

Wikipedia-artikkeli tekstilajina: koheesio keinot tekstin rakentajina

Maisterintutkielma

Heidi Paasikunnas

Suomen kieli

Kielten laitos, Jyväskylän yliopisto

2015

JYVÄSKYLÄN YLIOPISTO

Tiedekunta – Faculty Humanistinen	Laitos – Department Kielten laitos
Tekijä – Author Heidi Paasikunnas	
Työn nimi – Title Wikipedia-artikkeli tekstilajina: koheesion keinot tekstin rakentajina	
Oppiaine – Subject Suomen kieli	Työn laji – Level Maisterintutkielma
Aika – Month and year Toukokuu 2015	Sivumäärä – Number of pages 95 sivua
Tiivistelmä – Abstract Tässä tutkimuksessa tarkastelin suomenkielisiä Wikipedia-artikkeleita tekstilajin näkökulmasta. Wikipedia on nykyisin erittäin merkittävä tiedonhaun väline. Se sisältää valtavan määrän yhteistoiminnallisesti tuotettua tietoa ja poikkeaa kustannetuista tietosanakirjoista monin tavoin. Wikipedialle on leimallista, että kaikki voivat vapaasti osallistua sen artikkelien kirjoittamiseen ja muokkaamiseen. Tutkimukseni teoreettisena viitekehystenä toimi systeemis-funktionaalinen teoria. Pohjasin analyysini erityisesti M. A. K. Hallidayn ajatuksiin systeemis-funktionaalisesta tekstilajitutkimuksesta. Siinä kieltä tulkitaan sosio-kulttuurisessa kontekstissa. Teorian perusyksikkö on teksti: kieltä pyritään ymmärtämään tekstejä tarkastelemalla. Teksti nähdään teoriassa semanttisena kokonaisuutena, jolla on merkitys nimenomaan kokonaisuutena. Tutkimukseni tarkoitus oli saada käsitys Wikipedia-artikkelista tekstilajina. Kiinnitin erityisesti huomiota siihen, kuinka tekstien kohesiivinen rakenne rakentaa niiden yhtenäisyyttä. Tarkastelin aineiston tekstien yhtenäisyyttä myös tema- ja informaatorakenteen kautta. Pyrin lisäksi selvittämään, miten Wikipedia-artikkelit tyypillisesti rakentuvat ja millaiset tekstuaaliset piirteet ovat niille tunnusomaisia. Lisäksi analysoin funktioita, joita kyseisillä piirteillä on. Määrittelin myös Wikipedia-artikkelin rekisterin analysoimalla aineistosta tilannekontekstin kolmea muuttujaa: alaa, osallistujarooleja ja ilmenemismuotoa. Aineistoni sisälsi 30 Wikipedia-artikkelia, jotka jakautuivat aiheiltaan kaikkiin Wikipedian pääaiheluokkiin. Artikkelit vaihtelivat pituudeltaan huomattavasti. Aineiston pisin artikkeli oli 2928 sanaa pitkä, ja lyhin koostui ainoastaan 80 sanasta. Käsittelin artikkeleita siinä muodossa, jossa olin poiminut ne Wikipedia-sivustolta tammikuussa 2015. Wikipedia-artikkelit eivät kuitenkaan koskaan valmistu, vaan ne ovat aina muokkauksen ja päivityksen kohteina. Aineiston artikkelit osoittautuivat muotopiirteiltään ja koheesiokeinoiltaan varsin vaihteleviksi. Kaikkia aineiston artikkeleita kuitenkin yhdisti artikkelin käsittelemän aiheen määrittelemisen. Tämä tietosanakirjalle ominainen funktio osoittautui keskeiseksi tekstilajia määrittäväksi piirteeksi. Teksteille oli tyypillistä, etteivät tarkkaa lukemista vaativat koheesiokeinot nousseet suureen osaan, vaan teksteissä esiintyi paljon toistoa ja samaan merkityskenttään kuuluvia sanoja.	
Asiasanat – Keywords tekstintutkimus, tekstilaji, tietosanakirjat, koherenssi, koheesio	
Säilytyspaikka – Depository	
Muita tietoja – Additional information	

Sisällys

1 JOHDANTO	1
1.1. Tutkimuksen tausta ja tutkimuskysymykset	1
1.2 Metodi	3
1.3. Wikipedia-artikkelit aineistona	4
2 TIETOSANAKIRJAT	7
2.1 Yleisesti tietosanakirjoista	7
2.2 Wikipediasta	8
2.3 Aikaisempi tutkimus	12
3 SYSTEEMIS-FUNKTIONAALINEN TEORIA TEKSTILAJITUTKIMUKSESSA	15
3.1 Systeemis-funktionaalista teoriasta	15
3.2 Systeemis-funktionaalisen teorian käsitteitä	17
3.2.1 Kieli	17
3.2.2 Teksti	18
3.2.2 Koherenssi	20
3.2.3 Koheesio	21
3.2.3.1 Koheesiosta yleisesti	21
3.2.3.2 Kieliopillinen koheesio	23
3.2.3.3 Sanastollinen koheesio	29
3.2.4 Teema- ja informaatorakenne	33
3.2.5 Tekstilaji	34
3.2.6 Rekisteri	36
4 WIKIPEDIA-ARTIKKELIEN MUOTORAKENNE SEKÄ TEEMA- JA INFORMAATORAKENNE	39
4.1 Wikipedia-artikkelien muodon analyysi	39
4.2 Teema- ja informaatorakenne	44
5. KOHEESIO WIKIPEDIA-ARTIKKELEISSA	47
5.1 Kieliopillinen koheesio	47
5.1.1 Referenssi	47
5.1.2 Ellipsit ja substituutio	55
5.1.3 Kytkeä	58
5.2 Sanastollinen koheesio	65

5.2.1 Toisto	65
5.2.2 Yläkäsitteet	67
5.2.3 Yleismerkityksiset sanat	69
5.2.4 Merkityskentät	70
5.2.5 Paradigmaattiset merkityssuhteet	72
5.2.6 Parafraasit	77
6 PÄÄTÄNTÖ	79
6.1 Yhteenvedo Wikipedia-artikkelin tekstilajipiirteistä	79
6.2 Wikipedia-artikkelien rekisteri	81
6.3 Pohdintaa	86
LÄHTEET	90
Aineistolähteet	94

1 JOHDANTO

1.1. Tutkimuksen tausta ja tutkimuskysymykset

Tässä pro gradu -työssä tutkin suomenkieliseen Wikipediaan kirjoitettuja tietosanakirja-artikkeleita tekstilajin näkökulmasta. Koska päivitettyjä painettuja yleistietoteoksia ei enää Suomessa juuri julkaista, on perusteltua olettaa, että tiedon etsinnässä nojataan yhä enenevässä määrin internetin tarjontaan. Tästä tarjonnasta nousee selvästi esiin tietosanakirjasivusto Wikipedia, joka on laajimpia ja suosituimpia internetissä toimivia kaikille avoimia tietosanakirjoja. Dariusz Jemielniakin mukaan Wikimedia-projektit kuuluvat suurimpiin yhteistyöprojekteihin ihmiskunnan historiassa. Kun ihmiset kaikista kansallisuuksista rakentavat yhdessä maailman suosituimpiin kuuluvia verkkosivuja, he vaikuttavat suoraan siihen, mitä maailmassa tiedetään, uskotaan ja ajatellaan. (Jemielniak 2014: 4.)

”Täysin uudenlainen, lonkeroitaan joka suuntaan levittävä tietosanakirjaotus”, kuvailee puolestaan Kimmo Tuominen Wikipediaa (2006: 65). Lyhyesti määriteltynä Wikipedia on vapaa ja ilmainen verkossa toimiva yleistietosanakirja, jonka sisältö on tuotettu yhteistoiminnallisesti. Sivusto on nykyisin erittäin merkittävä tiedonhaun väline: se sisältää valtavan määrän tietoa, ja sitä käytetään tiedonhakuun ahkerasti. Englanninkielinen Wikipedia on laajuudessaan omaa luokkaansa, mutta myös suomenkielisestä Wikipediasta sisältöä löytyy todella suuri määrä. Wikipedia-sivustolta selviää, että suomenkielinen Wikipedia sisälsi yli 360 000 artikkelia joulukuussa 2014. Kyseisen joulukuun aikana sen artikkeleita luettiin yli 77 miljoonaa kertaa ja muokattiin yli 52 000 kertaa. (Wikipedia:Statistics 2015.) Hoda Baytiyeh ja Jay Pfaffman arvioivat, että syy Wikipedian suureen suosioon lienee sen sisältämän materiaalin valtava määrä ja se tosiasia, että hakukone Google yleensä ohjaa Wikipedia-sivuille (2010: 128).

Wikipediassa huomionarvoista ei ole ainoastaan sen houkutteleva lukijamäärä, vaan myös sen yhteistoiminnallinen tuotantotapa. Wikipedia on kerännyt miljoonia vapaaehtoisia kirjoittajia luomaan sisältöä. Sitä pidetään yhtenä suurimmista avoimen sisällön menestystarinoista. (Tereszkiewiz 2010: 66; Tuominen 2006: 65.) Wikipedian laaja suosio sekä tiedonhaun kohteena että osallistavana tiedon tuottajana tekee siitä varsin kiinnostavan ilmiön.

Tzvetan Todorov toteaa, että uudet genret syntyvät jo olemassa olevista genreistä (1990: 15). Wikipedia asettuu osaksi tietosanakirjojen jatkumoa, mutta eroaa perinteisistä tietosanakirjoista erityisesti tuotantotavoiltaan. Wikipedia-artikkeleita tutkittaessa on otettava huomioon niiden yhteisöllinen, perinteisesti julkaistuista tietosanakirja-artikkeleista poikkea-

va luomistapa. Kenellä tahansa on oikeus kirjoittaa Wikipediaan ja muokata sieltä jo löytyviä artikkeleita. Tämä ei kuitenkaan tarkoita, ettei sivuston sisältöä valvottaisi. Wikipedia-sivustolta löytyvät ohjeet artikkelien kirjoittamiseen, ja yhteisö myös pitää silmällä ja muokkaa artikkelien sisältöä.

Analysoin työssäni suomenkieliseen Wikipediaan kirjoitettuja artikkeleita tekstilajin näkökulmasta. Tutkimukseni teoreettisena viitekehysenä toimii systeemis-funktionaalinen teoria. Pohjaan tutkimukseni erityisesti M. A. K. Hallidayn käsityksiin systeemis-funktionaalisesta tekstilajianalyysistä. Pyrin aineistoni perusteella selvittämään, kuinka Wikipediaan kirjoitetut artikkelit tyypillisesti rakentuvat: millaiset piirteet ovat niille tunnusomaisia ja minkälaisia funktioita kyseisillä piirteillä on. Keskityn pääasiassa tarkastelemaan aineistoni rakennetta ja tekstuaalisia piirteitä, erityisesti sitä, miten tekstien koheesio ja temaattinen rakenne luovat artikkelien koherenssia. Analysoimalla aineiston teksteille ominaisia piirteitä pyrin muodostamaan käsityksen Wikipedia-artikkelien tekstilajista. Lisäksi määritän artikkelien rekisterin tarkastelemalla yllämainittujen asioiden lisäksi tilannekontekstin muuttujia. Haen vastauksia seuraaviin kysymyksiin:

1. Miten koherenssi muodostuu Wikipedia-artikkeleissa? Mitä koheesiokeinoja niissä käytetään?
2. Mitkä ovat Wikipedia-artikkeleille tyypillisiä tekstilajipiirteitä? Millaisia funktioita niillä on?
3. Millainen rekisteri Wikipedia-artikkeleissa muodostuu?

Ensimmäinen tutkimuskysymykseni käsittelee Wikipedia-artikkelien koherenssia. Muodostaakseni kuvan Wikipedia-artikkeleista teksteinä tutkin siis erityisesti tekstien koheesiota ja temaattista rakennetta. Saadakseni artikkelien tekstilajista kattavan kuvan tarkastelen myös muun muassa artikkelien rakennetta, jäsentelyä ja lähdeviittauskäytäntöjä. Näiden avulla pyrin vastaamaan toiseen tutkimuskysymykseen. Rekisterin määrittelyssä huomioin M. A. K. Hallidayta mukaillen tilannekontekstin muuttujat eli tekstin alan, ilmenemismuodon ja osallistujaroolit.

Koen tutkimukseni relevantiksi, sillä kuten aikaisemmin mainitut käyttäjäluvut osoittavat, Wikipedia on nykyisin merkittävässä osassa ihmisten tiedonhakua, ja käyttäjät myös muokkaavat runsaasti sen artikkeleita. Wikipediaa ei ole Suomessa aiemmin tutkittu tekstilajin näkökulmasta, minkä vuoksi tutkimukseni tarjoaa uutta tietoa Wikipedia-artikkelien koherenssista ja muista sille ominaisista tekstilajipiirteistä.

1.2 Metodi

Tutkimukseni metodisena viitekehyksenä käytän tekstilingvististä tutkimusta. Systemis-funktionaalisen teorian perusyksikkö on teksti (Halliday & Hasan 1976: 142), joten keskityn tutkimaan aineistosta erityisesti niitä kielellisiä ilmiöitä, jotka vaikuttavat tekstin muodostumiseen. Tarkastelen etupäässä leksikaalis-kieliopillisia ja tekstuaalisia ominaisuuksia. Päähuomioni on tekstin koherenssin tutkimisessa. Pyrin erityisesti selvittämään, mitä koheesion keinoja artikkeleissa käytetään.

Hiidenmaa huomauttaa, että rekisteriä ja tilanteista merkityksenantoa eivät ole vain poikkeukselliset kielen keinot, vaan myös ne keinot, jotka ovat samanlaisia kuin ”muualla”. Analyysia ei siis pitäisi kohdistaa pelkästään poikkeuksellisiin muotoihin. Tekstien perusteella voi kuitenkin todeta erilaisia tyypillisyyksiä eli kyseiselle rekisterille ominaisia valintoja, joita esiintyy muun muassa sanastossa, lauserakenteissa, tekstin jäsentelyssä ja muissa tekstin esitystavoissa. Hiidenmaan mukaan tekstien olemus on paikoin määritelty eksplisiittisesti, mutta paikoin odotukset ovat tekstilajin implisiittisiä käytänteitä, joihin lukijat ja kirjoittajat ovat sosiaalistuneet. (2001: 35, 37.)

Lingvistisen tekstianalyysin kannalta olennaista on teksti-käsitteen määrittelemineen. M. A. K. Halliday määrittää tekstin kieleksi, joka on funktionaalista, eli sillä on jokin merkitys tilanteessa (1978: 135–137). Hiidenmaa painottaa niin ikään, että teksti on käytössä olevaa kieltä, eli sillä on funktio, tehtävä ja tavoite. Tekstillä on myös hahmo ja olomuoto, mikä tarkoittaa, että se on sidoksissa aikaan, paikkaan, tilanteeseen ja henkilöihin. (Hiidenmaa 2000: 165–166.) Näistä näkökulmista pyrin käsittelemään aineiston tekstejä tutkimuksessani. Määrittelen tekstin käsitteen yksityiskohtaisemmin systemis-funktionaalisen teorian mukaan luvussa 3.2.2.

Tekstianalyysin tavoite on analysoida kolmea metafunktiota ja niiden kielellistä järjestymistä. Kieltä tarkastellaan tekstuaalisen metafunktion näkökulmasta, interpersoonaisen metafunktion eli vuorovaikutuksen näkökulmasta sekä ideationaalisen metafunktion eli sisällön ja representaation näkökulmasta. (Hiidenmaa 2000: 172.) Metafunktioihin palaan tarkemmin luvussa 3.1. Koska tarkastelen ensisijaisesti tekstien kielellisiä ominaisuuksia, tutkimukseni painotus on tekstuaalisessa metafunktiossa. Sen kautta kuitenkin muodostuu kielen suhde sekä kielelliseen ympäristöön että tilanteeseen ympäristöön (Halliday 1978: 112–113; Halliday & Hasan 1985: 23). Tekstuaalisen metafunktion kautta huomioidin analyysissäni myös interpersoonaisen ja ideationaalisen ulottuvuuden.

Analyysissä tarkastelen siis Wikipedia-artikkelien tekstuaalisia ominaisuuksia. Luvussa neljä tutkin artikkelien muotorakennetta ja temaattista rakennetta sekä sitä, miten tekstiä on jäsennellyt artikkeleissa. Erittelen sitä, kuinka muun muassa tiedon järjestäminen lukuihin ja lähdeviitekäytännöt sekä artikkelien teema- ja informaatorakenne rakentavat tekstilajia. Luvussa viisi analysoin koheesiokeinoja. Erittelen aineiston kieliopillista ja sanastollista koheesiota. Kieliopilliseen koheesioon sisällytän referenssin, ellipsit ja substituution sekä kytken. Sanastollista koheesiota tutkin puolestaan toiston, yläkäsitteiden, yleismerkityksisten sanojen, merkityskenttien, paradigmaattisten merkityssuhteiden ja parafrasien kautta. Keskityn tarkastelemaan erityisesti lause- ja virkerajat ylittävää koheesiota. Syntaktiset, lauseen sisäiset, ilmiöt jätän suurimmaksi osaksi analyysin ulkopuolelle. Määrittäessäni Wikipedia-artikkelien rekisteriä tarkastelen edellä mainittujen ominaisuuksien lisäksi myös tekstien alaa, osallistujarooleja ja ilmenemismuotoa.

Tutkimukseni on pääasiassa kvalitatiivista. Analyysini pohjaksi esitän joitakin lukuja siitä, kuinka monessa aineiston artikkelissa tai joissakin tapauksissa virkkeessä tietty ominaisuus tai koheesiokeino esiintyy, mutta muuten tarkastelen aineiston tekstuaalisia ominaisuuksia lähinnä laadullisesti. Hiidenmaa (2001: 36) tuo esiin, että merkkien merkitykset rakentuvat tilanteissaan, joten tilastolliset, muotoon perustuvat analyysit eivät tee rekistereille täyttä oikeutta, vaikka voivatkin toimia lähtökohtana.

1.3. Wikipedia-artikkelit aineistona

Aineistoni sisältää 30 suomenkielistä Wikipedia-artikkelia. Olen pyrkinyt saamaan kokoon mahdollisimman monipuolisen artikkelikokoelman, jotta voisin analysoida Wikipedia-artikkeleille yhteisiä tekstilajipiirteitä ja muodostaa niistä jonkinlaisen kokonaiskuvan. Tiedostan, että aineistoni on vain murto-osa suomenkielisen Wikipedian sisällöstä, mutta koska tarkoitus on keskittyä nimenomaan kielen piirteisiin, ei laajemman aineiston käsitteleminen pro gradu -tutkimuksessa olisi ollut mahdollista. Vaikka aineiston perusteella ei voi tehdä kovin vahvoja yleistyksiä, koen kuitenkin, että tämänkin laajuisen aineiston avulla onnistun valottamaan niitä ominaisuuksia, jotka ovat tyypillisiä Wikipedia-artikkelille tekstilajina.

Kun tutkimuskohteena on tiettyyn lajiin kuuluvia tekstejä, on aineiston rajaus jo oletus siitä, mitkä tekstit kuuluvat samaan tekstilajiin (Shore & Mäntynen 2006: 11). Valitsemalla satunnaisen otannan artikkeleita Wikipediasta osoitan samalla pitäväni tekstejä lähtökohtaisesti samaan tekstilajiin kuuluvina. Julkaisualusta on siis jo yksi keino määrittellä artikkeleiden edustavan samaa tekstilajia.

Wikipedian hakemisto on jaettu kahdeksaan pääluokkaan: *Kulttuuri ja taide*, *Politiikka ja yhteiskunta*, *Uskomukset*, *Humanistiset tieteet*, *Matematiikka ja luonnontieteet*, *Tekniikka ja sovelletut tieteet*, *Vapaa-aika ja viihde* sekä *Muut*. Nämä jakautuvat edelleen alaluokkiin, esimerkiksi pääluokka *Kulttuuri ja taide* jakautuu luokkiin *Arkkitehtuuri*, *Elokuva*, *Kirjallisuus*, *Kuvataide*, *Musiikki* ja *Populaarikulttuuri*. Olen valinnut artikkeleita satunnaisotannalla kaikista Wikipedian pääluokista ja monesta eri alaluokasta, jotta teksteissä käsiteltävät aiheet olisivat mahdollisimman erilaisia ja kuvastaisivat Wikipedia-artikkelien laajaa temaattista skaalaa.

Aineistoni artikkelit on poimittu Wikipedia-sivustolta tammikuussa 2015. Wikipedian luonteesta johtuen artikkeleita on todennäköisesti muokattu sen jälkeen, mutta tutkimuksessani käsittelen artikkeleita siinä muodossa, kun ne ovat sivustolla esiintyneet tammikuussa 2015. Artikkelit vaihtelevat pituudeltaan huomattavasti. Aineiston pisin, Bahai-uskosta kertova, artikkeli on 2928 sanaa pitkä ja sisältää 165 virkettä, kun taas lyhin, biologisesta antropologiasta kertova, artikkeli on vain 80 sanan pituinen ja koostuu ainoastaan neljästä virkkeestä. Keskimäärin artikkeleissa on 43,7 virkettä. Osa artikkeleista on jaettu lukuihin (ja mahdollisesti alalukuihin), ja niihin on luotu sisällysluettelo, osa taas koostuu vain johdantolukuun kirjoitetusta tekstiosasta.

Aineistoni käsittää seuraavat satunnaisotannalla Wikipedian eri pääkategorioista poimitut artikkelit: *Bahai-usko*, *Biologinen antropologia*, *Brasilia*, *Daniel Craig*, *Ensimmäinen ristiretki Suomeen*, *Fani*, *Fysioterapia*, *IKEA*, *Juutinrauma*, *Konsertto*, *Kuka pelkää mustaa miestä*, *Leidenin pullo*, *Liikenne*, *Loch Nessin hirviö*, *Magmakivi*, *Merkurius*, *Merlin (televisiosarja)*, *Metonymia*, *Mikrobiologia*, *Naisten äänioikeus*, *Näkymätön käsi*, *Pakina*, *Poro*, *Sammon taonta*, *Shetlanninlammaskoira*, *Siipi*, *Särkänniemi*, *Unisieppari*, *Vares – Sheriffi* ja *Yleinen kielitutkinto*.

Jo artikkelien otsikoita tarkastellessa voi havaita, että Wikipedian temaattinen kenttä on laaja ja poikkeaa jonkin verran painettujen sanakirjojen sisällöstä. Aineisto sisältää esimerkiksi näyttelijä Daniel Craigista, televisiosarja Merlinistä ja Vares – Sheriffi -elokuvasta kertovat, aihepiiriltään melko populaariset artikkelit, joiden kaltaisia ei tyypillisesti esiinny painetuissa yleistietosanakirjoissa. Myös Tereszkiewiczin mukaan Wikipediassa sisällön temaattinen skaala on painettujen sanakirjojen skaalaa laajempi (2010: 170).

Wikipedia-artikkelien ollessa tarkastelun kohteena täytyy huomata, että tekstit ovat erilaisissa vaiheissa riippuen siitä, koska ne on luotu ja miten niitä on luomisen jälkeen muokattu. Tämä koskee luonnollisesti sekä artikkeleiden sisältämää tietoa että niiden kieliasua. Mo-

lemmat elävät jatkuvan muokkaamisen myötä. En siis tutki joukkoa valmiiksi kirjoitettuja ja sitten julkaistuja artikkeleita, vaan aineistoni artikkelit ovat kirjoittamisen eri vaiheissa.

2 TIETOSANAKIRJAT

2.1 Yleisesti tietosanakirjoista

Ensyklopedioiden perinne on pitkä. Vaikka tietosanakirjojen muoto on luonnollisesti muuttunut vuosisatojen varrella, ajatus kaiken tietämyksen käsittävistä kirjoista syntyi jo tuhansia vuosia sitten. Vuosina 23–79 jKr. eläneen Plinius vanhemman luoman ensimmäisen länsimaisen ensyklopedian ja nykyisin englanninkielisen maailman suurimman tietosanakirjan, *Encyclopedia Britannican*, väliin mahtuu valtava määrä erilaisia yrityksiä tallettaa tietämys kansien väliin. Tiedon ja sen hallitsemisen oivallettiin jo varhaisessa vaiheessa olevan yhteiskunnan avainkysymyksiä: tieto ei useinkaan ollut neutraalia, vaan sillä ajettiin jonkun asiaa. (Häkli 1982: 14, 31, 40.)

Tietosanakirjat jaetaan tavallisesti yleisteoksiin ja johonkin rajattuun aiheeseen keskittyneisiin ensyklopedioihin. Perinteisesti tietosanakirjoja on kirjoittanut ryhmä asiantuntijoita, ja kustannustoimittajat puolestaan ovat valvoneet tietosanakirjan asetelua sekä tiedon valintaa ja järjestämistä. (Tereszkiewicz 2010: 33.) Esimerkiksi 2000-luvulla Suomessa ilmestynyt *Factum*-tietosanakirjasarja pohjautuu Werner Söderström osakeyhtiön (WSOY) tietosanakirjatoimituksen laajaan tietokantaan, johon ovat kirjoittaneet tekstejä sadat tiedemiehet. Tämän lisäksi kymmenet asiantuntijat ovat osallistuneet kyseisen tietokannan ajan tasalla pitämiseen. (*Factum* 2003.)

Kielitoimiston sanakirja määrittelee tietosanakirjan hakusanojen mukaan järjestetyksi kokoelmaksi lyhyehköjä kirjoituksia tietyiltä tiedonaloilta (KS 2012: s.v. *tietosanakirja*). *Facta*-tietosanakirja huomioi lisäksi, että tietosanakirja on hakuteos, jolle on tunnusomaista selitettyjen hakusanojen suuri lukumäärä (*Facta* 2006: s.v. *tietosanakirja*). *Factum* uuden tietosanakirjan avauspuheessa puolestaan tiivistetään tietosanakirjan tarkoitus seuraavasti: ”Tietosanakirjan tulee sisältää kaikki olennainen, yleissivistykseen kuuluva perustietämys, mutta samalla sen tulee tarjota lukijalle uusin tieto ja myös mahdollisuus jäsentää ja taustoittaa päivittäistä, joka taholta tulvivaa, usein vaikeasti hahmotettavaa informaatiota.” (*Factum* 2003.) Tätä tehtävää ovat Suomessa 1900- ja 2000-luvuilla toteuttaneet kustantamoista erityisesti Otava ja WSOY. Ne ovat vuosien varrella julkaisseet laajasti sekä yleistietosanakirjoja että erikoistietosanakirjoja. (Tarkemmin tietosanakirjojen kehityksestä Suomessa ks. Häkli 1982.)

Factumin avauspuheessa mainittu uusimman tiedon tarjoaminen lukijalle on kuitenkin osoittautunut painetuille tietosanakirjoille ylivoimaiseksi haasteeksi nykyisessä internetin ja

teknologian ympärille keskittyneessä maailmassa. Tietosanakirjojen kenttä on 2000-luvulla elänyt murrosvaihetta, sillä tietosanakirjat ovat paljolti siirtyneet verkkoon. Viime vuosina ei ole enää ilmestynyt päivitettyjä, painettuja suomenkielisiä tietosanakirjoja. WSOY:n Facta-tietosanakirjan uusin painos julkaistiin vuonna 2006. Weilin & Göösin Factum uusi tietosanakirja -sarjan kahdeksas ja viimeinen osa puolestaan julkaistiin vuonna 2005. Ruotsinkielisen Uppslagsverket Finlandin painetun version viimeinen osa ilmestyi vuonna 2007.

”Internet tappoi suomalaiset tietosanakirjat – myös netissä”, kirjoittaa Samuli Tiikkaja osuvasti Helsingin Sanomissa (Tiikkaja: 2012). Enemmän kuin perinteisten tietosanakirjojen verkkoversiot, ensyklopedioiden genreä ovat muuttaneet näiden rinnalle nousseet, vapaaseen tiedontuottamiseen perustuvat ensyklopediat. Tereszkiewicz määrittelee, että internetissä toimivat tietosanakirjat eroavat paperisista versioista erityisesti kustannusten, hinnan, julkaisutahdin, interaktiivisuuden ja lisäominaisuuksien osalta. Sen lisäksi, että uuden tiedon lisääminen artikkeleihin on helpompaa ja nopeampaa, voidaan niissä käyttää esimerkiksi kuvia, ääninäytteitä ja animaatioita. Artikkelien pituutta ei yleensä ole rajoitettu. Myös julkaisutahti on paljon nopeampi ja julkaisukanava interaktiivisempi kuin painetuissa tietosanakirjoissa. Lisäksi julkaiseminen on yleensä ilmaista, tosin internet-sivuston ylläpitämisellä on hintansa. (2010: 64.)

Tiedon etsimistapojen lisäksi internetissä toimivat ensyklopediat ovat vaikuttaneet ratkaisevasti myös ensyklopedisen tiedon tuottamistapoihin. Asiantuntijalähtöisyydestä on siirrytty avoimeen tiedon tuottamiseen, jossa kaikilla on mahdollisuus osallistua. Tässä kehityksessä Wikipedia on näytellyt merkittävää osaa.

2.2 Wikipediasta

”Wikipedia on tietosanakirja”, tiivistetään Wikipedian tehtävä yksinkertaisesti sen omalla sivustolla. Tarkemmin Wikipediaa kuvaillaan sivustolla seuraavasti:

Wikipedian tarkoitus on olla lukijalleen hyödyksi. Se yhdistää elementtejä yleisistä tietosanakirjoista, erikoistuneista tietosanakirjoista, [almanakoista](#) ja lehdistä. Wikipediassa [ei harjoiteta uutta tutkimusta](#). [Wikipedia ei ole](#) sekalainen kokoelma tietoa, [sanakirja](#), [triviakokoelma](#), keskustelualue, blogi, sanomalehti eikä kokoelma lähteitä.

Wikipedia on siis vapaa ja ilmainen verkossa toimiva tietosanakirja. Wikipedian perustivat internet-yrittäjä Jimmy Wales ja filosofi Larry Sanger vuonna 2001. Nykyisin se toimii 285 eri kielellä. Englanninkielisessä Wikipediassa oli yli neljä ja puoli miljoonaa artikkelia ja yli 2,6 miljoonaa sanaa joulukuussa 2014, mikä on yli 100 kertaa enemmän kuin suurimmassa

perinteisessä englanninkielisessä ensyklopediassa, Encyclopedia Britannicassa. Suomenkielisessä Wikipediassa puolestaan oli 363 244 artikkelia joulukuussa 2014. Joulukuun aikana suomenkielisen Wikipedian artikkeleita luettiin yli 77 miljoonaa kertaa ja niitä muokattiin 52 826 kertaa. Uusia artikkeleita luotiin keskimäärin 66 päivässä. (Wikipedia: Wikipedia 2015; Wikipedia:Viisi pilaria 2015; Wikipedia:Statistics 2015.)

Wikipedia eroaa perinteisesti kustannetuista tietosanakirjoista muutamalla merkittävällä tavalla. Wikipedian olennaisimmat piirteet, jotka samalla erottavat sen kustannetuista tietosanakirjoista voidaan määritellä sosiaalisen median piirteiden kautta. Niitä ovat muun muassa käyttäjälähtöisyys, vuorovaikutuksellisuus, muokattavuus ja multimediaisuus.

Perinteisistä tietosanakirjoista poiketen Wikipedian sisältö on tuotettu kokonaan vapaaehtoisten voimin. Kuka tahansa voi luoda uuden artikkelin sekä osallistua jo olemassa olevien artikkelien kirjoittamiseen ja muokkaamiseen. Wikipedia-artikkelien muokkaaminen on helppoa: muokkajan tarvitsee vain klikata *muokkaa*-painiketta, ja hän pääsee tekemään muutoksia sivulle. Osallistuminen ei vaadi edes käyttäjätunnuksen luomista Wikipediaan. Wikipedian suosion takana on osaltaan juuri Wiki-mekanismien helppo ja toimiva käyttö (Tereszkiewicz 2010: 67). Yhteistoiminnallinen tuottamistapa johtaa siihen, että kaikki käyttäjät ovat yhdessä vastuussa Wikipedian sisällöstä, jolloin muodostuu sosiaalinen luottamussysteemi (Emigh & Herring 2005: 2). Wikipedian toiminta perustuu oletukseen siitä, että käyttäjät toimivat sivuston parhaaksi eivätkä pyri vandalisoimaan sitä.

Wikipedia voidaan siis lukea osaksi sosiaalista mediaa. Lietsala ja Sirkkunen määrittävät sosiaalisen median kuvaavan verkkopalveluja, jotka saavat suurimman osan sisällöstä käyttäjiltään tai kokoavat sisällön muilta sivuilta. Sivustot pohjautuvat verkkoyhteisöihin ja niissä kuka tahansa voi olla sisällön tuottaja. (Lietsala & Sirkkunen 2008: 13, 18.) Erkkola (2009: 88) puolestaan määrittelee käsiteanalyysinsä perusteella sosiaalisen median vielä tarkemmin:

Sosiaalinen media on vuorovaikutteinen, teknologiasidonnainen ja rakenteinen prosessi, jossa verkottuneet yksilöt ja ryhmät rakentavat yhteisiä merkityksiä sisältöjen, yhteisöjen ja verkko-teknologioiden avulla vertais- ja käyttötuotannon kautta. Samalla sosiaalinen media on jälkiteollinen ilmiö, jolla on tuotanto- ja jakelurakenteen muutoksen takia vaikutuksia yhteiskuntaan, talouteen ja kulttuuriin.

Kun Wikipediaa katsotaan näiden määritelmien valossa, se näyttäytyy varsin erilaisena perinteisiin kustantajan kautta kulkeneisiin tietokirjoihin verrattuna. Perinteiselle medialle

vastakkaisesti käyttäjät ovat vastuussa sisällöstä: he rakentavat yhteisiä merkityksiä vertais-tuotannon kautta. Voidaan myös ajatella, että Wikipedialla (ja muilla internetissä vastaavan periaatteen mukaan toimivilla ensyklopedioilla) on ollut vaikutusta yhteiskuntaan, talouteen ja kulttuuriin laajemminkin. Ne ovat ensinnäkin syrjäyttäneet perinteiset, painetut tietosanakirjat. Lisäksi ne ovat vaikuttaneet – ja vaikuttavat – siihen, miten ylipäänsä näemme tiedon ja sen, kenellä on oikeus tuottaa tietoa.

Merkittävä ero painettuihin tietosanakirjoihin verrattuna on Wikipedian ajantasaisuus: vapaaehtoiset kirjoittajat päivittävät Wikipediaa jatkuvasti ja lisäävät sinne sisältöä. Laaja käyttäjäkunta mahdollistaa Wikipedian tiuhan päivittämisen. Lisäksi käyttäjien runsaus vaikuttaa Wikipedian temaattiseen skaalaan, joka on monipuolisempi kuin painetuissa tietosanakirjoissa (Tereszkiewicz 2010: 170). Käyttäjälähtöisyyteen liittyy olennaisesti vuoro-vaikutus. Sen lisäksi, että lukijat voivat antaa palautetta kirjoittajille ja kommentoida sisältöjä, Wikipediassa lukijoiden on mahdollista suoraan muokata artikkelia.

Yhteistuotanto on siis olennainen osa Wikipedian toimintaa. Wikipedia-artikkeleita tutkiessa täytyykin tiedostaa, että tarkastelee nimettömän kirjoittajajoukon tuotosta. On mahdollista tietää, kuinka moni ihminen ja missä vaiheessa on osallistunut tekstien työstämiseen. Kirjan muodossa julkaistu perinteinen tietosanakirja on painamisen jälkeen lopullisessa muodossaan toisin kuin Wikipedia-artikkeli, jonka sisältöä tarkistetaan, ajantasaistetaan ja laajennetaan jatkuvasti. Jotkut artikkelit ovat löytyneet sivustolta jo kauan, kun taas toiset on luotu vasta hiljattain, ja ne ovat siis prosessin alkuvaiheessa. Wikipedian prosessiluonne tulee esiin myös siinä, että artikkeleita ei missään vaiheessa julisteta valmiiksi, vaan niiden muokkaus- ja päivitystyö tulee mitä todennäköisimmin jatkumaan niin kauan kuin ne ovat olemassa.

Wikipedian vapaasta muokkaustavasta huolimatta sillä on kuitenkin käytäntöjä, joiden noudattamista kirjoittajilta edellytetään. Yhteisössä määritellyt säännöt on kuvattu sivustolla. Artikkelien pitää olla tyyliltään neutraaleja eikä niissä esiintyvä materiaali saa rikkoa tekijänoikeuksia. Sivustolta löytyy tyyliopas, johon on kerätty ohjeita siitä, miten Wikipedia-artikkeli luodaan. Ohjeita löytyy muun muassa artikkelin rakenteesta, artikkelin nimeämisestä ja kielioppiin liittyvistä asioista. (Wikipedia Ohje:Sisällys 2015.)

Sosiaalisen median ominaispiirteistä Wikipediaan liittyvät olennaisesti myös muokattavuus ja multimediaisuus. Artikkelien sisältöä on mahdollista muokata julkaisun jälkeen ja sen muuntelussa on monenlaisia mahdollisuuksia aina sisällön poistoon saakka. Multimediaisuus puolestaan näkyy esimerkiksi siinä, että sisällössä yhdistellään kirjoitettua tekstiä, videota ja kuvaa. (Erkkola 2009: 36–37.) Lisäksi Wikipedian kautta pääsee nopeasti ja helposti käsiksi

myös artikkelien ulkopuoliseen tietoon (Tereszkiewicz 2010: 171). Wikipediassa ei myöskään ole samanlaisia rajoituksia tilan suhteen kuin painetuissa tietosanakirjoissa.

Multimediaisuus liittyy olennaisesti myös Wikipedia-tekstin muotoon. Jokaisella tekstillä on aina muoto, jonka tutkimista ei voi sivuuttaa merkityksiä analysoidessa. Tekstissä on mukana kirjoitettu eli se on lineaarinen esitys, joka on tuotettu jollekin pinnalle. Siinä on muun muassa sanoja, välimerkkejä, otsikoita, kappaleita, kuvia ja kuvatekstejä. (Hiidenmaa 2000: 170–171.) Wikipedia-artikkeliä ei kuitenkaan voi enää kutsua pelkästään lineaariseksi esitykseksi. Se on syytä määritellä ennemminkin monilineaariseksi **hypertekstiksi** (Fritz 1999: 222). Anna Tereszkievicz näkeekin yhtenä Wikipedian keskeisenä erona painettuihin tietosanakirjoihin juuri sen, että se on päässyt eroon lineaarisuuden rajoituksista (2010: 171).

Fritz tuo esiin, että hyperteksti ja teksti eroavat toisistaan rakenteellisten ja käytännöllisten ominaisuuksiensa suhteen. Perinteisessä tekstissä kirjoittaja on tehnyt valintoja sen suhteen, mitkä tekstuaaliset elementit hän sijoittaa mihinkin kohtaan tekstiä, mutta lukijalle järjestys tulee annettuna. Hyperteksti sen sijaan on verkosto tekstuaalisia ominaisuuksia, joiden välillä lukija voi liikkua haluamassaan järjestyksessä. Kytkeviä eri aiheiden välillä kutsutaan linkeiksi: juuri ne luovat vaihtoehtoisia lukutapoja. (1999: 222–223.)

Myös Astrid Ensslin näkee hypertekstin epälineaarisenä. Hänen mukaansa hypertekstiä ei voi esimerkiksi tulostaa tuhoamatta sille tyypillistä rakennetta. Olennaista hypertekstille on se, että se antaa lukijalle mahdollisuuden valita oman polkunsä lukemattomista mahdollisista vaihtoehdoista. (2007: 5.) Wikipedia-artikkelin voi edelleen halutessaan lukea myös lineaarisesti, mutta ero perinteiseen tekstiin ilmenee muun muassa siinä, että lisätietoa löytyy linkkien takaa yleensä läpi koko artikkelin, jolloin lukijan on mahdollista valita oma polkunsä eri artikkelien välillä. Mikäli Wikipedia-artikkelin lähteenä on käytetty internetistä löytyvää tiedonlähdettä, pääsee lukija linkin kautta suoraan alkuperäiseen lähteeseen. Tulostettu Wikipedia-artikkeli on sekin edelleen kokonainen teksti, mutta se ei anna lukijalle samalla tavalla valinnanvaraa kuin artikkelin lukeminen suoraan Wikipedia-sivustolta.

Tereszkiewicz näkee olennaisimpana erona perinteisten genrejen ja internet-genrejen välillä toiminnallisuuden (functionality) nousemisen keskiöön. Tekstilajeja on perinteisesti määritelty tarkastelemalla niiden muotoa, sisältöä ja funktioita, mutta internet-genren ollessa kysymyksessä täytyy näiden lisäksi huomioida myös toiminnallisuus. (2010: 22.) Santini tuo esille, että internet-genret ovat muodoltaan internet-sivuja. Nämä sivut ovat tyypillisesti hyvin monimuotoisia: niissä kaikki osat eivät aina näytä kuuluvan yhteen, mutta siitä huolimatta ne muodostavat kokonaisuuden, joka ei tosin välttämättä ole lineaarinen. Hyperteksti sallii monenlaisten geneeristen rakenteiden ja muotojen käytön yhdellä sivulla. (Santini 2007: 2.)

Erkkola näkee sosiaalisen median ominaispiirteeksi myös henkilökohtaisuuden: sisältötuotannossa omat mielipiteet ja näkökulmat tuodaan usein esiin (Erkkola 2009: 27). Tämä sosiaaliselle medialle tyypillinen piirre on luonnollisesti vastakkainen tietokirjatekstin kanssa, jonka pitäisi sisältää neutraalia ja puolueetonta tietoa. Tereszkiewicz tuo esiin myös muita Wikipediaan liittyviä mahdollisesti haitallisia piirteitä. Näitä ovat esimerkiksi artikkelien vaihteleva taso, väärä informaatio, vandalismin mahdollisuus ja mahdollisuus tekijänoikeuksien rikkomiseen. Haitallisista piirteistä huolimatta sosiaalisessa luottamussysteemissä toimiva Wikipedia on sosiaalisesti tunnistettu ja laajalle levinnyt, mikä viittaa siihen, että sillä on tiedon tuottamisessa ja etsimisessä tärkeä rooli. (Tereszkiewicz 2010: 171.)

2.3 Aikaisempi tutkimus

Wikipediasta on tehty jonkin verran aikaisempaa tutkimusta. Tutkimus on kuitenkin suurimaksi osaksi keskittynyt englanninkieliseen Wikipediaan. Erityisesti Wikipedian luotettavuutta ja käyttäjäkuntaa on pyritty kartoittamaan, kielitieteellistä tutkimusta sen sijaan on tehty vähemmän. Suomessa Wikipediasta ei ole tehty tekstilajitutkimusta, mutta Anna Tereszkiewicz on analysoinut väitöskirjassaan *Genre analysis of online encyclopedias: the case of Wikipedia* englanninkielistä Wikipediaa nimenomaan tekstilajin näkökulmasta.

Tereszkiewicz päättelee analyysinsä perusteella, että ensyklopedia genrenä on liikkunut voimakkaasti eteenpäin perinteisistä, painetuista tietosanakirjoista. Internetin tietosanakirjat – erityisesti Wikipedia – ovat kompleksisia, dynaamisia ja audio-visuaalisia. Painettuun tietosanakirjaan verrattuna muutoksia on tapahtunut erityisesti tarkoituksen (function), sisällön (content), muodon (form), diskursiivisten ominaisuuksien (discourse properties) ja toiminnallisuuden (functionality) suhteen. Wikipedian funktio ei ole pelkästään antaa tietoa, vaan se toimii myös ilmaisukanavana. Sen sisältöön vaikuttaa tuottaja- ja vastaanottajaroolien muuttuminen, mikä puolestaan on johtanut siihen, että sisällön temaattinen skaala on laajentunut. Artikkelit ovat ulkoasultaan ja tyyllillisesti perinteisestä tietosanakirjamallista poikkeavia ja niiden välillä on enemmän kytköksiä. (Tereszkiewicz 2010: 170.)

Eroavaisuuksista huolimatta Wikipedialla on kuitenkin paljon yhteistä painettujen edeltäjiensä kanssa. Se ei ole täysin uusi tekstilaji, mutta se on selvästi enemmän kuin vain vanha genre siirrettynä uuteen ympäristöön. Tereszkiewicz kuvailee sitä jonkinlaiseksi välimuodoksi: konventionaaliset tekstilajin piirteet eivät ole kadonneet, mutta ne ovat muuttuneet merkittävästi internet-teknologian vaikutuksesta. Wikipedia on hyvä esimerkki siitä, miten genret muuttuvat sekä uuden tekniikan ja ympäristön vaikutuksesta että silloin, kun yksilöt kyseen-

alaistavat olemassa olevia konventioita, kuten tässä tapauksessa sen, kuka ensyklopedista tietoa voi tuottaa. (Tereszkiewicz 2010: 170.)

Myös William Emigh ja Susan Herring ovat tutkineet ensyklopedioita tekstilajin näkökulmasta. Artikkelissaan *Collaborative Authoring on the Web: A Genre Analysis of Online Encyclopedias* he vertasivat Wikipediaa ja niin ikään internetissä toimivaa, vapaasti muokattavissa olevaa Everything2-sivustoa sekä perinteistä ensyklopediaa, Columbia Encyclopediaa, keskenään. He analysoivat muodollisuutta kvantitatiivisesti, mutta ottivat huomioon, että olennaista oli myös tarkastella, miten artikkelit oli järjestetty ja minkälaista tietoa niistä löytyi. Tarkastelun ulkopuolelle jäivät kuitenkin tiedon oikeellisuus ja artikkelien teemat. (Emigh & Herring 2005: 5, 9.)

Emigh ja Herring saivat selville, että Everything2 oli kieliasultaan kahta muuta epämuodollisempi, mutta Wikipedia ja Columbia Encyclopedia eivät kielen muodollisuuden suhteen juuri eronneet toisistaan. Wikipedian ja Everything2-sivuston muodollisista eroavaisuuksista huolimatta tutkijat päätyivät kuitenkin luokittelemaan molemmat ensyklopediat saman genren alle, koska rakenteelliset ja funktionaaliset piirteet olivat molemmilla internet-tietosanakirjoilla yhteisiä. Molempien tavoite oli esittää yleistä tietoa, ne käyttivät hyperlinkkejä, olivat luettavissa verkossa ja niistä pystyttiin myös etsimään tietoa. Laaja vapaaehtois-ten, ei-asiantuntijoiden joukko on luonut molemmat, ja molemmista internet-käyttäjät etsivät tietoa laajalla skaalalla. (Emigh & Herring 2005: 5, 9.)

Wikipedian luotettavuutta on tutkittu vertaamalla sen artikkeleita perinteisessä painetussa tietosanakirjassa julkaistuihin. Nature-lehti vertasi vuonna 2005 perinteisen ensyklopedia Britannican artikkeleita Wikipediaan ja sai yllättäen selville, että molemmat tietosanakirjat olivat lähes samalla viivalla virheettömyyden suhteen, Britannica vain hiukan Wikipediaa edellä. (Baytiyehin ja Pfaffmanin 2010: 128 mukaan Giles 2005.)

Suomessa Helsingin Sanomat teetti vuonna 2013 tutkimuksen, jossa valittiin 134 Wikipedia-artikkelia eri aihepiireistä ja etsittiin kullekin artikkelille asiantuntija arvioimaan sen luotettavuutta. Asiantuntija (yleensä professori tai yliopistotutkija) arvioi artikkelia kuudella mittarilla: virheettömyys, kattavuus ja tasapainoisuus, lähteitys, tuoreus, neutraalius sekä selkeys. Tuloksista selvisi, että suomenkielisen Wikipedian artikkelit ovat (tiedoiltaan) pääosin virheettömiä. Laatu kuitenkin vaihteli enemmän kuin painetussa tietosanakirjassa: joukkoon mahtui myös joitakin artikkeleita, jotka sisälsivät erittäin runsaasti virheitä. Ote artikkeleissa oli pääosin neutraali, ja artikkelit saivat kohtuullisia pisteitä tuoreudesta ja selkeydestä. Puutteita sen sijaan löytyi runsaasti lähteiden käytössä: 38 prosenttia artikkeleista sai sen suhteen arvosanan välttävä tai heikko. (Koistinen 2013.)

Wikipediaa on lisäksi tutkittu sekä Wikipediaan kirjoittavien että sitä lukevien näkökulmasta. Hoda Baytiyeh ja Jay Pfaffman selvittivät 155 aktiivisen ylläpitäjän motivaatiota Wikipediaan kirjoittamiseen. Tärkeimmäksi motivaattoriksi nousi halu oppia ja toiseksi tärkeimmäksi julkisen aineiston luominen. Ylläpitäjät tunsivat vahvaa yhteenkuuluvuutta Wikipediaan ja identifioivat sen yhteisöksi, jolla on tärkeä rooli. He kokivat tärkeäksi laadukkaan tietosanakirjan olemassaolon sekä hyödyllisen ja kiinnostavan tiedon jakamisen koko maailmalle. (2010: 135–138.)

Isto Huvila puolestaan on tutkinut sitä, kuinka Wikipediaan kirjoittavat etsivät tietoa ja millä tavoin he osallistuvat Wikipedian sisällön luomiseen. Tutkimuksen perusteella Wikipediaan kirjoittava joukko näyttää olevan hyvin heterogeeninen. Osa kirjoittajista etsii tietoa nimenomaan Wikipediaa varten, toiset taas kirjoittavat etupäässä omista kiinnostuksen kohteistaan tai työhönsä liittyvistä asioista. Jotkut keskittyvät uuden tiedon tuottamiseen sijaan lähinnä korjailemaan kieliopillisia seikkoja ja epäloogisuuksia artikkeleista sekä kääntämään tietoa erikielisestä Wikipediasta toiseen. (Huovila 2010.)

Maiju Ihalainen on tutkinut Wikipedian laatua opiskelijoiden näkökulmasta Pro gradu -tutkielmassaan *Vikaa Wikipediassa? – Opiskelijoiden käsitykset Wikipedian laadusta*. Hänen kyselynsä vastasi 67 opiskelijaa, joista 46 prosenttia kertoi käyttävänsä Wikipediaa viikoittain, 27 prosenttia päivittäin ja 12 prosenttia useita kertoja päivässä. Luvut tukevat käsitystä Wikipedian suosiosta. Opiskelijat kokivat Wikipedian laadun, saatavuuden, tarkkuuden, ajantasaisuuden, tiiviyden, relevanssin sekä selkeyden varsin hyväksi. Kattavuuden, objektiivisuuden ja uskottavuuden suhteen opiskelijoiden käsitykset sen sijaan jäivät neutraaleiksi. Omiin kokemuksiin pohjautuen Wikipediaa pidettiin kuitenkin usein paikkansapitävänä ja ajantasaisena ja sen koettiin sisältävän usein käyttökelpoista tietoa. (2013: 29–30, 40.)

Tutkimustulokset Wikipedian laadusta näyttävät olevan suurimmaksi osaksi positiivisia. Niitä tarkasteltaessa on kuitenkin otettava huomioon, että ne on pääosin tehty englanninkielisestä Wikipediasta, jolla on väistämättä suomenkielistä Wikipediaa laajempi ja aktiivisempi käyttäjäkunta. Dariusz Jemielniak huomauttaa, että englanninkielisen Wikipedian artikkelit ovat usein paremmin kirjoitettuja ja sisältävät enemmän lähteitä kuin vastaavat artikkelit muunkielisissä Wikipedioissa (Jemielniak 2014: 12). Viégas, Wattenberg ja Dave saivat omassa tutkimuksessaan selville, että jo vuoden 2003 toukokuussa englanninkieliseen Wikipediaan lisätyt vandalismitapaukset huomattiin ja korjattiin jo muutaman minuutin sisällä (2004: 579).

3 SYSTEEMIS-FUNKTIONAALINEN TEORIA TEKSTILAJI-TUTKIMUKSESSA

3.1 Systemis-funktionaalista teoriasta

Nykyisen tekstilajitutkimuksen juuret ovat 1980-luvulla syntyneessä systemis-funktionaalissa teoriassa. 1900-luvun alkupuolella ja sen jälkeen kielitieteilijät ja kielestä kiinnostuneet antropologit, etnologit ja sosiologit kehittivät tilannesidonnaista kielen käytön tarkastelua ja siihen tarvittavia analyysivälineitä. Näiden tutkijoiden lähestymistavoista on muotoutunut useita tutkimustraditioita, joista yksi on systemis-funktionaalinen genre- ja rekisteriteoria. Se pohjautuu pääasiassa sosiaaliantropologi Bronislaw Malinowskin, kielitieteilijä J. R. Firthin ja Firthin oppilaan, kielitieteilijä M. A. K. Hallidayn ajatuksiin. (Ventola 2006: 96, 98–101.) Erityisesti Hallidayta voidaan pitää systemis-funktionaalisen kieliopin isänä (Luukka 2002: 91).

Systemis-funktionaalinen teoria on luonteeltaan avoin ja laajeneva, joten se tarjoaa mahdollisuuksia monenlaisiin sovelluksiin. Sen perusluonteeseen kuuluu valmius mukautua kielen muutoksiin. Erilaisista sovelluksista huolimatta teorian perusidea on kuitenkin säilynyt melko muuttumattomana, kenties siksi, että se ei ole tiukka teoria, vaan enemmän orientaatio kieleen ja kielen analysointiin. (Luukka 2002: 118.)

Systemis-funktionaalisen teorian kantavia ajatuksia ovat käsitys ihmisestä toimijana sosiaalisessa yhteisössä sekä käsitys kielestä yhtenä sosiaalisen toiminnan muodoista. Tutkimuskohteena on niin ihmisten merkityksenanto kuin kielisysteemin toimintatapojen selvittäminen. Kieli nähdään ennen kaikkea toimintana ja tekemisenä ja olennaisena pidetään sitä, että kieltä käytetään kussakin kielenkäyttötilanteessa kyseiselle tilanteelle ominaisella tavalla. (Luukka 2002: 89; Shore 2012: 131.) Systemis-funktionaalissa teoriassa kielenkäytön keskiössä on yksilön sijaan yhteiskunta. Yksilön kielipotentiali reaalistuu niissä sosiaalisissa suhteissa, joihin hän ottaa yhteiskunnassa osaa. (Halliday 1978: 16.)

Halliday kuvaa kieltä funktionaaliseksi: kielen luonne hahmotetaan sen funktioiden kautta. Systemis-funktionaalissa teoriassa mielenkiinto keskittyy erityisesti siihen, mitä kielellä voidaan tehdä. Halliday ei kuitenkaan viittaa kielen funktionaalisuudella ainoastaan siihen, kuinka kieltä käytetään ja tekoihin, joita sillä tehdään, vaan pitää funktionaalisuutta koko kielisysteemin perustavanlaatuisena ominaisuutena. Kielen funktioita ei siis voi tarkas-

tella ikään kuin kielen ulkopuolelta tai yläpuolelta, vaan funktionaalisuus määrittää kielenkäyttöä kokonaisuudessaan. (1978: 16–17.)

Hallidayn mukaan kieli on sellaista kuin on juuri niiden funktioiden vuoksi, joita sillä on ihmisten elämässä. Ihmiset jakavat merkityksiä jatkuvasti. Merkitysten vaihto on luova prosessi, jossa kieli on symbolinen resurssi. Kun kieli nähdään tässä valossa, lauseiden ja muiden kielenainesten rakenne selitetään funktiosta käsin. (Halliday 1978: 4.) Systemis-funktionaalisessa genre- ja rekisteriteoriassa pidetään periaatteena sitä, että kielen muotojen ja rakenteiden käyttö ei ole sattumanvaraista, vaan siihen vaikuttavat yhteisön sosiaaliset konventiot (Ventola 2006: 97). Kieltä on siis syytä tulkita sosiokulttuurisessa kontekstissa. Pohjimmiltaan kieli ei koostu lauseista, vaan teksteistä ja kontekstista – siis siitä, että ihmiset kommunikoivat keskenään. Konteksti määrittää sitä, mitä sanotaan, ja sanottu määrittää edelleen kontekstia. Jokainen teksti – niin kirjoitettu kuin puhuttu – toimii aina jossakin kontekstissa. (Halliday 1978: 2–3; 1985a: xiii.)

Halliday selittää kieltä funktionaalisten perustehtävien kautta, joita hän kuvaa puhuu metafunktioilla. Hän erottaa semanttisesta systeemistä kolme funktionaalista komponenttia, jotka ovat läsnä kaikessa kielenkäytössä. Ne reaalistuvat erilaisina valintoina kielisysteemissä. (1978: 112–113.)

Ideationaalinen eli kuvaava metafunktio käsittää kielen jostakin kertovana (Halliday 1978: 112). Kieltä käytetään sen mukaan kuvaamaan ja luomaan sekä kokemusmaailmaamme että mielikuvitusmaailmaamme (Shore 2012: 146). Ideationaalisen metafunktion kautta kieli ilmentää kulttuurisia kokemuksia ja puhuja puolestaan kuvaa oman yksilöllisen kokemuksensa. Ideationaaliset merkitykset voi nähdä tekstin sisältönä. Ideationaalinen metafunktio voidaan vielä jakaa eksperimentaaliseen ja loogiseen. Eksperimentaaliset suhteet liittyvät kokemusten esittämiseen. Eksperimentaalinen metafunktio näkyy tekstissä muun muassa nimeämisinä ja transitiivisuutena. Transitiivisuudella viitataan maailmaa konstruoiviin valintoihin, jotka reaalistuvat tekstissä erilaisina lausetyyppeinä. Loogiset suhteet taas juontuvat ainoastaan epäsuorasti kokemusmaailmasta. Ne ilmenevät tekstissä esimerkiksi alistussuhteina. (Heikkinen 2001: 132; Halliday 1978: 112–113; Halliday & Hasan 1985: 26; 1976: 26.)

Interpersonaalinen eli aktiivinen metafunktio taas käsittää kielen toimimisena. Sen kautta puhuja ilmaisee asennoitumisensa sanomaansa ja pyrkii vaikuttamaan muiden asenteisiin ja käytökseen. Se antaa puhujalle mahdollisuuden ilmaista rooleja ja asettaa muita rooleihin, esimerkiksi kysyjä–vastaaja-rooleihin. Se reaalistuu tekstissä muun muassa modusten ja modaalisuuden kautta. Nämä kaksi metafunktiota kattavat kaiken kielenkäytön: ympäristön

ymmärtämisen (ideationaalinen) ja siinä toisten kanssa toimimisen (interpersonaalinen). (Halliday 1978: 112–113; Halliday 1985a: xiii.)

Kolmas metafunktio eli **tekstuaalinen** metafunktio pitää sisällään puhujan tekstinmuodostuspotentiaalin: se tekee kielestä relevanttia. Se myös mahdollistaa ideationaalisen ja interpersonaalisen metafunktion toteutumisen tekstin tasolla. Tekstuaalisen metafunktion kautta muodostuu kielen suhde sekä kielelliseen ympäristöön että tilanteiseen ympäristöön. (Halliday 1978: 112–113; Halliday & Hasan 1985: 23.)

Halliday jakaa tekstuaalisen metafunktion rakenteelliseen ja kohesiiviseen osaan. Rakenteellinen osa näkyy tekstissä temaattisena rakenteena ja informaatorakenteena. Kohesiivinen osa käsittää luonnollisesti koheesion alalajit eli referenssin, ellipsit, substituution, kytkennän ja sanastollisen koheesion. (Halliday 1978: 112–113; Halliday & Hasan 1985: 23, 334.)

Hallidayn mukaan systeemis-funktionaalisisessa teoriassa kielioppi katsotaan enemmän valintojen kieliopiksi kuin sääntöihin perustuvaksi kieliopiksi (Halliday 1978: 4). Systeemis-funktionaalinen teoria käsittää kielen luonteeltaan holistiseksi. Kieli on samaan aikaan merkityksenannon resurssi, merkityspotentiaali ja itseorganisoituva systeemi. (Luukka 2002: 99–100.) Vaikka Halliday erottaa toisistaan puhutun ja kirjoitetun kielen, hän ei pidä kirjoitettua kieltä puhuttua tärkeimpänä. Puhuttua ja kirjoitettua kieltä käytetään ilmaisemaan eri merkityksiä, ja ne esiintyvät eri konteksteissa. Ne kummatkin ovat kuitenkin ensisijaisesti kieltä, ja niillä voidaan ilmaista koko kielen merkityspotentiaali. (Halliday 1985b: 92–93.)

Systeemis-funktionaalisisessa teoriassa tekstilajeja tutkitaan ja analysoidaan kolmella eri tasolla: kielen, tekstilajin ja rekisterin (Ventola 2006: 96). Käyn seuraavissa alaluvuissa tarkemmin läpi edellä mainittuja käsitteitä. Avaan myös systeemis-funktionaalisen teorian perusyksikköä eli tekstiä. Lisäksi käsittelen koheesion lajeja sekä tema- ja informaatorakennetta.

3.2 Systeemis-funktionaalisen teorian käsitteitä

3.2.1 Kieli

Systeemis-funktionaalisisessa teoriassa kielen käsitetään muuttavan ihmisellä oleva käyttäytymis- ja merkityspotentiaali merkityspotentiaalin muotoon. Merkityspotentiaali sisältää kaikki ne mahdollisuudet, joita ihminen voi ilmaista, ja ne taas saavat ilmiänsä sanaston ja kieliopin tasolla. (Halliday 1978: 21–22.)

Se, miksi kieli on muotoutunut juuri sellaiseksi kuin on, käy Hallidayn mukaan järkeen, kun ajattelee niitä tehtäviä ja toimintoja, jotka kielen täytyy kaikissa ihmisten kulttuureissa

täyttää. Kielellä täytyy pystyä tulkitsemaan koko olemassaolomme, niin ulkoinen kuin sisäisenkin maailmamme. Sen pitää pystyä ilmaisemaan loogisia suhteita esimerkiksi konjunktioiden avulla. Kielellä täytyy myös kyetä ilmaisemaan osallistumista, rooleja, tunteita, toiveita ja asenteita. Lisäksi kielen täytyy pystyä suoriutumaan kaikista tehtävistä samanaikaisesti, ja sen mitä sanotaan, täytyy liittyä johonkin kontekstiin ja heijastua aikaisemmin puhuttuun. Nämä edellytykset ovat osoittaneet suunnan sille, miten kieli on kehittynyt, ja niiden pohjalta rakentuu Hallidayn käsitys kielen semanttisesta systeemistä. (Halliday 1978: 21–22.)

Kieltä käsitellään Hallidayn teoriassa sosio-semioottisesta näkökulmasta. Halliday liittää kielen vahvasti sosiaalisiin rakenteisiin. *Sosio* hänen näkökulmassaan viittaa sekä kulttuuriin yleensä että tarkemmin suhteeseen kielen ja sosiaalisten rakenteiden välillä. Esimerkiksi lapsi sosiaalistuu kieltä oppiessaan ja sen kautta samanaikaisesti kulttuurin arvojärjestelmiin ja käyttäytymismalleihin. (Halliday & Hasan 1985: 4; Halliday 1978: 23.)

Semiotiikka määritetään merkkisysteemien tutkimukseksi, ei merkkien tutkimukseksi. Tällä Halliday haluaa tehdä eron merkkien tutkimiseen pelkästään suhteessa toisiinsa ja painottaa sitä, että hänen teoriassaan semantiikka on pohjimmiltaan merkityksen tutkimista. Halliday määrittelee kulttuurin joukoksi semioottisia systeemejä, joista kieli on yksi mutta ei ainoa. (Halliday & Hasan 1985: 3–4.)

3.2.2 Teksti

Systeemis-funktionaalisen teorian perusyksikkö on teksti. Kieli ei koskaan toimi tyhjiössä, vaan se reaalistuu aina tekstinä eli aitona käyttötilanteena (Halliday & Hasan 1976: 142). Systeemis-funktionaalisisessa teoriassa kieltä pyritään ymmärtämään tekstejä tarkastelemalla. Halliday määrittää tekstin kieleksi, joka on funktionaalista, eli sillä on jokin merkitys tilanteessa. Teksti on aina semanttinen yksikkö. Se näyttää siltä kuin se olisi tehty sanoista ja lauseista, mutta todellisuudessa se muodostuu merkityksistä. Se on samaan aikaan sekä tuotos että prosessi. (Halliday 1978: 135–137.)

Teksti voi olla minkä mittainen tahansa, kunhan sillä on tilannekontekstissaan merkitys (1978: 135–137). Monet tekstit, kuten varoitukset, otsikot ja sloganit, ovat virkettäkin lyhyempiä. Toisaalta teksti voi olla hyvin pitkä, esimerkiksi romaani käsitetään tyypillisesti yhdeksi tekstiksi. Teksti voi olla puhuttu tai kirjoitettu, mitä tekstilajia tahansa ja sisältää minkä tahansa määrän osanottajia. (Halliday & Hasan 1976: 293–294.)

Tekstistä ei löydy samanlaisia kieliopillisia osia kuin lauseista ja virkkeistä. Ei ole olemassa kieliopillisia rajoitteita sen suhteen, mitkä virkkeet tekstissä voidaan asettaa peräkkäin. Tekstille on kuitenkin tyypillistä, ettei sen virkkeiden järjestystä voi muuttaa tuhoamatta tai

ainakaan muuttamatta sen merkitystä. Tekstillä on merkitys nimenomaan kokonaisena tekstinä. (Halliday & Hasan 1976: 28.) Tämän pohjalta voidaan tehdä se johtopäätös, että on olemassa jotakin, joka tekee tekstistä yhtenäisen kokonaisuuden.

Halliday käyttää käsitettä tekstiys (texture) kuvaamaan tekstin yhtenäisyyttä. Tekstillä on tekstiyyttä, kun se on merkityksellinen kokonaisuus, joka ottaa huomioon ympäristönsä. (Halliday & Hasan 1976: 1–2.) Usein puhutaan tekstiuden sijaan koherenssista, niin teen myös tässä tutkimuksessa (koherenssista tarkemmin alaluvussa 3.2.2). Tekstin sisällä jokainen virke on riippuvainen tilanteisesta ympäristöstään ja koheisiivisista suhteistaan tekstin muihin virkkeisiin. Teksti voidaan siis määritellä tilannekontekstin ja koheesion avulla. Katkelman kieltä pitää olla koherentti molempien suhteen, jotta se käsitetään tekstiksi. Lisäksi tekstiuden käsitteeseen sisältyvät muut tekstuaalisen komponentin osat eli temaattinen rakenne ja informaatorakenne. Olennainen on myös tekstin diskursiivinen rakenne. (Halliday & Hasan 1976: 23, 28, 299, 327.) Ehyellä tekstin informaatorakenne tukee tutun ja uuden tiedon annostelemista, ja teemarakenne on tekstin puheenaiheen kuljettamisen suhteen relevantti (Ranta 2007: 62).

Tilannekontekstin käsite on myös olennainen tekstin määrittämisessä. Tilanne on se ympäristö, jossa teksti herää henkiin. Tilannekontekstiin sisällytetään osallistujat sekä se, mikä on kyseiselle sosiaaliselle kontekstille tyypillistä, ja mitkä ovat kielen funktiot tämän kontekstin sisällä. (Halliday 1978: 109, 227.) Halliday kuvaa tilannekontekstia kolmen muuttujan kautta. **Ala** (field) kertoo, mihin sosiaaliseen toimintaan osallistujat ottavat osaa. Tässä sosiaalisessa toiminnassa kieli toimii oleellisena komponenttina. **Osallistujarooleilla** (tenor) viitataan luonnollisesti niihin, jotka toimintaan osallistuvat, ja heidän rooleihinsa sosiaalisessa toiminnassa. Huomio kiinnittyy muun muassa siihen, minkälaisia pysyviä ja väliaikaisia rooleja osallistujilla on. Osallistujarooleista puhuessaan Halliday erottelee toisistaan puhujaroolit ja sosiaaliset roolit. **Ilmenemismuodolla** (mode) Halliday kuvaa kielen roolia tilanteessa. Kieli voi olla esimerkiksi puhuttua tai kirjoitettua. Nämä muuttujat – ala, osallistujaroolit ja kielen ilmenemismuoto – auttavat tulkitsemaan tekstin sosiaalista kontekstia eli sitä ympäristöä, jossa merkityksiä vaihdetaan. Tilannekonteksti määrittää tekstin sisällön, sen, mistä tekstissä on kysymys. (Halliday & Hasan 1985: 12; 1976: 26.)

Koheesio on olennainen osa koherenssia ja niin ollen myös tekstin oleellinen elementti. Koheesioon palaan tarkemmin luvussa 3.2.3, mutta lyhyesti määriteltynä sillä viitataan tekstissä esiintyviin merkityssuhteisiin. Vaikka niiden esiintymismuodot saattavat vaihdella paljonkin, niitä esiintyy (lähes) jokaisessa tekstissä. Toisin kuin tilannekonteksti, koheesio ei kerro tekstin sisällöstä, vaan siitä, kuinka teksti on muodostettu semanttisesti. Sekä tilanne-

kontekstia että koheesiota siis tarvitaan, jotta jotakin katkelmaa voidaan kutsua tekstiksi. Ne eivät kuitenkaan automaattisesti esiinny yhdessä: katkelma kieltä voi teoriassa olla tilanteisesti koherentti, mutta siinä ei esiinny koheesiota tai vastaavasti se voi olla täynnä koheesiokeinoja, mutta ei kyseisessä tilanteessa johdonmukainen. (Halliday & Hasan 1976: 23, 26.)

Vaikka tekstin määritelmä on melko selkeä, ei todellisuudessa ole aina helppo erottaa, kuinka monta tekstiä on jossakin pätkässä kieltä ja missä kohtaa teksti vaihtuu uuteen tekstiin. Koska kirjoittaja tai puhuja käyttää koheesiota luomaan tekstiin koherenssia, ja koheesio vaikuttaa myös kirjoittajan tai puhujan tulkintaan tekstistä, voidaan koheesiota käyttää kriteerinä tekstin rajojen määrittämiseen. Jos jossakin lauseessa tai virkkeessä ei ole minkäänlaista koheesiota, tämä muutos saattaa osoittaa uuden tekstin alkua. Läheskään aina yksittäinen lause ei kuitenkaan aloita uutta tekstiä. (Halliday & Hasan 1976: 294–295.)

Teksti (ja niin ollen myös koheesio) on osa prosessia. Halliday kuvaa tekstiä aktiiviseksi kumppaniksi todellisuuden luomisen ja muuttamisen prosessissa. (Halliday 1985: 339.) Tämä nousee erityisen hyvin esiin juuri Wikipediassa, jossa monet eri kirjoittajat muokkaavat artikkeleita ja näin osallistuvat kukin vuorollaan todellisuuden luomisen prosessiin.

3.2.2 Koherenssi

Edellisessä luvussa todettiin, että tekstiys (eli koherenssi) kuvaa tekstin yhtenäisyyttä. Tekstillä on koherenssia, kun se on merkityksellinen kokonaisuus, joka ottaa huomioon ympäristönsä. Yhtenäinen teksti muodostuu koheesiosta, teema- ja informaatorakenteesta sekä tekstin diskursiivisesta rakenteesta. (Halliday & Hasan 1976: 1–2, 299, 327.) Saukkosen mukaan koherenssin toteutumisen tärkein edellytys on yhtenäinen tekstiteema. Hän huomauttaa, että ilmaustason koheesiokeinot eivät ole syy siihen, että teksti on koherentti, vaan ne ovat sen sijaan seuraus koherenssin toteutumisesta. (Saukkonen 2001: 79.)

Koherenssi on osoittautunut varsin monimutkaiseksi käsitteeksi määritellä. Bublitz tuo esiin, että koherenssi koetaan edelleen käsitteeksi, jota ei täysin ymmärretä ja jonka olemuksesta kiistellään (1999: 1). Fritz käsittää koherenssin tekstintuottamisen tärkeimmäksi ohjenuoraksi ja tekstien ymmärtämisen perustaksi. Tekstin ymmärtämiseen sisältyy se, että näkee relevanttien sisäisten ja ulkoisten tekstuaalisten elementtien yhteydet. Vahvasti koherentti teksti on koherentti funktionaalisesti, topiikillisesti ja tiedollisesti, ja koherenssia ilmennetään siinä kohesiivisilla sidoksilla. (1999: 21–22.) Ventola puolestaan määrittelee koherenssin tilanteisesti päätellyksi suhteeksi, joka kytkeytyy tekstien kontekstien tulkintaan. Hänen mukaansa tekstuaalisesta koheesion ja koherenssin kuvauksesta pitäisi siirtyä eteenpäin sen tutkimukseen, kuinka tekstit kytkeytyvät semanttisesti toisiinsa. (1999: 101.)

Koherenssin käsittelyssä on liikuttu (muodollisesti kuvattun) koheesion ja (semanttisen) yhtenäisyyden kentiltä kohti käyttäjälähtöisempää ja kontekstipohjaisempaa tulkintatapaa. Bublitz näkee koherenssin kontekstista riippuvaisena, lukija- tai kuulijalähtöisenä ja ymmärrykseen perustuvana. Se ei siis ole niinkään riippuvainen tekstistä itsestään, vaan ennemminkin tulkitsijan ymmärryksestä tekstistä. Tämän vuoksi yhden tekstin koherenssilla voi olla monta tulkintaa riippuen siitä, kuka sitä tulkitsee. (Bublitz 1999: 1–2.)

Koherenssi voidaan nähdä prosessina, johon vaikuttavat kaikki kommunikointiin liittyvät asiat, kuten prosodinen vaihtelu ja tekstin järjestäytyminen. Se on riippuvainen tilanteesta ja tekstilajista tai -tyypistä. Kun koherenssia tarkastellaan prosessina, se nähdään tilanteisena ja väliaikaisena: tulkitsija muokkaa tulkintaansa aina uuden tiedon perusteella. Koherenssi voidaan käsittää myös yhteistoiminnalliseksi, koska se riippuu sekä kirjoittajan että lukijan halusta havaita koherenssia. (Bublitz 1999: 2–3.)

Huolimatta siitä, että koherenssia voi tulkita hyvin monella eri tavalla, pidän tässä tutkimuksessa lähtökohtanani Hallidayn ja Hasanin määritelmää ja pyrin analysoimaan koherenssia ensisijaisesti koheesion, teema- ja informaatorakenteen sekä tilannekontekstin kautta. Huomioin kuitenkin, että Wikipedia-artikkelin kaltaisessa tekstissä ei voi unohtaa vastaanottajan vaikutusta koherenssin luomiseen. Esimerkiksi artikkelien linkkien luoma vapaus valita lukemansa järjestys johtaa siihen, että lukija päätyy väistämättä muodostamaan käsitystään koherenssista prosessinomaisesti.

3.2.3 Koheesio

3.2.3.1 Koheesiosta yleisesti

Koheesiolla tarkoitetaan tekstissä esiintyviä merkityssuhteita. Koheesio luo koherenssia tekstiin ja on siis olennainen osa tekstin muodostumista. (Halliday & Hasan 1976: 10, 298.) Jotta teksti olisi koherentti kokonaisuus, jotta se ylipäätään olisi teksti, siitä täytyy löytyä koheesiota (Halliday 1985a: 339). Teoriassa voidaan väittää, että joissakin tapauksissa teksti voi olla koherentti ilman koheesiotakin, mutta tämänkaltaisia tekstejä ei käytännössä löydy. Mitä pidempi teksti on, sitä väistämättömämmin siinä esiintyy koheesiota. (Tanskanen 2006: 17.)

Koheesion voidaan tulkita käsittävän ne semanttiset resurssit, joiden avulla virke tai lause on mahdollista kytkeä johonkin aikaisemmin mainittuun. Koheesiota ilmenee, kun jonkin tekstin elementin tulkinta on riippuvainen tekstin toisesta elementistä. Kun elementtien välille muodostuu tällainen kohesiivinen suhde, se luo tekstiin yhtenäisyyttä. Yksittäinen ele-

menti ei synnytä koheesiota, vaan kysymys on aina useamman kielellisen elementin välille muodostuvasta semanttisesta suhteesta. (Halliday & Hasan 1976: 4–5, 10.)

Koheesio mahdollistaa eripituisten tekstinainesten kytkemisen toisiinsa riippumatta siitä, missä kohtaa tekstiä ne sijaitsevat. Koheesiota esiintyy siis sekä virkkeen eri osien välillä että virkkeiden välillä (Halliday 1985a: 310, 312). Koska koheesio on semanttinen ilmiö, ei virkkeiden kieliopillisen rakenteen tarkastelu ole sen kannalta olennaista. Koheesiota luovien elementtien väliset suhteet ovat samoja riippumatta siitä, esiintyvätkö ne samassa virkkeessä vai kauempana toisistaan. (Halliday & Hasan 1976: 9.)

Kuten alaluvussa 3.2.2 kävi ilmi, teksti on semanttinen kielenkäytön yksikkö, joten sitä ei voi käsitellä samojen kieliopillisten kategorioiden mukaan kuin yksittäisiä lauseita ja virkeitä. Lauseilla on omat kieliopilliset sääntönsä, joiden mukaan niiden osat linkittyvät toisiinsa, mutta virkkeiden välillä koheesio on ainut elementti, joka kytkee virkkeet kokonaisuudeksi ja luo niistä yhtenäisen tekstin. Tämän vuoksi koheesio on virkkeiden välillä suurempi kuin virkkeiden sisällä. (Halliday 1985a: 312.) Hallidayn ja Hasanin mallia mukailen keskityn tässä tutkimuksessa tarkastelemaan erityisesti virkerajat ylittävää koheesiota.

Lähes kaikista teksteistä löytyy koheesiota, olennaista onkin, että millaista ja kuinka paljon. Tyypillisesti tekstin jokaisessa lauseessa ensimmäistä lukuun ottamatta esiintyy jonkinlaista koheesiota. Tavallisesti kappaleiden sisällä koheesiota esiintyy enemmän kuin kappaleiden välillä. (Halliday & Hasan 1976: 23, 293, 296–297.)

Koheesiiviset suhteet ovat joko **anaforisia** tai **kataforisia**. Anaforisessa sidoksessa tarkoite, johon viitataan, on mainittu aikaisemmin tekstissä. Kataforisessa sidoksessa taas tarkoite on vielä edessäpäin. (Halliday & Hasan 1976: 17.) Kohesiiviset suhteet jaetaan myös **tekstiviitteisiin** (endoforisiin) ja **tekstin ulkopuolelle viittaaviin** (eksoforisiin). Tekstin ulkopuolelle viittaavat sidokset vaativat tulkintaansa tilannekontekstia, ne eivät siis rakenna tekstin koheesiota tekstiviitteisten sidosten tavoin. (Halliday & Hasan 1976: 18.) Tämän vuoksi jätän ne tässä tutkimuksessa koheesioanalyysin ulkopuolelle.

Vaikka koheesio on ennen kaikkea semanttinen ilmiö, se reaalistuu tekstissä kieliopillis-semanttisina suhteina. Koheesiota ilmaistaan sekä kieliopillisesti että sanastollisesti, minkä vuoksi se voidaan jakaa kieliopilliseen ja sanastolliseen koheesioon. Kohesiiviset sidokset ovat luonteeltaan ensisijaisesti joko kieliopillisia tai sanastollisia, mutta kaikessa koheesiossa on kuitenkin aina väistämättä läsnä sekä merkityksenvalinta että koheesioilmeneminen sanoina ja rakenteina. (Halliday & Hasan 1976: 5–6, 303, 323.)

Muun muassa edellä mainitusta syystä koheesiokeinojen kategoriat eivät ole tiukkaraajaisia, vaan monet kohesiiviset muodot ovat rajatapauksia, ja ne voidaan tulkita useaan ko-

heesiokeinojen luokkaan. Koska koheesio on sekä semanttinen että kieliopillinen ilmiö, saatavat semanttiset ja kieliopilliset kriteerit joissakin tapauksissa johtaa eri tulkintoihin. (Halliday & Hasan 1976: 88.) Lisäksi monessa tapauksessa useampi kohesiivinen kytkös esiintyy tekstissä yhtäaikaaisesti (Halliday 1985a: 311). Tämän vuoksi kohesiivisten keinojen selväpiirteinen jaottelu luokkiin on haastavaa. Tässä tutkimuksessa käsittelem erikseen kieliopillisia ja sanastollisia koheesiokeinoja.

3.2.3.2 Kieliopillinen koheesio

Halliday ja Hasan jakavat kieliopillisen koheesion **referenssiin, substituutioon, ellipseihin ja kytkentään**. Referenssi, substituutio ja ellipsi ovat kaikki koheesiokeinoja, jotka pyrkivät identifioimaan jotakin viittaamalla edellä mainittuun. (Halliday & Hasan 1976: 144–145.) Substituutio ja ellipsi poikkeavat referenssistä siinä, että niissä yhteys muodostuu lingvististen kielenainesten, kuten sanojen, lauseiden ja lausekkeiden välille, kun taas referenssi on merkitysten välinen suhde (Halliday & Hasan 1976: 88–89). Kuitenkin niissä kaikissa se, mihin viitataan, on tavallisesti esiintynyt aiemmin tekstissä, eli ne vaativat tekstin tarkempaa lukemista ja pyrkivät vähentämään toisteisuutta.

Kytkentä puolestaan on rajatapaus kieliopillisen ja sanastollisen koheesion välissä. Leksikaalisista komponenteistaan huolimatta se on kuitenkin pääasiassa kieliopillinen sidoskeino, joten se voidaan lukea kuuluvaksi kieliopilliseen koheesioon. (Halliday & Hasan 1976: 6.) Tässä tutkimuksessa luokittelem sen kieliopillisen koheesion keinoksi.

Referenssi on semanttinen ilmiö. Se voidaan määritellä suhteeksi asioiden tai ilmiöiden välillä. Tyypillisesti se liittyy yhteen yksittäisiä elementtejä (esimerkiksi osallistujia, prosesseja ja tilanteita). (Halliday & Hasan 1976: 144–145; Halliday 1985a: 310.) Halliday ja Hasan jakavat sen kolmeen tyyppiin: persoonalliseen, demonstratiiviseen ja komparatiiviseen. Persoonana nähdään persoonallisessa referenssityypissä eräänlaisena roolina. Se reaalistuu tekstissä persoonapronominien kautta. Sen katsotaan spesifioivan jonkin asian rooli tilanteessa, jolloin se voi viitata myös objekteihin, ei ainoastaan henkilöihin. Näin ollen esimerkiksi demonstratiivipronomini *se* (it) kuuluu Hallidayn ja Hasanin luokittelussa persoonalliseen referenssityyppiin. Olen kuitenkin tutkimuksessani päätenyt luokittelemaan kaikki suomen demonstratiiviset pronominit demonstratiiviseen referenssityyppiin. Demonstratiivinen referenssi ilmenee demonstratiivipronominien ja proadverbien kautta. Proadverbit (esimerkiksi *siellä, silloin, sinne*) ovat adverbeiksi kangistuneita pronominikantaisia ilmauksia.

Esimerkin (1) *hän*-pronomini on esimerkki virkerajan ylittävästä persoonallisesta referenssistä. Demonstratiivista referenssiä ovat puolestaan pernaruttoon viittaava demonstratiivi-

pronomini *sen* sekä Kochin postulaatteihin viittaava demonstratiivipronomini *nämä*. Myös koko edeltävään kappaleeseen viittaava *tästä* voidaan lukea demonstratiiviseen referenssiin. Jälkimmäisen esimerkin (2) *siellä* on esimerkki demonstratiivisesta referenssistä, jota ilmaistaan proadverbin avulla.¹

1) Modernin mikrobiologian isä yhdessä Pasteurin kanssa oli saksalainen lääkäri Robert Koch (1843-1910). *Hän* tutki pernaruttoa ja onnistui vuonna 1876 ensimmäisen kerran kuvaamaan infektiotaudin vaiheet aukottomasti sekä mikrobin osuuden *sen* synnyssä. Tutkimiensa tautien perusteella *hän* laati Kochin postulaatit. *Nämä* ja Pasteurin löydökset ovat edelleen mikrobiologian kaksi kulmakiveä. *Tästä* alkoi mikrobiologian kultainen kausi, jonka aikana useat taudit todettiin mikrobien aiheuttamiksi.

2) Britanniassa generatio spontanea -teoria eli noin kymmenen vuotta kauemmin kuin mantee-reella. *Siellä* vasta John Tyndal todisti ilmassa olevan hiukkasia, jotka heijastavat valoa ja aiheuttavat mikrobikasvua.

Komparatiivinen referenssityyppi puolestaan viittaa epäsuorasti identiteettiin ja samankaltaisuuteen. Se ilmenee positiivi- ja komparatiivimuotoina sekä eräiden vertailua ilmaisevien adjektiivien kautta. Näitä ovat pronominikantaiset adjektiivit eli proadjektiivit. Ne eivät kuvaa ominaisuutta perinteisten adjektiivien tavoin, vaan viittaavat yhteydestä ilmenevään ominaisuuteen. Siksi niitä käytetään myös itsenäisinä eivätkä ne yleensä saa komparaatiomuotoja. Proadjektiiveihin kuuluvat esimerkiksi *sellainen, tällainen, tuollainen, kaikenlainen* ja *silloinen*. Lisäksi komparatiivista referenssiä voidaan ilmaista taipumattomilla adjektiiveilla, kuten *sama* ja *eri*. (Halliday & Hasan 1976: 37–39; Hakulinen & Karlsson 1995: 313, 320; VISK § 610; § 715; § 726.)

Referentti on samantarkoitteinen jonkin edellä esiintyneen lausekkeen kanssa vain silloin, kun se muodostaa lausekkeen yksin. Jos se esiintyy määritteenä, sen tarkoite käy ilmi sen edussanasta. Silloin sen tehtävä on osoittaa, että se kertaa jotakin edellä mainittua. (Hakulinen & Karlsson 1995: 316–317.) Tämänkaltaisen lausekkeen edussana on hyvin usein aiemmin esiintyneen tarkoitteen parafrasi (VISK § 1429). Tämä ilmenee muun muassa aineiston Särkänniemi-artikkelista. Esimerkissä (3) demonstratiivipronominia *tämä* käytetään määritteenä ilmaisemaan, että kyseessä on edellisessä virkkeessä mainittu tema-alue.

¹ Olen kursivoinut aineistoesimerkeistä olennaisia tekstiaineksia. Olen myös merkinnyt ellipsin toistamatonta vastinetta esimerkeissä symbolilla Ø. Esimerkeissä esiintyvät lihavoinnit ja alleviivaukset sen sijaan ovat aineiston tekstien elementtejä. Aineistossa esiintyy toisinaan myös kursivointia. Tällöin olen merkinnyt sen työhön erottaakseni alkuperäiset kursivoinnit itse kursivoimistani tekstiaineksista. Virkkeiden perässä toisinaan esiintyvät numerot ovat Wikipedian lähdeviitemerkintöjä.

3) 2012 kesäksi Ipanaario sai kokea jättiuidistuksen, kun sen alueelle rakennettiin maailman ensimmäinen Angry Birds Land teema-alue. *Tämän* uuden teema-alueen tieltä poistettiin Orlochlessin torni.

Lauseen sisällä esiintyvää referenssiä ei pidetä samalla tavalla tekstuaalisen sidoksisuuden luojana kuin virkkeiden välillä esiintyvää (Hakulinen ja Karlsson 1995: 312). Halliday ja Hasan painottavat tekstuaalisuuden tarkastelussa etupäässä virkkeiden välistä referenssiä. Hakulinen ja Karlsson kuitenkin mainitsevat, että koska ero kahden virkkeen ja konjunktiolla yhdyslauseeksi kytketyn jakson välillä ei ole kovin selvä, voidaan myös yhdyslauseita käsitellä sidoksisuutta tarkastellessa (Hakulinen ja Karlsson 1995: 312). Omassa analyysissäni olen päättänyt tekemään juuri näin.

Kieliopillista koheesiota ilmaistaan referenssin lisäksi substituution ja ellipsien avulla. Substituutio ja ellipsit voidaan nähdä käytännössä samana prosessina, joka ilmenee kahdessa eri muodossa. Substituutio on kielellinen ja sanastollinen koheesiokeino. (Halliday & Hasan 1976: 6, 145.) Siinä kysymys ei siis ole kahden aineksen samantarkoitteisuudesta, vaan tietyn leksikaalisen aineksen korvaamisesta toisella. Suomen kielessä nominaalisten elementtien yleisin substituutti on *sellainen*. (Hakulinen & Karlsson 1995: 320.) Esimerkki (4) havainnollistaa, kuinka Konsertto-artikkelissa nominaalilauseke *Mozartin itsensä kadenssi* on korvattu substituutilla *sellainen*.

4) Oman kadenssin soittaminen klassisessa konsertossa on nykyään harvinaista, ja usein esimerkiksi Mozartin konsertoissa esitetään *Mozartin itsensä kadenssia* tai jos *sellaista* ei ole saatavilla, niin esimerkiksi Beethovenin kadenssia.

Ellipsi puolestaan voidaan kuvata joksikin, joka jätetään sanomatta (Halliday & Hasan 1976: 89, 143). Kyseessä on koheesiokeino, jossa kahdella perättäisellä lauseella tai lauseenosalla on jokin yhteinen jäsen, joka toisessa osassa jätetään toistamatta (VISK § 1177). Ellipsin voi myös luokitella substituution alalajiksi. Substituutiosta sana korvataan jollakin, joka pitäisi poistaa, jos aikaisemmin mainittu tekstinaines palautettaisiin paikalleen. Ellipsissä sen sijaan ei ole korvaavaa sanaa, vaan poistetun tekstiaineksen kohdalla on aukko. (Halliday & Hasan 1976: 89, 143, 145.) Esimerkiksi seuraavassa Liikenne-artikkelin virkkeessä lauseiden yhteinen jäsen, demonstratiivipronomini *se*, on jälkimmäisessä lauseessa jätetty toistamatta. Olen merkinnyt toistamatonta vastinetta symbolilla \emptyset .

5) Se nopeutti maaliikennettä ja \emptyset teki mahdolliseksi suurempien kuormien kantamisen kuin mitä ihminen itse pystyi kantamaan.

Ellipsejä käytetään toisteisuuden välttämiseen. Niiden käyttö on tavallisinta ja monipuolisinta rinnakkaisissa lauseissa: jotakin sanaa, sanajaksoa tai sananosaa ei toisteta lauseessa, koska se on löydettävissä vastaavalta paikalta viereisestä lauseesta. Jotkin ellipsityypit ovat mahdollisia myös yli virkerajojen, mutta ne ovat selvästi harvinaisempia. Ellipsejä käytetään myös lauseita pienemmissä yksiköissä. Esimerkiksi yhdyssanan alusta tai lopusta voi puuttua sen leksikaalinen osa. Tämä elliptinen osa on korvattu yhdysviivalla, ja toistamaton vastine esiintyy toisaalla rinnastusrakenteessa. (VISK § 1177; § 437.) Tässä tutkimuksessa en kuitenkaan tutki ellipsejä lauseita pienemmissä yksiköissä, vaan keskityn tarkastelemaan sitä, miten ne rakentavat koheesiota erityisesti yhdyslauseissa.

Ellipsien päätyypeihin yhdyslauseissa lukeutuvat nominaaliset ellipsit ja verbiosan ellipsit. Nominaaliset ellipsit esiintyvät lauseissa, joissa on jokin yhteinen nominaalinen jäsen: yhteinen subjekti, objekti, predikaatiivi, adverbialitäydennys tai lausekkeen edussana. Verbiosan ellipsejä yhdistää sen sijaan yhteinen finiittiverbi, mutta mahdollisesti lisäksi esimerkiksi yhteinen subjekti. (VISK § 1177.)

Nominaalisessa ellipsisissä poistetaan rinnasteisesta lauseesta substantiivilauseke, kun se on toisen osalauseen tarkoitteen kanssa samantarkoitteinen. Substantiivilausekkeella tarkoitetaan lauseketta, jonka edussana on substantiivi, joissakin tapauksissa myös pronomini tai numeraali. (Hakulinen & Karlsson 1995: 324–326; VISK § 442.) Verbinosan ellipsisissä poistetaan yleensä toisesta osalauseesta finiittiverbi. Mahdollista on myös joissakin tapauksissa poistaa verbilausekkeen oikeanpuoleinen osa, jolloin jäljelle jää finiittiverbi. VP eli verbilauseke katsotaan yleensä sellaiseksi verbiedussanaiseksi kokonaisuudeksi, jossa yksi täydennyksistä, subjekti, ei ole mukana. (Hakulinen & Karlsson 1995: 325–326; VISK § 443.)

Ellipsejä on kuitenkin vaikea jaotella tyhjentävästi, sillä lauseiden yhteisiksi osiksi tulkittavia elementtejä on monesti useita. Nominaalinen ellipsi ja verbinosan ellipsi esiintyvät usein yhdessä. (Hakulinen & Karlsson 1995: 325–326; VISK § 1177.) Tyypillisintä ja vähiten tulkintavaikeuksia aiheuttavaa ellipsien käyttöä on yhteisen subjektin puuttuminen jälkilauseesta. Tämä näkyy muun muassa esimerkistä 5, jossa subjektina toimiva pronomini *se* puuttuu jälkimmäisestä lauseesta. Eri ellipsityypit ovat kuitenkin eri tekstityypeissä eri määrin tavallisia. (VISK § 1177; § 874.)

Neljäs kieliopillinen koheesiokeino, **kytkentä**, spesifioi tavan, jolla jokin kielenaines liitetään edellä mainittuun. Huomio kiinnittyy semanttisten suhteiden funktioihin, kun ne liittyvät toisiin lingvistisiin elementteihin. Kytkeä pohjautuu siihen perusajatukseen, että lauseiden ja virkkeiden välillä on tiettyjä systemaattisia suhteita, jotka liittävät niissä ilmaistut merkitykset toisiinsa. (Halliday & Hasan 1976: 227; Halliday 1978: 300.) Lauseiden välisten

merkityssuhteiden tulkintaan vaikuttavat tietysti myös lauseiden sisältö ja niiden keskinäinen järjestys, mutta suhteiden suuntaa ja rajoja osoitetaan kytkennällä (VISK § 1121). Kytkennässä lauseita ja virkkeitä liitetään toisiinsa **konjunktioiden** ja **konnektiivien** avulla (Hakulinen & Karlsson 1995: 296).

Kytkennässä tarkastellaan ensinnäkin konjuktiollisten kytkösten semantiikkaa. Rinnastuskonjunktioilla on kullakin omat merkityksensä, joiden pohjalta konjunktioilla toisiinsa kytkettyjen lauseiden suhdetta voidaan tulkita. Konjunktio ennustaa, että kytketyillä lauseilla on jotakin yhteistä, minkä vuoksi ne on asetettu peräkkäin. Jos virkkeessä ei ole konjunktio(i)ta, suhde jää spefioimatta, jolloin lauseiden leksikaalinen sisältö ja niiden keskinäinen järjestys vaikuttavat tulkinnan muodostumiseen. Kytkevä tehtävä on myös konnektiiveilla: nekin osoittavat lauseiden välisiä semanttisia suhteita. (VISK § 820; § 1090; § 1091.)

Konjunktiot ovat taipumattomia sanoja, joilla on kiinteä paikka lauseessa. Ne kytkevät toisiinsa lauseita, lausekkeita, sanoja ja sananosia ja luovat näin tehdessään osien välille tietyn syntaktisen ja semanttisen suhteen. Ne eivät kuitenkaan asetu osaksi kumpaakaan kytkettävää elementtiä, vaan niiden tehtävät liittyvät lauseiden ja lausekkeiden välisiin alistus- ja rinnastussuhteisiin. Konjunktioita voidaan luokitella semanttisesti sen mukaan, minkä merkityssuhteen konjunktio luo kytkemiensä lauseiden välille. Konjunktiot jaetaan **additiivisiin, disjunctiivisiin kontrastiivisiin, kausaalisiin, konsessiivisiin, konditionaalisiin, finaaliisiin ja temporaaliisiin**. Konjunktiot ja konnektiivit ovat kuitenkin hyvin väljämerkityksisiä ja monikäyttöisiä, joten niiden lauseiden välille luomista semanttisista suhteista ei voi muodostaa tyhjentävää luokitusta. (VISK § 812; § 1121.)

Konnektiivit eroavat konjunktioista siinä, että niiden luokka on avoin eikä niillä ole lauseessa yhtä selvää sijaintipaikkaa. Suurin osa niistä on konnektiivisia partikkeleita, mutta konnektiiveihin kuuluu myös adverbeja sekä post- ja prepositioita. Konnektiiveja käytetään kytkemään lauseita ja laajempia tekstinosia toisiinsa. Ne voivat siis myös liittää yhteen lausetta laajempia tekstikappaleita. Myös konnektiivit luovat merkityssuhteita lauseiden välille. (VISK § 820; § 1121.)

Kytkennän keinoista additiivinen suhde on tekstuaalisesti keskeisimpiä. Sitä voidaan kutsua myös listaavaksi tai kopulatiiviseksi suhteeksi. Additiivisten merkitsemien kautta ilmaistaan, että lauseilla on esitysyhteydessään jotakin yhteistä, ja että ne ovat molemmat yhtä lailla paikkansapitäviä. Additiivisesti käytetty *ja* on konjunktioista neutraalein, niin sanottu prototyyppinen rinnastuskonjunktio. *sekä* on selvemmin rinnastettavien erillisyyttä korostava. Parikonjunktiot *sekä – että* ja *niin – kuin* ilmaisevat tavallisesti 'kumpikin mainitusta kahdesta, molemmat erikseen'. Kielteisiä rinnastuksia tehdään myös *eikä*-konjunktion avulla. Addi-

tiivisesti käytettäviä konnektiiveja puolestaan ovat esimerkiksi *lisäksi, myös* ja *ohella*. (VISK § 1092; § 1096; § 1097.)

Disjunkttiivista suhdetta voidaan ilmaista konjunktioilla *tai, vai, eli* ja *joko – tai* (VISK § 1091). Niitä käytetään merkitsemään vaihtoehtosuhdetta kahden elementin välillä. *tai* ilmaisee useimmiten vaihtoehtoista mahdollisuutta. *eli* puolestaan esiintyy tyypillisesti parafrasoin tai vaihtoehtoisen näkökulman merkitsimenä. Yleisin disjunkttiivisesti käytettävä konnektiivi on *muuten*. (VISK § 1098; § 1100.)

Finaalisessa suhteessa tyypillisin konjunktio on *jotta ~ jottei*. Sitä ilmaistaan myös konjunktioilla *että ~ ettei*. Finaalinen suhde ilmaisee tavan saada toinen asiantila aikaan, mutta esittää harvemmin toisen lauseen tavoitteen toteutuneena. Finaalisessa suhteessa käytettäviä konnektiiveja ovat ainakin *siksi että* tai *niin (että)*. (VISK § 1133.)

Kontrastiivisessa suhteessa lauseet esitetään toisilleen vastakkaisina. Käytettävä konjunktio vaihtelee sen mukaan, millaista kontrastiivista suhdetta halutaan ilmaista. Konjunktioilla *mutta, kun* ja *taas* kytketään tyypillisesti yhteen tapaukset, joissa korostetaan entiteettien (asioiden, joihin kielellä voi viitata) eroja. *vaan* ja *mutta* -konjunktioilla sen sijaan viitataan entiteetin tai asiantilan korvaamiseen toisella, ja *mutta*-konjunktioita käytetään lisäksi poikkeuksen ilmaisemiseen. Lisäksi kontrastiivisuutta merkitään muun muassa ilmauksilla *toisaalta, puolestaan, sen sijaan, paitsi, vain* ja *lukuun ottamatta*. (VISK § 1101.) Tässä tutkimuksessa en erottele suhteita kontrastiivisuuden sisällä tarkemmin, vaan niputan kaikki kontrastiiviset suhteet samaan ryhmään.

Kausaalista suhdetta voidaan ilmaista konjunktioilla *sillä, koska* ja *kun*. Lisäksi sen ilmaisemiseen käytetään partikkeleita, adverbejä sekä adpositioita. Näitä ovat esimerkiksi *siksi, sen tähden, näin ollen* ja *siten*. Kausaalinen suhde vallitsee kahden lauseen välillä, kun toisen lauseen asiantila esitetään syynä tai perusteluna toiselle. Lauseiden välillä vallitsee siis syyn ja seurauksen suhde. (VISK § 1128.)

Konsessiivisessa suhteessa kahden lauseen asiantilojen sanotaan pätevän, mutta samaan aikaan niiden yhteensopivuus voidaan käsittää kontekstuaalisesti epäodotuksenmukaiseksi. Tavallisin konsessiivisesti käytettävä konjunktio on *vaikka*. Myös *mutta*-konjunktioita käytetään toisinaan konsessiivisesti. Konsessiivista suhdetta ilmaistaan esimerkiksi konnektiiveilla *ainakin, kuitenkin, silti* ja *sittenkin*. (VISK § 1139.)

Temporaalisessa suhteessa on pohjimmiltaan kysymys aikasuhteesta. Siinä ilmaistaan kahden tilanteen ajankohtien ja kestojen välinen suhde useimmiten *kun*-konjunktin avulla. Tilanne voidaan nähdä toiseen verrattuna kokonaan tai osittain samanaikaisena, aiempänä tai myöhempanä. Aikasuhteen selventämiseen voi käyttää täsmentäviä adverbejä ja partikkeleita,

kuten *sen jälkeen (kun)*, *samalla (kun)*, *sitten (kun)*. (VISK § 1122.) *kun*-konjunktiota käytetään kausaalisesti, mutta kausaalisen tulkinnan saavat ilmaukset voidaan usein tulkita myös temporaalisiksi (VISK § 1130).

Konditionaaliset konjunktiot ja konnektiivit ilmentävät syyn ja seurauksen suhdetta. Toinen lause ilmaisee ehdon, jonka toteutuessa toisen ilmaisema asiantila voi toteutua. Konditionaalista suhdetta ilmaisee tyypillisesti *jos*, mutta se voidaan muodostaa myös käyttämällä esimerkiksi kielteisiä *ellei* ja *jollei* -konjunktioita ja kiteymiä *siinä tapauksessa että* sekä *sillä ehdolla että*. (VISK § 1134.)

Konnektiivit ja konnektorit kytkevät siis toisiinsa lauseiden ja virkkeiden lisäksi myös lausekkeita, sanoja ja sananosia. Tekstin koheesion kannalta on kuitenkin tarkoituksenmukaista keskittyä tarkastelemaan nimenomaan lauseiden välillä esiintyvää kytkentää. Disjunkttiivisten konjunktioiden kohdalla olen kuitenkin huomionut myös *eli*-konjunktion käytön lausekkeiden kytkijänä, sillä tämäntapaiset suhteet vaikuttavat olevan Wikipedia-artikkeleille erityisen olennaisia.

3.2.3.3 Sanastollinen koheesio

Sanastollisessa koheesiossa kohesiivinen suhde saadaan aikaan sanastollisilla valinnoilla (Halliday & Hasan 1976: 272). Se muodostuu kielenaineksista, jotka ovat sidoksissa aikaisemmin esiintyneisiin kielenaineksiin. Hallidayta mukaillen jaan tässä tutkimuksessa sanastollisen koheesion **toistoon, yläkäsitteisiin, yleismerkityksisiin sanoihin**, tyypillisesti lähikäin esiintyvien sanojen muodostamiin **merkityskenttiin** ja **paradigmaattisiin merkityssuhteisiin**. Paradigmaattisista merkityssuhteista käsitelen erityisesti **synonymiaa**, mutta kiinnitän huomiota myös **hyponymiaan, meronymiaan** ja **antonymiaan**. Useimmiten sanastolliset suhteet esiintyvät yksittäisten kielenainesten välillä, mutta sanastollisia koheesiosuhteita voi ilmetä myös pidempien ilmausten välillä. (Halliday 1985a: 311, 330–333.)

Sanastollinen koheesio on rakenteellisista seikoista riippumatonta. Kieliopillisessa koheesiossa kielelliset elementit (esimerkiksi kytkennässä konjunktiot) kantavat mukanaan merkitystä. Sanastollisessa koheesiossa sen sijaan jokainen leksikaalinen yksikkö voi olla osa kohesiivista sidosta, mutta kyseisen elementin ilmiäsu ei paljasta, että toimiiko se kohesivisesti vai ei. Tämä selviää vain kontekstin perusteella, tekstiviittauksia tarkastelemalla. (Halliday 1985a: 311; 1976: 288.) Sanastolliset elementit kantavat silti nekin mukanaan omia merkityksiään, omaa tekstuaalista historiaansa, joka on tulkittavissa niiden esiintymisympäristöstä. Kontekstin perusteella voidaan päätellä, mikä merkitys jollakin leksikaalisella elementillä on juuri kyseisessä tapauksessa. (Halliday 1976: 288–289.)

Toisto on suurin sanastollisen koheesion muoto. Siinä jokin sanastollinen kielenaines toistetaan eli se esiintyy tekstissä vähintään kahdesti. Näihin kielenaineksiin lasketaan kuuluvaksi myös ainekset, jotka eivät ole täysin samassa morfologisessa muodossa. (Halliday 1985: 331.) Esimerkiksi *pakinan* ja *pakinoitsijan* voidaan katsoa muodostavan kohesiivinen sidos. Toisto voi siis olla täydellistä tai osittaista.

Leksikaaliseen koheesioon lukeutuu lisäksi yläkäsitteiden käyttö. Yläkäsitteisiin kuuluvat kaikki kielenainekset, joiden merkitykseen aikaisemmin mainittu käsite sisältyy. (Halliday 1985a: 331–332; Halliday & Hasan 1976: 280.) Esimerkiksi Juutinraumasta kertovassa Wikipedia-artikkelissa puhutaan ensin *Juutinrauman sillasta*, jonka jälkeen siihen viitataan yläkäsitteellä *silta* (esimerkki 6).

6) Juutinrauman yli rakennettu 7 845 metriä pitkä [Juutinrauman silta](#) avattiin vuonna 2000. *Sillan* kaksikerroksisessa rakenteessa on väylä junille (kahdet kiskot) ja autoille (neljä kaistaa).

Sanastollista koheesiota luodaan myös yleismerkityksillä sanoilla. Yleismerkityksisten sanojen luokkaan kuuluvat sellaiset nominit, joihin voidaan viitata monenlaisissa tapauksissa. Ne luovat tekstiin yhtenäisyyttä, sillä niiden laajojen käyttömahdollisuuksien vuoksi ne lähes aina vaativat viittauskohteeseen aiemmin esiintyneen tarkoitteen. Yleissubstantiiveihin lukeutuvat muun muassa ihmiseen viittaavat substantiivit *yksilö*, *mies*, *nainen*, *lapsi*, *tyttö* ja *poika* sekä paikkaan tai asiaan viittaavat substantiivit *paikka*, *ajatus*, *asia*, *idea* ja *kysymys*. (Halliday & Hasan 1976: 274–276.)

Kaikki edellä mainitut sanastollisen koheesion lajit – osittainen tai täydellinen toisto, yleismerkityksiset sanat ja yläkäsitteet – toimivat saman periaatteen mukaan. Olennaista on se, että viitataan samaan tarkoitteeseen. Kaikki nämä sanastollisen koheesion lajit siis edellyttävät aiemmin esiintynyttä viittauskohdetta. (Halliday 1976: 278.) Esimerkissä 6 jälkimmäisen virkkeen *silta* viittaa samaan tarkoitteeseen kuin ensimmäisessä virkkeessä mainittu *Juutinrauman silta*. Näin luodaan yläkäsitettä käyttämällä kohesiivinen sidos. Mikäli esimerkin jälkimmäinen lause käsittelisi siltoja yleisesti, ei kohesiivista sidosta ylä- ja alakäsitteiden välille muodostuisi, sillä ilmauksilla ei olisi yhteistä tarkoitetta.

Jonkinlaista sanastollista koheesiota voidaan kuitenkin nähdä myös siinä tapauksessa, että ilmauksilta puuttuu yhteinen tarkoite. Tällöin huomio kiinnitetään niihin tekstin leksikaalisiin elementteihin, jotka säännöllisesti esiintyvät yhdessä. Joillakin sanoilla on merkityspohjainen tendenssi esiintyä yhdessä tietynlaisten toisten sanojen kanssa. Semanttisesti toisiaan lähellä olevat sanat muodostavat merkityskentän. (Karlsson 2006: 218, 232.) Leksikaalisen

koheesion kannalta ei ole olennaista, mikä yhdessä esiintyvien sanojen semanttinen suhde toisiinsa on, vaan niiden tendenssi esiintyä kielessä yhdessä (Halliday 1976: 285).

Merkityskenttään kuuluvien sanojen esiintyminen yhdessä ei luonnollisesti ole mikään joko–tai -ilmiö. Kysymys on ennemminkin sanojen välisen läheisyyden asteesta. Voidaan sanoa, että toiset sanat ovat läheisemmässä suhteessa toisiinsa, ja juuri tämä suhteen läheisyys määrittelee kohesiivisen voiman. Tekstin sanastolliseen ympäristöön kuuluu tietysti muitakin sanoja, jotka vaikuttavat tekstin tulkintaan, mutta nimenomaan semanttisesti lähellä toisiaan olevien sanojen esiintyminen synnyttää koheesiota. (Halliday 1976: 289.) Merkityskenttien yleisyys heijastaa sitä tosiseikkaa, että ulkomaailma jäsenyy hierarkkisesti osakokonaisuudeksi (Karlsson 2003: 219).

Merkityskentän muodostavien sanojen lisäksi voidaan tarkastella parafraaseja sekä sanojen välillä vallitsevia paradigmaattisia merkityssuhteita. Näistä käsitellen seuraavaksi synonymiaa, hyponymiaa, antonymiaa ja meronymiaa. Parafraasi puolestaan voidaan määrittää merkityksen vaihtoehtoiseksi, mutta samaa ilmaisevaksi merkitystavaksi (Karlsson 2006: 208).

Sanastollinen koheesio ilmenee siis synonymiassa. Hallidayn mukaan synonymiassa valittu kielenaines on jollakin lailla samanmerkityksinen jonkin edellä mainitun kielenaineksen kanssa. (Halliday 1985: 331–332.) Synonymialla tarkoitetaan siis kahden tai useamman muodoltaan eroavan kielenaineksen samanmerkityksisyyttä. Synonymia voidaan nähdä yhdeksi paradigmaattisen merkityssuhteen muodoista. (Jantunen 2004: 53.) Paradigmaattinen merkityssuhde on suhde saman käsitejärjestelmän muihin yksiköihin ja tieto viittaussuhteista ulkomaailmaan. Paradigmaattisia suhteita ovat synonymian lisäksi muun muassa hyponymia, meronymia ja antonymia. (Häkkinen 1996: 171.)

Synonyymit ovat lähes poikkeuksetta lähisyronyymejä eli niillä on semanttisia tai syntaktisia eroja (Karlsson 2006: 220). Crusen mukaan täydellinen synonymia vaatisi, että synonyymisten ilmausten pitäisi pystyä esiintymään yhtä uskottavasti kaikissa samoissa konteksteissa, eli ne pitäisi pystyä vaihtamaan toistensa paikalle esiintymisympäristöstä huolimatta ilman, että merkityksessä tapahtuisi muutoksia (Cruse 2004: 154). Tämä on käytännössä mahdottomuus. Vaikka ilmausten denotatiivinen merkitys olisikin sama, niiden sivumerkityksissä ja esiintymisympäristöissä on lähes aina eroja (Jantunen 2004: 55). Synonyymeissä kiinnostavaa onkin se, että niiden semanttinen yhdenmukaisuus on keskeisempää kuin niiden erot (Cruse 2004: 154).

Koska synonyymien sattumanvarainen vaihdettavuus ei ole mahdollista, on niiden kontekstuaalisissa piirteissä oltava piirteitä, joiden mukaan ne valikoituvat tiettyihin käyttöympä-

ristöihin (Jantunen 2004: 57). Aineiston shetlanninlammaskoirasta kertovassa artikkelissa käytetään esimerkiksi synonyymisesti substantiiveja *karva* ja *turkki*. Ilmaukset eivät ole synonyymisia kaikissa konteksteissaan, mutta kyseisessä tapauksessa (esimerkki 7) ne voidaan käsittää synonyymisiksi.

7) Shetlanninlammaskoiralla on pitkä *karva*, joka on kuitenkin suhteellisen helppohoitoinen: se ei tuuheudesta huolimatta takkuunnu eikä likaannu helposti. Shetlanninlammaskoiran *turkki* on hyvin pehmeä ja sileä.

Hyponymia tarkoittaa merkitysten hierarkkista alistussuhdetta. Alisteinen sana katsotaan yläkategorian hyponyymiksi, esimerkiksi ori nähdään hevosen hyponyymina. Yläkategorian sana on puolestaan hyperonyymi. (Hakulinen & Karlsson 1995: 312.) Sanojen tasosta riippuen sama denotaatio voi toimia sekä hyponyymina että hyperonyymina. Hyponymian avulla rakennetaan käsitehierarkioita, jotka ovat keskeisiä denotaatioiden määrittämisessä. Tieteellinen tutkimus tuottaa moniportaisempia, tarkempia järjestelmiä kuin arkikielen käsitejärjestelmät, niin sanottua ensyklopedista tietoa. (Karlsson 2006: 221–222.) Koska Wikipedia-artikkelit sisältävät ensyklopedista tietoa, on syytä myös olettaa, että niistä löytyy hyponymiaa.

Paradigmaattisiin merkityssuhteisiin lukeutuu lisäksi meronymia. Siinä kyse on osan ja kokonaisuuden välisestä suhteesta. Esimerkiksi käden voi jakaa sormiin ja kämmeniin, oksa taas on osa puuta. Edeltävässä tapauksessa oksaa kutsuttaisiin meronyymiksi ja puuta holonyymiksi. Meronymialla ja hyponymialla on yhtymäkohtia, mutta niitä ei kannata sekoittaa toisiinsa. Esimerkiksi sormi ei ole alakategoria kädelle, vaan käden osa. Koira puolestaan ei ole eläimen osa, vaan alakategoria eläimelle eli sen hyponyymi. (Karlsson 2006: 221–222, Cruse 2004: 150.)

Antonymia ilmentää vastakohtaisuutta. Tyypillisimmillään antonyymit on nähty astevastakohtina. Ne ovat yleensä adjektiiveja, joiden välillä on kysymys aste-erosta, esimerkiksi erosta kuuman ja kylmän välillä. (Karlsson 2006: 223.) Antonymiassa leksikaalisen parin merkitysten semanttinen vastakkaisuus voidaan määritellä oppositiosuhteeksi. Kahden sanan välillä on pienimmillään eroa vain yhdessä piirteessä. (Halliday & Hasan 1976: 284–286.) Vastakohtaisuus ilmenee myös joko–tai-tapauksissa. Jos jokin sisältyy yhteen kategoriaan, se ei voi sisältyä toiseen. Näin on muun muassa vastinparissa elävä – kuollut. (Karlsson 2003: 223.)

Analyysissa käsittelen sanastollisen koheesion osalta toistoa, yläkäsitteitä ja yleismerkityksisiä sanoja sekä merkityskenttien muodostumista, paradigmaattisia merkityssuhteita ja parafraseja.

3.2.4 Teema- ja informaatorakenne

Tekstin koherenssin muodostumiseen vaikuttavat koheesion lisäksi tekstin temaattinen rakenne ja informaatorakenne sekä yhtenäinen diskurssiteema eli topiikki (Halliday & Hasan 1976: 299, 327). Teemarakenteen tarkasteleminen on lauserakenteen tutkimista funktionaalilta kannalta. Lausetta analysoidaan enemmän tekstuaalisten kuin kieliopillisten jäsenten kannalta, jolloin sitä ei tarkastella itsenäisenä yksikkönä, vaan tekstin osana. Käytännössä tutkitaan lauseen informaatorakennetta. (Hakulinen & Karlsson 1995: 41.) Informaatorakenteeksi kutsutaan lauseen informaation annosteluun liittyvää jäsenystä. Keskiössä on se, mitä pidetään annettuna ja mitä esitetään varsinaisena uutisena. (VISK § 1370.)

Informaatorakenteeseen kuuluu jakautuminen temaattiseen osaan ja reemaan. Teema ilmaisee lauseen puheenaiheen, ja reema puolestaan on se osa, joka kertoo puheenaiheesta jotakin uutta. (VISK § 1370.) Halliday käyttää informaatorakenteesta puhuessaan käsitteitä annettu (given) ja uusi (new). Annettu on jotakin, mitä jo tiedetään, joko tilanteen perusteella tai hyvin usein siksi, että se on mainittu jo aikaisemmin. Uusi puolestaan on nimensä mukaisesti jotakin uutta. (Halliday 1985a: 298–299.)

Halliday erottaa toisistaan lauseen temaattisen rakenteen ja informaatorakenteen, mutta näkee niiden olevan vahvasti yhteydessä toisiinsa. Niitä erottaa pääasiassa lähtökohta: informaatorakennetta ilmaistaan englannin kielessä intonaatiolla eli se on puhutun kielen ilmiö. (Halliday 1985a: 298–299; 1976: 325.) Tässä tutkimuksessa olen mukaillut Ison suomen kieliopin sekä Hakulisen ja Karlssonin tulkintaa enkä siis tarkastele teema- ja informaatorakennetta toisistaan erillisinä. Myös Halliday toteaa, että hyvin usein teemaksi valikoituu se, mitä pidetään annettuna, ja reema puolestaan on ilmauksen uusi asia (Halliday 1985a: 298–299). Tekstin rakenteen tarkastelun kannalta teema on se elementti, joka luo tekstin temaattisen eheyden. Tunnettu tai tunnetuksi oletettu tieto (se, mitä vastaanottajan voi olettaa pitävän muistissaan) esitetään nimenomaan teemassa. (Ranta 2007: 77.)

Teema voidaan nähdä jatkuvana tai vaihtuvana. Jatkuva teema ilmenee, kun yhtenäisessä tekstissä kuljetetaan samaa diskurssireferenttiä teemapaikassa. Hiidenmaa käyttää käsitettä pysyvä teema, eli puheenaihe teemalausekkeessa pysyy samana ja toistuu virkkeestä toiseen (Hiidenmaa 2001: 47). Näin on esimerkiksi Pakina-artikkelin johdannossa (esimerkki 8), jossa *pakina* on peräkkäisissä virkkeissä teemapaikalla. Jatkuvasta teemasta puhuttaessa on kui-

tenkin huomioitava, että saman tarkoitteen uudelleenmaininnat voivat sijoittua lauseissa myös muihin asemiin kuin teemaksi.

8) *Pakina* on lyhyt, kevyehkö ja huvittava kirjoitus, joka yleensä ilmestyy säännöllisesti vaikiopaikalla sanoma- tai aikakauslehdessä. *Pakina* on usein ajankohtainen mielipideteksti. *Pakina* on [kolumnia](#) enemmän sanataiteellisia piirteitä.

Vaihtuva teema puolestaan joko mainitaan kontekstissaan ensimmäistä kertaa tai vaihtoehtoisesti se on mainittu aiemmin, mutta otetaan nyt uudestaan puheeksi. Hiidenmaa käyttää käsitettä etenevä teema, eli uuden lauseen teema liittyy edellisen virkkeen reemaan. (Hiidenmaa 2001: 47.) Sammon taonta -artikkelin johdannon ensimmäisessä virkkeessä teemana on *Sammon taonta* ja jälkimmäisessä lauseessa *Akseli Gallen-Kallela* (esimerkki 9).

9) *Sammon taonta* on [Akseli Gallen-Kallelan](#) tekemä realismia edustava [Kalevala](#)-aiheinen maalaus, joka kuvaa [Sammon](#) tekoa. *Akseli Gallen-Kallela* maalasi teoksen vuonna 1893 viettäessään perheensä kanssa kesää [Savossa](#).^[1]

Tavallisimpia sanajärjestyksiä kutsutaan tunnusmerkittömiksi (VISK § 1366). Näitä noudattavia teemoja voidaan myös kutsua tunnusmerkittömiksi. Tyypillisin teema suomen tekstilauseissa on subjekti, lisäksi passiivilauseissa objekti ja eksistentiaali- sekä omistuslauseissa adverbiaali. Jos lauseen teema on jokin vähemmän tyypillisessä teema-asemassa esiintyvä, sitä voidaan pitää tunnusmerkkisenä. Lauseen alussa voi olla tekstin aiheeseen liittyvän, ideationaalisen teeman lisäksi tekstuaalista ja interpersonaista merkitystä ilmaisevia kielinaineita. Suomen kielessä esimerkiksi konnektiivi ja adverbi voivat sijaita lauseen alussa ennen ideationaalista teemaa. (Ranta 2007: 74, 140–141.)

3.2.5 Tekstilaji

Tekstilajin käsitettä käytetään eri tekstintutkimuksen lähestymistavoissa eri tavoilla. Genre ja tekstilaji käsitetään usein synonyymeiksi. Tekstilajin käsitteeseen liittyvät läheisesti myös tekstityyppi ja rekisteri, joiden erottaminen tekstilajista on toisinaan hankalaa. Tekstityypillä viitataan yleensä kielen muotojen ja rakenteiden perusteella muodostettuun luokkaan, jolla pyritään kuvaamaan rajattua määrää kielellis-retorisia kategorioita. (Shore & Mäntynen 2006: 36; Taavitsainen 2006: 266.) Rekisterin määrittelyyn palaan tarkemmin luvussa 3.2.6.

Tässä tutkimuksessa pyrin muodostamaan käsityksen Wikipedia-artikkelin tekstilajista ja rekisteristä, mutta jätän tekstityypin määrittämisen tutkimuksen ulkopuolelle. Siksi otan tekstilajin määrittelyssä huomioon tekstinsisäisiä piirteitä. Wikipedia-artikkelien käsittelyssä

on kuitenkin tarpeen huomioida myös tekstinulkoiset piirteet, esimerkiksi artikkelien yhteistoiminnallinen tuotantotapa. Kalliokosken (2006: 240) mukaan tekstilaji on ensisijaisesti sosiokulttuurinen käsite. Sen konventiot syntyvät kielenkäyttäjien muodostamassa yhteisössä. Tekstilajit ja niiden konventiot myös muuttuvat jatkuvasti. Sosiokulttuurisen kontekstin muutokset vaikuttavat tekstilajeihin: lajit eriytyvät, sulautuvat yhteen ja uusia lajeja syntyy.

Tekstilajitutkimuksessa tekstin kokonaisrakenne on perinteisesti ollut keskeinen tekstilajia määrittävä ja tekstilajeja toisistaan erottava tekijä (Raevaara & Sorjonen 2006: 126). Tekstilajin käsitettä määriteltäessä tekstin rakenne mainitaan usein yhtenä keskeisenä kriteerinä. Rakennetta voi kuitenkin lähestyä monella tapaa; määritelmässä voidaan korostaa tekstin sisäisiä, kielellisiä piirteitä tai tekstin kontekstia. Tutkijoilla on eri käsityksiä siitä, voiko samassa tekstilajissa olla toisistaan poikkeavia rakenteita tai voiko samaan tekstilajiin jopa kuulua aivan erilaisia tekstejä, jolloin tekstilaji määriteltäisiin ainoastaan kontekstipiirteiden avulla. (Mäntynen 2006: 42, 46.)

Tekstilajitutkimuksen kannalta olennaista on se, kuinka kuvata kokonaisen tekstilajin rakennetta, jos saman lajin yksittäiset tekstit eivät ole toistensa kaltaisia (Mäntynen 2006: 42, 46). Haastavaa tekstien luokittelusta tekee siis se, jos juuri mikään ei toistu tekstistä toiseen tarkalleen samanlaisena. Eri esiintymillä voidaan kuitenkin nähdä niin paljon samankaltaisuuksia, että ne luokitellaan samaksi lajiksi. Lajien rajat eivät silti ole selviä, ja lajitkin limityvät ja sekoittuvat toisiinsa. Ne myös muuttuvat ajan kuluessa. (Shore & Mäntynen 2006: 11.) Bhatian mukaan nykyisessä teknologia- ja mediamaailmassa tekstilajien pitäminen erillään on entistä haastavampaa (2004: 88).

Ruqaiya Hasan lähestyy tekstilajin käsitettä rakennepotentiaalimallinsa kautta. Sen mukaan samaan tekstilajiin kuuluvilla teksteillä on yhteisiä komponentteja, jotka määrittävät tekstilajin. Pakolliset komponentit löytyvät jokaisesta samaan tekstilajiin kuuluvasta tekstistä, mutta valinnaiset komponentit voivat vaihdella. Hasan käyttää esimerkkinä myyntitapahtumaa, jossa muun muassa hinnasta kertominen tai myyntitapahtuman lopettaminen ovat pakollisia komponentteja, mutta tervehdys on valinnainen komponentti. (Halliday & Hasan 1985: 63–68.) Ventola (2006: 107–108) kuitenkin kritisoi Hasanin mallia sen suppeudesta. Hän kyseenalaistaa sen, voiko Hasanin mallilla uskottavasti kuvata tekstilajin kokonaisrakenteen.

Tekstilajin määrittelyä on Hasanin jälkeen lähestytty sosiaalisesta näkökulmasta niin sanotussa Sydneyn koulukunnassa. Hallidayn oppilas James Martin määritteli tekstilajin kussakin yhteisössä sosiaalisesti hyväksytyiksi tavoiksi toimia. Hän näki tekstilajin käyttävän ilmaisutasonaan rekisteriä, ja rekisterin puolestaan tulevan näkyväksi kielessä. (Ventola 2006: 108.)

John Swales puolestaan pitää tekstilajin tärkeimpänä piirteenä sen kommunikatiivista tavoitetta. Kaikki muut tekstilajipiirteet ovat määräytyneet tämän tavoitteen mukaan. Swalesin tekstilajin määritelmässä samaan genreen kuuluvilla teksteillä on samat päämäärät, ja ne on suunnattu samalle yleisölle. Yhteinen vuorovaikutukseen liittyvä tavoite yhdistää tekstejä, jotka voivat vaihdella kieleltään ja rakenteeltaan. Swalesin määritelmä poikkeaa aikaisemmista myös siinä suhteessa, että genreä lähestytään piirrekimpun näkökulmasta eikä kerrostuneina tasoina. Tärkeä termi Swalesin määritelmässä on prototyypisyys: jotkut tekstit ovat lajinsa prototyypisiä edustajia, toiset puolestaan epätyypillisempiä. Toisaalta Swalesin määritelmä rajaa diskurssiyhteisön käsite. Diskurssiyhteisö tarkoittaa tekstilajia tuottavaa ja kontrolloivaa yhteisöä. (Shore & Mäntynen 2006: 27–30, 287.)

Hypoteesini on, että Wikipedia-artikkeleista koostuva aineistoni sisältää varsin paljon variaatiota. En siksi koe Hasanin pakollisia elementtejä korostavaa rakennepotentiaalimallia tarkoituksenmukaiseksi tekstilajin määrittäjäksi. Yhteistoiminnassa kirjoitettujen tekstien määrittäjäksi ei tunnu suoraan soveltuvan myöskään Swalesin prototyypisyyden käsite. Pyrin analyysissäni kiinnittämään huomiota teksteissä toistuviin rakenteisiin ja ilmiöihin ja pääsemään tätä kautta kiinni myös niiden funktioihin. Lisäksi huomioin sosiaalisen median teksteille olennaisen vuorovaikutuksen. Käsittelen tekstuaalisen metafunktion lisäksi myös ideationaalista ja intersosiaalista metafunktiota.

3.2.6 Rekisteri

Aivan kuten tekstilajin käsitteen, myös rekisterin määrittelyssä on suuria eroja tutkijoiden välillä. Systemis-funktionaalisen teorian sisällä sitä on määritelty toisistaan poikkeavilla tavoilla. Hyvin yleisesti voidaan sanoa, että rekisterillä tarkoitetaan sitä, miten kielenkäyttö vaihtelee tilanteisesti. (Shore & Mäntynen 2006: 38.) Ventolan (2006: 100) mukaan alun perin J. R. Firth kiinnitti erityistä huomiota siihen, että kielentutkijan ei ole mahdollista kuvata kieltä kokonaisuudessaan, vaan hän kuvaa väistämättä jotakin sen rajattua osaa, esimerkiksi tietyn kirjailijan tai sanomalehtien otsikoiden kieltä. Tästä käsityksestä kehittyi myöhemmin rekisterin käsite.

Halliday määrittää rekisterin kokoelmaksi semanttisia resursseja, jotka tyypillisesti liitetään tiettyyn tilanteeseen ja jotka reaalistuvat sanastossa ja rakenteissa. Halliday käyttää rekisterien kuvaamiseen tilannekontekstin kolmea muuttujaa: alaa, osallistujarooleja ja kielen ilmenemismuotoa. Niistä riippuu, minkälaisia merkityksiä voidaan ilmaista ja missä muodossa niitä on mahdollista käyttää. (Halliday 1978: 31; Halliday & Hasan 1976: 22–23; 1985: 29, 36.)

Tilannekontekstin muuttajat puolestaan korreloivat semanttisen systeemin kanssa. Ala tulee näkyviin ideationaalisessa, osallistujaroolit interpersonaalisisessa ja kielen ilmenemismuoto tekstuaalisessa metafunktiassa. Alaan liittyvät siis esimerkiksi partisiippiirakenteet ja mennyt aika. Osallistujaroleista voidaan tarkastella esimerkiksi moduksia sekä monologisuutta ja dialogisuutta. Kielen ilmenemismuotoon sisältyvät muun muassa elliptisyys, toisto ja kollokaatio. (Halliday 1978: 31; Halliday & Hasan 1985: 29, 36.)

Rekisteripiirteitä tutkittiin aiemmin kiinnittämättä systemaattista huomiota siihen, missä tekstin kohdissa ne esiintyvät. Erityisesti Hasanin rakennepotentiaalimallin myötä tarkastelun kohteeksi kuitenkin nousivat myös tekstin vaiheet ja se, millaisia kielellisiä rakenteita tekstin eri jaksoissa käytetään. Hasanin mukaan toisiaan rakenteellisesti muistuttavat tekstit edustavat samaa tekstilajia eli kuuluvat samaan sosiaalisen toiminnan lajiin. Rekisteripiirteet ja tilannekonteksti (ala, osallistujaroolit, ilmenemismuoto) luovat kehykset, joiden sisällä tekstien voidaan nähdä kuuluvan samaan rekisteriin. (Ventola 2006: 102–103, 105.) Hasanin lineaarinen, välttämättömiin ja valinnaisiin jaksoihin perustuva rakennepotentiaalimalli sopii hyvin kuvaamaan esimerkiksi Hasanin tutkimaa asiakaspalvelutilannetta, mutta rakenteeltaan paljon vaihtelevampien Wikipedia-artikkelien kuvaamiseen se ei ole ihanteellinen.

Rekisterin käsitettä vei edelleen eteenpäin James Martin ja Sydneyn koulukunta (Shore & Mäntynen 2000: 21). Martin käsittää tekstilajin abstraktiksi tasoksi, johon sisältyvät ala, osallistujaroolit ja ilmenemismuoto eli siis rekisteri. Rekisteri puolestaan reaalistuu kielen (muun muassa semantiikan, leksikkokieliopin ja fonologian) kautta. (Martin 2008: 231.) Sydneyn koulukunta ei, toisin kuin Hasan, näe rekisteripiirteiden määräytyvän kontekstuaalisten piirteiden mukaan koko tekstille, vaan on sitä mieltä, että sosiaalinen kielellinen toiminta etenee vaiheittain ja ala, osallistujaroolit ja ilmenemismuoto on syytä kussakin vaiheessa määrittellä uudelleen. (Ventola 2006: 108–109.)

Mauranen tuo esiin, että mikäli rekisteri käsitetään tyypillisesti yhdessä esiintyvien kielipiirteiden muodostelmaksi, niin monet sosiaaliset tilanteet vaihtavat rekisteriä moneen kertaan. Systemis-funktionaaliseen teorialle tyypillinen ajatus siitä, että rekisteri voidaan tunnistaa pätkästä tekstiä rekisterille tunnusomaisten kielellisten piirteiden perusteella ja sitä kautta päätellä tekstille myös funktio, on siis soittautunut käytännössä monimutkaisemmaksi. (2006: 215.)

Hallidayn mukaan kysymykset, kuten kuuluvatko jotkin kaksi tekstiä samaan rekisteriin, eivät kuitenkaan ole kovin tarkoituksenmukaisia. Tilannekontekstit eivät koskaan ole kahden tekstin välillä täysin identtisiä, joten olennaista on sen sijaan kysyä, että mikä teksteille on yhteistä ja missä suhteessa ne eroavat toisistaan. (Halliday & Hasan 1976: 24–25.)

Halliday puhuu myös suljetuista ja avoimista rekistereistä tai oikeastaan jatkumosta niiden välillä. Esimerkiksi lentäjien kielen rekisteri kuuluu jatkumon suljettuun päähän. Kansainvälinen lentokieli on tiukasti säädeltyä ja merkitykset (lähes) yksiselitteisiä. Suurin osa rekistereistä on kuitenkin avoimia. Avoimuuden asteessa on eroja: virallisen lomakkeen teksti on säädellympää kuin esimerkiksi luokkahuoneen kieli. Täysin avointa rekisteriä ei kuitenkaan ole olemassa, sillä kaikkein spontaaneimmallakin keskustelulla on omat strategiansa ja tyylinsä. (Halliday & Hasan 1985: 39–40.)

Tässä tutkimuksessa seuraan Hallidayn tilannekontekstiin perustuvaa määritelmää rekisteristä. Hiidenmaa kiteyttää osuvasti Hallidayn käsityksen rekisteristä: ”miten tilanteen mukainen merkityksenanto toteutuu sanoina, kieliopillisina rakenteina ja tekstuaalisina valintoina.” Tekstit ovat tilanteita ja merkityksiä luovia rakenteita, joten näiden rakenteiden kautta tekstejä pitää myös tarkastella. Tämän vuoksi pelkkään tilastolliseen muotoon perustuvat analyysit eivät tee rekisterille täyttä oikeutta. (Hiidenmaa 2001: 36–37.)

Wikipedia-artikkelien rekisteri on luonnollisesti avoin. Hypoteesini on, että aineistoni rekisteri sijoittuu kohtalaisen avointen rekisterien joukkoon. Oletan artikkeleista löytyvän runsaasti variaatiota, koska niiden kirjoittajajoukko on hyvin laaja. On kuitenkin myös todennäköistä, että joiltakin osin tekstit ovat keskenään varsin samantyyllisiä, sillä ne nojaavat kaikki tietokirjatekstiin ja ne on kirjoitettu samalle alustalle.

4 WIKIPEDIA-ARTIKKELIEN MUOTORAKENNE SEKÄ TEEMA- JA INFORMAATIORAKENNE

4.1 Wikipedia-artikkelien muodon analyysi

Wikipedia-artikkelien rakenteeseen vaikuttavat sekä alustan asettamat raamit että kirjoittajien vapaus muokata artikkelien sisältöä. Monilta muodollisilta piirteiltään artikkelit ovat hyvin samankaltaisia, sillä ne on kaikki kirjoitettu samalle alustalle. Artikkelien kirjoittajat ovat kuitenkin voineet varsin vapaasti päättää, sisältyykö artikkeliin sisällysluettelo, miten teksti jaotellaan lukuihin ja alalukuihin ja kuinka lähteitä käytetään. Artikkelit myös vaihtelevat pituudeltaan huomattavasti. Biologisesta antropologista kertovassa artikkelissa sanoja on vain 80, kun taas Bahai-uskosta kertovassa artikkelissa niitä on 2928.

Aineiston artikkelit muistuttavat toisiaan ulkoisesti siinä, että ne kaikki noudattavat Wikipedialle tyypillistä muotoa fontin suhteen ja sen suhteen, miten teksti on aseteltu sivulle. Yhteistä on myös sivustolla toimiva linkitys, joka mahdollistaa sen, että lukija voi siirtyä sujuvasti Wikipedia-artikkelista toiseen. Linkitys mahdollistaa myös suoran siirtymisen artikkelissa käytettyihin internet-lähteisiin, mikäli ne on merkitty näkyviin. Tämä hypertekstuaalinen ominaisuus löytyy jokaisesta aineiston artikkelista pituudesta riippumatta.

Artikkelit kuitenkin eroavat toisistaan myös monin tavoin: variaation voikin sanoa olevan Wikipedialle ominainen piirre. Wikipedia-artikkelit voi asettaa jatkumolle, jonka toisessa päässä ovat lyhyet, lukuihin jaottele mattomat tekstit ja toisessa päässä pitkät, useita lukuja ja mahdollisesti alalukuja sisältävät, usein myös lähdeviitteiltään kattavammat artikkelit. Wikipedian variaatiosta kertoo kuitenkin myös se, etteivät artikkelit asetu kaikissa suhteissa samaan päähän tätä jatkumoa. Aineistossa on artikkeleita, jotka ovat lyhyitä, mutta niissä on käytetty lähdeviitteitä kattavasti ja toisaalta pitkiä, monipuolisia artikkeleita, jotka on järjestetty selkeästi lukuihin, mutta joista lähdeviitteitä ei käytännössä löydy lainkaan.

Lisäksi artikkelit poikkeavat rakenteen osalta jonkin verran toisistaan. Eroa on siis muun muassa siinä, miten artikkelien sisältö on jaettu lukuihin ja kuinka lähteitä käytetään artikkeleissa. Artikkelien kirjoittajilla on ollut vapaus valita, että jaetaanko artikkelia ensinkään lukuihin, löytyykö siitä sisällysluettelo ja käytetäänkö siinä lähdeviitteitä.

Valtaosassa artikkeleita teksti on järjestetty jonkinlaisiin lukuihin. 23 aineiston artikkelia eli noin 76 prosenttia aineistosta sisältää johdannon lisäksi yhden tai useamman aiheitaan kuvaavan, otsikoidun luvun. Olen laskenut artikkelit jaetuksi lukuihin, jos niissä on yksi tai

useampi itse aiheesta kertova luku. En siis ole laskenut mukaan niitä artikkeleita, joissa johdannon lisäksi on ainoastaan lähdeluettelo tai Aiheesta muualla -luettelo.

Kaikkia artikkeleita yhdistää Wikipedialle tyypillinen johdanto-osuus, eli jokainen artikkeli alkaa aiheensa määrittelevällä johdannolla. Mikäli artikkelia ei ole jaettu lukuihin, löytyy siis koko artikkelin sisältö johdanto-osasta. Tyypillisesti johdanto-osuutta seuraa kuitenkin lukuihin jaettuna aiheen tarkempi käsittely, jota mahdollisesti edeltää sisällysluettelo. Keskimäärin johdanto-osa muodostuu 7,1 virkkeestä, mutta vaihtelu artikkelien välillä on varsin suurta. Magmakivi-artikkelin johdanto on pisin. Se sisältää 24 virkettä. Lyhimmillään johdanto-osa on vain yhden virkkeen mittainen, kuten esimerkki (10) Merlin (televisiosarja) -artikkelista havainnollistaa. Siinä artikkelin aihe on määritelty ytimekkäästi ainoastaan yhdellä virkkeellä, jonka jälkeen seuraa jo Juoni-luku.

10) *Merlin*² on vuonna 2008 alkanut [BBC:n](#) koko perheen [fantasiadraamasarja](#).

Juoni

VAROITUS: Seuraava kirjoitus [paljastaa yksityiskohtia juonesta](#).

Jokaisen artikkelin johdannossa artikkelin aihe määritellään. Niissä artikkeleissa, joissa sisältö on jaettu lukuihin, johdanto-osa on tyypillisesti muutaman virkkeen mittainen. Mikäli kaikki artikkelin tieto on sisällytetty johdanto-osuuteen lukuihin jaottelemisen sijaan, on se yleensä pidempi.

Jos artikkelissa on sisällysluettelo, se sijoittuu aina välittömästi johdanto-osan jälkeen. 16 artikkelissa, eli noin 53 prosentissa aineiston artikkeleista, sisällysluettelo on merkitty näkyviin. Paitsi että sisällysluettelosta saa kokonaiskuvan artikkelin sisällöstä, se myös mahdollistaa linkkien kautta siirtymisen suoraan siihen lukuun tai alalukuun, jota lukija haluaa tarkastella tarkemmin. Näin artikkelit toimivat jouhevasti hyperteksteinä, joiden lukeminen ei ole enää välttämättä lineaarista.

Sisällön lukuihin jakamisen suhteen aineistossa esiintyy hyvin vaihtelevia käytänteitä. Kuten aikaisemmin tuli esille, ei seitsemässä aineiston artikkelissa ole lukuja lainkaan. Niissä artikkeleissa, joissa teksti on jaettu lukuihin, ei puolestaan esiinny yhtenäistä käytäntöä lukujen ja alalukujen suhteen. Joissakin artikkeleissa on runsaasti alalukuja, toiset koostuvat vain pääluvuista. Esimerkiksi Ikeasta kertovassa artikkelissa (esimerkki 11) sisältö on jaettu seitsemään lukuun, joista kaksi on edelleen jaettu alalukuihin.

² Kursivointi alkuperäinen

- 11) [1 Historiaa](#)
- [2 Mallisto](#)
- [3 Ikea Suomessa](#)
 - [3.1 Ikea Oy](#)
 - [3.2 Tavaratalot](#)
- [4 Ikea Family](#)
- [5 Keskustelu ja kritiikki](#)
- [6 Lähteet](#)
- [7 Aiheesta muualla](#)
 - [7.1 Kirjallisuutta](#)
 - [7.2 Dokumenttielokuvia](#)

Se, miten sisällysosa on jaoteltu lukuihin ja miten luvut on otsikoitu, riippuu paljon artikkelin aiheesta ja kirjoittajien näkemyksistä siitä, mikä on olennaista. Artikkelit voivat erota sisällön suhteen toisistaan paljonkin, joten niitä on myös jaoteltu lukuihin ja alalukuihin hyvin eri tavoin. Hyvin usein artikkeli kuitenkin sisältää aiheen määrittelyn lisäksi jonkinlaisen historiaosuuden.

Vesa Heikkinen määrittelee sisällysluettelon yhdeksi tehtäväksi myös idetionaalisesta merkityksestä eli sisällöstä kertomisen ja sen jäsentämisen lukijalle (Heikkinen 2001: 68). Wikipedia-artikkeleissa käsiteltävä aihe jäsentyy lukijalle sisällysluettelon kautta, jos sellainen on vain luotu artikkeliin. Artikkelin kirjoittajat ovat valinneet sisällysluetteloon nostamansa näkökulmat tekstin aiheen kannalta olennaisiksi ja näin ottaneet kantaa aiheeseensa ja kuvanneet maailmaa. Sama pätee tietysti myös lukuihin jaotteluun: sekin on kirjoittajien tapa jäsentää käsittelemäänsä ilmiötä ja samalla maailmaa laajemminkin. Lukijan tehtäväksi puolestaan jää päättää, missä järjestyksessä hän asioista lukee, sillä teksti tarjoaa tähän useita tapoja.

Wikipedialle on hyvin tyypillistä, että luvut ovat lyhyitä. Useimmiten ne sisältävät vain yhden tai kaksi kappaletta. Mikäli jostakin aiheesta on kirjoitettu pidemmin, se löytyy tavallisesti linkin takaa omana artikkelinaan. Esimerkiksi Brasilia-artikkelissa on kolmesta lyhyestä kappaleesta koostuva Väestö-luku, mutta luvun alussa on linkki *Pääartikkeli: Brasilian väestö*. Näin yleiskatsaus Brasilian väestöstä löytyy varsinaisesta Brasilia-artikkelista, ja tarkempi kuvaus Brasilian väestöstä on linkin takana omana artikkelinaan.

Sen lisäksi, että Wikipedia-artikkeleissa on lyhyitä lukuja, niissä on myös paljon lyhyitä kappaleita. 18 artikkelia sisältää ainakin yhden vain yhdestä virkkeestä koostuvan kappaleen, jotkut useampia. Tyypillisesti yhdestä virkkeestä koostuvat kappaleet liittyvät asiasisällöltään tavallista löyhemmin ympäröiviin kappaleisiin, joten ne on ollut helpompi jättää omiksi kappaleikseen. Näin on esimerkiksi Pakina-artikkelin johdannossa (12), jossa *pakinan* alkuperästä kertova virke on erotettu omaksi kappaleekseen:

12) **Pakina** on lyhyt, kevyehkö ja huvittava kirjoitus, joka yleensä ilmestyy säännöllisesti vaikiopaikalla sanoma- tai aikakauslehdessä. Pakina on usein ajankohtainen mielipideteksti. Pakinassa on [kolumnia](#) enemmän sanataiteellisia piirteitä. Pakinan kirjoittajaa kutsutaan pakinoitsijaksi.

Sana *pakina* on johdettu verbistä *pakista*³, joka on tarkoittanut 'puhisemista'.^[1]

Ilmiö kertoo myös ajoittaisesta artikkelien koheesion vähäisyydestä, sillä artikkeleissa esiintyviä yksittäisistä virkkeistä muodostuvia kappaleita ei ole pyritty koheesion keinoin liittämään jonkin toisen kappaleen yhteyteen. Koheesiota löytyy tyypillisesti enemmän kappaleiden sisältä kuin kappaleiden väliltä (Halliday ja Hasan (1976: 296–297)). Aineistosta onkin havaittavissa, että yksittäisten virkkeiden kohdalla kohesiivisiä sidoksia ei juuri esiinny. Koheesioon palaan laajemmin luvussa 5.

Lyhyissä kappaleissa tulee näkyväksi myös Wikipedian prosessiluonne. Tätä havainnollistaa muun muassa seuraava esimerkki (13). Brasilian urheilusta kertovan luvun loppuun on sijoitettu neljä lyhyttä kappaletta, jotka kaikki kertovat eri urheilulajeista. Luvun loppu vaatisi lisää tekstiä ja kohesiivisiä sidoksia, jotta siitä muodostuisi yhtenäinen kokonaisuus. Toisiinsa liittymättömien yksittäisten virkkeiden ympärille on kuitenkin helpompi alkaa kirjoittaa lisää sisältöä näin, kuin jos ne olisi pyritty kieliopillista koheesiota käyttämällä yhdistämään yhdeksi yhtenäiseksi kappaleeksi.

13) Brasiliassa on kehitetty oma [jujutsusuuntauksensa](#), [Brasilialainen jujutsu](#).^[56]

[Capoeira](#) on Brasilian kansallislaji.^[57]

Brasiliassa järjestetään [vuoden 2016 kesäolympialaiset](#) Rio de Janeirossa.^[58]

Brasilia on menestynyt myös hyvin judossa. Saavuttaen siinä maailmanmestaruuksia ja useita olympiamitaleita. Tunnettuja judokoitte ovat: Leandro Guilherme, Tiago Camilo ja Flavio Canto.

Wikipedia-artikkelien multimediaisuus tulee näkyviin niin, että varsinaisen tekstiosan lisäksi Wikipedia-artikkeleissa on paljon tietolaatikoita, taulukoita, luetteloita ja kuvia. Tietoa on siis havainnollistettu muutenkin kuin perinteisen tekstin muodossa, ja artikkelit ovat varsin visuaalisia. 20 artikkelista löytyy yksi tai useampi kuva. Aiheen kannalta olennaiset tiedot on tiivistetty tietolaatikkoon kymmenessä artikkelissa. Tietolaatikat sisältävät runsaasti linkkejä, joita klikkaamalla tiedon etsimistä voi helposti jatkaa.

Lähdeviitekäytäntö Wikipedia-artikkeleissa on vaihteleva. Koska artikkelien oikeellisuus ja uskottavuus eivät kuulu tutkimukseni piiriin, en kiinnitä huomiota siihen, minkälaisiin

³ Esimerkin kursivoinnit alkuperäisiä

lähteisiin artikkeleissa viitataan. Lähteiden merkitsemiskäytäntö sinällään on kuitenkin olennainen osa tietokirja-artikkelien tekstilajin määrittämistä, joten käsittelen lyhyesti aineistoa lähteiden osalta. Lähteiden merkitsemisessä näkyy Wikipedia-artikkeleille ominainen variaatio. Artikkelit voidaan jakaa lähdeviitteiden osalta karkeasti kolmeen ryhmään: niihin, joissa artikkeli on kauttaaltaan rakennettu lähdeviitteiden varaan, niihin, joissa osaan tekstistä on satunnaisesti merkitty lähdeviitteet sekä niihin, joissa lähdeviitteitä ei ole.

Ensimmäiseen ryhmään sisältyvät artikkelit, joihin lähdeviitteet on merkitty kattavasti kautta koko artikkelin. Tähän luokkaa kuuluu seitsemän artikkelia. Hieman yllättävästi lähdeviitteiden järjestelmällinen käyttö ei kuitenkaan tarkoita, että artikkelit olisivat muulta rakenteeltaan yhtä loppuun hiottuja. Bahai-uskoa ja Brasiliaa käsittelevissä artikkeleissa on kattava jako lukuihin, ja niihin on luotu myös sisällysluettelo, mutta Sammon taonnasta, Yleisestä kielitutkinnosta ja Vares – Sheriffi -elokuvasta puuttuu sisällysluettelo eikä metonymiasta ja naisten äänioikeudesta kertovien artikkelien sisältöä ole edes jaettu lukuihin.

Keskimmäinen luokka on lähdeviitteiden käytön osalta kirjavin. Aineistossa on viisitoista artikkelia, joista lähdeviitteet löytyvät osittain. Osassa artikkeleista lähdeviitteitä on käytetty melko säännönmukaisesti lähes koko artikkelissa, kun taas toisissa on vain yksi tai kaksi lähdeviitettä. Esimerkiksi Unisiepparia käsittelevässä artikkelissa lähdeviitteitä on runsaasti, mutta sen sijaan pitkässä Merkurius-artikkelissa ainoastaan tietolaatikkoon kirjoitettuihin lukuihin on merkitty lähteet.

Kolmanteen luokkaan kuuluvat tekstit, joissa lähdeviitteitä ei ole lainkaan. Niitä on aineistossa kahdeksan. Lähdekäytäntöä monimutkaistaa myös se, että vaikka artikkeliin ei olisi merkitty lähdeviitteitä, artikkelin lopussa on toisinaan Lähteet-osio, johon on merkitty, mistä artikkelin tiedot ovat peräisin. On kuitenkin mahdotonta tietää, mikä kaikki artikkelissa on peräisin tekstin lopussa mainituista lähteistä. Vaihtelevissa lähdekäytänteissä näkyvät Wikipedian prosessiluonne ja vuorovaikutteinen kirjoitustapa: kun useat kirjoittajat luovat ja korjailevat tekstiä, myös käytännöt lähdeviitteiden suhteen vaihtelevat.

Wikipedia-artikkeleista voi havaita hypertekstien piirteitä. Artikkelit on rakennettu niin, että ne voi kyllä lukea lineaarisesti alusta loppuun, mutta ne sisältävät myös runsaasti elementtejä, jotka osoittavat kohti ei-lineaarista lukemista. Näitä ovat esimerkiksi linkit toisiin Wikipedia-artikkeleihin sekä lähteiden kautta muualle internetiin ja sisällysluettelo, joka niinkään sisältää linkit eri lukuihin.

Vaikka osa teksteistä on hyvin lyhyitä eikä niissä ole juurikaan havainnollistavia kuvia tai tietolaatikoita, nekin ovat ainakin Wikipedialle tyypillisen linkityksen kautta yhteydessä muihin Wikipedia-artikkeleihin. Mikään artikkeleista ei ole olemassa tyhjiössä, vaan asetuu

osaksi Wikipedian tietoverkkoa. Nämä hypertekstien piirteet rakentavat kuvaa artikkeleista myös interpersonallisella tasolla. Niillä ohjataan lukijaa tietynlaiseen lukutapaan, jossa tekstin välillä on mahdollista liikkua ja kukin voi esimerkiksi sisällysluettelon perusteella lukea vain ne kohdat, jotka kiinnostavat tai joista tarvitsee tietoa. Wikipediassa lukija siis rakentaa oman polkunsä omien tarpeidensa tai kiinnostuksensa mukaan.

Tila ei ole Wikipedia-artikkeleille rajoite samalla lailla kuin perinteisille, painetuille yleistietosanakirjoille. Kaikkia Wikipedia-artikkeleita yhdistääkin se, että niissä ei käytetä lyhennettyjä sanoja ja lauseita, vaan ainoastaan kokonaisia virkkeitä. Tämä käytäntö erottaa artikkelit painetun yleistietosanakirjan artikkeleista. Lisäksi Wikipedia-artikkelien lähdeviitekäytäntö sekä lukurakenne muistuttavat enemmän tieteellistä artikkelia kuin yleistietosanakirja-artikkelia. Toisaalta Wikipedia-tekstit eroavat myös tieteellisistä artikkeleista muun muassa siinä, että niiden välinen linkitys mahdollistaa ei-lineaarisen lukemisen. Kaikkia artikkelissa esiintyviä termejä tai ilmiöitä ei ole tarpeen määrittellä tai avata kovin laajasti, sillä kyseiset termit voidaan vain linkittää ohjautumaan omille sivuilleen.

Wikipedian rakenteen analysoinnista käy ilmi, että siihen vaikuttaa teknologian mahdollistaman multimediaisuuden ja hypertekstin ominaisuuksien lisäksi vuorovaikutuksellinen ja prosessinomainen luomistapa. Artikkelien lähdeviitekäytännöt ja niiden jako lukuihin vaihtelevat tekstistä toiseen.

4.2 Teema- ja informaatorakenne

Teema- ja informaatorakenteella on oleellinen rooli tekstin koherenssin luomisessa. Informaatorakenteeksi kutsutaan lauseen informaation annosteluun liittyvää jäsenystä, eli huomio kiinnitetään siihen, mitä lauseessa pidetään annettuna ja mitä esitetään uutena (VISK § 1370). Kiinnitän huomioni aineistossa varsinaisesti ideationaalisiin teemoihin eli lauseiden aiheisiin. Lauseen alussa voi kuitenkin ideationaalisen teeman ohella esiintyä myös tekstuaalinen teema. Esimerkissä (14) *sana fani* on lauseen ideationaalinen teema, mutta tekstuaalinen teema. Esimerkissä (14) *sana fani* on lauseen ideationaalinen teema, mutta tekstuaalinen teema voidaan pitää partikkelia *toisaalta*.

14) Toisaalta sana *fani*⁴ juontuu latinankielisestä temppeliä tarkoittavasta sanasta *fanum* ja sen johdannaisesta *fanaticus*, joka voidaan kääntää *jumalan innoittamaksi* tai *temppelinpavelijaksi*.^[3]

Aineistossa useiden lauseiden teemana toimii jo heti otsikossa mainittu koko artikkelin aihe. Wikipedia-artikkeleille on melko tyypillistä, että niiden teemalausekkeen teema on py-

⁴ Kielenainesten kursivoinnit esimerkissä ovat alkuperäisiä

syvä. Usein samaa puheenaihetta – monessa tapauksessa koko artikkelin aihetta – toistetaan virkkeestä toiseen sen sijaan, että edeltävässä lauseessa uutena esitelty reema siirtyisi seuraavan virkkeen teemaksi. Hyvä esimerkki (15) tästä on Mikrobiologia-artikkelin johdanto. Siinä mikrobiologia toimii pysyvänä teemana koko kappaleen ajan.

15) *Mikrobiologia* on [mikro-organismeja](#) tutkiva [biologian](#) haara. Se tutkii [bakteerien](#), [arkkien](#), [sienten](#), [virusten](#), [levien](#) ja [alkueläinten](#) rakennetta, [elintoimintoja](#) ja niiden hyödyntämistä. *Mikrobiologialla* on runsaasti käytännön sovelluksia esimerkiksi [lääketieteen](#), [elintarviketeknologi-an](#) ja [ympäristötekniologian](#) aloilla. *Mikrobiologiaa* hyödynnetään myös bakteerien avulla tapahtuvassa [kompostoinnissa](#), joidenkin [kaivosten](#) mineraalituotannossa sekä [teollisuudessa](#) orgaanisten yhdisteiden tuotannossa.

Tietokirjatekstissä, jossa keskitytään määrittelemään pääasiassa yhtä asiaa tai ilmiötä, onkin järkeenkäypää, että uusien teemojen esittelemisen sijaan käsitellään yhtä teemaa. Juuri määrittelevät osuudet Wikipedia-artikkeleissa ovatkin usein niitä, joissa teema on pysyvä. Olennaista ei ole kertoa edellisen virkkeen reemasta jotakin uutta seuraavassa lauseessa, vaan käsitellä jo mainittua teemaa useammalta kannalta ja kertoa siitä enemmän asioita. Näin seuraavan virkkeen teema liittyy edellisen virkkeen teemaan eikä reemaan.

Harvemmin kappaleen teema on kuitenkaan vain yhtä tyyppiä. Vaikka pysyvä teemaa esiintyy aineistossa varsin paljon, usein teemat kuitenkin sekoittuvat. Tämä ilmenee myös Shetlanninlammaskoiraa käsittelevän artikkelin ulkonäöstä ja liikehdinnästä kertovassa luvussa. *Shetlanninlammaskoira* toimii teemana useissa kappaleen virkkeissä. Esimerkistä (16) löytyy kuitenkin myös etenevää teemaa. Esimerkiksi *häntä* toimii ensin reemana ja sitten seuraavassa virkkeessä teemana.

16) *Shetlanninlammaskoiralla* on pitkä karva, joka on kuitenkin suhteellisen helppohoitoinen: se ei tuuheudestaan huolimatta takkuunnu eikä likaannu helposti. *Shetlanninlammaskoiran* turkki on hyvin pehmeä ja sileä. *Shetlanninlammaskoirilla* on värityksestä riippuen joko mustat, ruskeat tai siniset silmät, jotka ovat mantelinmuotoiset. *Uroksille* tulee rintaan narttuja muhkeampi ”leijonaharjas”. *Sheltin* kuono on suippo ja häntä kaareutuu alas. *Häntä* on normaalisti erittäin tuuheaa. *Shetlanninlammaskoiran* korvat ovat rotumääritelmän mukaisesti hieman lurpallaan päistä, mutteivat saa mennä täysin alas asti. *Korvat* saattavat helposti nousta pystyyn, mutta sitä voidaan ehkäistä laittamalla korviin painot tai korvat voidaan liimata maitoliimalla karvoista koiran kasvuajaksi noin 2–4 kuukauden ikäisenä.

Edeltävä esimerkki havainnollistaa, että vaikka Wikipedia-artikkeleissa esiintyy varsin paljon pysyvää teemaa, ilmenee se kuitenkin harvoin kappaleen ainoana teemana. Kun aihetta, esimerkiksi edeltävässä esimerkissä shetlanninlammaskoiran ulkonäköä, käsitellään tarkemmin, esiintyy artikkelissa enemmän etenevää teemaa. Usein artikkelin pääaiheesta kuitenkin

kin tulee uudestaan pysyvä teema, kun siihen myöhemmin artikkelissa palataan uudemman kerran.

Edeltävissä esimerkeissä on varsin selkeä informaatorakenne. Joko artikkelin aihe toimii pysyvänä teemana tai edellisen lauseen reema siirtyy seuraavan lauseen temaksi. Välillä toisiaan seuraavien virkkeiden yhteyttä on Wikipedia-artikkeleissa kuitenkin vaikea nähdä. Muun muassa Liikennettä käsittelevän artikkelin ensimmäisen luvun alun lauseissa on kaikissa eri teema, joka ei kuitenkaan nouse edellisen lauseen reemasta. Tällöin teemarakenne ei merkittävästi rakenna artikkelin koherenssia, vaan tekstistä tulee helposti luettelomainen vaikutelma. Tämä korostuu esimerkissä (17), jossa on peräkkäin kolme pelkästään päälauseesta koostuvaa virkettä.

17) *Maanviljelyn ja käsityöläisten erikoistuminen* noin 12 000 vuotta sitten vaati tuotteiden kuljettamista tuottajalta kuluttajalle. *Liikenteen kehittymisen vallitsevat trendit* ovat olleet kuljetusten nopeutuminen, kuljetusyksiköiden suureneminen ja matkojen piteneminen. *Henkilöliikenteen historia* on keskeisiltä osiltaan samanlainen tavaraliikenteen historian kanssa.

Informaatorakenne ei liitä edellisen esimerkin virkkeitä yhteen. Tällöin yhtenäisyyttä rakentavat lähinnä yhteiseen merkityskenttään liittyvä sanasto ja se, että virkkeet on sijoitettu peräkkäin.

Huomattavaa on, että kaikkia aineiston artikkeleita yhdistää johdannon ensimmäisen lauseen yhtenäinen temaattinen rakenne. Jokaisen artikkelin ensimmäisen lauseen teemana toimii kyseisen artikkelin jo otsikossakin mainittu aihe ja reema puolestaan määrittelee tämän teeman, useimmiten parafraasia käyttämällä. Parafraaseihin palaan tarkemmin luvussa 5.2.6. Seuraava esimerkki (18) havainnollistaa artikkelien avauslauseiden yhtenäistä temaattista rakennetta:

18) **Fysioterapia** on [terapiaa](#), jonka keskeisenä tavoitteena on ylläpitää ja parantaa ihmisten toimintakykyä ja liikkumista.

Artikkelin otsikon käyttäminen teemana jokaisessa aineiston artikkelissa nopeuttaa tiedon löytämistä. Olenainen aiheen määrittely on löydettävissä heti ensimmäisestä virkkeestä, ja rakenne toistuu samanlaisena artikkelista toiseen.

5. KOHEESIO WIKIPEDIA-ARTIKKELEISSA

5.1 Kieliopillinen koheesio

5.1.1 Referenssi

Tässä luvussa tarkastelen keinoja, joilla tekstiin luodaan koheesiota referenssin avulla. Tutkin referenssiä tekstuaalisena ilmiönä eli keskityn vain niihin tapauksiin, joissa referenssin avulla sidotaan toisiinsa lauseita tai niitä laajempia kokonaisuuksia. Lauseiden sisäisen referenssin jätän huomiotta. Esimerkiksi Halliday painottaa nimenomaan virkkeiden välistä referenssiä tekstin koheesiota luovana keinona (Halliday 1985a: 312). Otan omassa analyysissäni kuitenkin huomioon myös yhdyslauseissa ilmenevän referenssin, sillä koen sen lisäävän aineiston tekstien koheesiota ja tätä kautta niiden sujuvuutta.

Aineiston artikkelit eroavat toisistaan esimerkiksi jo pituuden suhteen niin paljon, että koen laadullisen analyysin kertovan eniten niiden sisältämästä referenssistä. Analyysin pohjaksi olen kuitenkin laskenut aineiston sisältämien referenssitapausten määrän ja jakanut sen aineiston virkkeiden kokonaismäärällä. Näin saadaan selville, että keskimäärin 16 prosentissa aineiston virkkeistä ilmenee lauseiden tai virkkeiden välistä referenssiä. Kun referenssitapauksista ottaa huomioon ainoastaan ne, joissa referenssiä on ilmaistu virkkeiden välillä, saadaan tulokseksi prosenttiluku 8,9. Ainoastaan 8,9 prosentissa aineiston virkkeistä siis esiintyy referenssiä, jolla viitataan edeltävään virkkeeseen.

Kun tarkastellaan tarkemmin sitä, kuinka monessa prosentissa kunkin artikkelin virkkeistä ilmenee referenssiä, voidaan artikkelit jakaa karkeasti kolmeen kategoriaan. Ensimmäiseen kategoriaan sisältyvät ne artikkelit, joissa alle kymmenen prosenttia virkkeistä sisältää lauseiden tai virkkeiden välistä referenssiä ilmaisevan referenssikeinon. Tähän luokkaan kuuluvat lisäksi artikkelit biologisesta antropologiasta ja naisten äänioikeudesta, joista kummasakaan ei esiinny yhtään referenssitapausta. Toisen kategorian artikkeleissa referenssiä esiintyy jonkin verran, mutta se ei nouse olennaiseen asemaan koheesion luomisessa. Tähän luokkaan kuuluvat artikkelit, joiden virkkeistä 10–20 prosentissa ilmenee lauseiden tai virkkeiden välistä referenssiä. Kolmas luokka koostuu artikkeleista, joissa referenssiä on käytetty eniten eli yli 20 prosentissa virkkeistä.

Seitsemässä artikkelissa aineiston kolmestakymmenestä koheesiota ei luoda virkkeiden välisellä referenssillä juuri lainkaan. Yhdysvirkkeissäkään esiintyvää referenssiä näissä artikkeleissa ei paria poikkeusta lukuun ottamatta esiinny. Tämän kategorian tekstit ovat Särkän-

niemeä käsittelevää artikkelia lukuun ottamatta aineiston lyhimmästä päästä, ja niitä yhdistää luettelomaisuus ja rinnastaminen. Nämä ominaisuudet ilmenevät myös Särkänniemi-artikkelissa, jonka koheesiokeinojen vähyyteen osaltaan vaikuttanee artikkelin listausmainen kuvaus. Kategorian artikkeleissa käsiteltävästä ilmiöstä kerrotaan monesti luettelomaisesti ja asioita rinnastamalla ja pyritään näin mahdollisesti luomaan täsmällistä ja vakuuttavaa kuvaa. Tämä ilmenee esimerkiksi Siipeä käsittelevästä artikkelista, jossa virkkeet seuraavat toisiaan melko luettelomaisesti. Esimerkin (19) jälkimmäinen kappale koostuu ainoastaan yhdestä, lähinnä substantiiveja listaavasta virkkeestä.

19) **Siipi** on [lintujen](#) ja [hyönteisten](#) lenninelin. Lintujen siivet ovat [eturaajoja](#). Hyönteisten siivet, joita on tavallisesti kaksi paria, kiinnittyvät keskiruumiiseen. [Lepakoilla](#) on eturaajoissa siipimäiset lenninträpylät.

Lintujen siipiin kuuluvat [tertiaalit](#), [kyynärsulat](#), pienet [peitinhöyhenet](#), keskimmäiset peitinhöyhenet, isot peitinhöyhenet, [alula](#), [käsikulat](#) ja käsikulien peitinhöyhenet.

Yhdeksässä artikkelissa referenssiä ilmenee 10–20 prosentissa kyseisten artikkelien sisältämistä virkkeistä. Referenssin käyttäminen koheesion luomiseen on niissä satunnaista eikä sitä esiinny läheskään jokaisessa artikkelien luvussa, aluvussa tai kappaleessa. Kyseisiä artikkeleita yhdistääkin referenssin käytön satunnaisuus: sen sijaan, että referenssi olisi koko artikkelin läpi jatkuva koheesiokeino, sitä käytetään joitakin kertoja koheesion luomiseen. Tähän kategoriaan kuuluu pääasiassa pidempiä, useista luvuista koostuvia artikkeleita. Referenssi ei niissä siis ole vallitseva koheesiokeino, mutta tekstistä on pyritty tekemään kohesivista muilla keinoilla. Vaikuttaa siltä, että kyseisissä artikkeleissa korostuvat usein sanastollisen koheesion keinot. Esimerkiksi poroa käsittelevässä artikkelissa on käytetty erittäin runsaasti toistoa. Alla olevassa katkelmassa (esimerkki 20) hyödynnetään toistoa, mutta siinä esiintyy myös yksi tapaus virkkeiden välistä referenssiä: *poronsarveen* viitataan demonstratiivipronominilla *se*. Tämä tekee katkelman jälkimmäisestä kappaleesta heti ensimmäistä kappaletta sujuvamman.

20) Urosporon paino voi olla 90–180 kg ja lapakorkeus jopa 127 cm; naarasporo on jonkin verran pienempi, 60–100 kg.^[3] Vastasyntynyt vasa painaa 4–6 kg.^[3] Poron pituus on hännästä turvan kärkeen 150–210 cm.^[3] Poron elinikä voi olla 18–20 vuotta.^[3]

Molemmilla sukupuolilla on sarvet, jotka vaihtuvat vuosittain.^[4] Poronsarvi on maailman nopeimmin kasvavaa luuta. *Se* saattaa kasvaa pituutta vuorokaudessa 2 senttimetriä ja painoakin samassa ajassa puoli kiloa.^[3]

Kolmessatoista aineiston artikkelissa referenssiä ilmenee kunkin artikkelin virkkeistä yli 20 prosentissa. Tämän kategorian tekstit poikkeavat eniten toisistaan, sillä ne vaihtelevat jo pituudeltaan huomattavasti. Niihin kuuluu muun muassa televisiosarja Merlinistä kertova artikkeli, jossa on yksi ainoa referenssitapaus, mutta myös ainoastaan viisi virkettä. Toisaalta pidemmissä teksteissä, muun muassa Bahai-uskosta, Mikrobiologista ja Loch Nessin hirviöstä kertovissa artikkeleissa, referenssiä on käytetty koheesion luomiseen varsin monipuolisesti kautta koko tekstin. Niissä on hyödynnetty referenssiä lähes jokaisessa kappaleessa, lisäksi niissä on myös paljon virkerajat ylittävää referenssiä. Bahai-usko -artikkelista poimitussa esimerkissä (21) on käytetty monipuolisesti eri referenssikeinoja.

21) Bahai-usko pitää itseään erillisenä uskontona, mutta *siinä* näkyy islamin vaikutteita. Bahai-uskon mukaan Koraani tallettaa Jumalan sanan luotettavammin kuin Raamattu. Syynä *tähän* on, että Koraanissa on enemmän profeetta [Muhammadin](#) suoraa tekstiä kuin Raamatussa [Jeesuksen](#) omaa opetusta.^[22] *Toinen* syy on Koraanin myöhäisempi syntyäika. Baha'ullah viittasi kirjoituksissaan usein Koraanin jakeisiin, mutta muuten tekstisisältö erosi täysin. *Hän* viittasi myös [Uuteen testamenttiin](#), persialaisiin hengellisiin runoihin sekä [antiikin Kreikan](#) filosofiaan.^[26]

Aineiston artikkelit poikkeavat siis referenssin suhteen varsin paljon toisistaan. Joissakin referenssikeinoja ei ole käytetty lainkaan, mutta aineistoon mahtuu myös artikkeleita, joissa referenssiä on käytetty koheesion luomiseen runsaasti. Tähän väliin jää valtaosa aineiston artikkeleista: niissä referenssillä on luotu koheesiota joitakin kertoja. Osaksi variaatioon vaikuttavat jo mainitut erot artikkelien pituudessa: 165 virkkeen Bahai-artikkeliin luonnollisesti mahtuu enemmän referenssikeinoja kuin viidestä virkkeestä koostuvaan Merlin (tv-sarja) -artikkeliin. Tämä ei kuitenkaan selitä vaihtelua kokonaan, vaan joidenkin artikkelien kohdalla kysymys on enemmänkin siitä, että referenssin sijaan koheesion luomiseen on käytetty sanastollisen koheesion keinoja. Sanastollisen koheesion käyttö ilmenee jo esimerkissä 20 mainitun Poro-artikkelin lisäksi muun muassa Fani-artikkelista (esimerkit 22 ja 23), jossa on hyödynnetty erityisesti toistoa:

22) Sanaa *fani*⁵ käytettiin ilmeisesti ensimmäisen kerran yhdysvaltalaisessa lehdistössä 1800-luvulla kuvaamaan ammattilaisjalkapallojoukkueiden kannattajia.^[1] Nimitys on lähtöisin englannin sanasta *fanatic*⁶ eli kiihkoilija. Suomessa *fani*-termiä⁷ on käytetty lehdistössä ainakin 1960-luvulta lähtien kuvaamaan musiikkiyhtyeiden ihailijoita.^[2]

23) *Faniyhteisöjen* toiminnassa *fanituotannolla* on keskeinen merkitys: *fanit* luovat *fanitaidetta* ja kirjoittavat *fanifiktiota*, jotka perustuvat alkuperäistuotteen kierrättämiselle.

⁵ Alkuperäinen kursivointi

⁶ Alkuperäinen kursivointi

⁷ Alkuperäinen kursivointi

Olen analysoinut referenssiä tarkemmin mukaillen Hallidayn ja Hasanin jakoa kolmeen referenssityyppiin: persoonalliseen, demonstratiiviseen ja komparatiiviseen. Aineiston artikkeleista 15 sisältää persoonallista, henkilöön keskittyvää referenssiä eli referenssiä ilmaistaan persoonapronominilla. Se, että persoonallista referenssiä esiintyy vain puolessa aineiston artikkeleista, selittyyne osaltaan sillä, että artikkelien aiheet ovat paljolti asiakeskeisiä. Henkilöihin viittaamista esiintyy siis ylipäätään vähemmän. Toisaalta esimerkiksi näyttelijä Daniel Craigista kertovasta artikkelista persoonallista referenssiä löytyy henkilökeskeisyydestä huolimatta ainoastaan yksi tapaus (esimerkki 24). *hänellä* vaikuttaa tässä tapauksessa kaiuttavan Daniel Craigin omaa ääntä.

24) Daniel Craigin ilmoitettiin tehneen sopimuksen seuraavasta neljästä Bondista lokakuun loppulla 2007. Craig kuitenkin ilmoitti pian, ettei väite pidä paikkaansa ja että *hänellä* on sopimus vain seuraavasta Bondista. Craig kertoi aiemmin, ettei Ø aio näytellä viidessä James Bond -elokuvassa, mutta Ø on sittemmin muuttanut lausuntojaan ja Ø sanookin nyt haluavansa esittää kuuluisaa agenttia niin pitkään kuin mahdollista. Craigin kolmas James Bond -elokuva vahvistettiin loppuvuodesta 2008. Kolmas elokuva, [007 Skyfall](#)⁸, sai ensi-iltansa lokakuussa 2012. Craigilla on sopimus esittää Bondia vielä kahdesti.^[1]

Muuten näyttelijään viitataan artikkelissa leksikaalisen koheesion keinoilla, eniten toistoa käyttäen mutta myös yleismerkityksisillä sanoilla ja yläkäsitteillä. Lisäksi esimerkki (24) havainnollistaa nominaalisen ellipsin käyttöä. Näyttää siltä, että artikkelien asiakeskeiset aiheet eivät kokonaan selitä persoonallisen referenssin vähyyttä. Persoonallinen referenssi keskittyy artikkeleissa pääasiassa taustoittaviin osioihin, joissa kerrotaan jonkin asian tai ilmiön historiasta ja taustasta ja viitataan usein henkilöihin, jotka kyseistä asiaa tai ilmiötä ovat tutkineet tai kehittäneet. Sitä esiintyy esimerkiksi Mikrobiologiasta kertovan artikkelin historiaosuudessa (esimerkki 25).

25) Modernin mikrobiologian isä yhdessä Pasteurin kanssa oli saksalainen lääkäri [Robert Koch](#) (1843-1910). *Hän* tutki pernaruttoa ja onnistui vuonna 1876 ensimmäisen kerran kuvaamaan infektiotaudin vaiheet aukottomasti sekä mikrobin osuuden sen synnyssä. Tutkimiensa tautien perusteella *hän* laati [Kochin postulaatit](#).

Persoonallista referenssiä luodaan ylivoimaisesti eniten persoonapronominilla *hän* eri taivutusmuodoissaan. Monikon toista persoonaa käytetään vain kymmenkunta kertaa. Esimerkissä (26) viitataan *he*-pronominilla edellisessä lauseessa mainittuihin intiaaneihin. Seuraavassa esimerkissä (27) pronominin *he* avulla luodaan yhdyslauseiden välisen koheesion lisäksi virkkeiden välistä koheesiota.

⁸ Alkuperäinen kursivointi

26) Intiaanien [orjuus](#) kiellettiin vuonna 1758, mutta *heidän* sijaansa käytettiin mustia orjia, joita oli vuonna 1818 puolet asukasluvusta.

27) Eniten bahaita asuu [Intiassa](#), missä *heitä* on noin miljoona. Uskonnon synnyinmaassa Iranissa asuu noin 300 000–350 000 uskonnonharjoittajaa.^{[9][10]} Lisäksi *heitä* on huomattavasti [Latinalaisessa Amerikassa](#) ja osissa [Pohjois-Amerikkaa](#), mutta [Eurooppassa](#) suhteellisen vähän. Bahaita on vainottu monissa islamilaisissa maissa, erityisesti [Iranissa](#).^[8]

Demonstratiivista referenssiä esiintyy 24 aineiston artikkelissa. Se on selvästi yleisin aineistossa käytetty referenssin muodostuskeino. Ylivoimaisesti eniten käytetään demonstratiivipronominia *se* eri taivutusmuodoissaan. *Se*-pronominin viittauskohteena toimii useimmiten yksittäinen kielenaine. *Ne*-pronominia eri taivutusmuodoissaan esiintyy vähemmän, mutta sekin esiintyy aineistossa säännöllisesti. Alla olevat Shetlanninlammaskoira-artikkelista poimitut esimerkit (28 ja 29) havainnollistavat, miten demonstratiivipronomininit *se* ja *ne* tyypillisesti viittaavat aikaisemmin esiintyneeseen yksittäiseen kielenainekseen:

28) Shetlanninlammaskoira on luonteeltaan lempeä, mutta *se* voi olla vieraita kohtaan varautunut. Omaa perhettä kohtaan *se* on rakastavainen ja kiintyy perheeseensä voimakkaasti.

29) Shelteillä voi myös ilmetä joskus sydänsairauksia, mutta *niitä* on vain harvoilla.

Kuitenkin harvoissa tapauksissa *se*-pronominilla on aineistossa viitattu myös laajempaan kokonaisuuteen. Esimerkki (30) Unisieppari-artikkelista havainnollistaa, miten *se* viittaa edeltäviin lauseisiin.

30) Pajurenkaasta tehty unisieppari haurastuu ja hajoaa ajan myötä, ja *se* kuuluu asiaan, sillä myös lapsuus päättyy aikanaan.^[1]

Demonstratiivipronominin *tämä* käyttöalue on aineistossa *se*-pronominin käyttöaluetta laajempi. Sillä viitataan *se*-pronominin tavoin toisinaan yksittäisiin elementteihin (esimerkiksi osallistujiin, prosesseihin ja tilanteisiin), mutta hyvin usein myös niitä laajempiin kokonaisuuksiin. Hyvä esimerkki on Kuka pelkää mustaa miestä -artikkelin *tämä*-pronomini (esimerkki 31), joka viittaa koko edeltävään virkkeeseen.

31) Kiinniottajien huudettua "kuka pelkää mustaa miestä!" lähtevät vielä mukana olevat osantottajat jälleen juoksemaan viivalta toiselle. *Tämä* toistetaan niin monta kertaa, kunnes viimeinenkin osallistuja on jäänyt "mustaksi mieheksi".

Tämä-pronominia käytetään eri taivutusmuodoissaan artikkeleissa kuitenkin selvästi *se*-pronominia vähemmän. *Se*-pronominin käyttö on hyvin yleistä yhdyslauseissa ilmenevässä referenssissä, ja koska yli puolet aineiston referenssistä esiintyy juuri yhdyslauseissa, on luonnollista, että *se*-pronominia käytetään teksteissä runsaasti.

Demonstratiivista referenssiä ilmaistaan artikkeleissa myös proadverbien avulla, mutta niiden käyttö jää muutamaankin yksittäiseen tapaukseen. Proadverbeista eniten käytetään adverbejä *tällöin* ja *tuolloin*. Ison suomen kieliopin mukaan *tällöin* on temporaalinen adverbi, eli se viittaa aikaan. Selvä ajallinen viittaus on Loch Nessin hirviö -artikkelissa, mitä havainnollistaa seuraava esimerkki (32). *Tällöin* viittaa vuoteen 1975 eli edeltävää virkettä kauemmas.

32) Muutamat tutkijat ovat katsoneet valokuvien todistaneen Nessien olemassaolon. Vuonna [1975](#) brittiläinen luonnontieteilijä [Peter Scott](#) ehdotti sille [tieteellistä nimeä](#) *Nessiteras rhombopteryx*⁹. Nimi on [kreikkaa](#) ja tarkoittaa "timanttieväinen Nessin hirviö".^{[1][2]} *Tällöin* hänellä oli tarkoituksenaan myös saada laji Britanniassa [rauhoitettujen lajien](#) luetteloon.

Useimmissa aineiston *tällöin*-esiintymistä adverbillä on selvästi ajallinen merkitys. Seuraavassa esimerkissä (33) viittaussuhteen laatu on kuitenkin monimutkaisempi. Siinä *tällöin*-adverbillä alkava virke viittaa edelliseen virkkeeseen, mutta ajallinen jatkumo ilmenee siinä niin, että toimittaessa edellisessä virkkeessä kuvatulla tavalla, tapahtuu se, mitä *tällöin*-adverbillä alkava virke kertoo.

33) Perinteisesti MPN-menetelmä toteutettiin tavallisissa koeputkissa. Nykyään laskenta kuitenkin suoritetaan yleensä mikrotiiterilevyille. Muitakin vastaavia koeputkien korvaavia ratkaisuita on kehitetty. Yhteistä niissä on kuitenkin se, että näytetilavuutta ja työn määrää on vähennetty tuntuvasti. Monissa menetelmissä myös, kortti, kammio, levy tai muu kasvatusastia on kertakäyttöinen. *Tällöin* vältetään astioiden autoklavoinnilta ja tiskaukselta.

Myös *tuolloin* on temporaalinen adverbi. Sitä käytetään referenssikeinona neljää kertaa aineistossa, ja kaikissa tapauksissa sillä on selvästi ajallinen merkitys. Sillä viitataan joko vuosilukuun tai johonkin tiettyyn ajanjaksoon, kuten 1900-luvun alkuun seuraavassa esimerkissä (34):

34) Runoilija [Eino Leino](#) oli merkittävä 1900-luvun alun pakinoitsija. Suomalaisen pakinan päätyypeiksi olivat *tuolloin* nousseet kulttuuripakina, humoristinen pakina ja poliittinen pakina.

Lisäksi proadverbeista viittaamiseen käytetään aineistossa useampia kertoja myös paikkaan viittaavaa adverbiä *siellä*. Näin on muun muassa esimerkissä 35:

⁹ Alkuperäinen kursivointi

35) Gallen-Kallela etsi töilleen sopivia, aidon oloisia ympäristöjä kiertäen eri puolin Suomea. [Minna Canthin](#) neuvomana hän siirtyi perheineen kesäksi 1893 Savoan [Vehmersalmelle](#), Mustaharjun tilalle.^[3] *Siellä* hän maalasi Kalevalan kymmenennen runon inspiroimana Sammon taonta -teoksensa.

Kolmas tapa ilmaista referenssiä on komparaatio, mutta sen käyttö jää aineistossa erittäin vähäiseksi. Ainoastaan parissa tapauksessa proadjektiivia *sellainen* käytetään komparatiivisesti, kuten seuraavassa esimerkissä (36):

36) Oman kadenssin soittaminen klassisessa konsertossa on nykyään harvinaista, ja usein esimerkiksi Mozartin konsertoissa esitetään *Mozartin itsensä kadenssia* tai jos *sellaista* ei ole saatavilla, niin esimerkiksi Beethovenin kadenssia.

Tyypillisesti referenssin ajatellaan olevan koheesiokeino, joka identifioi jotakin viittamalla edellä mainittuun. Tällöin kyseessä on anaforinen sidos, eli viittauskohde on mainittu jo aikaisemmin tekstissä. Wikipedia-artikkelit sisältävät pääosin anaforisia sidoksia, mutta muutamassa artikkelissa esiintyy myös kataforinen sidos, jolloin viittauskohde on edessäpäin. Esimerkissä (37) Liikenne-artikkelin demonstratiivipronomini *se* viittaa sitä seuraavaan sivulauseeseen.

37) Maaliikenteen rajat asetti *se*, että lihasvoimaa parempia kuljetuskeinoja ei ollut vielä keksitty, joten maitse kuljetettiin edelleen vain pieniä yksiköitä ja lähinnä lyhyitä matkoja.

se toimii kataforisissa sidoksissa yleensä tukipronominina edellisen esimerkin mallin mukaisesti. Se voi joko esiintyä muodollisessa tehtävässä sijapäänteen kantajana tai, vaikkei sitä tarvittaisikaan ilmaisemaan sijapääntettä, se voi toimia definiittistäjänä eli viitata tuttuna (tai totena) esitettävään asiointilaan (VISK § 1145).

Wikipedia-teksteille vaikuttaa olevan tyypillistä myös se, että referentti on toisinaan epäselvä tai eri muodossa kuin sen viittauskohde. Tämänkaltaiset esiintymät nostavat esiin Wikipedia-tekstien prosessiluonteen. Esimerkissä (38) Leidenin pullosta kertovassa artikkelissa pronomini on yksikössä ja korrelaatti monikossa. Merkurius-esimerkissä (39) puolestaan pronominin *tämä* viittauskohdetta on vaikea tulkita. *tämä* näyttäisi viittaavan edellisiin virkkeisiin, mutta toisaalta *jopa noin 3 kilometriä* vaikuttaa semanttisesti vastakkaiselta ilmauksen *tasainen paikka* kanssa.

38) [PET](#)-muovipulloilla on 60 000 voltin läpilyöntikestokyky. Riittävän purkaustien säilyttämiseksi voidaan *se* täyttää vain noin 3/4 korkeuteensa.

39) Merkuriuksessa on myös vuoria, laaksoja, harjanteita, jyrkänteitä ja tasanko sekä eri tavoin heijastavia alueita. Kraattereista lähtee monesti säteittäisiä viiruja. Harjanteiden pituus on jopa satoja kilometrejä ja vuorten korkeus jopa noin 3 km. *Tämän* mukaan Merkurius on melko tasainen paikka.

Proadverbejä ja proadjektiiveja ei siis esiinny teksteissä systemaattisesti läpi aineiston, vaan ne jäävät joidenkin harvojen artikkelien referenssikeinoiksi. Erityisesti proadjektiivit toimivat kuitenkin aineistossa lisäksi määritteenomaisesti. Myös demonstratiivipronomineja käytetään aineistossa määritteenomaisesti. En laske näitä tapauksia varsinaiseen referenssiin kuuluviksi, sillä kyseisten määritteiden tehtävä on ensisijaisesti täydentää edussanaansa eikä niillä niin ollen ole selvää viittauskohdetta aikaisemmin tekstissä. Määritteinä käytetyillä pronomineilla ja proadjektiiveilla on kuitenkin oma selventävä funktionsa, ja koen niiden käytön lisäävän tekstin koheesiota. Siksi käsittelen niitä lyhyesti analyysissäni.

Kun demonstratiivipronominia tai proadjektiivia käytetään määritteenä, se ei siis ole samantarkoitteinen edellä esiintyneen lausekkeen kanssa, vaan sen tarkoite käy ilmi sen pääsanasta. Silloin sen tehtävä on useimmiten osoittaa, että se kertoo jotakin edellä mainittua. (Hakulinen & Karlsson 1995: 316–317.) Sillä voi olla täsmentävä tai tarkentava funktio (VISK § 564), mikä on erityisesti tietosanakirjatekstille ominaista. Artikkelimaisesti käytetyt *tämä* ja *tuo* voivat myös toisinaan nimetä uuden puheenaiheen tai näkökulman (VISK § 569), mikä korostaa niiden merkitystä koheesion luomisen kannalta. *tämä*-pronominia käytetään teksteissä varsin usein artikkelimaisesti. Esimerkissä (40) *tämä* korostaa artikkelimaisesti *törmääjää* ja samalla tarkentaa, että kysymys on edellisessä virkkeessä kuvatusta *suuresta kappaleesta*.

40) Merkuriuksen suuri rautaydin saattaa johtua siitä, että Aurinkokunnan syntyvaiheessa suuri kappale törmäsi planeettaan suurella nopeudella ja heitti osan planeetan silikaattivaipasta pois. *Tämän* törmäyksen massa olisi ollut arviolta kuudesosa Merkuriuksen massasta.

Määritteenomainen käyttö on tyypillistä myös teksteissä esiintyvälle proadjektiiveille. Proadjektiiveja käytetään aineistossa lähes ainoastaan määritteinä. Ainoa poikkeus on *sellainen*, joka esiintyy myös itsenäisesti. Seuraavassa esimerkissä (41) käytetään määritteinä proadjektiiveja *vastaavanlainen* sekä *tällainen*. *vastaavanlaisia ajatuksia* viittaa koko edeltävään kappaleeseen. Voidaan tulkita, että tässä tapauksessa *vastaavanlainen* täsmentää *ajatuksia*. Sama tarkoitus on proadjektiivilla *tällainen* myös *tällaisten ideoiden* kohdalla. Proadjektiivia *tällainen* käytetään määritteenä *ideoille*.

41) Näkymättömän käden mekanismi vaatii muutaman perusteen: hinnan, tehokkuuden, markkinat ja ahneuden. Yksilö tekee laadukasta työtä tehokkaasti, jolloin tuote on hyvä ja halpa. Hän saa myytyä sen markkinoilla kaikille osapuolille sopivaan hintaan. Ostaja saa tarvitsemansa hyödykkeen ja myyjä/tuottaja työstään korvauksen. Kaupassa kaikki saavat mitä haluavat.

Vastaavanlaisia ajatuksia olivat esittäneet jo muut valistusajattelijat kuten Antti Chydenius ennen Smithiä. Näistä tunnetuimpia on Bernard Mandeville. *Tällaisten ideoiden* laajempi poliittis-sosiaalinen konteksti on *sellainen*, että liberaalit ajattelijat halusivat osoittaa, että yhteiskunta toimii luhistumatta ilman vanhaa feodaaliajan hierarkkista järjestystäkin.

Fani-artikkelissa komparaatio-muoto esiintyy määritteenä, mutta se on siitä huolimatta välttämätön elementti merkityksen luomisen kannalta. Se kytkee toisiinsa edellisessä lauseessa määritellyn *faniuden* ja *uudemman selityksen*. Mikrobiologia-artikkelissa komparaatiota ilmaistaan määritteenomaisesti sanaa *muu* käyttämällä. Nämä esimerkit (42 ja 43) osoittavat selvästi artikkelinomaisen määritteen käytön kohesiivisen voiman.

42) Kun *fanius* nähdään oletetun sosiaalisen kyvyttömyden oireena, on "hulluja keräilijöitä" tai "villiintyneitä konserttiyleisöjä" helppo ymmärtää vaarallisina.^[3] *Uudempi* selitys korostaa faniutta kuluttajan aktiivisena kokemistapana.^[4]

43) Perinteisesti MPN-menetelmä toteutettiin tavallisissa koeputkissa. Nykyään laskenta kuitenkin suoritetaan yleensä mikrotiitterilevyille. *Muitakin vastaavia* koeputkien korvaavia ratkaisuita on kehitetty.

Referenssikeinojen määrä artikkelien välillä vaihtelee hyvin paljon. Ylivoimaisesti eniten referenssiä ilmaistaan aineistossa persoona- ja demonstratiivipronominien avulla. Proadverbejä käytetään joitakin kertoja demonstratiivisen referenssin luomiseen, mutta proadjektiivien käyttö jää muutamaankin tapaukseen. Komparaatiota ei aineistossa referenssin luomiseen juurikaan käytetä, jollei oteta huomioon tapauksia, joissa sitä on käytetty määritteenä.

5.1.2 Ellipsit ja substituutio

Ellipsi on koheesiokeino, jossa kahdella perättäisellä lauseella tai lauseenosalla on jokin yhteinen jäsen, joka toisessa osassa jätetään toistamatta (VISK § 1177). Poistettua kielenainesta ei ole korvattu toisella sanalla tai ilmauksella, vaan sen paikalla on aukko (Halliday & Hasan 1976: 145). Aineiston virkkeistä 12,2 prosenttia sisältää ellipsin. Laskutavasta on kuitenkin huomattava, että olen laskenut kaikki ellipsikeinot erikseen. Mikäli jossakin lauseessa ilmenee useampi ellipsikeino, olen laskenut kyseisessä lauseessa käytettävän ellipsiä useamman kerran. Näin on muun muassa esimerkissä (44), jossa on poistettu *antoi naisille äänioikeuden* eli yhdistetty nominaalinen ellipsi ja verbinosan ellipsi. Aineistoesimerkeissä olen merkinnyt aukko kohtaa symbolilla Ø.

44) Euroopan maista [Sveitsi](#) antoi naisille äänioikeuden vasta vuonna 1971 ja [Liechtenstein](#) Ø viimeisenä Euroopassa vuonna 1984.^[3]

Aineiston 28 artikkelissa esiintyy ainakin yksi ellipsikeino. Pääsääntöisesti yhdessä artikkelissa esiintyy kuitenkin vain muutama ellipsitapaus. Yli kymmenen ellipsikeinoa sisältäviä artikkeleita on aineistossa ainoastaan neljä. Tämän perusteella voidaan sanoa ellipsien olevan sikäli tyypillinen koheesiokeino aineistossa, että se esiintyy lähes jokaisessa artikkelissa, mutta ei ole missään niistä kovin vallitseva.

Virkkeiden välisiä ellipseja ei aineistossa käytännössä esiinny lainkaan, mutta virkkeiden väliset viittaussuhteet eivät ole ellipseille yleisestikään ottaen kovin tyypillisiä (VISK § 1177). Rinnasteisissa ja alisteisissa lauseissa ellipsit sen sijaan luovat aineiston teksteihin koheesiota. Nominilausekkeiden sisällä esiintyviä ellipseja en ole tutkinut, koska ne eivät ole olennaisia tekstuaalisia elementtejä.

Olen analysoinut ellipseja Ison suomen kieliopin jaottelun mukaan eli jakanut ne nominaalisiin ellipseihin ja verbinosan ellipseihin. Eniten aineistossa esiintyy nominaalisia ellipseja, eli rinnasteisesta lauseesta on poistettu substantiivilauseke. Näitä esiintymiä aineistossa on 110. Substantiivilauseke voi olla yksittäinen substantiivi (tai joissakin tapauksissa pronomini tai numeraali), mutta usein substantiivilausekkeen pääsanaan liittyy määritteitä, kuten esimerkiksi (45). Siinä koheesiota on luotu poistamalla substantiivilauseke jälkimmäisestä lauseesta, sillä subjekti on sama kuin edeltävässä lauseessa. Yhteisen subjektin poistaminen jälkimmäisestä lauseesta on tyypillisin tapa käyttää ellipseja. Näitä tapauksia esiintyy aineistossa runsaasti.

45) [Danzigilainen](#) fyysikko [Daniel Gralath](#) kytki ensimmäisenä useita Leidenin pulloja rinnakkain ja Ø sai näin korotettua tehoa.

Subjekti ei kuitenkaan ole ainoa lauseenjäsen, joka nominaalisissa ellipseissa jätetään toistamatta. Esimerkissä (46) on subjekti *osa* toistettu myös jälkilauseessa, mutta siihen kiinteästi liittyvä *onnettomuuksista* on sitä vastoin jätetty toistamatta:

46) Suurin osa *onnettomuuksista* johtuu inhimillisistä arviointi- ja havaintovirheistä, mutta osa Ø johtuu myös puutteellisista liikennejärjestelyistä ja huonoista kelioloista.

Pois jätettävänä substantiivilausekkeena voidaan käyttää myös pronominia. Tämä tapaus esiintyy aineistossa muutaman kerran, muun muassa Mikrobiologiaa käsittelevässä artikkelissa (esimerkki 47).

47) Siellä vasta [John Tyndall](#) todisti ilmassa olevan hiukkasia, jotka heijastavat valoa ja Ø aiheuttavat mikrobikasvua.

Ellipsi on toisteisuutta vähentävä koheesiokeino, mikä on selvästi nähtävissä myös edeltävistä esimerkeistä. Sen käyttö sujuvoittaa tekstiä, erityisesti edeltävän esimerkin (47) kaltaisissa tapauksissa, joissa on peräkkäin useita lyhyitä lauseita. Vaikka se vaatii toistoa tarkempaa tekstin lukemista, sen viittauskohde on kuitenkin aineistossa yleensä saman virkkeen sisällä, eikä se siis vaadi lauseita laajempien kokonaisuuksien kytkemistä toisiinsa.

Verbilausekkeissa esiintyviä ellipsejä löytyy aineistosta selvästi nominaalisia ellipsejä vähemmän. Verbinosan ellipsejä esiintyy aineistossa 45 eli puolet vähemmän kuin nominaalisia ellipsejä. Verbinosan ellipsinä toimii käytännössä aina finiittiverbin poistaminen rinnasteisen yhdyslauseen jälkimmäisestä osalauseesta. Esimerkissä (48) on poistettu finiittimuotoinen *olla*-verbi.

48) Syynä tähän on, että Koraanissa on enemmän profeetta [Muhammadin](#) suoraa tekstiä kuin Raamatussa Ø [Jeesuksen](#) omaa opetusta.^[22]

Aineistolle erittäin tyypillistä on, että nominaalinen ellipsi ja verbin ellipsi esiintyvät yhdessä. Esimerkin 45 mukaisia tapauksia, joissa käytetään ainoastaan nominaalista ellipsejä, löytyy aineistosta jonkin verran, mutta pelkkää verbinosan ellipsin käyttöä esiintyy selvästi vähemmän. Yleisestä kielitutkinnosta poimitussa esimerkissä (49) näkyy, miten nominaalisen ellipsin ja verbinosan ellipsin käyttäminen yhdessä vähentää tehokkaasti toiston tarvetta. Jälkimmäisessä esimerkissä (50) on niin ikään yhdistetty nominaalinen ellipsi ja verbinosan ellipsi. Tässä tapauksessa ei ole poistettu lauseen subjektia, vaan objekti.

49) Perustasoon kuuluvat tasot 1 ja 2, keskitasoon Ø 3 ja 4 sekä ylimpään tasoon Ø 5 ja 6. Kukin tasoon tarvittavat taidot on määritelty tasojen kuvauksissa.^[3]

50) [Tukkijoki \(1982\)](#) – laite on Suomen ensimmäinen vesivuoristorata ja vesilaite. Sen on valmistanut ja Ø rakentanut Tampereen kaupunki.

Edeltävistä esimerkeistä voi havaita, että aineistossa käytetään ellipsejä monesti rinnasteisissa ja luettelomaisissa virkkeissä. Muutamassa artikkelissa kokonaiset luvut on kirjoitettu luettelomaiseen muotoon, ja näissä luvuissa elliptisiä lauseita esiintyy runsaasti. Näin on muun muassa Mikrobiologiaa käsittelevän artikkelin Mikrobiologian osa-alueita -luvussa ja

Särkänniemi-artikkelin Kohteet-luvussa, joka on käytännössä lista Särkänniemen huvilaitteista. Jet Star -huvilaitteesta kertovassa esimerkissä (51) ellipsejä käytetään myös virkkeiden välillä. Ellipsien runsas käyttö ei kuitenkaan enää lisää sujuvuutta, vaan tekstistä muodostuu melko kömpelöä.

51) Jet Star (Laite kiihdytti heti lähdössä 60 km tuntinopeuteen. Aiemmin laitteessa ei ollut turvavyötä, sillä keskipakovoima piti kiinni penkissä. Turvavyöt lisättiin vuonna 2008 turvallisuussyistä. Jet Star reistaili kesällä 2012 kunnes Ø hajosi heinäkuun lopulla 2012 ja Ø oli poissa käytöstä koko loppukauden. Lopulta Ø purettiin huhtikuussa 2013 ja Ø romutettiin)

Additiiviset eli listaavat kytkökset (erityisesti lauseiden yhdistäminen toisiinsa *ja*-konjunktion avulla) ovat aineistossa hyvin yleisiä (katso tarkemmin luku 5.1.3), ja ne ovat kiinteästi yhteydessä ellipsien käyttöön. Ellipsit ilmenevät monesti yhdyslauseissa, jotka on yhdistetty toisiinsa nimenomaan *ja*-konjunktion avulla. Ellipsien melko runsas käyttö johtuu siis osittain Wikipedia-artikkeleille muutenkin tyypillisestä rinnasteisuudesta. Ellipsikeinoista aineistossa käytetään selvästi eniten nominaalista ellipsiä. Varsin yleistä on myös verbinosan ellipsin käyttäminen yhdessä nominaalisen ellipsin kanssa. Nämä keinot vähentävät aineiston toisteisuutta.

Neljättä kielipiillistä koheesiokeinoa eli substituutiota ei aineistossa esiinny kahta tapausta lukuun ottamatta lainkaan. Selvästikään jonkin aineksen korvaaminen toisella ei ole keino, jolla koheesiota Wikipedia-artikkeleissa muodostetaan. Yleisesti ottaen substituutiota luodaan proadjektiivilla *sellainen*, ja sitä on käytetty myös tässä aineistossa pari kertaa. Esimerkiksi Konsertto-artikkelissa on tapaus, jossa *sellainen* korvaa nominilausekkeen (katso esimerkki 4).

5.1.3 KytKentä

Koheesiota luodaan aineistossa varsin paljon myös kytkennän avulla, eli lauseita ja virkkeitä kytketään toisiinsa konjunktioita ja konnektiiveja käyttämällä. Tarkastelen analyysissäni, millaisia systemaattisia suhteita aineiston lauseiden ja virkkeiden välillä ilmenee. Jaottelen konjunktioita ja konnektiiveja semanttisesti sen mukaan, minkä merkityssuhteen ne luovat kytkemiensä lauseiden välille. KytKentäkeinot jaan Ison suomen kieliopin jakoa mukailleen additiivisiin, disjunktiiivisiin, kontrastiivisiin, kausaalisiin, konsessiivisiin, konditionaalisiin, finaaliin ja temporaalisiin.

Konnektiiveista on hyvä huomata, että niiden luokka ei ole samalla tavalla rajattu kuin konjunktioiden luokka, mikä hankaloittaa niiden luokittelamista. Olen nojannut konnektiivien luokittelussa Ison suomen kieliopin määrittelyyn niiden suhteista. Aineistossa esiintyy kuitenkin

kin lähinnä yksittäistapauksina konnektiiveja, joita Iso suomen kielioppi ei mainitse. Nämä konnektiivit olen sijoittanut oman harkintani mukaan siihen semanttiseen luokkaan, jonka merkityssuhdetta ne näyttävät ilmentävän.

Aineiston laaja variaatio ilmenee myös kytkentää tarkasteltaessa. Kaikista artikkeleista löytyy joitakin kytkenän keinoja, mutta muutamassa tekstissä niitä esiintyy vain yksi tai kaksi. Kuitenkin erityisesti pidemmissä teksteissä kytkentää on käytetty koheesion luomiseen varsin laajasti. Lähes kaikissa artikkeleissa kytkentää luodaan sekä konjunktioiden että konnektiivien avulla.

Iso suomen kielioppi määrittelee additiivisen suhteen yhdeksi tekstuaalisesti merkitsivimmäksi kytkenän keinoksi (VISK § 1092). Wikipedia-aineistossa esiintyykin selvästi eniten juuri additiivista kytkentää. Sitä luodaan erityisesti *ja*-konjunktin sekä *myös*-partikkelin avulla. *ja*-konjunktilla kytketään lauseita tai virkkeitä toisiinsa jokaisessa aineiston artikkeleissa. Leidenin pulloa käsittelevässä artikkelissa *ja*-konjunktilla on luotu rinnakkaista suhdetta kahdessa peräkkäisessä virkkeessä. Tätä havainnollistaa seuraava esimerkki (52):

52) Lopullisen muodon Leidenin pullolle antoivat [1748 lontolaiset](#) lääkärit [William Watson](#) ja [John Bevis](#). He luopuivat nesteistä *ja* päällystivät pullon sisältä ja ulkoa metallisella [foliolla](#). [Danzigilainen](#) fyysikko [Daniel Gralath](#) kytki ensimmäisenä useita Leidenin pulloja rinnakkain *ja* sai näin korotettua tehoa.

Additiivista kytkentää ilmaistaan *ja*-konjunktin lisäksi myös muilla konjunktioilla, mutta nämä esiintymät ovat aineistossa huomattavasti harvinaisempia. *sekä* esiintyy aineistossa kyllä usein sanatason listauksissa, mutta kokonaisia lauseita sillä ei juuri kytketä toisiinsa. Poikkeuksena toimii Bahai-usko-artikkeli, jossa kytkentäkeinoja muutenkin käytetään erittäin monipuolisesti. Siinä yhdistetään *sekä*-konjunktilla lauseita, joista on aukkoamalla poistettu finiittiverbi. Tämä ilmenee muun muassa seuraavassa esimerkissä (53):

53) Uskonto kärsi alkuvuosinaan ankarista vainoista, mutta ne laantuivat vähitellen. Muutos huonompaan tapahtui jälleen vuonna 1979, kun Iranissa tehtiin [islamilainen vallankumous](#). Tämän jälkeen satoja bahailaisia on [kidutettu](#) ja [teloitettu](#) *sekä* tuhansia [vangittu](#).^[47]

Bahai-usko -artikkelille vastakkainen on fysioterapiaa käsittelevä artikkeli, jossa kytkentää on puolestaan käytetty hyvin vähän. *ja*-konjunktia käytetään rinnasteisuuden luomiseen, mutta muuten tekstissä ei käytetä konjunktioita yhdistämään lauseita tai virkkeitä toisiinsa, ja konnektiiveistakin käytetään vain kerran additiivista *myös*-partikkelia. Koska referenssiäkään ei juuri esiinny, ilmaistaan koheesiota artikkelissa lähinnä teemarakenteella ja

virkkeiden järjestyksellä. Lisäksi fysioterapiaan ja terveydenhoitoon kiinteästi liittyvä sanasto muodostaa merkityskentän, joka sitoo virkkeitä yhteen. Tätä havainnollistaa seuraava esimerkki (54), joka on keskimäinen kappale Fysioterapia-artikkelista. Siitä ovat nähtävissä koko artikkelin ainoat kytkentäkeinot eli *ja*-konjunktio ja *myös*-konnektiivi, jotka molemmat ilmaisevat additiivista suhdetta. Lisäksi kieliopillista koheesiota ilmaistaan ellipsisillä. Hyvin kohesteisen tekstin eräs merkki on se, ettei sen virkkeiden järjestyksestä voi vaihtaa. Tässä kappaleessa koheesion keinoja on kuitenkin käytetty niin vähän, että virkkeiden uudelleen järjestäminenkin onnistuisi.

54) Fysioterapiassa arvioidaan kuntoutujan terveyttä, liikkumista, toimintakykyä ja toimintarajotteita hänen toimintaympäristössään. Arvioinnin perusteella laaditaan fysioterapiasuunnitelma, jota *myös* seurataan. Fysioterapeutti arvioi myös apuvälinetarvetta *ja* \emptyset ohjaa niiden käyttöä. Kuntoutujaa aktivoidaan *ja* \emptyset ohjataan ottamaan itse vastuuta omasta terveydestään, toimintakyvystään ja niiden edistämiseen liittyvistä valinnoista yhteistyössä muiden asiantuntijoiden kanssa.

Edellä käsitellyt esimerkkiartikkelit Bahai-uskosta ja fysioterapiasta toimivat hyvänä havainnollistuksena niin kytkennän kuin muidenkin koheesiokeinojen käytön variaatiosta aineistossa ylipäätään. Joissakin artikkeleissa kieliopilliset koheesiokeinot ovat lähinnä poikkeuksia tai jäävät muutamaa esiintymään, kun taas joissakin niitä käytetään erittäin monipuolisesti.

Additiivisiin konnektiiveihin kuuluva *myös* esiintyy lähes jokaisessa aineiston artikkelissa. Se on selvästi yleisin konnektiivi aineistossa. Niin ikään additiivinen *lisäksi* esiintyy kymmenessä artikkelissa. Sillä luodaan enemmän virkkeiden välistä koheesiota kuin konnektiivilla *myös*, eli se kytkee tekstin osia toisiinsa tätä vahvemmin. Muutama esiintymä aineistossa on myös additiivisilla *vielä*, *jopa* ja *edes*. Additiivista suhdetta kutsutaan myös listauksiksi suhteeksi (VISK § 1092). Rinnastus ja listaukset ovat Wikipedia-artikkeleille tyypillisiä, mikä selittää osaltaan additiivisen kytkennän yleisyyttä aineistossa.

Kontrastiivinen kytkentä on on additiivisen jälkeen aineiston yleisin kytkentäkeino. Lähinnä *mutta*-konjunktioilla kytketään yhteen lauseita, joiden eroja halutaan korostaa tai joista halutaan esittää jokin poikkeus, kuten seuraavan esimerkin (55) Unisiepparista kertovassa virkkeessä.

55) Unisiepparin alkuperäinen tarkoitus on suodattaa pahat [unet](#) ja ajatukset päästäen lävitseen hyvät unet ja ajatukset,^[4] *mutta* nykyisin se on usein koriste-esine.

Kontrastiivista suhdetta ilmaistaan myös *vaan*-konjunktioita käyttämällä, mutta aineistossa sitä käytetään ainoastaan kahdessa tekstissä. Bahai-uskosta kertovassa artikkelissa sitä kuitenkin

käytetään useasti ilmaisemaan vastakkaisuutta. Tämä on eräs Wikipedia-artikkeleille tyypillinen ilmiö: jokin kielen keino ilmenee vain yhdessä artikkelissa, mutta saattaa ilmetä siinä useita kertoja. Se lienee siis juuri kyseisen artikkelin kirjoittajalle tyypillinen. Ilmiö kertoo Wikipedian vapaasta tuotantotavasta johtuvasta variaatiosta. Bahai-uskoa käsittelevässä artikkelissa *vaan*-konjunktioita on käytetty kahdessa peräkkäisessä virkkeessä (esimerkki 56):

56) Bahait uskovat, että jokaisella on [kuolematon sielu](#). Se on ihmisen henkinen osa, joka ei kuole *vaan* jatkaa olemassaoloaan toisessa muodossa. Sielu ei asu ruumiissa, *vaan* se yhdistyy ihmiseen heti [hedelmöitymisvaiheessa](#).

Kontrastiivista suhdetta ilmaistaan myös konnektiiveilla, mutta mitään kontrastiivisista konnektiiveista ei käytetä aineistossa muutamaa tapausta enempää. Eniten käytetään fokuspartikkeliä *vain*. Se esiintyy ilmauksissa, joissa esitetään yleistys, johon *vain*-jakso esittää poikkeustapauksen tai rajauksen (VISK § 1110). Se on yleinen lauseketasolla aineistossa, mutta tapauksia, joissa *vain*-jakso on lauseketta pidempi, ei juuri esiinny. *vain* ei siis ole kovin vahvasti koheesiota luova keino teksteissä. Kontrastiivinen *taas* esiintyy aineistossa joitakin kertoja. Useamman kerran se ilmaisee kontrastiivista suhdetta yhdessä *kun*-konjunktin kanssa. Näin on muun muassa seuraavassa esimerkissä (57), jossa se esittää poikkeustapauksen kiteymänä *kun*-konjunktin kanssa. Esimerkki havainnollistaa myös sitä, miten listausta käytetään asioiden määrittelyyn Wikipedia-artikkeleissa.

57) Ultraemäksiset kivilajit eivät sisällä juuri ollenkaan piidioksidia. Vallitsevina mineraaleina ovat mustat taikka tummanvihreät, magnesium- ja rautarikkaat mineraalit kuten [pyrokseeni](#) ja [oliviini](#). [Pyrokseeniitit](#), [peridotiitit](#) ja [duniitit](#) ovat ultraemäksisiä syväkivilajeja, *kun taas* [komatiitit](#) ja [pikriitit](#) ovat ultraemäksisiä vulkaanisia [pintakiviä](#).

toisaalta esiintyy vain neljässä eri artikkelissa, mutta sillä luodaan lähes poikkeuksetta kontrastiivista, virkerajat ylittävää kytkentää. Seuraava esimerkki (58) on hyvä osoitus siitä, miten *toisaalta* luo tehokkaasti koheesiota ja sitoo kaikki kappaleen virkkeet yhteen. Esimerkissä esiintyy myös *kuitenkin*-partikkeli, mutta sen kohesiivinen voima on vähäisempi.

58) Monet nykytutkijat pitävät *kuitenkin* mahdollisena, ettei koko ristiretkeä koskaan tehty. Tällöin kyseessä olisi myöhemmin seipitetty taru, jolla pyhimyksenä palvotulle Eerik-kuninkaalle rakennettiin sankarin mainetta. *Toisaalta* on mahdollista, että ristiretkitaru perustuu johonkin todelliseen tapahtumaan, jota legendassa on paisuteltu ja dramatisoitu.

Muutamia kertoja aineistossa käytetään myös konnektiiveja *puolestaan*, *sen sijaan*, *paitsi* ja *päinvastoin* luomaan kontrastiivista kytkentää. *päinvastoin* voi muodostaa lausekkeen myös yksin – tästä on aineistossa kaksi esimerkkiä.

Vaihtoehtosuhde eli disjunkttiivinen suhde lauseiden tai virkkeiden välillä ei ole aineistossa tyypillinen kytkentäkeino. Lauseiden välistä *eli*-konjunktiolla ilmaistua disjunkttiivista kytkentää esiintyy ainoastaan neljässä artikkelissa. Ison suomen kieliopin mukaan *eli* esiintyy yleensä parafrasain tai vaihtoehtoisen näkökulman merkitsemänä (VISK § 1098), ja tämä funktio sillä on myös aineistossa. Aineistosta löytyy joitakin lauseiden välistä vaihtoehtosuhdetta ilmentäviä tapauksia. Esimerkissä (59) ensimmäinen lause ilmaistaan toisin sanoin *eli*-konjunktion jälkeisessä lauseessa.

59) Kalenteri perustuu aurinkovuoteen, eli siinä on 365 päivää, 5 tuntia ja noin viisitoista minuuttia. [34]

Vaikka lauseiden välinen *eli*-konjunktion käyttö on aineistossa melko harvinaista, käytetään *eli*-konjunktiota runsaasti sanojen ja lausekkeiden välillä. Artikkeleille on tyypillistä, että ilmiöiden määrittelemisessä hyödynnetään usein *eli*-konjunktiolla yhdistettyjä parafraseja. Ilmiö ei ole kovin oleellinen koheesion muodostumisen kannalta, mutta se on aineistossa hyvin tyypillinen ja havainnollistaa sitä, miten asioita ja ilmiöitä pyritään määrittelemään Wikipedia-artikkeleissa. Tätä voi pitää tietotekstilte tyypillisenä. *Eli*-konjunktiota on käytetty esimerkiksi Loch Nessin hirviön määrittelyssä artikkelin avauslauseessa (esimerkki 60):

60) Loch Nessien hirviö *eli* Nessie on järvihirviö *eli* tunnistamatonta lajia edustava suuri eläin, joka nykykansanperinteen mukaan elää Loch Ness -järvessä Skotlannissa.

Vain hieman *eli*-konjunktioiden lauseiden välistä käyttöä yleisempää on lauseiden välinen *tai*-konjunktiolla ilmaistu disjunkttiivinen kytkentä. Muutamassa tapauksessa konjunktiota seuraava lause on elliptinen. Näin on muun muassa esimerkissä (61), jossa *tai*-pronominilla ilmaistaan disjunkttiiviselle käytölle tyypillistä vaihtoehtoista mahdollisuutta.

61) Kaksisiipisillä sekä joillain päiväkorennoilla takasiivet ovat kehityksen aikana surkastuneet *tai* Ø kokonaan hävinneet.

Disjunkttiivista suhdetta voidaan ilmaista myös *muuten*-konnektiivilla. Kyseistä keinoa on käytetty kolmessa aineiston artikkelissa. Seuraavassa esimerkissä (62) on hyödynnetty disjunkttiivisen suhteen ilmaisemiseen sekä *eli*-konjunktiota että *muuten*-konnektiivia.

62) Voi olla mahdollista, että kehittynyt tietoliikenne voisi pitkälti korvata ihmisten viestintätarpeesta johtuvan liikkumisen *eli* [puhelimella](#), [faksilla](#) tai *muuten* sähköisesti kommunikoitaisiin henkilökohtaisen tapaamisen sijaan.

Kausaalisen suhteen ilmaiseminen on Wikipedia-artikkeleissa verrattain yleistä. Suhdetta ilmaistaan suunnilleen yhtä paljon *koska*- ja *sillä*-konjunktiolla. Myös *kun*-konjunktiota voidaan käyttää kausaalisen suhteen ilmaisemiseen, mutta tätä ei artikkelissa juuri esiinny. Syiden ja seurausten esittäminen kausaalisten kytkentöjen avulla vaikuttaa olevan yleistä erityisesti Leidenin pullosta, Merkuriuksesta ja Mikrobiologiasta kertovissa artikkeleissa eli aineiston pidemmissä luonnontieteellisissä artikkeleissa. Esimerkissä (63) *koska*-konjunktiolla alkavassa lauseessa esitetään perustelu sitä seuraavalle lauseelle.

63) *Koska* värjäys on voimakasta, sitä pitää laimentaa melko paljon ennen preparaatin valmistamista, jotta yksittäiset solut eivät ”huku hohtavansiniseen massaan”.

Bahai-usko-artikkelista poimitussa esimerkissä (64) puolestaan perustellaan jälkimmäisellä lauseella päälauseen asiasisältöä:

64) Käännytystyössä henkinen tai fyysinen painostaminen on kielletty, *sillä* jokaisen tulee ymmärtää totuus itsenäisesti. [39]

Syysuhteita ilmaistaan myös konnektiiveilla. Adpositioita *takia* ja *vuoksi* käytetään aineistossa melko paljon. Jonkin verran esiintyy myös adpositioita *johdosta* ja *ansiosta*. Kahdesti käytetään myös *siksi*-konnektiivia, jolla ilmaistaan edellisessä lauseessa esitetyn asiointilan seuraus syynä tai perusteluna (VISK § 1131).

Temporaalista suhdetta ilmaistaan aineiston teksteissä lähinnä *kun*-konjunktion avulla. Temporaalista kytkentää esiintyy kuudessa aineiston artikkelissa. Olen luokitellut *kun*-konjunktiiviset tapaukset pääasiassa temporaaliseen kytkentään, mutta aineistossa on joitakin rajatapauksia, jotka voi nähdä myös kontrastiiviseen tai kausaaliseen kytkentään kuuluviksi. *kun taas* -esiintymät olen luokitellut kontrastiivisiksi sidoksiksi. Esimerkissä (65) ensimmäisen lauseen asiantila esitetään selvästi ehtona toiselle, mikä puoltaa sen luokittelamista kausaaliseen kytkentään. *Kun*-konjunktio tuo kuitenkin virkkeeseen temporaalisuudelle olennaisen aikasuhteen, minkä vuoksi olen päättänyt luokittelemaan tapauksen temporaaliseen kytkentään.

65) Kun liikenteen valtiollisia rajoituksia vähennetään, mahdollistaa koko maapallon kattava nopea, laajakapasiteettinen ja kattava liikenneverkko muun muassa nykyisen globalisaation.

Selvästi temporaalinen suhde sen sijaan on seuraavassa virkkeessä. Virke on artikkelista Ensimmäinen ristiretki Suomeen, joka on aineiston ainoa varsinaisesti historiallinen artikkeli.

Historiallisia osuuksia, lukuja ja alalukuja tosin esiintyy muissakin artikkeleissa. *Kun*-konjunktioilla ilmaistua temporaalista suhdetta esiintyykin muutamassa muussakin artikkelissa nimenomaan historiallisista tapahtumista kerrottaessa. Muutoin sen käyttö jää aineistossa varsin vähäiseksi. Esimerkissä (66) aikasuhte tapahtumien välillä ilmenee selvästi.

66) *Kun* kuningas palasi takaisin Ruotsiin, piispa jäi järjestelemään Suomen oloja.

Konditionaalista suhdetta aineistossa ilmaistaan vähän. Ehdon ilmaiseminen ei selvästikään ole olennaista Wikipedia-artikkeleissa. Koska artikkelien teksti on pääosin määrittelyä, rinnastavaa ja kertovaa, ei ole yllättävää, että konditionaalisen kytkennän ilmaiseminen on vähäistä. Muutaman kerran aineistossa tosin ilmaistaan ehtoa *jos*-konjunktioilla. Näin on esimerkiksi (67), jossa *jos*-konjunktioilla alkava lause ilmaisee ehtoa, jonka toteutuessa päälauseen asiantila toteutuu.

67) Suurin osa liikennevälineistä polttaa hiilivetyjä. *Jos* nämä palavat vain osittain, syntyy saasteita.

Konditionaalista suhdetta ei ilmaista myöskään konnektiiveilla käytännössä lainkaan.

Konsessiivisessa suhteessa kahden lauseen asiantilojen sanotaan pätevän, mutta samaan aikaan niiden yhteensopivuus voidaan käsittää kontekstuaalisesti epäodotuksenmukaiseksi (VISK § 1139). Konsessiivista suhdetta ilmaistaan konjunktioilla aineistossa varsin vähän, mutta joissakin tapauksissa sitä ilmaistaan *vaikka*-konjunktioilla. Esimerkissä (68) olisi kontekstuaalisesti epäodotuksenmukaista, että vähäravinteisella tasangolla olisi rehevää kasvillisuutta, mutta kuumen ja kostean ilmaston takia näin kuitenkin on.

68) Amazonasin alangolla on erittäin rehevä kasvillisuus kuumen ja kostean ilmaston takia, *vaikka* maaperä on pääosin vähäravinteista ja hapanta.

Vaikka konjunktioilla ilmaistava konsessiivinen kytkentä on aineistossa harvinaista, ilmaistaan konsessiivista suhdetta kuitenkin varsin paljon konnektiivien avulla. Konsessiivisilla konnektiiveilla *silti*, *tosin*, *joka tapauksessa* ja *siitä huolimatta* kytkentää ilmaistaan joitakin kertoja, mutta *kuitenkin* esiintyy aineistossa 14 eri artikkelissa ja kielteinen *kuitenkaan* kuudessa eri artikkelissa. Kaksi kontekstuaalisesti epäodotuksenmukaista asiantilaa yhdistetään esimerkiksi (69) partikkelilla *kuitenkin*. Pitkää karva ja helppohoitoisuus ovat asiantiloja, joita ei yleensä mielletä yhteen, mutta *kuitenkin* liittävät nämä epäodotuksenmukaiset tilat toisiinsa.

69) Shetlanninlammaskoiralla on pitkä karva, joka on *kuitenkin* suhteellisen helppohoitoinen: se ei tuuheudestaan huolimatta takkuunnu eikä likaannu helposti.

Vähäistä aineistossa on niin ikään finaalin kytkentä. Sitä ilmaistaan joitakin kertoja *jotta*-konjunktiota käyttäen. Näin tapahtuu muun muassa Yleinen kielitutkinto -artikkelissa (esimerkki 70).

70) Ulkomaalaisen on saavutettava vähintään taso 3 suomen tai ruotsin kielessä, *jotta* hän voi hakea Suomen kansalaisuutta. ^[1]

Analyysin perusteella voidaan päätellä, että selvästi eniten aineistossa muodostetaan additiivista kytkentää, niin konjunktioiden kuin konnektiivienkin avulla. Seuraavaksi eniten esiintyy kontrastiivista ja kausaalista kytkentää. Rinnastuksien lisäksi siis poikkeuksista kertominen ja syy-seuraussuhteen esittäminen voidaan nähdä melko tyypillisiksi keinoiksi rakentaa Wikipedia-artikkelien tietotekstiä. Muut kytkentäsuhteet eivät ole vallitsevia aineistossa, vaan esiintyvät ainoastaan joissakin artikkeleissa. Niiden ajoittainen esiintyminen kuitenkin osoittaa, että Wikipedia-artikkeleissa käytetään melko monipuolisesti erilaisia systemaattisia suhteita liittämään ilmaistuja merkityksiä toisiinsa. Kytkentäkeinojen valinta riippuu myös tekstin aiheesta: esimerkiksi temporaalista kytkentää käytetään enemmän artikkelien historiallisissa osuuksissa.

5.2 Sanastollinen koheesio

Tässä luvussa erittelen kvalitatiivisesti aineistoni sanastollista koheesiota. Aineiston artikkeleista on löydettävissä useita sanastollisen koheesioilajeja. Analyysissäni kiinnitän huomiota toistoon, yläkäsitteisiin, yleismerkityksiin sanoihin sekä siihen, miten toisiinsa läheisesti liittyvät sanat muodostavat merkityskenttiä ja näkyvät aineistossa synonymian, antonymian, meronymian ja parafrasien kautta.

5.2.1 Toisto

Toisiinsa läheisesti liittyvien sanojen muodostamien merkityskenttien jälkeen aineiston yleisin leksikaalisen koheesioilajikeino on toisto. Siinä jokin sanastollinen kielenaines toistetaan eli se esiintyy tekstissä vähintään kahdesti. Aineksen ei kuitenkaan tarvitse toistua tekstissä täysin samassa morfologisessa muodossa. (Halliday 1985: 331.)

Toisto on siinä mielessä mielenkiintoinen keino koheesioilajikentällä, että sen käyttäminen lisää tekstin eksplisiittistä koheesiota, mutta sen liiallinen viljeleminen saa tekstin usein vaikuttamaan kömpelöltä ja jopa vaikeuttaa sen ymmärtämistä. Runsas toiston käyttäminen liukuu helposti tautofonian puolelle. Kielitoimiston sanakirja määrittää tautofonian samanlais-

ten äännejaksojen (häiritseväksi) toistoksi (KS s.v. *tautofonia*). Tautofoniassa samaa sanaa siis toistetaan useita kertoja samassa tai peräkkäisissä lauseissa, jolloin toiston käyttäminen lisää eksplisiittisesti koheesiota, mutta ei enää edistä tekstin vastaanottajan ymmärrystä. Toistoa voidaan välttää viittaamalla tarkoitteeseen referenssin avulla tai käyttämällä toiston sijaan synonyymejä, yleismerkityksisiä sanoja tai ylä- ja alakäsitteitä. Toiston sijaan käytettäviin sanastollisen koheesion keinoihin syvennyn tarkemmin seuraavissa alaluvuissa.

Toistoa ilmenee jokaisessa aineiston artikkelissa. Erittäin tyypillistä on artikkelin jo otsikossa mainitun aiheen toistuminen läpi tekstin. Monessa aineiston artikkelissa toisto on niin yleistä, että se saa tautofonian piirteitä. Toiston yleinen käyttö on todennäköisesti yhteydessä siihen, että synonyymejä käytetään varsin vähän, ja yläkäsitteitäkään ei esiinny läheskään jokaisessa aineiston artikkelissa. Lisäksi yleismerkityksisiä sanoja esiintyy aineistossa vain muutaman kerran. Hyvä esimerkki tekstistä, jossa esiintyy runsaasti tautofonista toistoa, on Poro-artikkeli. Siinä käytetään lähes pelkästään *poron* täydellistä toistoa, kuten esimerkki (71) osoittaa. Kuitenkin esimerkiksi referenssillä, yleismerkityksisellä substantiivilla *eläin* tai *poron* hyperonyymilla *hirvieläin* tekstiin saisi jo vaihtelua.

71) *Poro* kuuluu [märehtijöihin](#) ja sillä on neljä mahaa. Kesällä *porot* syövät enimmäkseen [koi-vun](#) lehtiä, [jäkälää](#), heinää ja ruohoa^{[51][6]}, mutta metsissä ne voivat käyttää myös monipuolisempaa ravintoa, kuten suomaiden [raatetta](#) ja luonnonniittyjen kasveja. Syksyllä *porot* keräävät vararavintoa talvea varten syömällä [sieniä](#), jäkälän ja heinien lisäksi. Talvella, kun ravintoa on luonnossa vähemmän, *porot* etsivät ensisijaisesti jäkälää. *Poro* pystyykin haistamaan jäkälän metrin paksuisen hangen läpi.^[3] Myös heinäkasvit ja varvut kelpaavat *porojen* ravinnoksi, jos jäkälää ei ole tarpeeksi saatavilla. Jäkälistä *poro* suosii palleroporonjäkälää. Poron ruuansulatus yksinkertaistuu talvella ja pötsin merkitys on erittäin vähäinen. *Poro* muistuttaa talvisin yksimahaista eläintä ja talvi merkitsee *porolle* laihtumista ja jopa vain hengissäselviämistä. *Poro* menettää talvella normaalistikin noin viidenneksen painostaan. Nykyisin yleistynyt lisäruokinta on muuttanut tilannetta merkittävästi.

Toinen esimerkki tautofonisesta toistosta on Liikenne-artikkeli. Sen johdannossa (esimerkki 72) sanaa *liikenne* käytetään jokaisessa virkkeessä ainakin kerran eikä liikenteeseen yhtä poikkeusta (*näiden*) lukuun ottamatta viitata millään muulla keinolla. Lisäksi *jakaa*-verbin toistaminen ja listaavien, pelkistä päälauseista koostuvien virkkeiden asettaminen peräkkäin lisää entisestään kömpelöä vaikutelmaa.

72) *Liikenne* on [ihmisten](#), [tavaroiden](#) ja [tiedon](#) kuljettamista paikasta toiseen. *Liikenne* jaetaan kuljetusteiden mukaan maa-, vesi- ja ilmaliikenteeseen. Näiden sisällä *liikenne* jaetaan matkustaja- ja tavaraliikenteeseen.

Toisaalta Juutinraumaa käsittelevä artikkeli on hyvä esimerkki siitä, miten tekstistä saadaan sujuvampaa korvaamalla toisto välillä jollakin muulla koheesiokeinolla. Juutinrauma-

artikkelin Maantiede-luvun alussa (esimerkki 73) Juutinraumaan viitataan myös yläkäsitteellä *salmi* ja demonstratiivipronominilla *se*.

73) Maantieteellisen määritelmän mukaan Juutinrauma käsittää Tanskan [Själlannin](#) ja Ruotsin [Skånen](#) välisen vesialueen, jonka rajana pohjoisessa on [Gillelejen](#) ja [Kullenin](#) niemen välinen linja, etelässä [Stevnsin](#) majakan ja [Falsterbon](#) välinen. Salmi on näin 118 kilometriä pitkä. Kaapeimmillaan se on 3,5 km leveä [Helsingørin](#) ja [Helsingborgin](#) välillä, leveimmillään 40 km.

Toistolla luodaan läpi koko aineiston tehokkaasti koheesiota. Toisto korostuu erityisesti niissä artikkeleissa, joissa kieliopillisia koheesiokeinoja ei juurikaan käytetä, ja saa tällöin myös helposti tautofonista sävyä.

5.2.2 Yläkäsitteet

Kun yläkäsitteitä analysoidaan sanastollisen koheesion kannalta, on olennaista huomata, että sanoilla täytyy olla viittauskohde tekstissä. Esimerkiksi Porosta kertovassa artikkelissa luetellaan poron eri nimityksiä iän ja sukupuolen mukaan. Ne ovat kiistatta poron alakäsitteitä, mutta koska niillä ei ole viittauskohdetta tekstissä, jätetään ne koheesiota luovien yläkäsitteiden tutkimisen ulkopuolelle. Sama koskee muun muassa esimerkissä (74) lueteltuja Magma-kivi-artikkelin eri kivilajeja. Ne ovat magmakiven alakäsitteitä, mutta ne eivät saa viittauskohdetta tekstissä, joten ne jäävät tässä luokittelussa yläkäsitteiden tutkimisen ulkopuolelle.

74) **Magmakivet** jaetaan kolmeen ryhmään muodostumistapansa perusteella. [Syväkivilajit](#) eli *plutoniset kivilajit* kiteytyvät kivilajista eli [magma](#)sta hitaasti syvällä maankuorella. *Juonikivet* syntyvät samalla tavoin, mutta kuoren ulkokerroksissa [magma](#)säiliössä, kun taas tulivuoresta maan pinnalle purkautunut magma muodostaa *vulkaanisia kivilajeja* eli [vulkaniitteja](#), joita kutsutaan yhteisellä nimellä myös [eruptiivikiviksi](#) tai *pintakiviksi*.

Luvussa 5.2.5 tarkastelen hyponymiaa laajemmin aineiston tekstuaalisena ja koheesiota rakentavana keinona ja huomioin silloin myös viittauskohteettomat hyponyymit.

Turhaa toistoa tekstissä voidaan siis välttää käyttämällä ylä- ja alakäsitteitä, synonyymejä sekä yleismerkityksisiä sanoja. Erotan analyysissäni yläkäsitteet ja synonyymit toisistaan, mutta on otettava huomioon, että niiden rajaa on toisinaan hankala määrittää. Esimerkiksi Särkänniemi-artikkelissa *elämyspuiston* voisi nähdä huvipuistoa laajempaan kategoriana, jolloin *huvipuisto* olisi sen hyponyymi. *Elämyspuistoa* on tekstissä kuitenkin käytetty synonyyminä *Särkänniemelle* samaan tapaan kuin *huvipuistokin*, joten katson tässä tapauksessa *elämyspuiston* ja *huvipuiston* rinnakkaisen käytön olevan synonyymiaa. Olennaista on se, että artikkelista voi päätellä kirjoittajan tarkoittaneen *elämyspuiston* *huvipuiston* synonyymiksi.

Wikipedia-artikkeleissa näyttää esiintyvän jonkin verran yläkäsitteitä. Niillä luodaan sanastollista koheesiota kahdessakymmenessä neljässä aineiston artikkelissa. Niitä käytetään viittaamiseen selvästi enemmän kuin synonyymejä. Synonyymeihin syvennyn enemmän paradigmaattisia merkityssuhteita käsittelevässä luvussa 5.2.5. Useassa artikkelissa jokin ilmaus mainitaan ensin kokonaan, ja sen jälkeen siihen viitataan yleisemmällä käsitteellä. Muun muassa Särkänniemi-artikkelin ensimmäisen luvun ensimmäisessä kappaleessa puhutaan ensin *Särkänniemestä*, jonka jälkeen siihen viitataan yläkäsitteellä *niemi* (esimerkki 75). Jälkimmäisessä esimerkissä (76) puolestaan *shetlanninlammaskoiraan* viitataan yläkäsitteellä *rotu*.

75) Särkänniemi syntyi Onkiniemestä ja viereisestä Särkänsaaresta, jotka yhdistettiin täytemaan avulla niemeksi 1880-luvulla^[4]. *Niemen* rajaama Mustanlahden satama oli 1960-luvulle saakka vilkas sisävesisatama, jonka läpi virtasi paljon tukkeja, halkoja ja rakennuskiviä eri puolilta Näsjärveä Tampereen tarpeisiin. *Niemestä* louhittiin myös kiveä.

76) *Shetlanninlammaskoiralle* riittää peruskoulutus. Koulutuksen pitää olla kuitenkin tavallansa pehmeää, eikä kovia otteita saa käyttää *rodun* herkkyyden vuoksi.

Yläkäsitteillä voidaan siis välttää täydellistä toistoa, mutta aineistossa niiden käyttö ei vähennä kovin paljon vaikutelmaa toistosta, sillä hyvin usein yläkäsite sisältyy leksikaalisesti aikaisemmin mainittuun alakäsitteeseen. Tätä keinoa käytetään esimerkiksi johdonmukaisesti läpi koko Merkurius-artikkelin, jossa puhutaan muun muassa *kiertoradasta* ja sen jälkeen *radasta*, *perihelikiertymästä* ja sitten *kiertymästä* sekä *Merkuriuksen pinnasta* ja tämän jälkeen *pinnasta*. Alla olevassa esimerkissä (77) *magneettikenttä* on korvattu *kentällä*.

77) Merkuriuksella uskotaan olevan valtava rautaydin, joka aiheuttaa *magneettikentän*, voimakkuudeltaan noin 1 % Maan *magneettikentästä*. *Kentän* kaltevuus pyörähdysakselia vastaan on 7 %.

Myös monessa muussa artikkelissa alakäsitteestä käytetään siihen leksikaalisesti sisältyvää yläkäsitettä. Esimerkiksi Liikenne-artikkelissa *auto-onnettomuuteen* viitataan *onnettomuutena*, *unisieppariin* viitataan *siepparina* ja *naisten äänioikeuteen* *äänioikeutena*. Myös erisnimen sisältävissä käsitteissä on tyypillistä pudottaa ensimmäisen esiintymän jälkeen nimiosa pois ja viitata siihen yleisemmällä käsitteellä. *Pyhän Eerikin Legendaan* viitataan *legendana*, *Piispa Henrikin surmavirrestä* puhutaan vain *surmavirtenä*, *Brasilian ylänkö* on *ylänkö* ja *Juutinrauman silta* on *silta*.

Yllä lueteltuja tapauksia enemmän elävyyttä yläkäsitteet tuovat tekstiin silloin, kun ne eivät sisälly leksikaalisesti aikaisemmin mainittuun alakäsitteeseen. Nämä tapaukset ovat aineistossa selkeästi vähemmistössä, mutta niitä löytyy kuitenkin jonkin verran. Esimerkiksi

Juutinraumaan viitataan myös *salmena* ja *shetlanninlammakoiraan* viitataan *rotuna*. Lisäksi esimerkiksi Merkurius-artikkelissa puhutaan Merkuriuksesta *planeettana* (esimerkki 78) ja Daniel Craigista kertovassa artikkelissa *James Bondiin* viitataan *kuuluisana agenttina* (esimerkki 79).

78) Merkuriuksen sijainti on haasteellinen myös tutkimusluotaimen rataa ajatellen. *Planeetta* ei ole kiinnostanut tutkijoita eikä rahoittajia yhtä paljon kuin vaikkapa Mars, Venus ja Saturnus.

79) Craig kertoi aiemmin, ettei aio näytellä viidessä James Bond -elokuvassa, mutta on sittemmin muuttanut lausuntojaan ja sanookin nyt haluavansa esittää *kuuluisaa agenttia* niin pitkään kuin mahdollista.

Vaikka monipuolisuutta tekstiin luovat yläkäsitteet eivät nouse merkittäväksi keinoksi aineistossa keskimäärin, käytetään yksittäisessä Bahai-uskosta kertovassa artikkelissa yläkäsitteitä monipuolisesti: *Bahaihin* viitataan *uskonnonharjoittajana*, *baabilaisuuteen uskonnollisena liikkeenä* ja *bahai-uskoon uskontona*. Ne tuovat tekstiin vaihtelua ja luovat artikkelista kohesiivisesti erittäin tiiviin.

5.2.3 Yleismerkityksiset sanat

Yleismerkityksiset sanat eivät ole Wikipedia-artikkeleissa tyypillinen keino muodostaa koheesiota. Niitä esiintyy aineistossa vain hyvin vähän siitä huolimatta, että yleismerkityksisten ilmausten käyttökonteksti sinänsä on hyvin laaja. Aineistossa käytetään ihmiseen viittaavaa yleissubstantiivivia *mies*. Lisäksi paikkaan viittaavat substantiivit *maa* ja *alue* esiintyvät muutaman kerran. Seuraava esimerkki (80) osoittaa, miten Brasilia-artikkelissa toistoa vähennetään käyttämällä substantiivivia *alue* viittaamaan *Amazonasin alankoon*. Katkelmassa muodostuu moninkertainen sidos, sillä *Amazonasin alankoon* viitataan ensin persoonapronominilla *se* ja tämän jälkeen seuraavassa virkkeessä yleissubstantiivilla *alue*.

80) Amazonasin alanko on 3500 kilometriä pitkä alanko [Andien](#) ja [Atlantin](#) välissä. Se on leveimmillään pohjois-eteläsuunnassa 2000 kilometriä leveä. *Alue* peittää Etelä-Amerikasta viidenneksen eli 3 600 000 neliökilometriä, josta 2 300 000 neliökilometriä kuuluu Brasilialle.

Asiaan viittaavaa substantiivivia *ongelma* käytetään koheesioon luomiseen virkkeiden välille konserttoa käsittelevässä artikkelissa (esimerkki 81). Se kiteyttää koko edellisen virkkeen reeman ja sidostaa näin tekstiä tehokkaasti. Sanalla *ongelma* on aivan eri sävy, kuin jos edellisen virkkeen ilmaus olisi korvattu esimerkiksi demonstratiivipronominilla *tämä*.

81) Kirkossa eri soitinryhmät ja kuorot sijoituivat eri paikkoihin ja kaiun aiheuttamien viiveiden takia tämä teki täydellisen yhteissoitin vaikeaksi. *Ongelma* ratkaistiin säveltämällä teokset niin useat kuorot lauloivat vuorotellen kirkon eri osista aikaansaaden lähes stereofonisen vaikutelman samalla, kun instrumentit ja [urut](#) oli sijoitettu kirkkoon siten, että kaikki kuorot kuulivat ne yhtä hyvin.

Yleismerkityksisten sanojen puuttuminen lähes kokonaan Wikipedia-artikkeleista selittää osittain niiden koheesion rakentumista paljolti toiston ja merkityskenttien varaan. Se kertoo toisaalta siitä, että artikkeleissa on pyritty täsmällisyyteen toistamalla mieluummin käsiteltävää aihetta kuin korvaamalla se yleismerkityksellisellä sanalla. Toisaalta artikkeleihin ei ole myöskään koettu tarpeelliseksi hakea vaihtelevampaa kieltä yleismerkityksisillä sanoilla.

5.2.4 Merkityskentät

Sanastollista koheesiota muodostetaan myös keinoilla, joissa ilmaukset eivät vaadi yhteistä tarkoitetta. Samaan merkityskenttään kuuluvien sanojen käyttäminen on merkittävä keino leksikaalisen koheesion muodostamisessa Wikipedia-artikkeleissa. Vaikka osa artikkeleista sisältää vain vähän muita kohesiivisiä keinoja, niiden sanasto on kuitenkin lähes aina kautta koko tekstin semanttisesti yhtenäistä. Sanasto kietoutuu artikkelin aiheen ympärille: jokaisesta artikkelista löytyy runsaasti otsikon muodostamaan merkityskenttään kuuluvia sanoja, jotka rakentavat artikkelin yhtenäisyyttä.

Fysioterapiasta kertovaan artikkeliin koheesiota luo vahvasti fysioterapiaan sekä myös yleisemmin terveydenhoitoon liittyvä sanasto. Seuraavasta esimerkistä (82) löytyvät *fysioterapian* lisäksi samaan merkityskenttään kuuluvat *terapia, ylläpitää, parantaa, toimintakyky, toimenpiteet, synnynnäinen vamma, sairaus, tapaturma, alentunut toimintakyky, sosiaali-, terveys- ja kuntoutuspalvelujärjestelmä, fysioterapeuttinen ohjaus sekä kipu- ja oirehoito*. Myös artikkelin aihetta tietämättä sen sisällöstä voi muodostaa kuvan vain näitä samaan merkityskenttään kuuluvia sanoja tarkastelemalla. Tämä on tyypillistä aineiston artikkeleille yli-päättään.

82) **Fysioterapia** on [terapiaa](#), jonka keskeisenä tavoitteena on ylläpitää ja parantaa ihmisten toimintakykyä ja liikkumista. Yleensä fysioterapian toimenpiteet kohdentuvat synnynnäisen vamman, sairauden, tapaturman tai ikääntymisen aiheuttamaan alentuneeseen toimintakyvyn edistämiseen fysioterapeuttisen ohjauksen ja neuvonnan, terapeuttisen harjoittelun sekä [manuaalisen](#) terapian keinoin. Kipu- ja oirehoitoina käytetään mm. fysikaalisia hoitoja. Fysioterapia on osa julkista ja yksityistä sosiaali-, terveys-, ja kuntoutuspalvelujärjestelmää.

Hallidayn mukaan tekstille on tyypillistä, ettei sen virkkeiden järjestystä voi vaihtaa muuttamatta tai jopa tuhoamalla tekstin merkitystä (Halliday & Hasan 1976: 28). Wikipedia-artikkeleissa osa virkkeistä ja kappaleista sisältää kuitenkin niin vähän koheesiokeinoja, että niiden paikkoja luvun sisällä voisi vaihdella merkityksen suurestikaan muuttumatta. Tästä huolimatta yhteinen merkityssisältö liittyy artikkelien kappaleet toisiinsa. Merkityskentät ovat siksi Wikipedia-artikkeleissa hyvin merkittävä keino muodostaa koheesiota.

Saukkonen määrittelee koherenssin toteutumisen tärkeimmäksi edellytykseksi yhtenäisen tekstiteeman (Saukkonen 2001: 79). Merkityskenttiä tarkastelemalla voidaan nähdä, että jokainen artikkeli keskittyy oman aiheensa tarkasteluun eli tekstin topiikin käsittelyyn. Edellisessä esimerkissä (82) toisiaan seuraavat virkkeet liittyvät varsin heikosti toisiinsa pintatason koheesiokeinoin, mutta ne käsittelevät kaikki samaa teemaa, ja niissä esiintyvät terveydenhoitoon ja fysioterapiaan liittyvät sanat ovat läheisessä suhteessa toisiinsa. Näin ne rakentavat tehokkaasti artikkelien koherenssia.

Sanojen sisältyminen samaan merkityskenttään sitoo monesti artikkeliin nekin yksittäisistä virkkeistä koostuvat kappaleet, jotka muuten vaikuttaisivat irrallisilta. Niitä ei aina ole liitetty muuhun tekstiin kieliopillisin tai viittaussuhteita vaativin leksikaalisin koheesiokeinoin, mutta niiden sanasto on yhdenmukaista artikkelin muun sanaston kanssa. Näin on esimerkiksi Juutinrauma-artikkelin yhden virkkeen kappaleessa (esimerkki 83), jonka sanasto kuuluu edellisen kappaleen kanssa samaan merkityskenttään. Katkelmassa mereen ja rannikkoon liittyvä sanasto korostuu, ja samaa sanastoa esiintyy myös viimeisessä virkkeessä, joka ei liity sitä edeltävään kappaleeseen näkyvien kieliopillisten koheesiokeinojen avulla.

83) Pohjan laatu ja veden suolaisuus vaihtelevat paljon salmessa, matalissa lahdissa on hiekka-pohja ja vähäsuolaista Itämerestä peräisin olevaa murtovettä, syvemällä on savi- ja mutapohjia ja painavampaa Kattegatista peräisin olevaa vettä. Veden virtaussuunta on useimmiten pohjoiseen, ulosvirtaava Itämeren murtovesi on merivettä kevyempää ja ulosvirtaus tapahtuu näin pintakerroksissa. Veden kerrostunut suolapitoisuus säätelee vastaavasti meri- ja murtovesieliöiden esiintymistä salmen eri osissa. Salmen pohjoisosassa 30 metrin syvyydessä on Knähakenin vedenalainen luonnonsuojelualue, jonka mereistä lajirunsautta on verrattu koralliriuttoihin. Myös Falsterbon niemen ympäristö salmen kaakkoiskulmassa on vastaavasti merensuojelualue.

Juutinrauman saaria ovat [Ven](#), [Amager](#), [Saltholm](#) sekä keinotekoinen, siltaan liittyvä [Peberholm](#).

Wikipedia-artikkeleissa suositaan määrittelevää ja täsmällistä kieltä, joten on luonnollista, että niistä löytyy monipuolisesti kunkin artikkelin merkityskenttään sisältyvää sanastoa. Artikkelit eivät myöskään ole laajoja ja polveilevia, vaan päinvastoin varsin lyhyitä ja tiiviitä, ja ne pysyvät tiukasti asiassa. Ne keskittyvät aiheensa jäsentämiseen, eli käsittelevät monesti

yhtä tiettyä teemaa, mikä sekin osaltaan sitoo niiden sanastoa samaan merkityskenttään. Wikipedialle on tyypillistä, että tarkempi määrittely jollekin artikkelissa mainitulle, mutta ei siihen ensisijaisesti liittyvälle käsitteelle, löytyy linkin takaa kyseisen käsitteen omalta Wikipedia-sivulta. Tämä mahdollistaa sen, että artikkelit voivat pysyä tiiviisti oman aiheensa sanastossa.

5.2.5 Paradigmaattiset merkityssuhteet

Merkityskenttien lisäksi sanastollista koheesiota voi tarkastella tarkemmin paradigmaattisten merkityssuhteiden kautta. Näistä käsittelen seuraavaksi synonymiaa, hyponymiaa, antonymiaa ja meronymiaa. Synonyymejä käytetään artikkeleissa jonkin verran, mutta niitä hyödynnetään vähemmän kuin yläkäsitteitä. Synonyymejä esiintyy varsin monessa aineiston artikkelissa, mutta niitä on yleensä käytetty koheesiokeinona vain kerran tai pari artikkelissa. Yksi syy synonyymien vähäisempään määrään on mahdollisesti se, että Wikipedia-artikkeleihin pyritään tuottamaan eksaktia tietoa eikä artikkelien aiheille näin ollen erityisesti yritetä löytää synonyymejä vain tekstin elävöittämiseksi. Lisäksi synonymian käyttöön vaikuttanee kussakin artikkelissa käsiteltävä aihe. Joillekin termeille tai ilmiöille voi olla vaikea keksiä vaihtoehtoisia ilmaisutapaa. Esimerkiksi Kuka pelkää mustaa miestä -artikkelissa *leikkiin* on helppo viitata välillä *pelinä*, mutta muun muassa *Leidenin pulloa* tai *äänioikeutta* on vaikeampi korvata synonyymilla.

Artikkeleille on siis tyypillistä, että synonymiaa löytyy joitakin yksittäistapauksia. Esimerkiksi pitkästä Brasiliasta kertovasta tekstistä löytyy vain muutama synonyymitapaus. *Sademetsästä* on käytetty synonyymia *trooppinen metsä*. Daniel Craig -artikkelissa puolestaan verbejä *näytellä* ja *esittää* on käytetty synonyymisesti läpi artikkelin. Nämä yksittäiset keinot eivät kuitenkaan riitä ratkaisevasti lisäämään kummankaan artikkelin vivahteikkuutta.

Loch Nessin hirviöstä kertova artikkeli on hyvä esimerkki siitä, miten synonyymien ja yläkäsitteiden käyttäminen artikkelin aiheesta elävöittää tekstiä. Vaikka synonymiaa ei olisi käytetty tekstissä kovin monipuolisesti, niin sujuvuutta lisää huomattavasti se, jos usein käytettävästä termistä puhutaan synonyymeillä. Sen lisäksi, että Loch Nessin hirviöön viitataan yläkäsitteellä *hirviö*, siitä käytetään useasti myös synonyymia *Nessie*. Lisäksi hirviöön viitataan kuvailevalla ilmauksella *pitkäkaulainen otus*, mikä sekin kertoo artikkelin rikkaasta kielestä (esimerkki 84). Ero on selvä esimerkiksi porosta kertovaan artikkeliin, jossa synonyymejä ei käytetä lainkaan. Esimerkissä 71 *poro* toistuu yhdessä kappaleessa niin monta kertaa, että artikkelin kieli vaikuttaa töksähtelevältä ja kömpelöltä.

84) *Loch Nessin järvi* mainitaan ensimmäisen kerran vuonna [565](#), jolloin [Pyhän Columban](#) väitetään torjuneen "Paholaisen" hyökkäyksen. Huhtikuussa [1933](#) John Mackay teki ensimmäiseksi nykyaikaiseksi kutsutun havainnon *Nessistä*. Vasta tämän jälkeen väitetyjä näköhavaintoja *Loch Nessin hirviöstä* alettiin tehdä säännöllisemmin. 1. huhtikuuta 1934 [gynecologi](#) Robert Wilson otti *pitkäkaulaisesta otuksesta* kuuluisaksi tulleen valokuvan, jonka hän itse tunnusti vuonna 1994 huijaukseksi.

Bahai-uskosta kertova artikkeli antaa esimerkin siitä, miten tekstistä saadaan sujuvaa ja vaihtelevaa käyttämällä monipuolisesti synonyymejä sekä lisäksi yläkäsitteitä ja referenssiä. Useimmissa aineiston artikkeleissa synonymiaa esiintyy satunnaisesti, mutta kyseisessä artikkelissa sitä käytetään läpi koko artikkelin. Muun muassa *uskonnonharjoittajaan* viitataan sanoilla *uskova ja uskonnon kannattaja*, *uskonnosta* käytetään välillä ilmausta *uskonnollinen liike* ja *bahai* saa toisinaan synonyymien *bahailainen*. *Kirjalliseen tuotantoon* puolestaan viitataan *julkaisuna*. Synonymian monipuolinen käyttö tekee tekstistä sujuvaa luettavaa. Toisaalta Bahai-uskoa käsittelevä teksti vaatii Poro-artikkelia tarkempaa lukemista, sillä käytetyt koheesiokeinot edellyttävät edellisten virkkeiden pitämistä mielessä.

Synonyymit eivät juuri koskaan ole täysin samanmerkityksisiä, vaan niillä on sävyeroja. Shetlanninlammaskoirasta kertovassa artikkelissa shetlanninlammaskoiran synonyyminä käytetään rodun lempinimeä *sheltti*. Tämä tuo artikkeliin vaihtelevuutta, mutta samalla tekstistä syntyy vähemmän formaali vaikutelma. Myös edellisen esimerkin (84) *Loch Nessin hirviöön* viittaaminen *Nessienä* tuo artikkeliin samankaltaista huoletonta sävyä. Näkymättömästä kädestä kertovassa artikkelissa puolestaan *Näkymättömästä kädestä* puhutaan niin *metaforana*, *käsitteenä* kuin *mekanisminakin*. Nämä ilmaukset ovat merkitykseltään jo melko kaukana toisistaan, mutta tekstin kirjoittajat ovat pitäneet kyseisiä ilmauksia synonyymisina toisilleen.

Synonyymien analyysi paljastaa, että vaikka synonymiaa esiintyy yksittäistapauksina jonkin verran, koheesiota luodaan Wikipedia-artikkeleissa kuitenkin valtaosin muilla keinoilla. Muutamaa artikkelia lukuun ottamatta synonymian käyttö on melko rajoittunutta: sitä käytetään vain kerran tai kaksi artikkelissa ja/tai ainoastaan yhdestä tai kahdesta sanasta. Synonyymien vähyys ja satunnaisuus kertovat myös siitä, että Wikipedia-artikkeleissa ei pyritä erityiseen vivahteikkaaseen tai sanastoltaan vaihtelevaan kieleen.

Synonyymien vähyys ei kuitenkaan automaattisesti tarkoita monotonista kieltä ja runsaasta toiston käyttöä. Jos artikkelissa on käytetty viittaamiseen synonymian sijasta yläkäsitteitä ja referenssiä, teksti voi olla hyvinkin sujuva. Monotonista ja kömpelöä tekstistä tulee siinä vaiheessa, kun koheesion luominen jää suurelta osin vain toiston käytön varaan. Mikäli tekstissä usein toistuvaa aihetta (yleensä tekstin otsikko) ei korvata synonyymeillä eikä yläkäsitteillä, tulee tekstistä helposti toisteinen vaikutelma. Hyvä esimerkki tästä on Pakina-artikkeli.

Esimerkin (85) kappale myös koostuu lähes pelkistä päälauseista, mikä lisää töksähtelevyyttä ja kömpelyyttä.

85) *Pakinatyyli* on tuttavallista. *Pakinoitsija* ikään kuin juttelee lukijan kanssa. *Pakinointiin* kuuluu olennaisesti normipoikkeama: *pakina* rikkoo kielen ja tyylin normeja. Muista lehtiteksteistä poiketen *pakinassa* voidaan liioitella ja vääristellä. Esimerkiksi poliitikko voidaan kuvata täysin kuvitteellisessa tilanteessa.

Artikkelien aiheiden toistaminen aina samassa muodossa tekee siis tekstistä usein kömpelöä mutta toisaalta myös varsin helppoa luettavaa. Tarpeellinen tieto on nopeasti löydettävissä artikkelista.

Hyponymiassa kysymys on merkitysten hierarkkisesta alistussuhteesta (Karlsson 2006: 221–222). On odotuksenmukaista, että ensyklopedisessä tekstissä asioita ja ilmiöitä pyritään määrittelemään hyponyymejä (alakäsitteitä) ja hyperonyymejä (yläkäsitteitä) käyttämällä. Yläkäsitteiden analyysissä luvussa 5.2.2 pyrin käsittelemään yläkäsitteitä, joilla oli viittauskohdetekstissä. Tässä luvussa analysoin tarkemmin hyponymiaa artikkeleissa ylipäänsä. Hyponymien ja hyperonyymien kautta artikkeleissa muodostuu näin käsittekkäitä ja luodaan käsitystä maailmasta. Tässä tekstuaalinen ja ideationaalinen merkitys kohtaavat. Esimerkiksi Biologisesta antropologiasta kertovassa artikkelissa aihetta jäsennetään luettelemalla sen piiriin kuuluvia tieteenaloja eli *biologisen antropologian* hyponyymejä (esimerkki 86).

86) Biologisen antropologian piiriin kuuluvia suppeampia tieteenaloja ovat muun muassa [fossilieita](#) tutkiva [paleoantropologia](#), [DNA](#)-tutkimusta hyödyntävä [molekyyliantropologia](#), ihmisten mittasuhteita ja koostumusta analysoiva [antropometria](#) sekä kädellisiä tutkiva [primatologia](#).

Rinnastuksilla pyritään listaamaan yhteenkuuluvia tekoja, toimijoita ja kohteita. Se saa tekstin vaikuttamaan tyhjentävältä ja näin ollen myös vakuuttavalta. Rinnastus tuottaakin vaikutelman täsmällisyydestä ja maailmanjaottelusta. (Hiidenmaa 2001: 51–53.) Kuten aikaisemmissa luvuissa on jo käynyt ilmi, rinnastuksia käytetään Wikipedia-artikkeleissa yleisesti. Monessa artikkelissa käsiteltävää aihetta jäsennetään luettelomaisesti. Esimerkiksi Poro-artikkelissa luetellaan poron nimitykset iän, sukupuolen ja värin mukaan. Samalla luodaan koheesiota hyponymiaa käyttämällä. Tätä havainnollistaa seuraava esimerkki (87):

87) Poroilla on useita nimityksiä iän ja sukupuolen mukaan. Porouroksista käytetään nimitystä *hirvas*, ja naaraille vastaava termi on [vaadin](#). Pohjoissaameksi vastaavat nimitykset ovat *sarvvvis* sekä *áldu* [7]. Jälkeläisiä kutsutaan *vasoiksi*. Toisella ikävuodella olevasta naaraasta käytetään nimitystä *vuorso* ja uroksesta *urakka*. Kun poro täyttää kuusi vuotta, nimitykset loppuvat, jolloin nimitys on *nimiloppu*. [8]

Poroilla on nimitykset myös värin mukaan. Täpläkäs poro on kirjakko, vaalea on *suivakko* ja täysin valkoinen poro on *valkko*.^[3]

Myös esimerkiksi Siivestä kertovassa artikkelissa hyponymiaa esiintyy runsaasti. Esimerkissä (88) *selkärangaton* toimii hyperonyymina *hyönteiselle* ja katkelmassa mainitut hyönteislajit ovat puolestaan *hyönteisen* hyponyymejä. Hyponymian avulla luodaan kyseisessä Siivet hyönteisillä -nimisen luvun katkelmassa koheesiota melko vahvasti.

88) [Selkärangattomista](#) eläimistä vain [hyönteisillä](#) on siivet. Hyönteisistä alkeellisimmilla tosin siipiä ei ole koskaan ollutkaan kuten *hyppyhäntäisillä*, *alkuhyönteisillä* ja *kaksi- ja kolmisukahäntäisillä*. Joiltakin lajeilta siivet taas ovat surkastuneet kehityksen aikana. Näitä ovat *kirput*, *täit*, eräät *kärpäset*, *muurahaistyöläiset*, eräät *koskikorennot* ja *suorasiipiset*, jotkut *kovakuoriaiset* ja *luteet* sekä jotkut *perhosnaaraat* (*halla-* ja *pakkasmittari*) jne. Useimmilla siivekkäillä hyönteisillä siipiä on kaksi paria. *Kaksisiipisillä* sekä joillain *päiväkorenoilla* takasiivet ovat kehityksen aikana surkastuneet tai kokonaan hävinneet. Siivet ovat pingottuneina siipisuoniston avulla.

Kuten edeltävät esimerkit osoittavat, hyponymiaa esiintyy aineistossa paljon yhdessä rinnastusten ja listauksen kanssa. Kysymys on tällöin osaltaan maailman ja todellisuuden jäsentämisestä: artikkeleissa käsiteltävät asiat ja ilmiöt asetetaan muun muassa hyponymian avulla osaksi isompaa kuvaa. Toisaalta asioiden määrittelemisen kannalta on myös olennaista pilkkoa asioita pienempiin kokonaisuuksiin, erityisesti jos artikkelin käsittelemä aihe on hyvin laaja. Näin on esimerkiksi Liikenne-artikkelissa, jossa *liikenne* saa runsaasti erilaisia hyponyymejä. Sen alakäsitteenä mainitaan esimerkiksi *vesiliikenne*, joka puolestaan saa hyponyymit *meriliikenne*, *joki-* ja *kanavaliikenne* sekä *vedenalainen liikenne*. Se, millaisiin kokonaisuuksiin kirjoittaja on käsittelemänsä ilmiön jaotellut, kertoo hänen näkemyksistään maailmasta.

Myös osan ja kokonaisuuden suhdetta eli meronymiaa hyödyntämällä muodostetaan aineistossa koheesiota. Se ei kuitenkaan ole yhtä merkittävä koheesiokeino kuin edellä käsitelty hyponymia. Joissakin aineiston artikkeleissa käsiteltävää aihetta pyritään osittain jäsentämään osa-kokonaisuus-suhteen avulla. Merkurius-artikkelissa esiintyy paljon meronymiaa luvuissa, joissa käsitellään Merkuriuksen osia ja fyysisiä ominaisuuksia. Nämä määrittävät tekstiä laajalla tasolla, sillä ne ovat löydettävissä jo sisällysluettelosta. *Fyysiset ominaisuudet* -luku jakautuu alalukuihin *Kaasukehä*, *Pinnanmuodot* ja *Sisäinen rakenne*. Näissä alaluvuissa meronymiaa esiintyy runsaasti. Artikkelissa Merkuriusta on siis lähdetty määrittelemään osien ja kokonaisuuden kautta. Vähäisemmässä määrin meronymiaa hyödynnetään hyönteisten siipiä käsittelevässä esimerkissä (88), kun siipien toiminnasta kerrottaessa käytetään meronymia *siipisuonisto*.

Seuraavassa esimerkissä (89) puolestaan hyödynnetään koheesio tuottamisessa vieruskäsitteitä. Kyseisen naisten äänioikeudesta kertovan artikkelin koheesio rakentuu hyvin vahvasti vieruskäsitteiden varaan. Naisten äänioikeuden historiaa selostetaan siinä eri maiden kautta läpi koko artikkelin. Useat artikkelin virkkeistä alkavat maan nimellä. Ensimmäisessä kappaleessa rinnastetaan *Uusi-Seelanti* ja *Australia*. Toisessa kappaleessa äänioikeuden historiaa jäsennetään puolestaan *Suomen* ja *Euroopan* välisen suhteen avulla. *Suomi* toimii *Euroopan* meronyymina.

89) *Uusi-Seelanti* antoi ensimmäisenä naisille äänioikeuden vuonna 1893. *Australiassa* myönnettiin vuonna 1902 rajoitettu äänioikeus ja vaalikelpoisuus, jotka kuitenkin koskivat vain valkoista väestöä.^[11]

Suomi antoi kolmantena maailmassa ja ensimmäisenä *Euroopassa* naisille äänioikeuden vuonna 1906 osana yleistä ja yhtäläistä äänioikeutta. Samalla sallittiin naisille myös oikeus asettua ehdolle vaaleissa ja vuonna 1907 *eduskuntaan* valittiin maailman ensimmäiset naisparlamentaarikot.^[12]

Yhdysvalloissa naisten äänioikeus oli *New Jersey*n osavaltion perustuslaissa jo vuonna 1776, mutta se kumottiin 1807. Myöhemmin naiset saivat äänioikeuden *Yhdysvalloissa* vuonna 1920.^[13]

Ranskassa Pariisin kommuuni antoi naisille äänioikeuden 1871, mutta päätös kumottiin kommuunin kaatuessa.^[14]

Brasiliasta kertovassa artikkelissa koheesiota muodostetaan osassa artikkeleita niin ikään meronymian avulla. Artikkelissa puhutaan osavaltioista, joista Brasilia koostuu sekä siitä, että Brasilia on osa Etelä-Amerikkaa.

Antonymiaa eli vastakohtaisuutta hyödynnetään artikkeleissa jonkin verran, mutta kovin merkittävää osaa koheesio luomisessa sillä ei aineistossa ole. Toisin kuin hyponymia ja meronymia se ei vaikuta laajasti minkään artikkelin koheesioon, vaan ilmenee ainoastaan joissakin yksittäisissä kohdissa aineistoa. Shetlanninlammaskoira-artikkelissa sidostetaan antonymian avulla peräkkäiset virkkeet toisiinsa. Esimerkissä (90) havainnollistetaan, miten *vieraat* ja *oma perhe* toimivat vastakkaisessa suhteessa.

90) Shetlanninlammaskoira on luonteeltaan lempeä, mutta se voi olla *vieraita* kohtaan varautunut. *Omaa perhettäin* kohtaan se on rakastavainen ja kiintyy perheeseensä voimakkaasti.

Käsitlemistäni paradigmaattisista merkityssuhteista hyponymiaa käytetään koheesio luomiseen eniten. Joissakin artikkeleissa esiintyy myös jonkin verran meronymiaa. Synonymiaa käytetään monessa artikkelissa, mutta ei kovin merkittävässä määrin. Sitä ei ole juuri käytetty artikkelien sanaston rikastamiseen. Antonymian käyttö jää muutamaan yksittä-

seen tapaukseen, joten sen ei voi sanoa rakentavan merkittävästi artikkelien koheesiota. Paradigmaattisten merkityssuhteiden käytössä korostuvat artikkelien luettelomaisuus ja määrittelevä luonne.

5.2.6 Parafraasit

Edellisessä luvussa kävi ilmi, että artikkelien käsittelemiä aiheita ja niihin sisältyviä käsitteitä pyritään määrittelemään jonkin verran ylä- ja alakäsitteiden avulla eli hyponymiaa käyttämällä sekä meronymian avulla. Tämän lisäksi niitä määritellään myös ilmaisemalla asia toisin eli käyttämällä parafraaseja. Esimerkiksi Siipi-artikkelissa *siiven* parafraasina käytetään määritelmää *lintujen ja hyönteisten lenninelin*. Parafraaseja esiintyy aineistossa erityisesti artikkelien ensimmäisissä, määrittävissä virkkeissä. Muun muassa Mikrobiologia-artikkelin ensimmäinen virkkeessä *mikrobiologia* ilmaistaan parafraasilla (esimerkki 91).

91) **Mikrobiologia** on [mikro-organismeja](#) tutkiva [biologian](#) haara.

Edeltävä esimerkki havainnollistaa myös artikkelien avausvirkkeille erittäin tyypillistä rakennetta. Kahta artikkelia lukuun ottamatta kaikki aineiston artikkelit alkavat aihetta määrittävällä predikatiivilauseella. Artikkelin otsikko toimii teemana, sitä seuraa olla-verbi ja sen jälkeen aihe tavallisesti määritellään parafraasin avulla.

Artikkelien ensimmäisten virkkeiden määritelmät pyrkivät usein olemaan mahdollisimman tyhjentäviä. Osan artikkelien ensimmäisistä virkkeistä voi nähdä alkavan identifioivalla ekvatiivilauseella. Ekvatiivilause muodostuu olla-verbistä ja kahdesta nominatiivimuotoisesta NP:sta. Tyypillisimmillään ekvatiivilauseisiin kuuluvat lauseet, jotka samaistavat kaksi yksiselitteisesti tunnistettavaa tarkoitetta. Väljän määritelmän mukaan ekvatiivilauseisiin voidaan kuitenkin lukea myös tyhjentävät määritelmät. Niissä ei pyritä vain luonnehtimaan tai kategorisoimaan subjektia, vaan osoittamaan kahden NP:n tarkoitteet samaksi (VISK § 950.) Käsiteltävästä asia kerrotaan siis toisin sanoin mahdollisimman tarkasti. Esimerkissä (92) *aurinkokunnan pienin ja Aurinkoa lähin planeetta* on nimenomaan Merkurius, muita vaihtoehtoja ei ole. Voidaan siis sanoa, että jälkimmäinen substantiivipredikatiivi määrittää Merkuriuksen tyhjentävästi. Ekvatiivilauseessa ei ole muodollisia tuntomerkkejä, joista ilmeni, kumpi NP:istä on subjekti ja kumpi objekti, mutta tavallisesti verbinetinen NP katsotaan subjektiksi (VISK § 950).

92) **Merkurius** ([symboli: ☿](#)) on [aurinkokunnan](#) pienin ja [Aurinkoa](#) lähin [planeetta](#).

Nämä ensimmäisten virkkeiden tyhjentävät määritelmät ovat usein varsin pitkiä ja nimutkaisia, joten on odotuksenmukaista, ettei ensimmäisessä virkkeessä mainittu parafraasi esiinny enää myöhemmin tekstissä, vaan artikkelin aiheeseen viitataan referenssillä, toistolla, yläkäsitteellä ja joskus myös synonyymeilla. Esimerkissä (93) määritellään ensimmäisessä virkkeessä parafraasilla *naisten äänioikeus* ja seuraavassa virkkeessä *vaalikelpoisuus*.

93) **Naisten äänioikeus** on taloudellinen ja poliittinen uudistus, joka antaa naisille [oikeuden äänestää](#). Äänioikeus ei kuitenkaan aina merkinnyt [vaalikelpoisuutta](#) eli naisten mahdollisuutta asettua ehdolle vaaleissa.

Edeltävä esimerkki valottaa hyvin myös Wikipedian toimimista hypertekstinä. Sekä *oikeuden äänestää* että *vaalikelpoisuutta* johtavat linkkeinä omille Wikipedia-sivuilleen, joiden kautta niistä voi lukea lisää. Siksi käsitteitä ei ole tarpeen varsinaisessa artikkelissa avata enempää.

Läheskään kaikki artikkelien ensimmäisten virkkeiden parafraasit eivät kuitenkaan ole tyhjentäviä määritelmiä. Muun muassa seuraavassa esimerkissä (94) Ikea ei ole ainut ruotsalaislähtöinen huonekalumyymäläketju eivätkä kaikki ruotsalaislähtöiset huonekalumyymäläketjut ole Ikeoita. Kyseinen parafraasi asettaa Ikean ruotsalaislähtöisten huonekalumyymäläketjujen luokkaan, mutta ei määrittele sitä tyhjentävästi.

94) **Ikea** on [ruotsalaislähtöinen](#) huonekalumyymäläketju. Vuonna 2011 sillä oli omistuksessa 26 maassa 287 tavarataloa, joissa vieraili vuoden aikana 655 [miljoonaa](#) kävijää.^[1] Suurin osa tavarataloista on [Euroopassa](#). Tavarataloja on lisäksi [Yhdysvalloissa](#), [Kanadassa](#), [Australiassa](#), [Aasiassa](#) ja [Latinalaisessa Amerikassa](#). Vuonna 1982 Ikean omistus siirrettiin hollantilaiselle Ingka Holding -konsernille, jonka omistaa Stichting Ingka Foundation -säätiö.

Parafraasit ovat tyypillistä artikkeleille, joissa pyritään antamaan asioista tietokirjamääritelmä. Wikipedia-artikkeleissa lähes poikkeuksetta ensimmäisessä lauseessa määritellään artikkelin aihe kertomalla siitä toisin sanoin. Artikkelit on siis muodostettu niin, että ensimmäisestä virkkeestä saa heti yleiskäsityksen artikkelin käsittelemästä aiheesta. Huolimatta siitä, käsitelläänkö aihetta artikkelissa hyvin suppeasti vai monen alaluvun verran, on artikkelien ensimmäisistä lauseista yleensä artikkelista riippumatta löydettävissä suunnilleen sama määrä tietoa yhdenmukaisessa muodossa. Tämä nopeuttaa tiedon etsimistä Wikipediasta.

Edeltävä esimerkki Ikeasta (94) käsittää koko artikkelin johdannon. Siitä käy ilmi, että vaikka artikkelin aiheesta saa lisää tietoa ensimmäisen virkkeen jälkeen, tulee olennaisin informaatio esille heti ensimmäisen virkkeen määritelmässä. Tämä ominaisuus yhdistää vahvasti aineiston artikkeleita.

6 PÄÄTÄNTÖ

6.1 Yhteenveto Wikipedia-artikkelin tekstilajipiirteistä

Aineiston artikkelit osoittautuivat rakenteeltaan ja tekstuaalisilta ominaisuuksiltaan varsin vaihteleviksi. Toisaalta aineistossa esiintyi monipuolisesti eri koheesiokeinoja sisältäviä, tyyppillisesti pitkiä ja useisiin lukuihin jaettuja artikkeleita. Aineisto sisälsi kuitenkin yhtä lailla artikkeleita, joissa koheesiota rakensivat lähinnä toisiinsa läheisesti liittyvistä sanoista muodostuvat merkityskentät ja toisto, ja joiden sisältö ei useinkaan ollut jaoteltu eri lukuihin.

Aineiston vaihtelevuudesta huolimatta siitä oli kuitenkin löydettävissä säännöllisesti esiintyviä piirteitä, jotka voi nähdä kyseiselle tekstilajille tyyppillisenä. Lisäksi variaatio itsessään kertoo tekstilajista paljon. Se osoittaa, että vaikka Wikipedia-artikkelit pohjautuvat tietoa antaviin teksteihin, niiden tuottamistapa antaa kirjoittajille vapauksia tekstilajin suhteen. Artikkelien kirjoittajilla on ollut vapaus esimerkiksi valita, että jaetaanko artikkeli lukuihin, muodostetaanko siihen sisällysluettelo ja kuinka siinä käytetään lähdeviitteitä. Vaihtelevuus kyseisissä piirteissä kertoo sekä Wikipedian käyttäjälähtöisestä ja vuorovaikutuksellisesta tuottamistavasta että myös sen prosessimaisuudesta. Voidaankin ajatella, että esimerkiksi artikkeleita, josta lähteet puuttuvat tai, johon ne on vaillinaisesti merkitty, ei vain kukaan ole vielä korjannut. Se ei kuitenkaan estä kyseistä artikkeleita olemasta osa Wikipediää.

Wikipedia-artikkelien ideationaaliseen merkitykseen päästään käsiksi, kun huomioidaan kirjoittajien vapaus rakentaa artikkeleita ja niiden sisältöä ja tätä kautta ilmentää omaa käsitystään maailmasta. Samalla kirjoittajat ottavat kantaa myös käsitykseen tiedosta.

Wikipedian tuottamistapa on hyvin vuorovaikutuksellinen, mikä puolestaan nostaa esiin interpersonaalista funktiota. Kirjoittajat rakentavat merkityksiä yhdessä. Kirjoittajan rooli ei ole ylhäältä päin määrätty, vaan lukijalla on yhtä lailla mahdollisuus asettua kirjoittajan rooliin. Lisäksi Wikipedia-artikkelit sisältävät hypertekstin elementtejä, jotka mahdollistavat sen, että lukija voi valita, missä järjestyksessä hän lukee tekstiä ja millaisen kokonaisuuden hän lukemastaan rakentaa.

Wikipedian rakenteelliset ja tekstuaaliset ominaisuudet osoittavat siihen suuntaan, että nopea tiedon löytäminen on Wikipedia-artikkelille olennainen funktio. Artikkelien luvut ja alaluvut ovat pääasiassa lyhyitä. Useimmiten ne koostuvat vain yhdestä tai kahdesta kappaleesta. Lyhyiden lukujen lisäksi artikkeleissa on myös lyhyitä kappaleita. 18 aineiston artikkeleita sisältää ainakin yhden yhdestä virkkeestä koostuvan kappaleen. Myös toistuvat tekstuaaliset piirteet, kuten kaikissa artikkeleissa olennaisen tiedon löytäminen johdanto-osasta ja mää-

ritelmän löytyminen heti ensimmäisestä virkkeestä helpottavat nekin tiedon nopeaa löytämistä.

Silmäiltävyyteen ja nopeaan tiedonhakuun viittaa myös artikkelien temaattinen rakenne. Pysyvä teema on artikkeleille tyypillinen piirre. Monessa tapauksessa ei siis ole olennaista kertoa edellisen lauseen reemasta jotakin uutta seuraavassa lauseessa, vaan sen sijaan kertoa lisää edellisen lauseen teemasta eli kuljettaa samaa teemaa virkkeestä toiseen. Etenevän teeman vähäisempi käyttö pysyvään teemaan verrattuna mahdollistaa niin ikään nopean, silmäilevän lukemisen.

Nopean tiedonhaun funktiota korostaa myös se, että koheesiota on aineistossa pääosin riittävästi, jotta teksti on helposti luettavaa, mutta sellaisten koheesiokeinojen käyttäminen, jotka vaativat tekstin tarkempaa lukemista, on harvinaisempaa. Kun niitä käytetään, ne esiintyvät useimmiten yhdysvirkkeen sisällä, eivät virkkeen välissä, eli niiden luoma koheesio ei ole yhtä vahvaa kuin virkkeiden välissä esiintyvien koheesiokeinojen.

Kieliopillisten koheesiokeinojen käyttöä artikkeleissa yhdistää se, että niitä kyllä esiintyy läpi aineiston, mutta melko harvassa artikkelissa niitä käytetään kattavasti läpi koko tekstin. Ne rakentavat siis artikkelien yhtenäisyyttä, mutta eivät ole sen muodostumisesta kovin suurelta osin vastuussa. Demonstratiivipronomineilla ilmaistua referenssiä ilmenee aineistossa varsin paljon ja jonkin verran ilmenee myös persoonallista referenssiä, mutta komparaatiota ei käytännössä esiinny lainkaan. Keskimäärin 16 prosentissa aineiston virkkeistä ilmenee lauseiden tai virkkeiden välistä referenssiä, mutta ainoastaan 8,9 prosentissa aineiston virkkeistä esiintyy referenssiä, jolla viitataan edeltävään virkkeeseen.

Ellipsejä esiintyy 12,2 prosentissa aineiston virkkeistä. Ellipsikeinoista pelkistys ja verbini finitiiseen osaan kohdistuva aukkoaminen ovat aineiston selvästi suosituimmat keinot muodostaa ellipsejä ja ne myös esiintyvät usein yhdessä. Substituutio on aineistossa marginaalista, sen avulla artikkelien yhtenäisyyttä ei rakenneta käytännössä lainkaan.

Referenssi, ellipsit ja substituutio ovat koheesiokeinoja, jotka erityisesti virkkeiden välisinä vaativat tekstin tarkempaa lukemista, jotta niiden viittauskohde selviää. Ne ovat myös toistoa vähentäviä keinoja. Se että niiden käyttö jää aineistossa ainoastaan kohtuulliseksi, viittaa siihen, että toisto on aineiston koheesion rakentamisessa melko suuressa osassa. Aineistossa onkin artikkeleita, joissa ei juuri esiinny referenssiä, ellipsejä tai substituutiota mutta sen sijaan runsaasti toistoa, ja jotka näin antavat helposti kömpelön vaikutelman.

Aineistossa käytetään varsin monipuolisesti eri kytkentäkeinoja, mutta additiivinen kytkentä on aineistossa ehdottomasti yleisin kytkentäkeino, mikä nostaa esiin Wikipedia-artikkeleille luonteenomaisen rinnasteisuuden. Kytkennän analyysin tulokset heijastavat sitä,

että Wikipedia-artikkelit ovat tietoa antavia tekstejä, joissa korostuvat rinnasteisuus ja asioiden sekä ilmiöiden määrittely. Rinnasteisuus ilmenee myös ellipsien käytön kautta, joita käytetään usein yhdessä additiivisen kytkennän kanssa. Lisäksi rinnasteisuus näkyy joissakin artikkelissa selkeästi niin, että tieto on osaksi luetteloiden muodossa.

Myös sanastolliset koheesiokeinot rakentavat kuvaa tekstistä, jota on helppo lukea ei-lineaarisesti ja joka ei välttämättä vaadi tarkkaa syventymistä. Sanastollisia koheesiokeinoja kyllä ilmenee aineistossa, mutta niillä ei yleensä pyritä erityisen rikkaaseen tai ilmaisuvoimaiseen sanastoon. Analyysi osoittaa, että toisto, yläkäsitteet ja synonyymit luovat artikkeleihin koheesiota, mutta eivät merkittävästi rikastuta tai monipuolistuta artikkelien sanastoa. Näistä leksikaalisen koheesion keinoista aineistossa korostuvat ne, joissa käytetään samoja tai toisiinsa muistuttavia sanoja ja ilmauksia. Synonyymejä esiintyy selvästi vähemmän kuin yläkäsitteitä. Artikkeleissa ei siis ole koettu tarpeelliseksi käyttää synonyymejä tekstin rikastamiseksi. Toisaalta Wikipedia-artikkeleissa käsitteisiin pyritään viittaamaan melko täsmällisesti, mikä myös vaikuttanee siihen, ettei synonyymien käyttäminen ole tyypillinen koheesiokeino.

Toisteisuuden välttäminen on neutraalia verrattuna toistamiseen (VISK § 1177), minkä vuoksi toisto nousee aineistosta esille. Aineistosta löytyvä runsas toisteisuus on kuitenkin linjassa sen oletuksen kanssa, että yksi Wikipedian olennaisista funktioista on silmäiltävyys ja nopea tiedonhaku. Monotoninen sanasto ja toisteisuus tekevät tekstin silmäilemisen ja lukemisen aloittamisen keskeltä helpommaksi ja palvelevat siten nopeaa tiedonhakua ja artikkelien eri osien välillä ei-lineaarisisessa järjestyksessä liikkumista.

Wikipedia-artikkeleissa maailmaa pyritään jäsentämään hierarkkisesti. Hyponymian, parafrasien ja artikkelia jäsentävien merkityskenttien runsas esiintyminen kertoo ilmiöiden ja asioiden määrittelyn tärkeydestä artikkeleissa. Myös vähemmän esiintyvän meronymian käyttö liittyy asioiden ja ilmiöiden määrittelyyn. Parafraseja käytetään usein artikkelien ensimmäisissä lauseissa, kun käsiteltävästä asiasta halutaan antaa mahdollisimman tyhjentävä määritelmä.

6.2 Wikipedia-artikkelien rekisteri

Wikipedia-artikkelien rekisterin määrittelyssä huomioon jo esille tulleiden tekstuaalisten ja rakenteellisten piirteiden lisäksi tilannekontekstin muuttujat: alan, osallistujaroolit ja ilmene-mismuodon. Niiden avulla on mahdollista tulkita tekstin sosiaalista kontekstia eli sitä ympäristöä, jossa merkityksiä vaihdetaan (Halliday & Hasan 1985: 12). Tarkastelen seuraavassa tarkemmin, millainen rekisteri Wikipedia-artikkeleissa muodostuu.

Tekstin alan voi Hallidayn mukaan määritellä siksi sosiaalisesti toiminnaksi, johon osallistujat ottavat osaa ja jossa kieli toimii oleellisena komponenttina (Halliday & Hasan 1985: 12). Wikipedia-artikkelit ovat tietoa antavia tekstejä, joiden muoto on alustan etukäteen määräämä, mutta joiden sisällön tuottamistapa on hyvin vapaa. Niiden alaksi voi määritellä tiedon jakamisen ja tuottamisen. Kaikille artikkeleille on yhteistä se, että ne keskittyvät antamaan tietoa jo otsikossa kerrotusta aiheesta. Jokaisessa tekstissä määritellään käsiteltävä aihe ja pysytään tiiviisti kyseisen aiheen käsittelyssä.

Wikipedia-artikkelien takana on anonyymi kirjoittajajoukko. Tekstien osapuolia ei siis tuoda esiin, vaan sekä kirjoittajat että lukijat jäävät implisiittisiksi. Heikkinen kuitenkin huomauttaa, että tästä huolimatta osallistujien rooleja voidaan tekstintutkimuksen keinoin tarkastella. Tietynlaiset tekstuaalistamiset rakentavat interpersoonaista merkitystä. Yksi kontekstin ulottuvuus on se, mitä tekstin kielelliset valinnat kertovat tekstin osapuolista ja toisaalta kirjoittajuuden ja lukijuuden erityispiirteistä tekstin rakentamassa sosiokulttuurisessa todellisuudessa. (2001: 203–204.)

Osallistujaroolien tarkasteluun linkittyy myös ilmenemismuodon tarkastelu. Wikipedia-artikkeleissa pyritään kuvaamaan ja määrittelemään asioita ja ilmiöitä, samalla rakennetaan tietynlaista kuvaa maailmasta ja todellisuudesta. Artikkeleissa ei juurikaan spekuloida, vaan tieto esitetään totuutena. Totuus jostakin asiasta riippuu tietysti aina tarkastelijasta, mutta Wikipedia-artikkeleissa totuus esitetään varsin yksioikoisena. Asioita ja ilmiöitä määritellään preesensissä ja niiden historiasta ja taustoista kerrotaan pääosin imperfektissä. Näin on esimerkiksi Unisiepparista kertovan artikkelin ensimmäisessä kappaleessa (esimerkki 95).

95) **Unisieppari** on pajurenkaasta, lankaverkosta ja höyhenistä tehty [Pohjois-Amerikan intiaanien](#) taikakalu, joita myydään matkamuistoiksi ja koriste-esineiksi. Se on alun perin [ojibwa-intiaanien](#) perinnettä^[1], mutta 1960-luvun kansalaisuusliike nosti unisiepparin yleisen intiaankulttuurin tunnukseksi. Nykyisin myös [new age](#) -ryhmät valmistavat ja myyvät unisieppareita.^{[2][3]} Unisiepparin alkuperäinen tarkoitus on suodattaa pahat [unet](#) ja ajatukset päästämällä lävitseen hyvät unet ja ajatukset,^[4] mutta nykyisin se on usein koriste-esine.

Aineiston artikkeleissa kirjoittaja näyttäytyy asioiden tietäjänä ja määrittelijänä: eräänlaisena todellisuuden jäsentäjänä. Useimmiten ei oteta kantaa siihen, mistä tieto tulee tai kuka väittää, että joku asia on tietyllä tavalla. Mikäli artikkeliin on merkitty lähdeviitteet, voidaan tiedon alkuperä nähdä, mutta monessa artikkelissa lähdeviitteitä on puutteellisesti tai ei lainkaan, mikä asettaa kirjoittajat auktoriteeteiksi. Asioista kerrotaan monesti passiivissa tai muuten tiedon lähdeä henkilöimättä. Tämä havainnollistuu hyvin liikennettä käsittelevässä artikkelissa (esimerkki 96).

96) *On havaittu*, että välineitä hyväksikäyttävä [kommunikaatio](#) itse asiassa luo lisää vuorovai-
kutusta, myös henkilökohtaisella tasolla. Liikenteen kasvuun liittyy kommunikaatiota, joka on
elintärkeää kehittyneille liikennejärjestelmille. Junien ja lentokoneiden *on tiedettävä* toisten ju-
nien ja lentokoneiden sijainti, ja niiden on pystyttävä kommunikoimaan. *On myös havaittu*, että
tietoliikenteen lisääntyminen lisää liikennettä ja päinvastoin.

Silloinkin, kun Wikipedia-artikkeleissa viitataan johonkin auktoriteettiin, on tämä usein
varsin epämääräinen. Kuten esimerkistä (97) voi havaita, muun muassa Merkurius-
artikkelissa nojataan *toisten tutkijoiden* käsityksiin kyseisiä tutkijoita sen tarkemmin erittele-
mättä.

97) Merkuriuksen suuri rautaydin saattaa johtua siitä, että Aurinkokunnan syntyvaiheessa suuri
kappale törmäsi planeettaan suurella nopeudella ja heitti osan planeetan silikaattivaipasta pois.
Tämän törmääjän massa olisi ollut arviolta kuudesosa Merkuriuksen massasta. *Toisten tutkijoi-
den mukaan* alkuaurinko olisi kuumentanut Merkuriuksen korkeaan lämpötilaan ja höyrystänyt
siitä osia pois.

Myös Daniel Craigista kertovassa artikkelissa (esimerkki 98) tukeudutaan hyvin epämääräi-
sesti määritellyn tahon mielipiteeseen, mutta sen sijaan Vares-elokuva-artikkelin elokuvan
vastaanottoa käsittelevässä luvussa mainitaan nimeltä kriitikot, joiden arvioihin viitataan
(esimerkki 99).

98) Osa kansainvälisestä lehdistöstä pitää Craigia jopa parhaana Bondina.

99) [Keskisuomalaisen](#) kriitikko [Marko Ahonen](#) ja HS:n Avola nostavat molemmat esiin eloku-
van yhtäläisyydet [Quentin Tarantinon](#) ja [Robert Rodriguezin](#) tyyliin. Ahonen kritisoi elokuvan
heikkoa käsikirjoitusta ja ”tarinankerronnallista sekasotkua”. Hänen mukaansa ”lopputulos on
lähempänä freakshowta kuin rikosjännäriä”.^[8]

Pääosin artikkeleissa muodostuu tiedosta kuitenkin hyvin selvä kuva, jota ei kyseen-
alaisteta. Tämä yksioikoisuus peittää alleen sen, että artikkeleista löytyvä tieto on aina jonkun
valikoimaa eikä suinkaan neutraalia, vaikka kertovat, tietoa antavat lauseet ja usein käytetty
passiivimuoto luovatkin vaikutelman neutraaliudesta. Niissäkin tapauksissa, joissa lähdeviit-
teet on merkitty artikkeliin kattavasti, ovat sen sisältämä informaatio ja asioiden järjestys silti
kirjoittajien valitsemia. Esimerkiksi Daniel Craigista kertova artikkeli koostuu lähes ainoas-
taan luvusta *Bondina*, jossa kerrotaan miehen urasta James Bondin näyttelijänä. Sen lisäksi
artikkelissa on kahden virkkeen johdanto ja kahden virkkeen pituinen Yksityiselämästä-luku.
Kyseisen artikkelin kirjoittajat ovat nähneet olennaiseksi käsitellä lähinnä Daniel Craigin uraa
James Bondina. He ovat siis valinneet näkökulmansa ja toteuttaneet sitä artikkelissa. Tietokir-

jatekstissä pyritään määrittelemään ilmiö ja kertomaan siitä olennaiset asiat, joten artikkeliin valitut tiedot nousevat varsin suureen osaan. Ne määrittelevät kyseisen asian tai ilmiön. Näin kirjoittajien maailmankuva heijastuu väistämättä heidän kirjoittamissaan Wikipedia-artikkeleissa.

Joissakin artikkeleissa myös rikotaan neutraaliuden vaikutelmaa, ja kirjoittajan suhtautuminen käsittelemäänsä asiaan tulee selvemmin esille. Brasilia-artikkelin intiaaneista kertovan kohdan sävystä käy ilmi, että kirjoittaja paheksuu intiaanien kohtelua Brasiliassa (esimerkki 100). Tähän viittaavat esimerkiksi käytetyt kytkentäkeinot, *mutta*-konjunktio ja konnektiivit *edes* ja *vasta*. Intiaanien heikkoa asemaa korostetaan myös selvästi sävyä ilmaisevilla *varsin heikot* ja *kohtuuttomasti*.

100) Intiaanien oikeudet ovat olleet *varsin heikot* pitkälle 1990-luvulle saakka. *Vielä* 1950-luvulla intiaaneilla ei ollut juuri mitään oikeuksia edes maihinsa, missä he asuivat. Vuoden 1988 perustuslaki tunnusti intiaanien oikeuden perinteisiin maihinsa, *mutta* määrätty viiden vuoden siirtymisaika venyi monin paikoin *kohtuuttomasti*.^[39] *Vasta* 2000-luvulla intiaanien elinoloja on alettu parantaa.

Kuka pelkää mustaa miestä -artikkelissa kirjoittaja on katsonut aiheelliseksi ottaa kantaa kyseisen leikin kokemiseen rasistiseksi ja pyrkii leikin alkuperää avaamalla epäsuorasti puolustamaan leikkiä. Tämä ilmenee seuraavasta esimerkistä (101):

101) Jotkut kokevat ”kuka pelkää mustaa miestä” -leikin rasistiseksi, koska leikissä ”musta mies on varottava asia”. Alkuperäinen ”musta mies” ei kuitenkaan viittaa etniseen alkuperään vaan katoliseen pyhimykseen Pyhään Jaakkoon, joka myös heittää kylmän kiven jaakonpäivänä 25. heinäkuuta.^[1]

Edellisten esimerkkien kaltaisissa kohdissa korostuu kirjoittajien vapaus käsitellä aihetta ja rikkoutuu tekstien luoma neutraali vaikutelma. Ne saavat huomion kiinnittymään siihen, että artikkelien kirjoittajat ovat olleet vapaita päättämään, miten aihetta artikkelissa käsitellään ja mitä asioita siitä tuodaan esille. Tietosanakirjatekstille on tyypillistä neutraalius ja puolueettomuus, kun taas sosiaalisen median teksteissä perinteisesti nostetaan esiin omia mielipiteitä ja näkökulmia. Koska Wikipedia on sekoitus molempia, hypoteesini oli, että nämä kaksi aspektia sekoittuvat aineistossa. Neutraalista vaikutelmasta poikkeavaa ainesta on artikkeleissa kuitenkin hyvin vähän eikä se esiintyessään ole räikeää.

Wikipedia-artikkeleissa huomionarvoista on, että vaikka ne on kirjoitettu tavalla, joka asettaa kirjoittajat auktoriteettiasemaan oikean tiedon jakajiksi ja lukijat vastaanottajiksi, ei tämä roolijako ole pysyvä. Wikipedian tapauksessa lukijat eivät ole passiivisia vastaanottajia, vaan voivat aina halutessaan muokata lukemaansa. Näin kirjoittajat ja lukijat konstruoivat

todellisuutta yhdessä. Sosiaalisen median määritelmään nojautuen voidaankin sanoa, että kirjoittajat ja lukijat rakentavat yhteisiä merkityksiä sisältöjen, yhteisöjen ja verkkoteknologioiden avulla vertais- ja käyttötuotannon kautta.

Sivusto myös kutsuu muokkaamaan. Leidenin pullo -artikkelin alussa huomautetaan:

102) Tähän artikkeliin tai osioon ei ole merkitty lähteitä, joten tiedot kannattaa tarkistaa muista tietolähteistä.

Voit lisätä artikkeliin [tarkistettavissa olevia](#) lähteitä ja merkitä ne [ohjeen](#) mukaan.

Osallistujaroolit siis sekoittuvat Wikipedia-artikkeleissa, kun osallistujat voivat olla sekä tiedon tuottajia että vastaanottajia. Roolien sekoittuminen näkyy artikkeleissa esimerkiksi niin, että niissä ei ole tekstistä toiseen toistuvaa samankaltaista jaottelua lukuihin eivätkä niiden lähdeviittauskäytännöt toistu samankaltaisina artikkelista toiseen. Nämä ovat asioita, joihin kirjoittajat ovat saaneet itse vaikuttaa. Myös koheesiokeinot vaihtelevat varsin paljon eri artikkelien välillä.

Ilmenemismuodoltaan teksti pyrkii siis olemaan yleistajuista ja ymmärrettävää. Asioita ei juurikaan kyseenalaisteta. Toisaalta Wikipedia-artikkeleissa ei kuitenkaan pyritä avaamaan jokaista aiheeseen liittyvää käsitettä, vaan usein tämä käsite on linkki uuteen artikkeliin. Esimerkiksi Shetlanninlammaskoira-artikkeleissa lonkkaniveldysplaasiasta sanotaan seuraavaa (esimerkki 103):

103) Shetlanninlammaskoira on suhteellisen terve rotu. Sillä esiintyy jonkin verran lonkkavikaa eli [lonkkaniveldysplasiaa](#). Sen yleisin sairaus on silmänsairaus CEA (Collie Eye Anomaly), joka voi vakavimmillaan johtaa jopa sokeutumiseen.

Mikäli lukija haluaa tietää lonkkaniveldysplaasiasta enemmän, voi hän linkkiä klikkaamalla siirtyä seuraavaan artikkeliin Wikipediassa ja palata sen jälkeen takaisin lukemaansa artikkeliin. Wikipedia-artikkelit toimivat siis hyperteksteinä. Niissä lukijalla on vapaus valita lukemansa järjestys ja se, kuinka paljon hän haluaa asioihin syventyä. Näin lukijaa osallistetaan eikä hän jää ainoastaan tiedon passiiviseksi vastaanottajaksi.

Rekisteripiirteitä tarkastelemalla voidaan havaita, että tilannekontekstin muuttujat korreloivat metafunktioiden kanssa. Ne tuovat näkyviin sitä, kuinka kieltä käytetään kuvaamaan ja luomaan maailmaamme (ideationaalinen metafunktio) ja kuinka kirjoittaja ilmaisee asennoitumisensa sanomaansa ja pyrkii vaikuttamaan muiden asenteisiin (interpersonaalinen metafunktio). Näiden kahden metafunktion voi nähdä kattavan kaiken kielenkäytön: ympäristön

ymmärtämisen ja siinä toisten kanssa toimimisen (Halliday 1978: 112–113; Halliday 1985a: xiii).

Wikipedia-artikkelien rekisteri on joiltakin ominaisuuksiltaan varsin avoin. Artikkelit vaihtelevat paljon siinä, miten ne on järjestetty lukuihin ja miten lähteitä on käytetty. Myös koheesion keinoja on käytetty toisissa artikkeleissa huomattavasti enemmän kuin toisissa. Artikkelien alusta kuitenkin asettaa muodolle rajoituksia, joten ulkoisesti ne kaikki muistuttavat toisiaan. Linkitys on kaikissa artikkeleissa samanlainen. Jokainen artikkeli myös alkaa samankaltaisella määrittävällä lauseella. Edellä käsitellyt tilannekontekstin muuttujat tuovat esiin, että artikkelien taustalla toimivat yhtenäiset funktiot. Tiedon antaminen ja asioiden määrittelemine yhdistää kaikkia tekstejä.

6.3 Pohdintaa

Tutkimukseni tarkoitus oli aineiston rakenteellisia ja tekstuaalisia ominaisuuksia tarkastelemalla muodostaa käsitys Wikipedia-artikkelista tekstilajina. Pohjasin tutkimukseni systeemifunktionaaliseen teoriaan, erityisesti M. A. K. Hallidayn käsityksiin funktionaalisesta tekstilajitutkimuksesta. Tulkitsin siis kieltä sosiokulttuurisessa kontekstissa eli pyrin ottamaan huomioon myös yhteisön sosiaaliset konventiot ja selittämään tekstien rakennetta funktiosta käsin.

Ensimmäinen tutkimuskysymykseni käsitteli artikkelien koherenssia. Koin tekstien yhtenäisyyden tutkimisen tärkeäksi, sillä tekstin täytyy olla ympäristössään merkityksellinen kokonaisuus, jotta se ylipäänsä on teksti. Myös Wikipedialle ominainen tekstien tuottamista pa, se, että Wikipedia-artikkeleja kirjoitetaan ja muokataan useiden kirjoittajien toimesta, teki koherenssin analysoimisesta olennaista. Keskityin erityisesti tarkastelemaan sitä, kuinka artikkeleissa ilmenevä kieliopillinen ja sanastollinen koheesio rakentaa tekstien yhtenäisyyttä. Tämän lisäksi tarkastelin artikkelien teema- ja informaatorakennetta, joka on olennainen osa koherenssia ja niin ollen myös tekstin välttämätön elementti. Yhtenäisessä tekstissä informaatorakenne tukee tutun ja uuden tiedon annostelemista ja on tekstin aiheen kuljettamisen kannalta relevantti.

Kaikki artikkelit osoittautuivat odotetusti kontekstissaan koherenteiksi, mutta eroja oli siinä, miten paljon tekstin yhtenäisyyttä rakennettiin koheesion avulla. Myös käytetyt koheesiokeinot vaihtelivat artikkelien välillä. Teksteille oli tyypillistä, että niissä ei pyritty koheesiokeinojen avulla erityisen vivahteikkaaseen tai vaihtelevaan kieleen. Toistoa vähentävät ja tarkkaa lukemista vaativat koheesiokeinot, kuten referenssi, ellipsit, synonymia ja yleismerkityksiset sanat eivät nousseet vallitsevaan osaan. Yleiskäsitteitä esiintyi jonkin verran,

mutta nekään eivät varsinaisesti lisänneet tekstin vivahteikkuutta. Samaan viittasi myös tekstin informaatorakenne, sillä etenevän teeman vähäisempi käyttö pysyvään teemaan verrattuna ei lisännyt tekstin jouhevuuatta.

Toisessa tutkimuskysymyksessäni pyrin määrittelemään Wikipedia-artikkelin tekstilajille ominaisia piirteitä analysoimalla jo mainitun koherenssin lisäksi sitä, miten artikkelit on muodostettu, ja mitkä ovat niille olennaisia rakenteellisia ominaisuuksia. Lisäksi analysoin aineistosta esiin nostamieni piirteiden funktioita eli tekstuaalisen metafunktion lisäksi pyrin huomioimaan myös ideationaalisen ja interpersonaalisen metafunktion.

Tekstilajille tyypillisiksi piirteiksi nousivat analyysin perusteella prosessinomaisuus ja avoimesta tuotantotavasta johtuva variaatio. Ne näkyvät aineistossa tietynlaisena keskeneräisyytenä. Jotkut artikkelit ovat hyvin lyhyitä, toisia ei ole jaettu lukuihin ja lähdeviitteiden merkinnässä on paljon puutteita. Monet kappaleet koostuvat vain yhdestä virkkeestä. Artikkelien sisältö ja asioiden esittämisjärjestys vaihtelevat tekstistä toiseen, mikä kertoo siitä, että artikkelien kirjoittajat ja muokkaajat ovat voineet varsin vapaasti jäsentää artikkeleita. Artikkeleista nousee esiin heidän käsityksensä maailmasta sekä siinä, mitä aiheesta kerrotaan että siinä, miten kyseinen tieto on järjestetty.

Tekstilajille tyypilliseksi funktioksi nousi rakenteellisten ja tekstuaalisten piirteiden analyysin perusteella mahdollisuus nopeaan, silmäilevään lukemiseen. Teema- ja informaatorakenne, koheesiokkeinot, artikkelien yhdenmukaiset aloitukset, lyhyet kappaleet, tiiviys ja linkitys osoittavat kaikki tähän suuntaan. Nämä ominaisuudet edesauttavat myös artikkelien ei-lineaarista lukemista ja tekevät niiden muokkaamisen helpommaksi.

Kolmas tutkimuskysymykseni käsitteli aineiston rekisteriä. Analysoin sitä tutkimalla tilannekontekstin muuttujia. Tekstin määrittämisessä myös tilannekonteksti on olennainen. Tilannekontekstin kolme muuttujaa – ala, osallistujaroolit ja ilmenemismuoto – auttoivat tulkitsemaan ympäristöä, jossa merkityksiä vaihdetaan.

Wikipedia-artikkelien alana toimii tiedon tuottaminen ja jakaminen. Kaikille artikkeleille on yhteistä se, että ne keskittyvät antamaan tietoa jo otsikossa kerrotusta aiheesta. Jokaisessa tekstissä määritellään käsiteltävä aihe ja pysytään tiiviisti kyseisen aiheen käsittelyssä. Ilmenemismuodoltaan teksti on pääosin yleistajuista ja ymmärrettävää. Asioita ei juurikaan kyseenalaisteta. Tämä saa kirjoittajan näyttäytymään asioiden tietäjänä ja määrittäjänä ja lukijan melko passiivisena vastaanottajana. Nämä roolit eivät kuitenkaan ole pysyviä, vaan lukija voi halutessaan siirtyä muokkaajaksi. Näin osallistujaroolit artikkeleissa sekoittuvat.

Aineistoni analyysi toi esiin Wikipedia-artikkelien variaation rakenteellisissa ja tekstuaalisissa piirteissä. Jo kolmenkymmenen artikkelin laajuisen aineiston tarkastelu paljastaa,

että kiveen hakattujen päätelmien tekeminen Wikipedia-artikkeleista olisi yksinkertaistamista. Aineistosta nousee esiin tekstilajille tyypillisiä piirteitä, mutta käytännössä mitkään niistä eivät esiinny aivan kaikissa artikkeleissa ja toisaalta esiintyvät joissakin tietyissä teksteissä hyvin laajasti.

Kakkuri-Knuutila ja Kylänpää ottavat artikkelissaan kantaa siihen, voiko tekstejä ylipäätään jakaa johonkin luokkaan niiden yhteisten piirteiden perusteella eli onko olemassa tekstilajille, essentiaalisia, välttämättömiä piirteitä (Kaakkuri-Knuutila & Kylänpää 2012: 160). Mielenkiintoinen kysymys onkin se, onko Wikipedia-artikkeleilla välttämättömiä piirteitä. Onko joitakin ominaisuuksia, joita kyseisistä teksteistä on pakko löytyä, jotta ne voidaan nähdä tilanteessaan yhtenäisiksi merkityksellisiksi kokonaisuuksiksi? Sellaisia ominaisuuksia eivät selvästikään ole aiheen jakaminen lukuihin, kattava lähteidenkäyttö tai tietynlaiset koheesiokeinot. Kaikista artikkeleista kuitenkin nousi esiin artikkelin käsittelemän aiheen määrittäminen, yleensä jo heti ensimmäisessä virkkeessä. Voidaan siis ajatella, että tietokirjatekstin olennaisin funktio, sen käsittelemän aiheen määrittäminen, yhdistää myös Wikipedia-artikkeleita.

Kaakkuri-Knuutila ja Kylänpää huomattavat myös tekstilajin historiallisesta muutoksesta puhuessaan, että useimmiten korkeintaan osa ominaisuuksista on välttämättömiä eikä useimpia ominaisuuksia ole tarvetta pitää muuttumattomina (2012: 160). Wikipedia-artikkelit ovat siis säilyttäneet tietosanakirja-artikkelin tärkeimmän ominaisuuden, artikkelin käsittelemän asian määrittelyn. Muut analyysissa esille nostamani aineistolle tyypilliset ominaisuudet toki määrittelevät Wikipedia-artikkeliä tekstilajina, mutta eivät rajaa joukosta pois artikkeleita, joissa näitä ominaisuuksia ei esiinny.

Wikipedia voidaan määrittää osaksi sosiaalista mediaa. Sosiaaliselle medialle olennaiset käyttäjälähtöisyys ja vuorovaikutuksellisuus, muokattavuus ja multimediaisuus näkyvätkin aineistossa. Näiden piirteiden valossa tarkasteltuna Wikipedia-artikkelit ovat selvästi liikkuneet eteenpäin julkaistuista tietosanakirja-artikkeleista. Sosiaaliselle medialle niin ikään tyypillinen henkilökohtaisten mielipiteiden näkyminen tuotoksessa on tietosanakirja-artikkelille kuitenkin vastakkaista. Tässä suhteessa Wikipedia-artikkelit ovatkin lähempänä painettua tietosanakirjaa, sillä niissä ilmaistaan asennetta kielellisesti varsin vähän. Pääosin ne antavat neutraalin vaikutelman.

Erkkola (2009: 88) määrittelee sosiaalisella medially olevan vaikutusta yhteiskuntaan ja kulttuuriin. Wikipedia on selvästi muuttanut tietosanakirjaa tekstilajina. Internet-alusta ja yhteistoiminnallinen tuotantotapa ovat vaikuttaneet tekstilaji- ja rekisteripiirteisiin. Lisäksi ne ovat vaikuttaneet siihen, miten ylipäänsä näemme tiedon ja sen, kenellä sitä on oikeus tuottaa

ja jäsenellä. Tere Vadén näkee muutoksen pysyvistä kappalemaisesta tietokäsityksestä muuttuvaan ja koko ajan elävään tietokäsitykseen eräänä Wikipedian olennaisimmista vaikutuksista (2007: 11). Wikipedia-artikkeleista esiin noussut variaatio kiinnittää huomion siihen, että samalla, kun jotakin artikkelia muokataan, muokataan myös sen sisältämää tietoa.

Wikipedia on nykyisin niin merkittävässä asemassa myös suomalaisten tiedonhaussa, että koin sen tekstilaji- ja rekisteripiirteiden tarkastelun relevantiksi. Tutkimukseni antoi uutta tietoa Wikipedia-artikkelien koherenssista, tekstilajista ja rekisteristä. On kuitenkin olennaista huomata, että tutkimukseni on vain raapaisu suomalaisten Wikipedia-artikkelien kentässä. Valitsin tutkimukseeni satunnaisotannalla 30 artikkelia Wikipedian pääluokista. Toisella otannalla olisi todennäköisesti päätyttyä ainakin osittain erilaisiin tuloksiin. Keskityin analyysissäni etupäässä artikkelien rakenteen, koherenssin ja tilannekontekstin tarkasteluun ja tekstilajin funktioiden määrittämiseen näiden kautta. Joku toinen tutkija olisi todennäköisesti tehnyt erilaisia ratkaisuja ja mahdollisesti päätyttyä poikkeaviin tuloksiin.

Wikipedia tarjoaa mahdollisuuksia monenlaiselle lisätutkimukselle tekstilajitutkimuksen kentällä. Tekstilaji on sosiokulttuurinen käsite, jonka konventiot syntyvät kielenkäyttäjien muodostamassa yhteisössä (Kalliokoski 2006: 240). Näin on ehdottomasti myös Wikipedian kohdalla. Wikipedian yhteistoiminnallisen ja vapaan tuottamistavan tarkempi tutkimus tarjoaisi mielenkiintoisen näkökulman Wikipedia-artikkelien prosessinomaiseen rakentumiseen. Olisi kiinnostavaa selvittää, miten Wikipedia-artikkeleita kirjoitetaan ja muokataan: saako Wikipedia-artikkeli muokkauksien myötä esimerkiksi lisää sosiaaliselle medialle tyypillisiä piirteitä. Mielenkiintoista olisi myös verrata suomenkielisiä Wikipedia-artikkeleita sekä painettujen tietosanakirjojen teksteihin että tieteellisiin artikkeleihin. Wikipedia-artikkeleita hyperteksteinä olisi myös havainnollistavaa analysoida tarkemmin. Tässä tutkimuksessa kiinnitän huomiota Wikipediasta löytyviin hypertekstin piirteisiin, mutta ne eivät nouse kovin suureen osaan analyysissäni.

LÄHTEET

- Baytiyeh, Hoda & Pfaffman, Jay 2010: Volunteers in Wikipedia: Why the community matters. *Educational Technology & Society*. 13 (2) s. 128–140.
- Bhatia, Vijay K. 2004: *Worlds of written discourse: A genre-based view*. London: Continuum.
- Bublitz, Wolfram 1999: Introduction: Views of Coherence. – Wolfram Bublitz, Uta Lenk & Eija Ventola (toim.), *Coherence in spoken and written discourse: how to create it and how to describe it: selected papers from the International Workshop on Coherence, Augsburg, 24-27 April 1997* s. 1–8. Amsterdam: John Benjamins Publishing.
- Cruse, Alan D. 2004: *Meaning in language. An introduction to semantics and pragmatics*. Oxford: Oxford University Press.
- Emigh, William & Herring, Susan 2005: Collaborative Authoring on the Web: A Genre Analysis of Online Encyclopedias. *Proceedings of the 38th Hawaii International Conference on System Sciences*. –
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.226&rep=rep1&type=pdf>
 20.03.2015.
- Ensslin, Astrid 2007: *Canonizing hypertext: explorations and constructions*. New York: Continuum.
- Erkkola, Jussi-Pekka 2009: *Vuorovaikutteisuus sosiaalisessa mediassa: sosiaalisen median käsitemallia*. Pro gradu -tutkielma. Jyväskylän yliopiston viestintätieteiden laitos. –
<https://jyx.jyu.fi/dspace/handle/123456789/24942> 12.04.2015.
- Factum. Uusi tietosanakirja*. 2003. WSOY (toim.). Helsinki: WSOY.
- Fritz, Gerd 1999: Coherence in hypertext. – Wolfram Bublitz, Uta Lenk & Eija Ventola (toim.), *Coherence in spoken and written discourse: how to create it and how to describe it: selected papers from the International Workshop on Coherence, Augsburg, 24-27 April 1997* s. 221–234. Amsterdam: John Benjamins Publishing.
- Giles, Jim 2005: Special Report Internet encyclopaedias go head to head. – *Nature* 438. –
<http://www.nature.com/nature/journal/v438/n7070/full/438900a.html>
- Hakulinen, Auli & Karlsson, Fred 1995: *Nykysuomen lauseoppi*. Helsinki: Suomalaisen kirjallisuuden seura.
- Halliday, M. A. K. 1978: *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.
- 1985a: *An Introduction to Functional Grammar*. London: Edward Arnold.
- 1985b: *Spoken and written language*. London: Edward Arnold.

- Halliday, M. A. K. & Hasan, Ruqaiya 1976: *Cohesion in English*. London: Longman.
- 1985: *Language, context, and text: Aspects of language in a social-semiotic perspective*. Geelong: Deaking University Press.
- Heikkinen, Vesa 2001: Teksteihin tunkeutuvat todellisuudet. – Vesa Heikkinen, Pirjo Hiidenmaa & Ulla Tiililä (toim.), *Teksti työnä, virka kielenä* s. 116–214. Helsinki: Gaudeamus.
- 2001: Tekstuaalinen pirunnyrkki. – Vesa Heikkinen, Pirjo Hiidenmaa & Ulla Tiililä (toim.), *Teksti työnä, virka kielenä* s. 63–115. Helsinki: Gaudeamus.
- Heikkinen, Vesa & Voutilainen, Eero 2012: Genre – monitieteinen näkökulma. – Vesa Heikkinen, Eero Voutilainen, Petri Lauerma, Ulla Tiililä & Mikko Lounala (toim.), *Genreanalyysi-tekstilajitutkimuksen käsikirja* s. 17–47. Helsinki: Gaudeamus.
- Hiidenmaa, Pirjo 2000: Lingvistinen tekstintutkimus. – Kari Sajavaara & Arja Piirainen-Marsh (toim.), *Kieli, diskurssi ja yhteisö* s. 161–190. Jyväskylä: Soveltavan kielentutkimuksen keskus.
- 2001: Poimintoja virkakielen rekisteristä. – Vesa Heikkinen, Pirjo Hiidenmaa & Ulla Tiililä (toim.), *Teksti työnä, virka kielenä* s. 35–62.
- Huvila, Isto 2010: "Where does the information come from? Information source use patterns in Wikipedia" *Information Research*, 15 (3). – <http://www.informationr.net/ir/15-3/paper433.html> 22.03.2015 18.01.2015.
- Häkkinen, Kaisa 1996: *Kielitieteen perusteet*. Helsinki: Suomalaisen kirjallisuuden seura.
- Häkli, Esko 1981: *Tiedon taika: ensyklopedioista ennen ja nyt*. Helsinki: Otava.
- Ihalainen, Maija 2013: *Vikaa Wikipediassa? - opiskelijoiden käsitykset Wikipedian laadusta*. Pro gradu -tutkielma. Tampereen yliopiston informaatiotutkimuksen ja interaktiivisen median laitos. – <http://tampub.uta.fi/bitstream/handle/10024/94852/GRADU-1389945813.pdf?sequence=1> 02.02.2015.
- Jantunen, Jarmo 2004: *Synonymia ja käännessuomi: korpusnäkökulma samanmerkityksisyyden kontekstuaalisuuteen ja käännöskielen leksikaalisiin erityispiirteisiin*. Joensuu: Joensuun yliopisto.
- Jemielniak, Dariusz 2014: *Common Knowledge? An Ethnography of Wikipedia*. Stanford: Stanford University Press.
- Kakkuri-Knuuttila, Marja-Liisa & Kylänpää, Antti 2012: Tieteellisen artikkelin rakenne: historiaa ja esihistoriaa. – Vesa Heikkinen, Eero Voutilainen, Petri Lauerma, Ulla Tiililä & Mikko Lounala (toim.), *Genreanalyysi – tekstilajitutkimuksen käytäntöä* s. 157–185. Helsinki: Kotimaisten kielten keskus.

- Kalliokoski, Jyrki 2006: Tekstilajin taju ja toisella kielellä kirjoittaminen. – Anne Mäntynen, Susanna Shore & Anna Solin (toim.), *Genre – tekstilaji* s. 240–265. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Karlsson, Fred 2006: *Yleinen kielitiede*. Helsinki: Helsinki University Press.
- Koistinen, Olavi 2013: HS selvitti: Näin luotettava Wikipedia on. – *Helsingin Sanomat* 30.11.2013 – <http://www.hs.fi/tiede/a1305754303586> 02.03.2015.
- KS = *Kielitoimiston sanakirja*. Kotimaisten kielten keskuksen julkaisuja 166. Helsinki: Kotimaisten kielten keskus ja Kielikone 2012. – <http://mot.kielikone.fi/mot/jyu/netmot.exe?motportal=80> 01.02.2015.
- Kuosmanen, Riitta-Liisa, Halinen, Ari & Haataja, Lauri 2006: *Facta-tietosanakirja*. Helsinki: WSOY.
- Lietsala, Katri & Sirkkunen, Esa 2008: *Social media: Introduction to the tools and processes of participatory economy*. Tampere: Tampere University Press.
- Luukka, Minna-Riitta 2002: M. A. K. Halliday ja systeemis-funktionaalinen kielitiede. – Hannele Dufva & Mika Lähteenmäki (toim.), *Kielentutkimuksen klassikoita*. Jyväskylä: Soveltavan kielentutkimuksen keskus.
- Martin, J. R. & Rose, David 2008: *Genre relations. Mapping culture*. London: Equinox Publishing Ltd.
- Mauranen, Anna 2006: Genre, käänös ja korpus. Elämäntaito-oppaat tarkastelussa. – Anne Mäntynen, Susanna Shore & Anna Solin (toim.), *Genre – tekstilaji* s. 214–239. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Raevaara, Liisa & Sorjonen, Marja-Leena 2006: Vuorovaikutuksen osanottajien toiminta ja genre. – Anne Mäntynen, Susanna Shore & Anna Solin (toim.), *Genre – tekstilaji* s. 122–150. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Ranta, Tuula 2007: *Kirjoittamisprosessi teksteinä. Tekstilingvistinen näkökulma abiturienttien tekstintuottamismenettelyihin*. Joensuu: Joensuun yliopisto.
- Santini, Marina 2007: Characterizing Genres of Web Pages: Genre Hybridism and Individualization. *Proceedings of the 40th Hawaii International Conference on System Sciences*. – <http://www.computer.org/csdl/proceedings/hicss/2007/2755/00/27550071.pdf> 20.01.2015.
- Saukkonen, Pauli 2001: *Maailman hahmottaminen teksteinä. Tekstirakenteen ja tekstilajien analyysia*. Helsinki: Yliopistopaino.
- Shore, Susanna 2012: Kieli, kielenkäyttö ja kielenkäytön lajit systeemis-funktionaalisessa teoriassa. – Vesa Heikkinen, Eero Voutilainen, Petri Lauerma, Ulla Tiililä & Mikko Louhala (toim.), *Genreanalyysi-tekstilajitutkimuksen käsikirja* s. 17–47. Helsinki: Gaudeamus.
- Shore, Susanna & Mäntynen, Anne 2006: Johdanto – Anne Mäntynen, Susanna Shore & Anna Solin (toim.), *Genre – tekstilaji* s. 9–41. Helsinki: Suomalaisen Kirjallisuuden Seura.

——— 2006: Tekstilajien risteyksessä. – Anne Mäntynen, Susanna Shore & Anna Solin (toim.), *Genre – tekstilaji* s. 283–302. Helsinki: Suomalaisen Kirjallisuuden Seura.

Taavitsainen, Irma 2006: Genredynamiikkaa aikaperspektiivissä. – Anne Mäntynen, Susanna Shore & Anna Solin (toim.), *Genre – tekstilaji* s. 266–282. Helsinki: Suomalaisen Kirjallisuuden Seura.

Tanskanen, Sanna-Kaisa 2006: *Collaborating towards coherence: lexical cohesion in English discourse*. Amsterdam: John Benjamins Publishing.

Tereszkiewicz, Anna 2010: *Genre analysis of online encyclopedias: the case of Wikipedia*. Kraków: Jagiellonian University Press.

Tiikkaja, Samuli 2012: Internet tappoi suomalaiset tietosanakirjat - myös netissä. – *Helsingin Sanomat* 26.03.2012
<http://www.hs.fi/paivanlehti/arkisto/Internet+tappoi+suomalaiset+tietosanakirjat+-+my%C3%B6s+netiss%C3%A4/aaHS20120326SI1KU01gcr?src=haku&ref=arkisto%2F>
 12.01.2015.

Todorov, Tzvetan 1990: *Genres in discourse*. Cambridge: Cambridge university press.

Tuominen, Kimmo 2006: *Tiedon partaalla: kuinka hallita informaatiotulvaa*. Helsinki: BTJ kirjastopalvelu.

Vadén, Tero 2007: Wikipedia ja filosofit: huomioita wikipedioiden epistemologisesta potentiaalista. *Tieteessä tapahtuu*, 25 (7) s. 11–19. – <http://ojs.tsv.fi/index.php/tt/issue/view/54>
 03.05.2015.

Ventola, Eija 1999: Semiotic Spanning at Conferences: Cohesion and coherence in and across conference papers and their discussions. – Wolfram Bublitz, Uta Lenk & Eija Ventola (toim.), *Coherence in spoken and written discourse: how to create it and how to describe it: selected papers from the International Workshop on Coherence, Augsburg, 24-27 April 1997* s.101–124. Amsterdam: John Benjamins Publishing.

Ventola, Eija 2006: Genre systeemis-funktionaalisessa kielitieteessä. Esimerkkinä asiointitilanteet. Anne Mäntynen, Susanna Shore & Anna Solin (toim.), *Genre – tekstilaji* s. 96–121. Helsinki: Suomalaisen Kirjallisuuden Seura.

Viégas, Fernanda B., Wattenberg, Martin & Dave, Kushal 2004: Studying Cooperation and Conflict between Authors with history flow Visualizations. Proceedings of the SIGCHI conference on Human factors in computing systems. 6 (1). – http://alumni.media.mit.edu/~fviegas/papers/history_flow.pdf 28.03.2015.

VISK = Hakulinen, Auli, Vilkuna, Maria, Korhonen, Riitta, Koivisto, Vesa, Heinonen, Tarja Riitta & Alho, Irja 2004: *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura. – <http://scripta.kotus.fi/visk> 11.01.2015.

Wikipedia 2015: Wikipedia:Statistics – <http://en.wikipedia.org/wiki/Wikipedia:Statistics>.
14.02.2015

Wikipedia 2015: Ohje:Sisällys – <http://fi.wikipedia.org/wiki/Ohje:Sis%C3%A4llys>
02.03.2015.

Wikipedia 2015: Wikipedia – <http://fi.wikipedia.org/wiki/Wikipedia>. 02.03.2015.

Wikipedia 2015: Wikipedia:Viisi pilaria – http://fi.wikipedia.org/wiki/Wikipedia:Viisi_pilaria 02.03.2015.

Aineistolähteet

Bahai-usko: <http://fi.wikipedia.org/wiki/Bahai-usko> 15.01.2015

Biologinen antropologia: http://fi.wikipedia.org/wiki/Biologinen_antropologia 15.01.2015

Brasilia: <http://fi.wikipedia.org/wiki/Brasilia> 15.01.2015

Daniel Craig: http://fi.wikipedia.org/wiki/Daniel_Craig 15.01.2015

Ensimmäinen ristiretki Suomeen:

http://fi.wikipedia.org/wiki/Ensimm%C3%A4inen_ristiretki_Suomeen 15.01.2015

Fani: <http://fi.wikipedia.org/wiki/Fani> 15.01.2015

Fysioterapia: <http://fi.wikipedia.org/wiki/Fysioterapia> 15.01.2015

IKEA: <http://fi.wikipedia.org/wiki/IKEA> 15.01.2015

Juutinrauma: <http://fi.wikipedia.org/wiki/Juutinrauma> 15.01.2015

Konsertto: <http://fi.wikipedia.org/wiki/Konsertto> 15.01.2015

Kuka pelkää mustaa miestä:

http://fi.wikipedia.org/wiki/Kuka_pelk%C3%A4%C3%A4_mustaa_miest%C3%A4
15.01.2015

Leidenin pullo: http://fi.wikipedia.org/wiki/Leidenin_pullo 15.01.2015

Liikenne: <http://fi.wikipedia.org/wiki/Liikenne> 15.01.2015

Loch Nessin hirviö: http://fi.wikipedia.org/wiki/Loch_Nessin_hirvi%C3%B6 15.01.2015

Magmakivi: <http://fi.wikipedia.org/wiki/Magmakivi> 15.01.2015

Merkurius: <http://fi.wikipedia.org/wiki/Merkurius> 15.01.2015

Merlin (televisiosarja): http://fi.wikipedia.org/wiki/Merlin_%28televisiosarja%29 15.01.2015

Metonymia: <http://fi.wikipedia.org/wiki/Metonymia> 15.01.2015

Mikrobiologia: <http://fi.wikipedia.org/wiki/Mikrobiologia> 15.01.2015

Naisten äänioikeus: http://fi.wikipedia.org/wiki/Naisten_%C3%A4%C3%A4nioikeus
15.01.2015

Näkymätön käsi:

http://fi.wikipedia.org/wiki/N%C3%A4kym%C3%A4t%C3%B6n_k%C3%A4si 15.01.2015

Pakina: <http://fi.wikipedia.org/wiki/Pakina> 15.01.2015

Poro: <http://fi.wikipedia.org/wiki/Poro> 15.01.2015

Sammon taonta: http://fi.wikipedia.org/wiki/Sammon_taonta 15.01.2015

Shetlanninlammaskoira: <http://fi.wikipedia.org/wiki/Shetlanninlammaskoira> 15.01.2015

Siipi: <http://fi.wikipedia.org/wiki/Siipi> 15.01.2015

Särkänniemi: <http://fi.wikipedia.org/wiki/S%C3%A4rk%C3%A4nniemi> 15.01.2015

Unisieppari: <http://fi.wikipedia.org/wiki/Unisieppari> 15.01.2015

Vares – Sheriffi: http://fi.wikipedia.org/wiki/Vares_%E2%80%93_Sheriffi 15.01.2015

Yleinen kielitutkinto: http://fi.wikipedia.org/wiki/Yleinen_kielitutkinto 15.01.2015

