
lippu	52	112	31.495
anteeksi	53	147	31.317
maahanmuuttovirasto	54	23	30.982
turvapaikanhakija	55	76	30.978
koulutus	56	166	30.822
säättää	57	70	30.801
yhteisö	58	141	30.461
rikollisuus	59	117	30.360
paperiton	60	57	30.138
äidinkieli	61	59	30.018
pakkoruotsi	62	45	29.515
serbi	63	34	29.256
määräaika	64	27	29.169
mahdollisuus	65	172	28.304
romaninuori	66	21	28.288
palsta	67	293	27.815
liittyä	68	271	27.568
tausta	69	185	27.496
hyväksyä	70	292	27.388
israel	71	131	27.075
ohjelmatarjonta	72	20	26.941
romaniyhteisö	73	40	26.817
kokea	74	228	26.559
rasismi	75	525	26.287
pyytää	76	273	26.103
viittomakielinen	77	24	25.359
lindberg	77	24	25.359
kansanryhmä	79	56	25.348
selvitys	80	41	25.334
saapua	81	144	24.936
liikkuvuus	82	38	24.695
johtaja	83	142	24.637
kielilaki	84	18	24.247
meänkieli	84	18	24.247
yhdysvallat	86	35	24.137
romaniveri	86	35	24.137
kunnioittaa	88	259	24.108
yhtäläinen	89	23	24.095
ylläpitää	90	83	23.807
auttaa	91	295	23.758
kieliryhmä	92	32	23.695
romaniaasia	93	26	23.466
vammaisuus	93	26	23.466
kinnunen	93	26	23.466
romanisuku	96	29	23.432

