

**Korpusavusteinen diskurssintutkimus *Kouvola*-paikannimen
diskurssiprosodiasta Suomi24-keskustelufoorumilla**

**Kandidaatintutkielma
Veera Männistö
Suomen kieli
Kieli- ja viestintätieteiden
laitos
Jyväskylän yliopisto
2020**

JYVÄSKYLÄN YLIOPISTO

Tiedekunta – Faculty Humanistis-yhteiskuntatieteellinen tiedekunta	Laitos – Department Kieli- ja viestintätieteiden laitos
Tekijä – Author Veera Männistö	
Työn nimi – Title Korpusavusteinen diskurssintutkimus <i>Kouvola</i> -paikannimen diskurssiprosodiasta Suomi24-keskustelufoorumilla	
Oppiaine – Subject suomen kieli	Työn laji – Level kandidaatintutkielma
Aika – Month and year 04/2020	Sivumäärä – Number of pages 24
<p>Tiivistelmä – Abstract</p> <p>Tutkimuksessa tarkastellaan Kouvola-paikannimeen liittyvää diskurssiprosodiaa Suomi24-keskustelupalstalla. Tarkoituksena on selvittää, mitä ennako-oletuksia ja asenteita keskustelijat tyypillisesti liittävät sanaan. Tutkimus kuuluu korpusonomastiikan alalle, joka on suhteellisen uusi suuntaus suomalaisessa tutkimuksessa. Niinpä tutkimus täydentää tätä tutkimussuuntausta. Laajemmin tutkimus sijoittuu korpusavusteisen diskurssintutkimuksen alueelle. Metodina siinä on hyödynnetty kollokaatioanalyysia. Tutkimuskysymykset ovat: Mitä kollokaatteja <i>Kouvola</i> saa Suomi24-korpuksessa? Millaisia diskurssiprosodioita kollokaateista nousee esiin?</p> <p>Tutkimuksen aineistona on käytetty Suomi24-keskustelufoorumin kommentteihin pohjautuvaa Suomi24-korpusta vuosilta 2015–2017. Korpus on koottu hakusanalla <i>Kouvola</i> Kielipankin Korp-palvelussa. Kaksoiskappaleiden poiston jälkeen aineisto on siirretty AntConciin, jonka avulla on laskettu <i>Kouvolan</i> kollokaatit tilastollisen merkitsevyyden perusteella. Yhteensä kollokaatteja ilmeni 246, joista analysoin tarkemmin sataa merkittävintä. Analyysi on tehty luokittelemalla kollokaatit merkitysryhmiin aineistoperusteisesti ja siitä edelleen diskurssiprosodioihin.</p> <p>Tutkimuksessa tuli esille viisi merkittävää diskurssiprosodiaa: politiikka, palvelut ja tapahtumat, <i>Kouvola</i> tapahtumapaikkana, kritisointi sekä puolustaminen. <i>Kouvola</i> tapahtumapaikkana -diskurssiprosodia oli sävyltään neutraali, kun taas politiikka, palvelut ja tapahtumat sekä kritisointi sisälsivät paljon kaupungin kritisointia. Puolustamisen diskurssiprosodia puolestaan oli sävyltään positiivinen. Tutkimuksen tuloksia voi hyödyntää vastaavissa korpusonomastisissa tutkimuksissa, jolloin esimerkiksi eri paikannimien diskurssiprosodioiden vertailu olisi mahdollista.</p>	
Asiasanat – Keywords diskurssiprosodia, keskustelupalstat, kollokaatio, korpusonomastiikka, Kouvola	
Säilytyspaikka – Depository JYX-julkaisuarkisto	
Muita tietoja – Additional information	

SISÄLLYS

1	JOHDANTO.....	1
2	KORPUSONOMASTIIKKA.....	3
2.1	Korpuslingvistiikka korpusonomastisessa tutkimuksessa.....	3
2.2	Korpusavusteinen diskurssintutkimus.....	5
2.3	Diskurssiprosodia.....	6
3	AINEISTO JA MENETELMÄT.....	8
3.1	Suomi24-keskustelufoorumi korpuksessa.....	8
3.2	Korp-konkordanssityökalu aineistonkeruumenetelmänä.....	8
3.3	AntConc-konkordanssiohjelma analyysin apuna.....	9
4	ANALYYSI JA TULOKSET.....	11
4.1	Kollokaattien luokittelu ja rajaus tässä tutkimuksessa.....	11
4.2	<i>Kouvolaan</i> liittyvien kollokaattien diskurssiprosodiat.....	13
4.2.1	Politiikka.....	13
4.2.2	Palvelut ja tapahtumat.....	15
4.2.3	<i>Kouvola</i> tapahtumapaikkana.....	17
4.2.4	Kritisointi ja puolustaminen.....	18
4.2.5	Diskurssiprosodioiden ulkopuolelle jäävät kollokaatit.....	20
5	PÄÄTÄNTÖ.....	22
5.1	Yhteenveto ja johtopäätökset.....	22
5.2	Arviointi ja jatkotutkimusmahdollisuudet.....	23
	LÄHTEET.....	25

1 JOHDANTO

Kouvola on Kymenlaaksossa, Kymijoen varrella sijaitseva kaupunki, jossa asuu noin 83 000 asukasta. Kunta on perustettu vuonna 2009, jolloin vanhaan Kouvolan kaupunkiin liitettiin sen viisi lähikuntaa: Elimäki, Jaala, Kuusankoski, Anjalankoski ja Valkeala. (Kouvolan kaupunki 2020.) Pinta-alaltaan nykyinen Kouvola kattaa noin 2880 neliökilometrin kokoisen alueen (Maanmittauslaitos 2020). Maisemaltaan kunta on pitkälti maaseutua, joskin vanhat kuntakeskukset muodostavat nykyisellään taajama-alueita loitommalla keskustasta. Kouvolan keskusta-alue puolestaan koostuu vanhan Kouvolan kaupungin keskustasta, jonka arkkitehtuuri on pitkälti 1970-luvun betoniarkkitehtuuria. Keskustan ulkonäköä on verrattu etenkin sosiaalisessa mediassa usein Neuvostoliiton ja Itä-Berliinin arkkitehtuuriin (ks. esim. Salmi 2019).

Kandidaatintutkielmassani tutkin *Kouvola*-nimeen liittyviä taustaoletuksia ja assosiaatioita. Aiheeni onkin se, miten *Kouvolasta* puhutaan kansalaiskeskustelussa: minkä tyyppisten sanojen kanssa se useimmiten esiintyy ja millaisissa diskursseissa siitä puhutaan. Aineistonani hyödynsin Suomi24-keskustelufoorumien kommentteista koottua korpusta. Aineiston keruu- ja analyysimenetelmänä sovelsin korpusavusteisen diskurssitutkimuksen periaatteita. Koska tutkimukseni kohteena on paikannimeen liittyvät diskurssiprosodiat, sijoittuu sen teoreettinen viitekehys onomastiikan ja tarkemmin korpusonomastiikan alalle. Tutkimuskysymykseni ovat:

1. Mitä kollokaatteja *Kouvola* saa Suomi24-korpuksessa?
2. Millaisia diskurssiprosodioita kollokaateista nousee esiin?

Tutkimusaiheeni on siksi merkittävä, että Kouvolalla koetaan olevan poikkeuksellisen huono maine, ja kaupungin imagon pelastamiseksi onkin tehty runsaasti töitä (ks. Niemi 2018; Turunen 2018). Esimerkiksi muutama vuosi sitten kunnan mainetta pyrittiin nostattamaan #munbetonihelvetti-kampanjan avulla sosiaalisessa mediassa (Heikura 2015). Siksi onkin tähdellistä tutkia, puhutaanko kaupungista todella niin negatiiviseen sävyyn kuin yleisesti oletetaan, ja mitkä kunnan piirteet tai ominaisuudet nousevat keskusteluissa merkittäviksi. Näin tutkimus saattaa tarjota kaupungille eväitä tehdä alueesta houkuttelevampi. Vastaavia tutkimuksia kaupunkien nimiin liittyvistä diskurssiprosodioista ei juurikaan ole tehty Suomessa, vaan viime vuonna julkaistu Ainialan ja Jantusen *Hesa* ja *Stadi* -slanginimiä käsittelevä tutkimus on ensimmäinen laatuaan (Ainiala & Jantunen 2019). Oma tutkimukseni täydentää Ainialan ja Jantusen avaamaa tutkimussuuntausta.

Tämä tutkielma etenee siten, että luvussa 2 esittelen tarkemmin tutkimukseni teoreettista viitekehystä sekä aiemmin erityisesti korpusonomastiikan alalla tehtyä tutkimusta. Seuraavaksi käsittelen tutkimuksessa hyödyntämäni aineistoa ja sen analyysimenetelmiä tutkimuksen teon kannalta kronologisessa järjestyksessä. Luvussa 4 avaan syvällisemmin tekemääni analyysia sekä sen kautta saamiani tutkimustuloksia. Lopuksi pohdin työtäni suhteessa muihin korpusonomastiikan alan tutkimuksiin ja tuon esille oman tutkimukseni mahdollisia tulevaisuusnäkyviä.

2 KORPUSONOMASTIIKKA

Korpusonomastiikka on kielentutkimuksen osa-alue, joka sijoittuu laajemmin onomastiikan alaan. Onomastiikka eli nimistöntutkimus on kielitieteen ala, joka tutkii erisnimiä. Tutkimuksen fokuksessa voivat olla niin henkilön- kuin paikannimetkin, ja lisäksi onomastiikassa voidaan keskittyä vaikkapa eläinten tai automerkkien nimiin – käytännössä siis mihin tahansa monoreferentiaalisiin nimiin. Olennaista tutkimussuuntauksessa on se, mitä nimetään ja millä perustein nimiä annetaan. Siksi onomastiikka onkin tutkimusalana tiiviisti yhteydessä myös muihin tieteenaloihin, kuten maantieteeseen ja historiaan. Nimenannon lisäksi nimistöntutkijat ovat kiinnostuneita myös nimien funktioista ja siitä, miten ne toimivat osana monipuolista kielenkäyttöä. (Ainiala, Saarelma & Sjöblom 2008: 12–20.) Onomastiikan alalla onkin tutkittu esimerkiksi epävirallisia paikannimiä, tietynä historiallisena ajankohtana valittuja sukunimiä sekä etunimen vaikutusta lapsen identiteettiin (ks. esim. Holm 2012; Mikkonen 2013; Aldrin 2017). Lisäksi onomastiikan alalle kuuluvat myös erisnimiin liittyviä asenteita ja diskursseja tarkastelevat tutkimukset, joihin oma työni lukeutuu.

Korpusonomastiikka puolestaan viittaa nimistöntutkimukseen, jonka aineisto saadaan korpuksista ja sitä analysoidaan korpuslingvistiikalle tyypillisin menetelmin, kuten konkordanssi- tai avainsana-analyysin avulla. Korpusonomastiikassa tutkimuksen kohteena on nimien yleisyys ja käyttö sekä paikallinen tai alueellinen variaatio. Terminä korpusonomastiikka on uusi: sitä ovat käyttäneet ainoastaan Ainiala ja Jantunen uudessa, *Hesa* ja *Stadi* -slanginimiä ja niiden käyttöön liittyviä diskurssiprosodioita käsittelevässä tutkimuksessaan. (Ainiala & Jantunen 2019: 58, 60.) Myös korpusonomastista tutkimusta on tehty suhteellisen vähän – Suomessa Ainialan ja Jantusen tutkimus on alallaan ensimmäinen. Ulkomaisia esimerkkejä korpusonomastisista tutkimuksista sen sijaan löytyy: esimerkiksi Liettuassa on tutkittu *America*-nimen käyttöä (Vaičenonienė 2001) ja Norjassa *Maastricht* ja *Schengen* -nimien esiintymistä (Halverson & Engene 2010). Nämä aiemmat esimerkit kuitenkin perustuvat sanomalehtikirjoituksista tehtyihin korpuksiin, kun taas oma tutkimukseni sekä Ainialan ja Jantusen tutkimus tarkastelevat paikannimien diskurssiprosodioita internet-keskusteluissa. Jälkimmäisissä fokuksessa on siis kansalaiskeskustelu, kun taas edellä mainituissa tarkastelun kohteena ovat virallisemmat mediatekstit.

2.1 Korpuslingvistiikka korpusonomastisessa tutkimuksessa

Korpuslingvistiikka on kielentutkimuksen osa-alue, jossa tarkastellaan tutkimusaineistossa esiintyviä säännönmukaisuuksia tilastollisten todennäköisyyksien avulla. Tutkimuksen aineistona käytetään niin kutsuttuja korpuksia. (Partington, Duguid & Taylor 2013: 8–9.) Korpus tarkoittaa laajaa tekstien

tai tekstin osien kokoelmaa, joka koostuu datasta. Data puolestaan merkitsee käytännössä kaikkea käytettävissä olevaa informaatiota. Korpus voi sisältää sekä luonnollista että keinotekoisia kieltä. Luonnollinen kieli on reaali maailmassa tuotettua, puhuttua tai kirjoitettua kieltä, jota ei alun perin ole tarkoitettu tutkimusaineistoksi. Keinotekoinen kieli taas voi olla esimerkiksi tiettyä tutkimusta varten kerättyä kieltä, eikä se siksi ole luonnollista. (Lounela & Heikkinen 2012: 120–121.) Useimmiten korpuksia muodostetaan etukäteen luonnollisesta, dokumentoidusta kielestä, esimerkiksi julkisista puheista, historiallisista teksteistä tai nykyaikana internet-keskusteluista (Jantunen 2012: 360; Partington, Duguid & Taylor 2013: 6). Korpuksia ovat kooltaan suuria: usein ne koostuvat miljoonista sanoista (Lounela & Heikkinen 2012: 121). Tästä syystä niitä käsitellään nykyaikana pääsääntöisesti mekaanisesti erilaisten tietokoneohjelmien avulla, mikä osaltaan mahdollistaa suurten aineistojen analyysin (Partington, Duguid & Taylor 2013: 5). Yksi tunnettu korpustyökalu on myös tässä tutkimuksessa hyödyntämäni Korp, jonka toimintaa esittelen tarkemmin luvussa 3.2 (Aller Media Oy 2014).

Tutkimusperinteeltään korpuslingvistiikka on kvantitatiivista, ja siinä tutkimuksen kohteena voi olla esimerkiksi tietyn sanan esiintymisfrekvenssi tietyssä aineistossa tai muut tilastolliseen merkitsevyyteen perustuvat löydökset (Jantunen 2012: 361). Perinteisesti korpuslingvistiikassa tutkimuksen pääpaino on kuitenkin ollut kieliopin ja sen käytön tutkimuksessa (Partington, Duguid & Taylor 2013: 5). Korpustutkimus voi keskittyä myös esimerkiksi tekstin variaatioon tai morfologisiin tai muuten rakenteellisiin ominaisuuksiin – esimerkiksi Padula, Panza ja Muñoz ovat tutkineet englanninkielisen pronominin *this* käyttöä 130 teknillisen alan artikkelin korpuksessa (Padula, Panza & Muñoz 2020). Koska korpuslingvistiikan pääpaino on enemmän aineistosta esiin nousevien säännönmukaisuuksien ilmentämisessä kuin pelkkien numeroiden raportoinnissa, yhdistyy siihen usein myös kvalitatiivisen tutkimuksen – esimerkiksi diskurssintutkimuksen – piirteitä (Biber, Conrad & Reppen 2007: 4–5). Yksi esimerkki kvantitatiivisen korpuslingvistiikan ja kvalitatiivisen diskurssianalyysin yhdistämisestä onkin tässä tutkimuksessa hyödyntämäni korpusavusteinen diskurssianalyysi, jota esittelen seuraavaksi.

Korpusonomastisessa tutkimuksessa korpuslingvistiikka on merkittävässä osassa, sillä sen tehtävä on tarjota tutkimukselle niin tutkimuksen aineisto kuin sen analyysimenetelmäkkin. Korpusonomastiikalle onkin tyypillistä, että siinä nimistöntutkimuksen ja korpustutkimuksen tavoitteet ja lähtökohdat yhdistyvät. (Ainiala & Jantunen 2019: 58.)

2.2 Korpusavusteinen diskurssintutkimus

Korpusavusteinen diskurssintutkimus on tutkimussuuntaus, joka yhdistelee korpustutkimuksen ja diskurssianalyysin periaatteita. Omalle työlleni korpusavusteinen diskurssintutkimus tarjoaa aineiston keruu- ja analyysimenetelmät.

Diskurssilla viitataan näkyviin tai läpinäkymättömiin ennakko-oletuksiin ja arvoihin, jotka voidaan paljastaa tekstistä tarkastelemalla sitä käyttökontekstissaan. Diskurssianalyysi onkin tutkimusta, jossa tarkastelun kohteena ovat tekstin sisältämät diskurssit. Diskurssintutkimukselle on tyypillistä, että siinä kieltä tutkitaan osana sosiaalista toimintaa enemmän kuin pelkinä vakiintuneina merkityksinä tai rakenteina: esimerkiksi sanojen merkitykset ovat riippuvaisia niiden käytöstä ja voivat siksi muuttua käyttöyhteyksien muuttuessa. Lisäksi kielen ajatellaan olevan resurssi, jota kielenkäyttäjä voi hyödyntää kussakin tilanteessa eri tavoin. (Pietikäinen & Mäntynen 2009: 11–12, 15.) Diskurssintutkija kiinnittää siis huomionsa tekstin sanavalintojen ja rakenteiden lisäksi ennen kaikkea niiden taustalla toimiviin psykologisiin tekijöihin, kuten odotuksiin, uskomuksiin ja lukijan oletettuun taustatietoon (Yule 1996: 84). Diskurssintutkimus nojaa vahvasti kvalitatiiviseen tutkimusperinteeseen, ja siinä olennaista osaa näyttelevät tutkijan omat arviot ja päätelmät. Tyypillinen diskurssianalyttinen tutkimus onkin tekstitutkimusta, jossa pyritään selvittämään esimerkiksi tiettyyn sanaan liittyviä sosiaalisia merkityksiä valitussa kontekstissa. (Pietikäinen & Mäntynen 2009: 139, 146.)

Korpusavusteiselle diskurssintutkimukselle on ominaista, että siinä yhdistyvät kvalitatiivinen ja kvantitatiivinen tutkimus. Tutkimussuuntauksen tavoite onkin tuottaa tilastolliseen näyttöön perustuvaa informaatiota ja syventää siitä saatavaa tietoa diskurssianalyysin keinoin. Pelkistetysti voidaan siis ajatella, että korpusavusteisessa diskurssianalyysissä korpuslingvistiikka tarjoaa tutkimuksen aineiston, kun taas analyysin periaatteet tulevat diskurssintutkimuksen alalta. (Partington, Duguid & Taylor 2013: 10–11.) Tutkimuksessa korpusavusteinen diskurssianalyysi toimii samanaikaisesti sekä teoria- että metodilähteenä (McEnery, Xiao & Tono 2006: 7). Tutkimussuuntaus ei kuitenkaan ole sidottu mihinkään tiettyyn diskurssianalyysin alalajiin, vaan tutkija voi hyödyntää diskurssianalyysia haluamallaan tavalla. Diskurssintutkimuksen tavoin myös korpusavusteisessa diskurssianalyysissä tutkimuksen päämäärä on nostaa esille muuten läpinäkymättömiin jääviä diskursseja, arvoja ja oletuksia. (Partington, Duguid & Taylor 2013: 10–11.) Käytännössä korpusavusteisessa diskurssintutkimuksessa voidaan tarkastella esimerkiksi sanalistoja, genreavainsanoja tai konkordansseja (Jantunen 2018a: 27).

Koska diskurssianalyysin ja korpuslingvistiikan toimintatavat poikkeavat toisistaan vahvasti, on niiden yhdistäminen uudeksi tutkimussuuntaukseksi suhteellisen tuore ilmiö (Partington 2004:

11). Korpusavusteisessa diskurssintutkimuksessa diskurssianalyysin rooli on kuitenkin merkittävä, sillä siinä tutkimuksen kohteena ovat juuri diskurssit, vaikka aineisto ja sen analyysimenetelmä tulevatkin korpuslingvistiikan alalta. Niinpä korpusavusteisessa diskurssintutkimuksessa kieli nähdään samalla tavalla kompleksisena ja tilannesidonnaisena resurssina kuin diskurssi-analyysissäkin. (McEnery, Xiao & Tono 2006: 111.) Sekä diskurssianalyysia että korpuslingvistiikkaa – ja siten myös korpusavusteista diskurssintutkimusta – yhdistää lisäksi se, että molemmissa tutkimuksen kohteena on aito kielenkäyttö (Biber, Conrad & Reppen 2007: 1). Lisäksi korpusavusteiseen diskurssintutkimukseen yhdistyy fraseologinen kielikäsite, jonka mukaan kieli on toistuvaa ja säännönmukaista, ja sen käyttöä ohjaavat sosiaaliset tarpeet. Korpusavusteiselle diskurssintutkimukselle ominainen fraseologinen kielikäsite näkyy tutkimuksessa erityisesti siinä, että analyysissä keskitytään tyypillisesti leksikaalisiin myötäesiintymiin, semanttisiin preferensseihin tai diskurssiprosodioihin. (Jantunen 2018a: 27.) Tässä tutkimuksessa olen syventynyt *Kouvola*-sanaa koskeviin diskurssiprosodioihin.

2.3 Diskurssiprosodia

Tietyle sanalle tai muulle kielelliselle ilmaukselle – vaikkapa paikannimelle – voi olla ominaista esiintyä sille tyypillisessä diskursiivisessa kontekstissa kielenkäyttötilanteesta toiseen. Ilmauksen ominaisuus liittyä toistuvasti tietynlaisiin diskursseihin puheessa ilmentää siihen liitettäviä asenteita ja assosiaatioita sekä kertoo kielenkäytön funktioista. Ilmiötä kutsutaan diskurssiprosodiaksi. (Stubbs 2001: 65–66.) Käytännössä diskurssiprosodia näkyy siis esimerkiksi siinä, miten englanninkielinen verbi *happen* 'tapahtua' assosioituu pääsääntöisesti negatiivisiin diskursseihin, kuten onnettomuuksiin (Lehto 2018: 81).

Diskurssiprosodian avulla voidaan siis tutkia, millaisiin diskursseihin kielenkäyttäjät tyypillisesti sitovat tietyn sanan tai ilmauksen. Useimmiten tutkimusta tehdään tarkastelemalla sanan käyttöyhteyksiä suurissa aineistoissa, kuten korpuksissa. Suomessa diskurssiprosodiaa onkin tutkinut muun muassa Jantunen sanoja *homo* ja *hetero* tarkastelevassa tutkimuksessaan (Jantunen 2018b). Myös Johansson, Jantunen, Heimo, Ahonen ja Laippala ovat tutkineet *kansa*-sanon diskurssiprosodiaa (2018).

Kuten diskurssintutkimuksen, myös diskurssiprosodian tutkiminen on vahvasti kvalitatiivista: diskurssien esille nostamiseksi on tarkasteltava niiden esiintymiskonteksteja, ja varsinainen tulkinta on puhtaasti analysoijan vastuulla. Diskurssiprosodiatutkimus eroaa kuitenkin perinteisestä diskurssianalyysistä juuri siten, että diskurssiprosodiaa tarkasteleva tutkimus perustuu pienemmän

tekstimäärän sijaan laajan tekstiaineiston analysointiin. (Jantunen 2004: 26.) Siksi diskurssiprosodiatutkimus soveltuukin erinomaisesti korpusavusteisen diskurssintutkimuksen alalle.

3 AINEISTO JA MENETELMÄT

Tässä luvussa esittelen tutkimuksessani hyödyntämiä aineistoja ja menetelmiä sekä avaan tarkemmin tutkimukseni kulkua näiden seikkojen valossa. Aluksi kerron Suomi24-korpuksesta, jota hyödynsin tutkimukseni aineistona. Sitten käsittelen aineiston tarkemmassa keruussa käyttämäni Korp-konkordanssityökalua. Lopuksi esittelen AntConc-konkordanssiohjelmaa, jonka avulla olen seulonut *Kouvolan* kollokaatit muusta aineistosta.

3.1 Suomi24-keskustelufoorumi korpuksesta

Suomi24-keskustelufoorumi on Suomen suosituin verkkoyhteisö, jolle kuka tahansa voi kirjoittaa keskustelunaloituksia ja kommentteja anonyymisti rekisteröitymättä. Keskustelunaiheet on jaoteltu sivustolla 21 aihealueen alle, jotka jakautuvat hierarkkisesti vielä tarkemmin omille alatasoilleen. Foorumilla voikin käydä keskusteluja lähes mistä tahansa – vaikkapa politiikasta, kasvatuksesta tai käärmeiden ääntelystä. (Aller Oy 2020.) Sivusto on Aller Oy:n omistama ja se on ollut käytössä vuodesta 2001 alkaen. Tutkimusaineistona Suomi24-foorumia on hyödynnetty vuodesta 2015, jolloin sen sisältämistä keskusteluista koostettiin laaja korpus. Yhteensä se sisältää yli 70 miljoonaa kommenttia ja 4,3 miljardia sanaa lähes 20 vuoden ajalta. Korpus on tallennettuna Kielipankkiin. (Lagus, Pantzar, Ruckenstein & Ylisiurua 2016: 5, 11; Borin, Forsberg & Roxendal 2012.)

Tutkimuskäytössä Suomi24-korpuksen etuna on se, että aineisto sisältää aitoa, luonnollista kielenkäyttöä, jota tuottavat eri taustoista tulevat suomalaiset ympäri maan. Koska sivusto on laajalti käytössä, voidaan sen avulla saada selville suhteellisen yleistettävää tietoa. Lisäksi Suomi24-korpus on helposti saatavilla Kielipankin verkkosivuilta ja sitä voi suoraan hyödyntää tutkimuksessa, sillä aineistosta on korpuksen kokoamisvaiheessa karsittu käyttäjien IP-osoite ja muut yksilöinnin mahdollistavat tekijät (Lagus, Pantzar, Ruckenstein & Ylisiurua 2016: 9). Niinpä informanttien yksityisyys on hyvin suojattu, mikä puolestaan edistää tutkimuksen eettisyyttä. Tälle tutkimukselle Suomi24-korpus on tarjonnut aineiston. Tutkimuksessa olen hyödyntänyt vuoden 2017 versiota Suomi24-korpuksesta.

3.2 Korp-konkordanssityökalu aineistonkeruumenetelmänä

Tutkimukseni aineisto on kerätty Kielipankin ylläpitämällä Korp-konkordanssityökalulla. Kyseessä on selainpohjainen työkalu, jonka avulla on mahdollista tehdä hakuja aineistoista morfosyntaktisen jäsennyksen perusteella (Kielipankki 2020). Työkaluun on tallennettu laajoja korpuksia niin

internetin keskustelupalstoilta – kuten tässä tutkimuksessa hyödyntämäni Suomi24 – kuin kirjallisuudesta ja sanomalehdistäkin. Aineisto on annotoitu. Työkalulla voi tehdä hakuja sanalla, sanan osalla tai useamman sanan jonolla, ja hakutulokset näytetään konkordansseina eli lyhyinä tekstikonteksteina. (Lagus, Pantzar, Ruckenstein & Ylisiurua 2016: 15–16.) Käytännössä Korpin avulla voidaan siis rajata pienempiä, tiettyyn aiheeseen liittyviä korpuksia kaikista Kielipankin sisältämistä korpuksista.

Tämän tutkimuksen aineisto on koottu Korp-konkordanssityökalulla hakusanalla *Kouvola*, joka on asetettu sanan perusmuodoksi. Haussa isot ja pienet alkukirjaimet on samaistettu, jolloin aineistoon päätyvät kaikki Kouvola-sanan sisältämät kommentit nimen kirjoitusasusta ja -muodosta riippumatta. Haun aikaväliksi on asetettu 1.1.2015–31.12.2017, jolloin tutkimukseen päätyy aineistoa yhteensä kolmen vuoden ajalta. Aikaväli on rajattu näin, sillä tutkimuksen tarkoituksena on saada mahdollisimman tuoretta tietoa ja toisaalta aineiston tulee olla riittävän laaja tutkimuksen reliabiliteetin turvaamiseksi. Yhteensä konkordansseja tältä aikaväliltä tuli noin 19 000. Korpissa tehdyn haun jälkeen aineisto on siirretty kokonaisuudessaan Exceliin, missä siitä on automaattisesti poistettu kaksoiskappaleet. Viimeiseksi olen poistanut jäljelle jääneet kaksoiskappaleet manuaalisesti Wordissä. Kaksoiskappaleiden poiston jälkeen analysoitavaksi jäi yhteensä noin 11 200 konkordanssia ja 135 000 sanetta. Yhden konkordanssin keskimääräinen pituus oli 12 sanaa.

Kollokaattien löytymisen mahdollistamiseksi aineiston jokainen sana on lemmatisoitu eli muutettu perusmuotoon Korpissa. Asetus on välttämätön, sillä muuten kollokaattien seulomiseen käyttämäni konkordanssiohjelma olisi tulkinut kunkin sanan taivutetut muodot eri tuloksiksi. Korpin tekemässä lemmatisoinnissa on kuitenkin se ongelma, ettei ohjelma tunnista kaikkia puhekielisiä ilmauksia. Niinpä esimerkiksi sähköpostin lähettämiseen viittaava verbi *mailata* kääntyi lemmatisoinnissa substantiiviksi *maila* ja substantiivi *persujen* jäi kollokaatiolistalle genetiivimuodossa. Näitä ongelmia pyrin korjaamaan käsin ennen siistityn aineiston siirtämistä AntConc-konkordanssiohjelmaan.

3.3 AntConc-konkordanssiohjelma analyysin apuna

AntConc-konkordanssiohjelma on ilmaiseksi ladattava tietokonesovellus, jonka avulla voi laskea tekstin osien tilastollisia todennäköisyyksiä ja frekvenssejä. Ohjelmaa voi käyttää monipuolisesti korpustutkimuksessa, sillä sen avulla on mahdollista tehdä niin kollokaatio-, sanalista- kuin avainsana-analyysiäkin. Tässä tutkimuksessa konkordanssiohjelmaa on hyödynnetty kollokaatioanalyysissä. Käytössä on ollut AntConcin versio 3.5.8. (Anthony 2015.)

Kollokaatioanalyysi on yksi korpusavusteisen diskurssintutkimuksen alaisista tutkimusmetodeista, jossa tutkimuksen kohteena on kahden tai useamman sanan toistuva myötäesiintyminen (Sinclair 1991). Siinä tutkitaan niin kutsuttuja kollokaatteja, eli tietyssä korpuksessa, tietyn noodin eli tutkittavan sanan ympärillä tyypillisesti toistuvia sanoja. Kollokaation tyypillisiä piirteitä ovatkin toistuvuus, tilastollinen merkitsevyys ja ulottuvuus. Ulottuvuudella viitataan noodin ja tutkittavien kollokaattien välimatkaan, lukumäärään, kieliopilliseen suhteeseen ja tarkastelualueen symmetrisyyteen. (Jantunen 2004: 16–18.) Kollokaatin asema suhteessa noodiin vaikuttaa siis siihen, miten merkitsevä se tilastollisesti on. Omassa tutkimuksessani valitsin sanojen tarkasteluväliksi neljä sanaa oikealle ja neljä vasemmalle noodista (4L–4R), sillä tällöin sanamäärä kattaa noin yhden tyypillisen suomenkielisen virkkeen (Jantunen 2018a: 29). Kollokaatiksi laskettavan sanan esiintymisfrekvenssiksi puolestaan olen valinnut 50, mikä varmistaa sen, että kollokaateiksi päätyy vain useimmiten aineistossa esiintyviä sanoja. Tarkasteluväli onkin melko yleinen korpustutkimuksessa (McEnery, Xiao & Tono 2006: 56).

Sanojen myötäesiintymisen todennäköisyys lasketaan tietokoneohjelmalla useimmiten joko MI-testin tai t- tai z-pisteiden avulla (McEnery, Xiao & Tono 2006: 56). Tässä tutkimuksessa todennäköisyys on laskettu MI-testiä hyödyntäen. Kyseessä on informaatiotieteestä lainattu tilastollinen formula, joka laskee logaritmeja ympäröivien sanojen oletetuista frekvensseistä. Sen avulla voidaan mitata kollokaation voimakkuutta. Mitä korkeamman MI-pisteytyksen kollokaatti saa, sitä voimakkaampi yhteys noodin ja kollokaation välillä vallitsee. Jos taas MI-pisteytys on lähellä nollaa, esiintyvät sanat todennäköisesti vain sattumalta yhdessä. (McEnery, Xiao & Tono 2006: 56; Baker 2010: 24–25.) Hunstonin (2002) mukaan sanan voidaan tulkita olevan kollokaatti, jos sen saama MI-pisteytys on kolme tai enemmän. Tässä tutkimuksessa kaikkien sadan merkitsevimmän kollokaatin MI-pisteytys oli välillä 3.19–3.93, jolloin ne voidaan tulkita kollokaateiksi.

4 ANALYYSI JA TULOKSET

Tässä luvussa luokittelen aineistoni sata merkitsevintä kollokaattia merkitysryhmiin ja jaottelen kollokaatit sitten suurempiin ryhmiin niitä koskevien diskurssiprosodioiden mukaan. Käytännössä olen siis tutkinut kommentteja, joissa kollokaatit esiintyvät, ja pyrkinyt etsimään niissä toistuvia diskursseja. AntConcin tekemän tilastollisen analyysin perusteella korpuksesta seuloitui yhteensä 246 kollokaattia hakusanalle *Kouvola*. Näistä perehdyin tarkemmin sataan merkitsevimpään kollokaattiin.

Tämä luku etenee siten, että aluksi esittelen analysoimiani kollokaatteja yleisellä tasolla sekä avaan analyysini kulkua ja sen aikana havaitsemiani ongelmakohtia. Seuraavaksi käsitelen *Kouvolan* liittyviä diskurssiprosodiota yhteisesti, minkä jälkeen avaan kutakin havaitsemaani diskurssiprosodiaa tarkemmin yksitellen. Lopuksi tuon tarkasteltavaksi diskurssiprosodioiden ulkopuolelle jääneet kollokaatit ja pohdin syitä sille, mikseivät nämä kollokaatit sopineet mihinkään diskurssiprosodiaan.

4.1 Kollokaattien luokittelu ja rajausta tässä tutkimuksessa

Ensisilmäyksellä *Kouvolan* kollokaatit näyttivät listalta pääosin neutraaleja sanoja, vaikka joukkoon mahtuikin myös sellaisia sävyiltään värittyneitä sanoja kuin *paska*, *homo* ja *persujen*. Listalta löytyi niin verbejä, substantiiveja kuin adjektiivejakin. Lisäksi kollokaateiksi oli päätynyt muutama pronomini ja adverbi. Suurin osa tutkimuksen kollokaateista oli kuitenkin substantiiveja. Taulukkoon 1 on koottu kymmenen merkitsevintä kollokaattia.

Taulukko 1. *Kouvolan* kymmenen merkitsevintä kollokaattia. MI-luku kertoo kollokaatin merkitsevyydestä. Frekvenssi sen sijaan osoittaa, montako kertaa kollokaatti esiintyy aineistossa.

MI	frekvenssi	kollokaatti
3.92622	50	homo
3.89844	57	lastensuojelu
3.88176	76	paras
3.82630	952	sanoma
3.78400	60	seurata
3.69680	151	taas
3.64223	101	tänään
3.64184	81	hovioikeus
3.63866	217	seutu
3.63478	53	veturi

Tutkimuksessa sata merkitsevintä kollokaattia on jaoteltu merkitysryhmiin aineistoperustaisesti. Aineistosta on siis pyritty löytämään tiettyä kollokaattia koskeva yleinen merkitys. Vaikka yksi kollokaatti voikin saada kommentteissa useita erilaisia merkityksiä käyttötilanteesta riippuen, enemmistöön kollokaateista näytti kuitenkin liittyvän aina yksi tai joskus kaksi selvästi erottuvaa yleismerkitystä. Kollokaatit on siis jaoteltu ryhmiin näiden merkitysten mukaisesti. Merkitysryhmien kautta pyrin tekemään päätelmiä *Kouvola*-sanon diskurssiprosodiasta. Diskurssiprosodiakohtaiset merkitysryhmät on esitelty tarkemmin kunkin diskurssiprosodian kohdalla (ks. taulukot 3–7). Merkitysryhmien ja siten myös diskurssiprosodiatutkimuksen ulkopuolelle jäivät sellaiset kollokaatit, joiden tavallista merkitystä ei ollut mahdollista päätellä aineistosta. Nämä kollokaatit on ryhmitelty luokkaan *muut*.

Kollokaattien jakaminen merkitysryhmiin tai niihin liitettäviin diskurssiprosodioihin ei ole yksinkertaista, sillä kutakin sanaa voidaan käyttää eri tavoin erilaisissa konteksteissa, ja paikoin samalla kirjoitusasulla voidaan viitata jopa täysin eri merkityksiin. Homonymia ja polysemia ovatkin suhteellisen yleisiä ilmiöitä suomen kielessä. (Laalo 1989.) Tässäkin tutkimuksessa esiintyi useampia sanoja, joiden ensisilmäyksellä voisi olettaa viittaavan eri merkitykseen, kuin mihin Suomi24-foorumin keskustelijat kommentteissaan viittasivat: esimerkiksi kollokaatilla *veturi* tarkoitettiin rautateillä liikkuvan kulkuneuvon sijaan Kouvolassa sijaitsevaa kauppakeskus Veturia. Hajontaa kollokaattien käytön tulkinnassa aiheutti myös ei-kieltosanan käyttö aineistossa, sillä kieltosana voi muuttaa kollokaatin merkityksen täysin päinvastaiseksi – *on huono* ja *ei ole huono* sisältävät molemmat saman *huono*-kollokaatin, vaikka kommenttien sävyt ja sanomat poikkeavat toisistaan huomattavasti. Koska pelkästä kollokaattilistasta ei ole nähtävissä, millaisten lausekkeiden osina sanat kontekstissaan ovat, ei kollokaateista yksinään voida tehdä luotettavia päätelmiä niiden yhteydestä noodiin. (Baker 2010: 21.) Siksi olikin tärkeää tarkastella aineiston konkordanssirivejä ja tutkia niistä, missä yhteydessä kollokaatti esitetään suhteessa tutkittavaan sanaan.

Haasteita kollokaattien analysointiin aiheuttivat tekniset vaikeudet: kollokaattihaussa AntConc näytti harvakseltaan jättävän välilyönnit huomiotta, minkä takia esimerkiksi *suuri*-kollokaatin konkordanssiriveillä oli sana *sensuuri* ja *nainen*-kollokaatin riveillä *punainen*. Ilmiö on kuitenkin onneksi harvinainen, eivätkä *sensuuri* tai *punainen* toistuneet montaa kertaa konkordansseissa. Sen sijaan AntConcin laskemaa kollokaattia *no* tarkastellessani havaitsin, että kirjainyhdistelmä esiintyy aineistossa useammin muiden sanojen osana kuin itsenäisesti partikkelina: kollokaatin konkordanssiriveillä toistui sellaiset sanat kuin *huono* ja *hieno*. Tutkimuksessani tällaiset epämääräisyydet jätettiin pois analyysistä, jolloin kollokaattien lopulliseksi määräksi jäi 245.

4.2 Kouvolaan liittyvien kollokaattien diskurssiprosodiat

Aineistoperäisessä ryhmittelyssä kollokaatit jakaantuivat yhteensä viiteen diskurssiprosodiaan, joista selkeimpiä olivat *politiikka* sekä *palvelut ja tapahtumat*. Muut prosodiat olivat *Kouvola tapahtumapaikkana*, *kaupungin yleinen kritisointi* ja *puolustaminen*. Lisäksi diskurssiprosodioiden ulkopuolelle jäi muutamia kollokaatteja, jotka on sijoitettu luokkaan *muut*. Taulukosta 2 ilmenee sadan merkitsevimmän kollokaatin sijoittuminen merkitysryhmiin prosentuaalisesti. Prosenttituloksiin vaikuttaa se, että muutamia kollokaatteja on sijoitettu yhden sijaan kahteen eri diskurssiin.

Taulukko 2. Kollokaattien sijoittuminen diskurssiprosodioittain. Prosenttiluku ilmaisee tiettyyn diskurssiprosodiaan kuuluvien kollokaattien osuuden suhteessa kaikkiin kollokaatteihin.

Diskurssiprosodia	Kollokaattien prosentuaalinen osuus
<i>politiikka</i>	25 %
<i>palvelut ja tapahtumat</i>	17 %
<i>Kouvola tapahtumapaikkana</i>	37 %
<i>kaupungin yleinen kritisointi</i>	8 %
<i>puolustaminen</i>	5 %
<i>muut</i>	27 %

4.2.1 Poliitiikka

Ehkäpä selkein Kouvolaan liittyvä diskurssiprosodia oli politiikka, johon kuului poliittisia ryhmiä ja elimiä, paikkoja, politiikkaan liittyviä verbejä, henkilöitä sekä abstrakteja substantiiveja kuvaavia kollokaatteja. Vaikka tämä diskurssiprosodia ei ollutkaan prosentuaalisesti suurin aineistosta esiin noussut prosodia, oli se helpommin havaittavissa aiheen tarkkarajaisuuden vuoksi suhteessa esimerkiksi *Kouvola tapahtumapaikkana* -diskurssiprosodiaan, johon palaan tarkemmin luvussa 4.2.3. Poliitiikkaan kytkeytyvät kollokaatit on jaoteltu merkitysryhmittäin taulukkoon 3.

Taulukko 3. Diskurssiprosodia *politiikka*.

<i>poliittinen ryhmä tai elin</i>
KOKOOMUS, KESKUSTA, LASTENSUOJELU, HOVIOIKEUS, PERSUJEN, KAUPUNKI, KAUPUNGINHALLITUS, VALTUUSTO
<i>paikannimi</i>
VETURI, KESKUSTA, KOULU, VASTAANOTTOKESKUS

politiikkaan liittyvä verbi

JÄÄDÄ, PERUSTAA, PÄÄTTÄÄ, MAKSAA

henkilö

PÄÄTTÄJÄ, ASUKAS, KORPIVAARA, PEKKA, PERUSSUOMALAINEN, VERONMAKSAJA

abstrakti substantiivi

IHME, HALLINTO, SUUNTA, ONGELMA

Kuten taulukosta 2 tulee ilmi, tutkimuksen kollokaateista neljäsosa (25%) kytkeytyi politiikka-diskurssiprosodiaan. Yhteensä politiikkaan liittyviä kollokaatteja oli 25, joista KESKUSTA on ryhmitelty sekä poliittiseksi ryhmäksi että paikannimeksi.

Kouvola ja politiikka nivoutuivat Suomi24-verkkokeskusteluissa yhteen siten, että keskusteluissa kommentoitiin Kouvolan kaupungin harjoittamaa kunnallispolitiikkaa sekä sen vaikutuksia kaupunkilaisten arkeen. Kommenteissa puhuttiin esimerkiksi koulujen lakkauttamisesta (esimerkki 1), vastaanottokeskusten perustamisesta (esimerkki 2) sekä kaupunginvaltuuston sisäisistä tapahtumista (esimerkki 3). Keskustelu sisälsi paljon myös toteamuksia kaupungin päätäntäelinten tekemistä päätöksistä (esimerkit 4 ja 5).

- 1) Kouvolasta pilattiin terveystalvelut ja suljettiin **kouluja** säästösyistä.
- 2) Kouvolaan **perustettavan vastaanottokeskuksen** ensimmäinen vaihe saadaan toimintaan todennäköisesti jo ensi viikolla.
- 3) Ei Kouvolan **valtuustossa** ole ollut mitään seksuaalista häirintää, ei siitä ainakaan mitään ole missään mainittu
- 4) Kouvolan **kaupunginhallitus** hyväksyi hetki sitten pelisäännöt luottamushenkilöiden käyttäytymiselle
- 5) asiasta **päätti** Kouvolan lasten ja nuorten lautakunta keskiviikon kokouksessaan

Vaikka monen politiikkaa käsittelevän kommentin sävy näytti olevan toteavan neutraali, ilmeni myös kaupungin päätäntäelimiä tai niiden tekemiä päätöksiä kritisovia kommentteja runsaasti (esimerkit 6, 7, 8; ks. myös 1). Sen sijaan päätöksiä tai päätäntäelimiä koskevia positiivisia kommentteja aineistosta ei juurikaan tullut ilmi.

- 6) näin Kouvola-tyyliin, näin Kouvolassa valitaan tärkeysjärjestyksen **valtuuston** käsiteltävät asiat
- 7) suosittelen mamujen tekevän joukkoraiskauksen Kouvolan **lastensuojelun** elukoille
- 8) Kouvolan **veronmaksajana** ihmettelen teidän päätöksenteon laatua

Kaupungin poliittisia ryhmiä ja elimiä kritisoidessaan monet kommentoijat kritisoivat samalla myös Kouvolan kaupunkia niin hallinnollisena alueena kuin asuinpaikkanakin. Tämä tulee esille esimerkissä 6, jossa kritisoinnin kohteena oleva valtuuston toimintajärjestys nimetään yleisemmin ”Kouvola-tyyliksi”. Kunnallispolitiikka ja sen esitetty kehnous toimivat joissakin kommenteissa

myös ikään kuin esimerkkeinä siitä, mikä kaikki Kouvolan kaupungissa on vialla (esimerkit 9, 10, 11). Näissä kommentteissa poliittisten elinten toimimattomuuden syyksi esitettiin se, että ne ovat osa Kouvola ja sen hallintoa. Niinpä vika vieritettiin pois yksittäisten organisaatioiden kontolta ja asetettiin yleisesti koko kaupungin syyksi. Ilmiö tulee hyvin ilmi erityisesti esimerkistä 11, jossa kritisoidun toiminnan tekijää ei nimetä laisinkaan, vaan subjektiksi asetetaan koko kaupunki.

- 9) ei **ihme** että Kouvolna **lastensuojelu** on perseellään kun katsoo kuka siellä hyvinvointi/lastensuojelun palvelupäällikkönä toimii
- 10) esimerkiksi Kouvola **kaupunginhallituksen ongelma** on se, että sieltä löytyy jatkuvasti sellaista suunnittelua ja keskustelua, joka ei julkisuutta siedä
- 11) 2013 pelkkiin pakolaisten tulkkiin upposi 738600 € ja Kouvolan **kaupunki** laski sen omiksi tuloiksi todistettavasti tilinpäätöksessään

Aineistossa kunnallispolitiikka sidottiin siis tiukasti osaksi Kouvola, ja sille annettiin paikon myös arvottava asema suhteessa kaupunkiin. Poliitiikan monimuotoinen rooli kommentteissa lieneekin syy siihen, miksi politiikka-diskurssiprosodia näytti olevan niin vahvasti esillä Kouvola koskevassa keskustelussa.

4.2.2 Palvelut ja tapahtumat

Toiseksi eniten *Kouvolan* diskurssiprosodioista esille nousi *palvelut ja tapahtumat*, johon on ryhmitelty 17% noodiin liittyvistä kollokaateista, kuten taulukosta 2 tulee ilmi. Tähän diskurssiprosodiaan kuului yhteensä 17 eri kollokaattia. Kollokaatit on esitetty merkitysryhmittäin taulukossa 4.

Taulukko 4. Diskurssiprosodia *palvelut ja tapahtumat*.

<i>palvelu</i>
KOULU, POLIISI
<i>tapahtumapaikka</i>
KESKUSTA, SEUTU, VETURI
<i>palveluntuottaja</i>
OY
<i>verbi</i>
LÖYTYÄ, JÄÄDÄ, TARVITA, JÄRJESTÄÄ, SIIRTYÄ
<i>adjektiivi</i>
PARAS, HUONO, HYVÄ
<i>abstrakti substantiivi</i>

Kuten taulukosta 4 ilmenee, palvelut ja tapahtumat -diskurssiprosodiaan sisältyi kollokaatteja yhteensä kuudesta eri merkitysryhmästä: palvelu, tapahtumapaikka, palveluntuottaja, verbi, adjektiivi ja abstrakti substantiivi. Näitä kollokaatteja olivat esimerkiksi KOULU, KESKUSTA, OY, TARVITA, HYVÄ ja HINTA. Käytännössä kollokaatit esiintyvät aineistossa useimmiten Kouvolan alueen palveluita arvioivassa kontekstissa. Kommenttien fokuksessa näytti olevan yleensä pikemminkin tietty palvelu kuin varsinaisesti kunnan alueella tarjottavat palvelut tai niiden laatu yleisesti (esimerkit 12, 13). Aineistoon kuului myös huomattavan monta konkordanssiriviä, joilla kommentoijat pyysivät suosituksia palveluista (esimerkki 14).

- 12) löytyykö Kouvolasta **huonompaa** asiakaspalvelua kuin Torniomäen Tokmannilta.
- 13) **Paras** karaokepaikka Suomessa on Kouvola Pyöreä torppa!
- 14) mikä mahtaa olla Kouvolan **paras** karaokemesta?

Esimerkeissä Kouvolan rooli on toimia pelkästään tapahtumapaikkana: kommentoijat olivat kiinnostuneita ennemminkin tietyistä palveluista ja niiden laadusta kuin Kouvolasta itsestään. Sen sijaan kaupungin merkitys tuli selvemmin esille sellaisista konkordansseista, joissa keskustelun aiheena olivat kunnan tarjoamat palvelut, kuten viranomaistoiminta ja opetus. Näissä kaupungin rooli on toimia palveluntarjoajana, jolloin kommentit useimmiten sisälsivät kannanoton myös kaupungista itsestään (esimerkit 15, 16; ks. myös 1). Erityisesti koulu-kollokaattiin liittyvät konkordanssit sisälsivät usein avointa kritiikkiä kaupunkia kohtaan. Esimerkissä 1 koulujen lakkauttaminen rinnastetaan terveyspalvelujen ”pilaamiseen”, mikä viestii kommentoijan negatiivisesta suhtautumisesta kaupungin toimintaa kohtaan. Esimerkissä 16 puolestaan viattoman kysymyksen taakse piiloutuu taustaoletus, jonka mukaan kaupunki ei tarjoa koululaisille riittävän ravitsevaa lounasta. Kaupunkia siis kritisoitiin sen tarjoamien palvelujen kautta.

- 15) ammattimaisia varkauksia vanhuksiin kohdistuen tehtailaan Kouvolassa ja **poliisi** vain hysyyttelee
- 16) Miksi Kouvolan **kouluissa** lapset nälässä?

Kaupungin arvioiminen palvelujen ja tapahtumien varjolla tuli ilmi erityisesti palvelut ja tapahtumat -diskurssiprosodiaan kuuluvien kollokaattiverbien kautta (esimerkit 17, 18). Myös abstrakteihin substantiiveihin liittyvät konkordanssit sisälsivät useimmin pohdintaa Kouvolan seudun palvelutarjonnasta yleensä (esimerkit 19, 20).

- 17) kiva kun Kouvola **järjestää** Kuusankoskella ilotulituksen
- 18) ei kait Kouvolassa **tarvitse** tai edes ole pysäköintivalvontaa.

- 19) tämän rahastamisen yhteydessä on Kouvolan asukkaiden **palveluja** karsittu
 20) sentään jotain **toimintaa** Kouvolaissa.

Esimerkeissä kaupunkia arvioidaan palvelujen ja tapahtumien kautta. Vaikka suuri osa palvelut ja tapahtumat -diskurssiprosodian konkordansseista painottuikin yksittäisten palvelujen kommentoimiseen, oli myös kaupunkia tai sen toimintaa arvottavia kommentteja huomattavasti. Niinpä myös palveluilla oli vahvasti vaihteleva rooli *Kouvolasta* puhuttaessa.

4.2.3 *Kouvola tapahtumapaikkana*

Kouvola tapahtumapaikkana -diskurssiprosodiaan on koottu kollokaatteja useista toisiinsa liittymättömistä merkitysryhmistä. Näitä kollokaatteja yhdisti kuitenkin se, että niiden alle lukeutuvilla konkordanssiriveillä Kouvolan pääsääntöinen rooli oli toimia pelkästään tapahtumapaikkana. Niinpä siinä missä palvelut ja tapahtumat -diskurssiprosodiaan sisältyi runsaasti kanta-aottavia konkordansseja, tässä ryhmässä *Kouvola* esitettiin neutraalina paikanilmauksena, jota ei ole tarpeen kommentoida. Diskurssiprosodiaan kuuluvat merkitysryhmät olivat media, paikat ja paikanilmaukset, asuminen, seuranhaku, liikkuminen sekä rikokset ja onnettomuudet. Yhteensä diskurssiprosodiaan kuului 37 kollokaattia, jotka kattoivat 37% kaikista tutkimuksen kollokaateista. Kollokaatit on esitetty tarkemmin taulukossa 5.

Taulukko 5. Diskurssiprosodia *Kouvola tapahtumapaikkana*.

<p><i>media</i> SANOMA, SEURATA, NÄIN, TOIMITTAJA, UUTISOIDA, EPÄILLÄ, KATSOA, UUTINEN</p> <p><i>paikat ja paikanilmaukset</i> INKEROINEN, TIE, LAHTI, ASEMA, TÄÄLLÄ, KUUSANKOSKI, SAVONLINNA, KYMENLAAKSO, KOTKA, LAPPEENRANTA</p> <p><i>asuminen</i> KESKUSTA, SEUTU, ALUE, ASUA, MUUTTAA, ASUKAS, HINTA</p> <p><i>seuranhaku</i> SEURATA, LÖYTYÄ, NAINEN, KAVERI</p> <p><i>liikkuminen</i> MATKA, SUUNTA, MENNÄ, AJAA</p> <p><i>rikokset ja onnettomuudet</i> EPÄILLÄ, POLIISI, NAINEN, TAPAHTUA</p>

Näille kollokaateille on siis yhteistä, ettei niihin liittyvissä konkordansseissa käsitelty niinkään *Kouvolaa*, vaan kaupunki esiintyi kommentteissa adverbiaalinen roolissa. Niinpä myös suhtautuminen

kuntaan vaikutti neutraalilta, eikä selkeitä positiivisia tai negatiivisia mielle yhtymiä ollut havaittavissa. Sen sijaan konkordanssien fokuksessa olivat esimerkiksi kaupunkiin sijoittuvat tapahtumat, kommentoijien henkilökohtaiset halut ja kokemukset, paikallisuutiset sekä etäisyydet eri kaupunkien välillä (esimerkit 21, 22, 23, 24, 25, 26).

- 21) Hennalan mielenosoituksella ja samana yönä Kouvolassa **tapahtuneella** polttopulloiskulla ei ole yhteyttä toisiinsa
- 22) **löytyiskö** halukasta nuorta misua Kouvola Kuusankoski
- 23) Kouvolaan **muutin** pääkaupunkiseudulta
- 24) elimäkeläinen yrittäjä kertoo Kouvolan **Sanomille**, että kaikki hänen kesätöihin palkkaamansa nuoret tuottivat pettymyksen
- 25) Kouvolasta matkaa 6 kilsaa
- 26) onneksi ensi viikolla on vapaata ja voin käydä Lahdessa, Kouvolassa ja/tai Jyväskylässä

Kouvola tapahtumapaikkana -diskurssiprosodia on siinä suhteessa erityisen kiinnostava, että vaikka prosodiaan sisältyi monenlaisia kommentteja ja jopa erilaisia puheenaiheita, prosodia oli määrällisesti kaikkein merkittävin tässä tutkimuksessa havaitsemistani diskurssiprosodioista. Sen rajat olivat siis tietyllä tapaa löysimmät, ja prosodian sisälle mahtui siksi paljon erilaisia konkordansseja. Muista diskurssiprosodioista se erottui lisäksi neutraaliutensa kautta, mikä teki prosodiasta tavallaan myös vaikeammin havaittavan. Erityisen paljon *Kouvola* tapahtumapaikkana -diskurssiprosodiasta erottuivat kritisoinnin ja puolustamisen diskurssiprosodiat, joita esittelen seuraavaksi.

4.2.4 Kritisointi ja puolustaminen

Tässä alaluvussa käsitellään kahta vastakkaista aineistosta esiin noussutta diskurssiprosodiaa: kritisoinnin ja puolustamisen diskurssiprosodioita. Prosodiat erosivat siten muista, edellä esitellyistä diskurssiprosodioista, että niissä keskustelun aiheena oli nimenomaan Kouvola kokonaisuudessaan – siinä missä politiikka- ja palvelut ja tapahtumat -diskurssiprosodioissa kaupunkia kommentoitiin sen yksittäisten ominaisuuksien kautta, kritisoinnin ja puolustamisen diskurssiprosodioissa puhe oli kaupungista yleisesti. Lisäksi kritisointi- ja puolustamisdiskurssiprosodioille oli ominaista, että niissä Kouvola esitettiin mustavalkoisesti joko pelkästään hyvänä (diskurssiprosodia puolustaminen) tai pelkästään huonona (diskurssiprosodia kritisointi).

Kritisoimisen diskurssiprosodiaan kytkeytyi yhteensä kahdeksan eri kollokaattia, mikä tarkoittaa 8% kaikista tutkimuksen kollokaateista. Kollokaatit on jaoteltu neljään eri merkitysryhmään, jotka olivat verbit, adjektiivit, abstraktit substantiivit ja toistuvuutta ilmaisevat sanat. Tällaisia sanoja olivat AINA ja KAIKKI. Kollokaatit on esitetty tarkemmin taulukossa 6.

Taulukko 6. Diskurssiprosodia *kritisointi*.

<i>verbi</i>
NÄHDÄ, MUUTTAA, OSATA
<i>adjektiivi</i>
PASKA, HUONO
<i>abstrakti substantiivi</i>
IHME
<i>toistuvuutta ilmaiseva sana</i>
AINA, KAIKKI

Puolustamisen diskurssiprosodiaan puolestaan lukeutui verbejä, adjektiiveja, abstrakteja substantiiveja ja vertailua ilmaisevia sanoja. Yhteensä kollokaatteja oli viisi, mikä tekee 5% kaikista kollokaateista. Kollokaatit on ilmaistu merkitysryhmittäin taulukossa 7.

Taulukko 7. Diskurssiprosodia *puolustaminen*.

<i>verbi</i>
NÄHDÄ
<i>adjektiivi</i>
HYVÄ
<i>abstrakti substantiivi</i>
ONGELMA
<i>vertailua ilmaiseva sana</i>
MUUALLA, KAIKKI

Käytännössä kritisoinnin diskurssiprosodiaan sisältyi runsaasti Kouvolan avointa arvostelemista, mutta konkordanssiriveiltä ei siitä huolimatta ollut nähtävissä juurikaan perusteluja näille väitteille. Useimmat konkordanssit olivat siis pelkkiä toteamuslauseita (esimerkit 27, 28, 29).

- 27) Mitä sun elämässä on tapahtunut että olet joutunut **muuttamaan** Kouvolaan?
- 28) Kouvola on muuten ihan **paska** kaupunki, jos sitä nyt kaikki ei jo tiedä.
- 29) Ei **ihme** että teitä vituttaa kun masentavimmat rakennukset Kouvolaasta!

Tyypillistä kritisointi-diskurssiprosodian kollokaateille näytti olevan myös se, että niissä painottui toistuvuuden ilmaisu. Selkeästi tämä tuli ilmi kollokaattien AINA ja KAIKKI kohdalla, sillä sanat jo itsessään ilmaisevat toistuvuutta: AINA ajallista ja KAIKKI nomineihin viittaavaa toisteisuutta (esimerkit 30, 31). Lisäksi esimerkiksi kollokaatin OSATA kohdalla toisteisuuden painottaminen tuli

esille konkordansseilla ilmaistujen subjektien tai niiden puutteen kautta: tarkan tekijän nimeämisen sijaan kommentoijat käyttivät esimerkiksi passiivia, joka yleistää tekijän kaikiksi ihmisiksi (esimerkit 32, 33).

- 30) Kouvola on ollut **aina** juoruajien mekka.
- 31) Kouvolaassa **kaikki** kyttävät toisiaan
- 32) Kouvolaassa ei **osata** lukea liikennesääntöjä
- 33) ei kouvolaalaiset **osaa** tehdä mitään oikein eikä opettaa varsinkaan

Puolustamisen diskurssiprosodia taas toimi ikään kuin vastauksena kritisointi-diskurssiprosodialle. Siinä missä kriittiset konkordanssit olivat pääosin yksinkertaisia väittämiä, puolustavat olivat analyttisempia ja sävyiltään usein sovinnollisia. Käytännössä tämä tuli esille etenkin vertailua ilmaisevissa kollokaateissa MUUALLA ja KAIKKI, joissa Kouvola verrattiin muihin suomalaiskuntiin (esimerkit 34, 35). Huomion arvoista on myös se, että kehuviin kollokaattien (esimerkki 36) lisäksi aineistosta erottui myös päämerkitykseltään kielteisiä kollokaatteja, kuten ONGELMA. Tällaisia sanoja oli konkordansseissa käytetty kielteisessä lauseyhteydessä, jolloin kollokaatin merkitys muuttui neutraaliksi tai jopa positiiviseksi, sillä kieltosana kumoaa substantiivin negatiivisen merkityksen (esimerkit 37, 38).

- 34) Ei täällä Kouvolaassa ole poliiseja liikenteessä yhtään sen enempää kuin missään **muuallakaan**.
- 35) Kouvolaassa on **kaikki** palvelut
- 36) Kouvola on **hyvä** kunta
- 37) **Ongelmia** ei ole pelkästään Kouvolaassa.
- 38) Ei maahanmuutto ole mikään Kouvolan **ongelma!**

Puolustamisen ja kritisoinnin diskurssiprosodiat erosivat toisistaan siis runsaasti, mutta samalla ne olivat ainoat prosodiat, joissa keskustelijat ottivat kantaa suoraan *Kouvolaan*. Kiinnostavaa onkin se, ettei näitä konkordansseja ollut kovinkaan paljoa koko aineistossa, vaan aineistossa painottui enemmän kunnan yksittäiset palvelut, tapahtumat ja päätökset. Ilmiöstä viestii myös kollokaattien suhteellisen pieni määrä: yhteensä puolustamis- ja kritisointi-diskurssiprosodioihin lukeutui 10% tutkimuksen sadasta merkittävimmästä kollokaatista.

4.2.5 Diskurssiprosodioiden ulkopuolelle jäävät kollokaatit

Edellä esiteltyjen diskurssiprosodioiden ulkopuolelle jäi yhteensä 27 kollokaattia, mikä tarkoittaa 27% sadasta merkittävimmästä kollokaatista. Kollokaatit on esitetty taulukossa 8.

Taulukko 8. Diskurssiprosodioiden ulkopuolelle jäävät kollokaatit.

<i>adverbi</i> TÄNÄÄN, TAAS, MIKSI, KUINKA
<i>partikkeli</i> KYLLÄ, VARMAAN, VAI, MUUTEN, AINAKIN, MYÖS
<i>pronomini</i> MINÄ, MÄ, KUKA, MIKÄ
<i>verbi</i> KUULLA, LIITTYÄ, OLLA, PÄÄSTÄ, TULLA, TAITAA
<i>substantiivi</i> HOMO, TULO, NIMI
<i>adjektiivi</i> SUURI, UUSI, VANHA, SEURAAVA

Kuten taulukosta 8 tulee ilmi, suuri osa diskurssiprosodioiden ulkopuolelle jääneistä kollokaateista on taipumattomia tai vajaasti taipuvia sanoja, kuten partikkeleita ja adverbeja. Näitä sanaluokkia leimaa se, että niihin kuuluvat sanat ovat tyypillisesti lauseen sisällön ulkopuolella, jolloin sanoja käytetään tuomaan lausumaan jonkin lisämerkityksen. Niinpä sanoja voidaan käyttää millaisessa keskustelussa tahansa, jolloin niiden runsas esiintyminen on luonnollista. (VISK § 794.) Myös taulukossa 8 esitetyt pronominit, verbit, substantiivit ja adjektiivit ovat yleisiä suomen kielen sanoja, joita käytetään runsaasti eri konteksteissa (ks. esim. Laalo 1989). Siksi on luontevaa, että nämä kollokaatit esiintyivät aineistossa eri merkityksissä, jolloin niiden sijoittaminen tiettyyn diskurssiprosodiaan ei ollut mahdollista. Kollokaatit jäivät siis tutkimuksessa esiin nousseiden diskurssiprosodioiden ulkopuolelle niiden luonnollisen variaation vuoksi.

5 PÄÄTÄNTÖ

5.1 Yhteenveto ja johtopäätökset

Tämän tutkielman tarkoitus on ollut selvittää, millaisia diskurssiprosodioita *Kouvola*-sanaan liittyy Suomi24-keskustelupalstalla. Ensimmäinen tutkimuskysymykseni oli, mitä kollokaatteja *Kouvola* saa. Tutkimuksessa *Kouvola* sai yhteensä 246 kollokaattia, joista analysoin tarkemmin sataa merkittävintä. Kollokaattien kirjo oli laaja: joukkoon mahtui sanoja lähes kaikista sanaluokista. Suurin osa kollokaateista oli perusmerkitykseltään neutraaleja, mutta konkordansseja tarkastellessa kävi ilmi, että monia niistä käytettiin toistuvasti kritisoiivassa merkityksessä. Osa *Kouvolan* kollokaateista oli sekä perusmerkitykseltään että käyttöfunktioiltaan positiivisia.

Toinen tutkimuskysymykseni oli, millaisia diskurssiprosodioita kollokaateista nousee esiin. Analyysissa kollokaateista erottui viisi diskurssiprosodiaa: *politiikka, palvelut ja tapahtumat, Kouvola tapahtumapaikkana, kritisointi ja puolustaminen*. Osa tutkimuksen kollokaateista jäi näiden diskurssiprosodioiden ulkopuolelle. Keskeistä tutkimuksessa esiin nousseissa diskurssiprosodioissa oli se, että suuri osa myös niistä vaikutti ensinäkemältä neutraaleilta – vain kritisoinnin ja puolustamisen diskurssiprosodiat olivat selkeästi negatiivis- tai positiivissävytteisiä. Kollokaattien käyttöä aineistossa tutkailtaessa tuli kuitenkin ilmi, että niin politiikka-, palvelut ja tapahtumat- kuin *Kouvola* tapahtumapaikkana -diskurssiprosodianakin sisälle sijoittui runsaasti etenkin kriittisiä konkordansseja neutraalien sijaan. Suurin osa tutkimuksessa esiin nousseista diskurssiprosodioista oli siis sävyiltään kriittisiä, vaikka myös positiivisia ja neutraaleja konkordansseja oli paljon.

Kritiikin suuren määrän syyhyn tämä tutkimus ei tarjoa vastausta, mutta huomion arvoista lienee se, että suurin osa kommentoijista vaikutti olevan paikallisia: kommenteissa mainittiin tarkkoja sijainteja kaupungin sisällä, siteerattiin paikallista sanomalehteä ja etsittiin seuraa. Nämä seikat kielivät monien kommentoijien olevan kouvoolalaisia, mikä saattaa vaikuttaa keskustelujen sävyyn.

Koska keskustelijat vaikuttavat olevan pääosin paikallisia, ei tutkimuksen aineistosta voi tehdä päätelmiä myöskään siitä, millaisia diskurssiprosodioita ulkopaikkakuntalaiset mahdollisesti liittäisivät *Kouvolaan*. Tämä eroaa runsaasti siis esimerkiksi Ainialan ja Jantusen tutkimuksesta, jossa *Stadi* pystyttiin liittämään paikallisiin kommentoijiin ja *Hesa* ulkopaikkakuntalaisiin (2019). Tämän tutkimuksen tulokset erosivat muutenkin paljon Ainialan ja Jantusen tuloksista, sillä paikkojen ja suuntien sijaan omassa tutkimuksessani korostuivat politiikka sekä palvelut ja tapahtumat.

5.2 Arviointi ja jatkotutkimusmahdollisuudet

Tämän tutkimuksen tutkimusasetelma oli siitä yksinkertainen, ettei minun tarvinnut juuri miettiä Korpissa käyttämäni hakusanaa: *Kouvola* itsessään on neutraali paikannimi, joka ei tietävästi sisällä arvottavaa aspektia. Siksi olikin erityisen kiinnostavaa tutkia, mitä toistuvia diskursseja nimeen liitetään. Ongelmallisempaa sen sijaan oli AntConcin laskemien kollokaattien analysoiminen, sillä analyysiprosessissa jouduin pitkälti tukeutumaan omaan tuntumaani siitä, missä yhteydessä tiettyä kollokaattia oli konkordansseissa eniten käytetty. Epävarmuutta aiheuttivat sekä AntConcissa näytettävien konkordanssien lyhyys että sanojen lemmatisoitu muoto, sillä molemmat pakottivat minut ainakin joiltakin osin päättelemään konkordanssin sanajärjestyksestä, mitä kommentin alkuperäinen kirjoittaja on kirjoittaessaan tarkoittanut. Tästä johtuen on mahdollista, että olen analyysissani tehnyt satunnaisia tulkintavirheitä. Virheiden määrän uskon kuitenkin jääneen pieneksi, sillä konkordanssin sanajärjestys ohjaa tulkintaa pääasiassa oikeaan suuntaan (ks. esim. Hakulinen 1975). Lisäksi konkordanssien suuren määrän takia jouduin luottamaan omaan tuntemukseeni siitä, missä merkityksessä tiettyä kollokaattia oli eniten käytetty. Kvalitatiiviselle tutkimukselle tutkijan oman näkemyksen painottuminen on kuitenkin ominaista, jolloin myös virheitä saattaa sattua (Pietikäinen & Mäntynen 2009: 139). Tässä tutkimuksessa pyrin minimoimaan mahdollisesti tekemiäni virheitä lukemalla konkordansseja useampaan otteeseen, jolloin väärintulkintojen mahdollisuus pienenee. Näin olen pyrkinyt edistämään tutkimukseni reliabiliteettia.

Tutkimuksen validiteettia sen sijaan lisää korpusavusteisen diskurssintutkimuksen hyvä soveltuvuus aiheeseeni. Tutkimussuuntaus sopii työhöni hyvin, sillä siinä tavoitteena on tilastolliseen näyttöön perustuvan informaation tarkempi analysoiminen kvalitatiivisin keinoin (Partington, Duguid & Taylor 2013: 10). Tämä pätee myös omaan työhöni, jonka tarkoituksena oli tuoda näkyväksi noodiin yleisesti liitettäviä ennako-oletuksia ja asenteita. Tämä onkin tyypillinen asetelma korpusavusteisessa diskurssintutkimuksessa (Partington, Duguid & Taylor 2013: 10–11). Lisäksi nettikeskusteluissa toistuvia diskursseja tutkittaessa aineiston on oltava riittävän laaja, jotta merkityskenttien toistuvuus voidaan luotettavasti todeta. Siksi tilastollinen analyysi sopi tutkimukseeni mainiosti. Aineistona Suomi24-korpus tarjosi siis tämänkaltaiselle tutkimukselle vahvemman reliabiliteetin kuin esimerkiksi itse koottu, pienempi aineisto olisi voinut tarjota. Käyttämässäni korpusaineistossa on myös se hyvä puoli, että se sisältää luonnollista kielenkäyttöä, jolloin diskurssiprosodian tutkiminen on mahdollista. (Lounela & Heikkinen 2012: 120–121; Ainiala & Jantunen 2019: 60.) Jos taas olisin teettänyt esimerkiksi kyselytutkimuksen, olisi tutkimuksen kohteeksi muodostunut ennemmin informanttien käsitykset Kouvolasta kuin se, miten kaupungista todella puhutaan. Korpusavusteisen diskurssintutkimuksen etu on lisäksi se, että siinä kielenkäyttöä

tarkastellaan kontekstisidonnaisesti osana ympäröivää yhteiskuntaa (Partington, Duguid & Taylor 2013: 11). Tämä sopi työhöni hyvin – onhan tutkimukseni kohde osa yhteiskunnallista keskustelua.

Tilastollinen analyysi ja erityisesti aineiston suuri määrä nostivat myös tutkimustulosteni yleistettävyyttä. Kyseenalaista on kuitenkin se, voidaanko Suomi24-keskustelufoorumille perustuvia tutkimustuloksia yleistää nettikeskustelujen ulkopuolelle: koska keskustelupalstan osallistujien ei ole välttämätöntä rekisteröityä, ei keskustelijoiden ikä- tai sukupuolijakaumasta ole luotettavaa tietoa. Lisäksi foorumilla on ala-arvoisen keskustelukulttuurin maine (Saarikoski 2013: 150). Niinpä ei voida myöskään sanoa varmaksi, kuinka suuri osa tutkimuksen aineistona käytetyistä kommentteista on kirjoitettu tarkoituksellisesti provosoiden. Siksi tutkimustulosten yleistäminen Suomi24-keskustelufoorumien ulkopuolelle ei ehkä ole perusteltua.

Tässä tutkimuksessa analysoin tarkemmin sataa merkitsevintä kollokaattia, mikä saattoi olla liian vähän: jos tutkittavia kollokaatteja olisi ollut enemmän, olisi myös diskurssiprosodioita saattanut paljastua suurempi määrä. Tämä näkyy erityisesti *Kouvola* tapahtumapaikkana -diskurssiprosodiassa, johon kerääntyi useita, löyhästi toisiinsa liittyviä merkitysryhmiä. On mahdollista, että kollokaattien suurempi määrä olisi osoittanut näiden merkitysryhmien kuuluvan todella omiksi diskurssiprosodioikseen. Siksi *Kouvolan* tutkiminen suuremmassa aineistossa voisi olla tarpeen.

Myös muita jatkotutkimusmahdollisuuksia on useita. Ensinnäkin olisi kiinnostavaa jatkaa korpusonomastista tutkimusperinnettä myös muihin suomalaiskuntiin ja verrata niitä koskevia diskurssiprosodioita *Kouvolaan*: onko esimerkiksi alueen koolla yhteyttä siihen liittyviin diskurssiprosodioihin? Toisekseen *Kouvolan* saamia diskurssiprosodioita voisi olla tähdellistä tutkia esimerkiksi sanomalehtikorpuksista, jolloin aineiston lähde olisi voimakkaasti erilainen Suomi24-korpuksien verrattuna. Tällöin olisi mahdollista pohtia myös sitä, antavatko asiatekstit ja kansalaistekstit kenties keskenään erilaisen kuvan kaupungista.

LÄHTEET

- Ainiala, Terhi & Jantunen, Jarmo Harri 2019: Korpusonomastinen tutkimus slanginimistä *Hesa* ja *Stadi* digitaalisissa diskursseissa. – *Sananjalka* 61 (61) s. 57–79.
- Aldrin, Emilia 2017: Creating identities through the choice of first names. – Terhi Ainiala & Jan-Ola Östman (toim.), *Socio-onomastics: the pragmatics of names* s. 45–68. Amsterdam: John Benjamins Publishing Company.
- Aller Media Oy 2014: *Suomi 24 -korpus (2017H2)* [tekstikorpus]. Kielipankki. – https://korp.csc.fi/#?stats_reduce=word&cqp=%5B%5D 3.12.2019.
- Anthony, Laurence 2015: *AntConc. Version 3.5.8. Computer Software*. Tokyo: Waseda University. – <http://www.laurenceanthony.net/> 5.12.2019.
- Baker, Paul 2010: *Sociolinguistics and Corpus Linguistics*. Edinburgh: Edinburgh University Press Ltd.
- Biber, Douglas; Conrad, Susan & Reppen, Randi 2007: *Corpus linguistics. Investigating language structure and use*. Cambridge: Cambridge University Press.
- Borin, Lars; Forsberg, Markus & Roxendal, Johan 2012: Korp – the corpus infrastructure of Språkbanken. Versio 5.0.10. – https://korp.csc.fi/#?stats_reduce=word&cqp=%5B%5D&corpus=ftb3europarl,ftb3_jrcacquis,ftb2,reittidemo 10.12.2020.
- Hakulinen, Auli 1975: Sanajärjestyksen eri tehtävistä. – *Virittäjä* 79 (1), s. 85–93. <https://journal.fi/virittaja/article/view/36538> 18.3.2020.
- Halverson, Sandra & Engene, Jan Oskari 2010: Domains and dimensions in metonymy: A corpus-based study of Schengen and Maastricht. – *Metaphor and Symbol* 25 (1), s. 1–18. <https://doi.org/10.1080/10926480903538456>
- Heikura, Minna 2015: Suomen harmainta kaupunkia väritetään uusin kuvin – "Viesti tulee ihmisiltä sydämistä". *YLE uutiset* 14.11.2015 – <https://yle.fi/uutiset/3-8453480> 7.12.2019.
- Holm, Vesa 2012: *Hirvenmetsästäjien epävirallista paikannimistöä*. Pro gradu -tutkielma. Jyväskylän yliopiston kielten laitos. – <http://urn.fi/URN:NBN:fi:juu-201205301760> 25.1.2020.
- Hunston, Susan 2002: *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Jantunen, Jarmo Harri 2004: *Synonymia ja käännessuomi. Korpusnäkökulma samanmerkityksisyyden kontekstuaalisuuteen ja käännesskielen leksikaalisiin erityispiirteisiin*. Vaitöskirja. Joensuun yliopisto.
- 2012: Korpusvetoinen tekstilajianalyysi: sanalistat ja genreavainsanat. – Vesa Heikkinen, Eero Voutilainen, Petri Lauerma, Ulla Tiililä & Mikko Lounela (toim.), *Genreanalyysi – tekstilajitutkimuksen käsikirja* s. 360–371. Helsinki: Gaudeamus.
- 2018a: Korpusavusteinen diskurssitutkimus (CADS): analyysiesimerkki homouden ja heterouden digitaalisista diskursseista. *AFinLA-E: Soveltavan Kielitieteen Tutkimuksia* 11 (2018) s. 20–44. – <https://doi.org/10.30660/afinla.69259> 20.1.2020.
- 2018b: Homot ja heterot Suomi24:ssä: analyysi digitaalisista diskursseista. – *Puhe ja kieli* 38 (1) s. 3–22.
- Johansson, Marjut; Jantunen, Jarmo H.; Heimo, Anne; Ahonen, Mirka & Laippala, Veronika 2018: Verkko keskustelujen kansa: Korpusavusteinen diskurssianalyysi Suomi24-keskustelupalstasta. – *Sananjalka* 60 (60) s. 96–117.
- Kielipankki 2020: Työkalut. – <https://www.kielipankki.fi/tyokalut/> 27.2.2020.

- Kouvola kaupunki 2020: Kouvola-tietoa. – <https://www.kouvola.fi/kouvolankaupunki/kouvola-tietoa/> 28.1.2020.
- Laalo, Klaus 1989: Homonymiasta ja polysemiasta. – *Virittäjä* 93 (2) s. 220–235. – <https://journal.fi/virittaja/article/view/38284> 18.3.2020.
- Lagus, Krista; Pantzar, Mika; Ruckenstein, Minna & Ylisiurua, Marjoriikka 2016: *Suomi24. Muodonantoa aineistolle*. Helsinki: Helsingin yliopisto. – https://helda.helsinki.fi/bitstream/handle/10138/163190/257383_HY_VALT_suomi24_muodonantoa_aineistolle.pdf?sequence=1 21.2.2020.
- Lehto, Liisa-Maria 2018: *Korpusavusteinen diskurssianalyysi japaninsuomalaisten kielipuheesta*. Väitöskirja. Oulun yliopiston suomen ja saamen kielen laitos. – <http://urn.fi/urn:isbn:9789526219097> 27.1.2020.
- Lounela, Mikko & Heikkinen, Vesa 2012: Korpus. – Vesa Heikkinen, Eero Voutilainen, Petri Lauerma, Ulla Tiililä & Mikko Lounela (toim.), *Genreanalyysi – tekstilajitutumuksen käsikirja* s. 120–127. Helsinki: Gaudeamus.
- Maanmittauslaitos 2020: Suomen pinta-ala kunnittain 1.1.2020. – https://www.maanmittauslaitos.fi/sites/maanmittauslaitos.fi/files/attachments/2020/01/Vuoden_2020_pinta-alatilasto_kunnat_maakunnat.pdf 28.1.2020.
- McEnery, Tony; Richard Xiao & Yukio Tono 2006: *Corpus-based language studies: An advanced resource book*. New York: Routledge.
- Mikkonen, Pirjo 2013: *"Otti oikean sukunimen": vuosina 1850-1921 otettujen sukunimien taustat*. Väitöskirja. Helsingin yliopisto.
- Niemi, Petri 2018: Kouvola suuren tuskan edessä – sadat ihmiset muuttavat pois. *YLE uutiset* 17.4.2018. – <https://yle.fi/uutiset/3-10158131> 7.12.2019.
- Padula, Malena; Panza, Carolina & Muñoz, Verónica 2020: The pronoun this as a cohesive encapsulator in engineering semi-popularization articles written in English. – *Journal of English for Academic Purposes* 44 (2). – <https://www.sciencedirect.com.ezproxy.jyu.fi/science/article/pii/S1475158518305575?via%3Dihub> 18.1.2020.
- Partington, Alan 2004: Corpora and discourse, a most congruous beast. – Alan Partington, John Morley & Louann Haarman (toim.), *Corpora and discourse* s. 11–20. Bern: European Academic Publishers.
- Partington, Alan; Duguid, Alison & Taylor, Charlotte 2013: *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies (CADS)*. Amsterdam: John Benjamins Publishing Company.
- Pietikäinen, Sari & Mäntynen, Anne 2009: *Kurssi kohti diskurssia*. Tampere: Vastapaino.
- Saarikoski, Petri 2013: 2008 – Kaiken kansan Facebookista Obaman Twiitteihin. – Jaakko Suominen, Sari Östman, Petri Saarikoski & Riikka Turtiainen (toim.), *Sosiaalisen median lyhyt historia* s. 146–170. Helsinki: Gaudeamus.
- Salmi, Sara 2019: Pasilan Triplasta tuli järkälemäinen laatikko, jota verrataan Itä-Berliiniin ja Kouvola arkkitehtuuriin – ”Se, mitä rakennetaan, kuvastaa yhteiskunnan arvoja” *YLE uutiset* 9.9.2019. – <https://yle.fi/uutiset/3-10949871> 28.1.2020.
- Sinclair, John 1991: *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Stubbs, Michael 2001: *Words and phrases: corpus studies of lexical semantics*. Oxford: Blackwell.
- Turunen, Jyri 2018: Kouvola on tekemistä tunnettuuden kanssa. – *Kouvola Sanomat* 30.8.2018 s. 18.

- Vaičėnonienė, Jurate 2001: Using corpora to obtain social and cultural information: A case study of America. – *Kalbu Studijos / Studies about Language* 1, s. 6–9.
- 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> 3.4.2020.
- Yule, George 1996: *Pragmatics*. Oxford Introductions to Language Study. Oxford: Oxford