JYX JYVÄSKYLÄN YLIOPISTO UNIVERSITY OF JYVÄSKYLÄ

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Karkulehto, Sanna; Laippala, Veronika; Launis, Kati; Märsynaho, Jaana; Saviniemi, Maija; Sääskilahti, Minna

Title: Selkosten Proust taipuu moneen : Iijoki-korpus ja digitaalisen tekstilouhinnan mahdollisuudet

Year: 2022

Version: Published version

Copyright: © 2022 Kirjoittajat ja SKS

Rights: CC BY-NC-ND 4.0

Rights url: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the original version:

Karkulehto, S., Laippala, V., Launis, K., Märsynaho, J., Saviniemi, M., & Sääskilahti, M. (2022). Selkosten Proust taipuu moneen : Iijoki-korpus ja digitaalisen tekstilouhinnan mahdollisuudet. In M. Saviniemi (Ed.), Kalle Päätalo tutkijoiden silmin (pp. 170-189). Suomalaisen Kirjallisuuden Seura. Tietolipas, 271. https://oa.finlit.fi/site/books/e/10.21435/tl.271/

Kalle Päätalo tutkijoiden silmin

Toimittanut
Maija Saviniemi

Suomalaisen Kirjallisuuden Seura • Helsinki • 2022

TIETOLIPAS 271

Teos on Suomalaisen Kirjallisuuden Seuran nimeämien asiantuntijoiden tarkastama *:llä merkittyä lukua lukuun ottamatta.

VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

© 2022 Maija Saviniemi ja SKS

Lisenssi CC BY-NC-ND 4.0 International, ellei toisin mainita.

Kannen suunnittelu: Eija Hukka
Kannen kuva: Päätalo-instituutti / Professori Kalle Päätalon arkisto.
EPUB: Tero Salmén

ISBN 978-951-858-434-9 (nid.)
ISBN 978-951-858-435-6 (EPUB)
ISBN 978-951-858-436-3 (PDF)

ISSN 0562-6129 (nid.)
ISSN 2670-2584 (verkkojulkaisut)

DOI https://doi.org/10.21435/tl.271

Teos on lisensoitu Creative Commons CC BY-NC-ND 4.0 International -lisenssillä. Tutustu lisenssiin osoitteessa https://creativecommons.org/licenses/by-nc-nd/4.0/deed.fi.

Teos on avoimesti saatavissa osoitteessa https://doi.org/10.21435/tl.271 tai lukemalla tämä QR-koodi mobiililaitteella.

TEHTY SUOMESSA
PRODUCERAD I FINLAND

Hansaprint Oy, Turenki 2022

Selkosten Proust taipuu moneen – *Iijoki*-korpus ja digitaalisen tekstilouhinnan mahdollisuudet

Sanna Karkulehto
https://orcid.org/0000-0001-7267-9612

Veronika Laippala
https://orcid.org/0000-0002-7635-429X

Kati Launis
https://orcid.org/0000-0002-8498-7361

Jaana Märsynaho

Maija Saviniemi
https://orcid.org/0000-0003-2977-1143

Minna Sääskilahti

”Kun kokeilee kädellä, tuntee, että tässä kohtaa on harjanne”, siteeraa *Helsingin Sanomien* pitkäaikainen kulttuuritoimittaja Antti Majander (2019) kolumnissaan talonsa kellarin sementtilattiaa esittelevää Kalle Päätaloa ja jatkaa: ”Siis valuvika.” Tämä oman kellarin sementtilattian harjanteen huolellinen harmittelu kuvastaa Majanderin mukaan sitä, että Päätalo on Suomen uskottavin kirjailija. Vastaus kysymykseen, miksi valuvika tekee Päätalosta uskottavan, edellyttää kuitenkin taustoitusta. Miten *Helsingin Sanomien* kulttuuritoimittaja päätyi Kalle Päätalon kellariin Tampereelle todistamaan tämän lattiaansa aikaansaamaa harjannetta?

Kun Majander oli lähdössä haastattelemaan pian 79 vuotta täyttävää Päätaloa, oli hänen toimittajakollegansa Ilkka Malmberg – ainakin anekdootin mukaan – ehdottanut, että pikkutarkaksi ja huolelliseksi tiedetyltä rakennusmestari-kirjailijalta kysyttäisiin, oliko tämä tyytyväinen omaan, itse omilla käsillään ja omin avuin rakentamaansa taloon. Vastaus kysymykseen on arvattava: aivan tyytyväinenhän Kalle ei ollut. Kellarin sementtilattiasta löytyy nimittäin yksi harjan-

ne, jonka hän muistaa ottaa esille vielä elämänsä ehtoopuolellakin. Tämän rikkumattoman muistin sekä rehellisyyden yhdistelmän toimittaja nimeää osoitukseksi Päätalon uskottavuudesta.

Lähdemme tämän teoksen päätännössä Päätaloon usein liitetyn rehellisyyden tavoittelun, pikkutarkan muistin sekä muistamisen merkityksen motivoimina tarkastelemaan, mitä mahdollisuuksia jo edellä esitelty Oulun yliopiston ja valtakunnallisen Kielipankin julkaisema *Iijoki*-tekstikorpus (2019) voi tarjota. Tarkastelemme, mitä annettavaa sillä on yhtäältä Päätalo-tutkimukselle ja toisaalta niin kirjallisuuden, kielen ja kulttuurin kuin myös niiden kulttuuristen ja historiallisten kontekstien tutkimukselle. Pohdimme, millaisia tutkimusaiheita korpusavusteinen, korpusvetoinen ja korpuspohjainen lähestymistapa mahdollistavat.

On sinänsä selvää, että laaja, koko Päätalon *Iijoki*-sarjan sisältävä tekstikorpus avaa aivan uudenlaisia mahdollisuuksia tutkia sellaisia aiheita, joiden analyysiä tutkijoiden rajallinen muisti ja kapasiteetti käsitellä suuria aineistoja eivät ole aiemmin mahdollistaneet. Tässä kokoomateoksessa *Iijoki*-korpusta on jo hyödynnetty Saviniemen ja Sääskilahden luvuissa, ja korpus olisi soveltunut varsin hyvin muidenkin kirjoittajien aineistoksi.

Voi olla, että tähänastista, jopa yllättävän niukkaa Päätalo-aiheista tutkimusta[1] (ks. myös Saviniemi & Mantila tässä teoksessa) selittää juuri tutkittavan aineiston massiivisuus. Tutkijalle voi syntyä kokemus siitä, että luotettavan analyysin tekeminen aineistosta edellyttäisi sellaista Päätalon koko tuotannon käsittävää, syvällistä asiantuntemusta ja mittavaa perehtymistä koko *Iijoki*-sarjaan, jota hän ei voi rehellisyyden nimissä väittää edustavansa. Korpus mahdollistaa nyt sen, ettei koko sarjaa tarvitse ottaa lähilukuun, vaan on mahdollista syventyä tarkemmin tutkimusaiheen kannalta keskeiseen teokseen ja hakea muista sarjan teoksista olennaista tietoa erillisten hakulausekkeiden avulla. Ilman korpusta olisi lähes mahdotonta tehdä pikkutarkkaa tutkimustyötä, jossa tutkijan täytyisi hakea analyysikohteet muistinsa varassa.

Ryhdymme nyt – Päätaloon liitettyihin rehellisyyden tavoitteluun ja pikkutarkkaan muistiin kiinnittyen – tarkastelemaan *Iijoki*-korpuksen tutkimusmahdollisuuksia. Pohdimme *Iijoki*-sarjan omaelämäkerrallista luonnetta ja käsittelemme samalla sitä, miten digitaali-

https://doi.org/10.21435/tl.271

sen tekstikorpuksen avulla voitaisiin tutkia tarkemmin *Iijoki*-sarjaan, teosten tekijyyteen sekä faktan ja fiktion välimaastoon asettumiseen kytkeytyviä kirjallisuudentutkimuksellisia peruskysymyksiä.

Kirjallisuudentutkimuksen perusteiden äärellä: *Iijoki*-sarjan laji faktan ja fiktion välissä

Kuten edellä toteamme, Päätalosta puhuttaessa hänen pikkutarkka, ilmiömäiseksi kuvattu muistinsa nousee usein esiin. Todennäköisesti lähes yhtä monesti päädytään kysymykseen Päätalon teosten lajista eli siitä, onko hänen tuotantonsa faktaa vai fiktiota – onko Päätalon tuotannossa kyse laajennetusta omaelämäkerrasta vai laajamittaisesta sarjasta kaunokirjallista proosaa? Päätalon teosten vastaanotosta väitöskirjansa laatinut ja Kalle Päätalon elämäkerran kirjoittanut Ritva Ylönen (2017: 12) linjaa, että suurin osa Päätalon tuotannosta on omaelämäkerrallista. Tämä muotoilu, tekstin määrittely *omaelämäkerralliseksi*, korostaa teosten olevan Päätalon omaa elämää kertaavia ja siitä kertovia, siis *kerronnallisia* teoskokonaisuuksia, mikä tuo ne väistämättä kertomusten lainalaisuuksien piiriin. Eletty elämä ei sinänsä ole kertomus eikä varsinkaan fiktiota, mutta kun elämästä aletaan kertoa, se *kerronnallistuu* eli muuttuu kertomukseksi, jota kerrotaan kertojansa muistin ja tarinallisuuden tajun varassa.

Tästä huolimatta lukijoiden ja joidenkin elämäkertureiden on paikoin ollut vaikeaa tehdä eroa, onko *Iijoki*-sarjan päähenkilö fiktiivinen Kalle Päätalo vai kirjailija itse. Karonen (1999: 201) kärjistää, kuinka jopa itsestään selvänä pidetään sitä, että *Iijoki*-sarjan minäkertoja ja päähenkilö Kalle Päätalo on sen kirjoittaja, kirjailija Kalle Päätalo. Näin toki onkin, mutta siitä huolimatta meillä ei ole Päätalon kaunokirjallisen tuotannon kautta pääsyä todellisten henkilöiden elämään vaan ainoastaan muistin varassa kerronnallistettujen ja fiktiivisiksi muodostuneiden hahmojen kirjoissa kuvattuun todellisuuteen. Edes pikkutarkaksi, ilmiömäiseksi kuvattuun muistiinsa nojaava Päätalo ei olisi voinut tuottaa ulkomuistista esimerkiksi dialogia kymmenien vuosien jälkeen (ks. esim. Rajala 1999: 176). Tämän Päätalo on itsekin todennut: kaikkien repliikkien muistaminen olisi hänenkin mielestään suorastaan luonnotonta (Vehviläinen 1978: 166). Roos

https://doi.org/10.21435/tl.271

huomauttaa, että koska dialogi on tässä mielessä keksittyä, siis sepitettä ja näin aina lähempänä fiktiota kuin faktaa, merkittävä osa – Päätalon itsensä arvioimana noin 95 prosenttia – omaelämäkerrallisesta tekstistä on luotu kirjailijan vapaudella. (Roos 1999: 27–28.)

On todennettu, että kirjailija Päätalon ja fiktiivisen Kalle Päätalon elämät eivät kaikin osin ole identtiset, mikä sekin osoittaa, että kaunokirjallisuus aina käyttää niin kerronnallisia kuin muitakin kaunokirjallisia keinojaan tarinoiden elävöittämiseen ja fiktiivisen maailman luomiseen. Vuonna 2017 ilmestynyttä elämäkertaa *Kalle Päätalo. Kirjailijan elämä* kirjoittaessaan Ylönen on pitänyt lähtökohtanaan sitä, että *Iijoki*-sarja kuvaa todellisia tapahtumia. Hän on kuitenkin korjannut *Iijoki*-sarjan luomaa kuvaa niiltä osin kuin romaanissa kuvattu ei ole yhdenpitävä muiden tietolähteiden, kuten arkistoaineistojen, kanssa. (Ylönen 2017: 13.) Meillä on nyt siis todellisen henkilön elämänvaiheista entistä enemmän tietoa, joka todentaa, ettei Kalle Päätalon elämässä kaikki mennyt niin kuin kirjoissa.

Paikoin liiankin kärkevään ja mustavalkoiseen kahtiajakoon perustuva spekulaatio faktan ja fiktion tai toden ja valheen välillä johtunee osaltaan Päätaloon liitetystä rehellisyyden vaatimuksesta. Päätalo kirjasi itselleen muistiin johtoajatuksensa: ”Ja kaikki REHELLISESTI. En saa sääliä enkä lelliä myöskään itseäni” (Heikkinen 2019: 132; Päätalon tuotannon rehellisyydestä ks. esim. Vehviläinen 1978: 99; Pietilä 1980: 11; Marttinen 1996: 147; Vahtola 1999: 151, 2001: 30). Päätalo on tuonut rehellisyyden ajatusta vahvasti esiin itse haastatteluissa, joissa hän on toistanut elämänkulustaan samoja asioita, joita kirjojen Kalle kertoo. Ei siis ole mikään ihme, että kirjailijan elämä sekä lukijoiden muista lähteistä päivittämä tieto sekoittuvat keskenään. (Ylönen 2013: 253.) Tästäkin huolimatta Päätalo on kertonut värittäneensä ja liioitelleensa asioita, jos draaman kaari on sitä vaatinut (Janatuinen 2005: 19; ks. myös Roos 1999: 29), ja vaikka hänen intentiotaan pyrkiä totuudellisuuteen ja rehellisyyteen ei sinänsä ole syytä epäillä, hänen ymmärrystään omasta toiminnastaan nimenomaan kaunokirjallisena kertojana ei sitäkään pidä vähätellä. Kuten Ylitalo (2010) on osoittanut, Päätalon minäkertoja ottaa toisinaan myös kerronnallista etäisyyttä itseensä. Näin tapahtuu esimerkiksi teoksessa *Loimujen aikaan* (1976, sit. Ylitalo 2010: 121):

https://doi.org/10.21435/tl.271

> Monissa erilaisissa tunnelmissa siinä itsensä näköinen päähenkilö oli nousemassa postiautoon ja lähtemässä sotaväkireissulle. Nyt on poika taas lähdössä, ja tämä lähtö on totta. Pojan perhe on hajonnut ja kotiselkonen on joutunut sodan jalkoihin.
>
> Tätä lähtöä en koskaan pystynyt kuvittelemaan.

Vaikka Päätalo on ollut tietoinen totuudellisuuden tavoittelun sekä kuvitteellisen ja sepitteen välisistä ristiriidoista omassa tuotannossaan, hänen teostensa varsinaisen toiminnan kuvauksen aitoutta ei Räisäsen (1981: 6) mukaan ole silti syytä epäillä: Räisänen pitää sitä niin ”oikeana” kuin mihin hyvämuistinen kirjailija on pystynyt. Päätalo piti faktan ja fiktion välisen rajan vetämistä vaikeana, mutta linjasi, että *Iijoki*-sarjassa kaikki vähänkin merkittävät tapahtumat ovat totta (Vehviläinen 1978: 166; ks. myös Ylönen 2013: 286, 2017: 12) – tai ainakin mahdollisimman totta.[2] Voitaneen ajatella, että *Iijoki*-sarjan henkilöhahmoihin on saatu ainekset kahdesta hyvin erilaisesta näkökulmasta. Lapsuuden tuntemukset, muistelukset, väärin muistamiset ja määrittelemättömät tunnelmat on suhteutettu aikuisen ja tietoisen kirjailijan tietoon ja ennen muuta siihen, millaisia hahmoja hän on elämänsä henkilöistä teoksissaan rakentanut. Kaunokirjallisuudessa on kyse korkeintaan *totuudellisuudesta* ja sen tuottamisesta. Kuitenkin Päätalon teosten käyttö menneisyyden tutkimuksen ja tulkinnan, esimerkiksi viime aikoina mielenkiintoisia tutkimusaiheita näkyväksi nostaneen muistitietotutkimuksen ja muistitietohistorian (ks. Fingerroos & Haanpää 2006; Leiviskä tässä teoksessa), aineistona on vähintäänkin kiinnostava mahdollisuus. Kaunokirjallisena, faktan ja fiktion väliin asettuvana aineistona se saattaisi avata uudenlaisia teoreettis-metodologisia näkökulmia siihen eri tieteiden väliseen kriittiseen keskusteluun ja osittaiseen ohipuhuntaan, jota on käyty muistitiedon käytön luotettavuudesta historiantutkimuksessa (ks. esim. Teräs 2018).

Toisaalta digitaalinen *Iijoki*-korpus tarjoaa uusia mahdollisuuksia tutkia aina läsnä ollutta mutta tähän mennessä analysoimatta jäänyttä peruskysymystä sekä Päätalon tuotannon lajista faktan ja fiktion välissä että siihen kiinnittyvästä tekijyyden suhteesta teosten minäkertojaan. Kaunokirjallisuuden digitaalisista tekstikorpuksista tähän mennessä tehty tutkimus ja digitaalinen tekstilouhinta on ni-

https://doi.org/10.21435/tl.271

menomaan aloitettu lajia, tekijyyttä ja henkilöhahmoja koskevilla analyyseilla. On tutkittu, pystyvätkö algoritmit ”tunnistamaan” kirjallisuuden lajeja, miten erottaa kirjailijan tyyli lajityypille tyypillisistä piirteistä ja miten lukijat hahmottavat fiktiivisiä henkilöhahmoja (Allison ym. 2011; Mahlberg ym. 2016; Schöch 2017; Santos 2018: 91–93).

Kirjallisuudentutkimuksen uusia mahdollisuuksia: *Iijoki*-sarja digitaalisen tekstilouhinnan esimerkkiaineistona

Päätalon tuotannon mittavuuden kannalta korpustutkimus tuntuu luontevalta tavalta syventyä hänen teoksiinsa. Pelkästään *Iijoki*-sarjan vuonna 2019 digitoiduissa 26 osassa on peräti 17 014 sivua (Roos 1999: 25): lähes puoli miljoonaa virkettä ja liki 5,3 miljoonaa sanetta. *Iijoki*-korpus rinnastuu laajuudessaan muun muassa Nottinghamin ja Birminghamin yliopistojen CLiC-yhteisprojektin digitaalisiin Dickens-, ChiLit- ja ArTs-teoskorpuksiin, jotka kaikki jäävät sanamääriltään kuitenkin sitä pienemmiksi.[3] Kuten Saviniemi ja Mantila toteavat tässä teoksessa, kyseessä on ensimmäinen näin laaja kokoelma yhden kirjailijan tuotantoa Kielipankissa. Esimerkiksi Suomalaisen Kirjallisuuden Seuran Aleksis Kivi -korpuksessa on tämän teoksen ilmestyessä yhteensä 413 735 sanetta.[4] Päätalon tuotantoa on verrattu kilvan Marcel Proustin *Kadonnutta aikaa etsimässä* -kirjasarjaan (*À la recherche du temps perdu* 1913–1927), jossa on noin kolmetuhatta sivua (ks. esim. Janatuinen 2005: 329–330, 341; Kirstinä 2006; Kirves 2005; Niemi 1999: 36; Roos 1999: 25; Rajala 1999: 179; Sinnemäki 1998; Ylönen 2013: 13–14).[5] Noin kolmetuhatsivuinen on myös kuusiosainen omaelämäkerrallinen romaanisarja *Taisteluni* (*Min kamp*), jonka on kirjoittanut toisinaan leikkisästi Norjan Päätaloksikin kutsuttu Karl Ove Knausgård vuosina 2009–2011 (Knausgårdin ja Päätalon eli ”selkosten Proustin” vertaamisesta ks. esim. Koivisto 2017).

Päätalon tuotannosta ei ole juuri tehty kokonaista, kokoavaa tutkimusta. Kustannustoimittaja Jaakko Syrjä arvioi jo 1970-luvulla sen johtuvan aineiston suuresta koosta: sen analysoimiseen ei ole ollut

https://doi.org/10.21435/tl.271

soveltuvia välineitä. Toisaalta kirjallisuudentutkimuksen ja kritiikin keinot eivät Syrjän mukaan ole tuntuneet toimivan Päätalon teosten kohdalla muutenkaan. (Vehviläinen 1978: 199.) Ylösen (2013: 331) mukaan argumentointia sarjan hyvyydestä tai huonoudesta hankaloitti se, ettei lajissa ollut vertailukohtaa. Rajala (1999: 184) taas havaitsi uuden kerronnan genren alkavan erottautua, ja sen analysoimiseksi olisi hänen mukaansa kehitettävä tieteidenvälisiä menetelmiä. *Iijoki*-korpus voisi tarjota siihen yhden tutkimusmahdollisuuden.

Päätalo dokumentoi ja arkistoi yksityiskohtia hyvin huolellisesti ja millintarkasti. Nyt vireillä oleva korpustutkimus ja digitaalinen tekstilouhinta avaavat uudenlaisia mahdollisuuksia tehdä yksityiskohtaistakin tutkimusta sekä Päätalon teksteistä että niiden kulttuurisista ja historiallisista konteksteista, mikseipä psykohistoriallisestakin näkökulmasta. Nyt se pikkutarkka yksityiskohtaisuus, josta Päätalo sai kritiikkiä aikalaisiltaan, nousee uudenlaiseksi tutkimusmahdollisuudeksi. Korpuksen digitaalinen louhinta mahdollistaa määrällisen tutkimuksen, jonka avulla yksityiskohtien merkitys voidaan todentaa ja osoittaa validiksi kvantitatiivisen datan sekä sen laadullisen analyysin avulla. Tämä taas tarjoaa aivan uusia määrällisen ja laadullisen tutkimuksen yhdistämisen mahdollisuuksia.

Datarikasteisessa kirjallisuudentutkimuksessa pitäisikin pyrkiä määrällisen ja laadullisen analyysin yhdistämiseen: kaukolukeminen, laajojen kokonaisuuksien tarkastelu ja lukeminen etäältä kirjallisuudentutkimukselle tyypillisen lähilukemisen sijaan (ks. Moretti 2013 [2000]), kaipaa tulkinnan tueksi syvällisen kirjallisuuden ja kirjallisuushistorian sekä kirjallisuudentutkimuksen menetelmällisen tuntemuksen (Parente-Čapková 2017: 40; Navarro-Colorado 2018; Santos 2018). *Iijoki*-korpuksen avulla Päätalon teosten kokonaisrakenteesta pystyttäneen tulevaisuudessa tuomaan esiin sellaisia seikkoja, joita kirjailija ei ehkä itsekään hahmottanut valtavan tekstimassansa äärellä. Nyt nimenomaan sarjan kokonaisrakenteen tutkiminen alkaa olla tähänastista helpompaa: käsityönä 26 romaanin analysoiminen olisi huomattavasti työläämpää ja todennäköisesti rajoittuneempaa. Korpustyökalujen avulla pystytään tarkastelemaan tietyille kirjoille ominaisia tendenssejä, joita ei käsin tekstejä käsiteltäessä välttämättä havaita. Toisaalta näitä tendenssejä voi lähestyä erottavien tai tyypillisten vertailutekstien kautta, jolloin voidaan nostaa esiin kysymyksiä

https://doi.org/10.21435/tl.271

siitä, mikä on tyypillistä juuri *Iijoki*-sarjan teoksille, mikä erottaa niitä muista tai mikä vaihtoehtoisesti yhdistää niitä muihin teksteihin tai teoksiin. Tämänkaltaiset ominaisuudet eivät välttämättä hahmotu pelkällä lähiluvulla, eikä sellaisten tarkastelu isosta aineistosta olisi aina mahdollistakaan. Yhden sanan käytön tarkastelu koko *Iijoki*-sarjasta olisi valtava työ ilman korpustyökalujen apua. Korpuksen käyttö vähentää myös inhimillisen virheen mahdollisuuksia.

Suomessa kirjallisuustieteellistä korpustutkimusta ei ole tiettävästi tehty vielä kovin laajasti. Digitaalisesta kirjallisuudentutkimuksesta kiinnostuneet tutkijat ovat Suomessa keskittyneet esimerkiksi suomalaisten lukutottumusten ja kirjallisen maun tutkimukseen laajan kirjaston lainausdatan pohjalta. Toiseksi on tutkittu kirjallisuushistoriaa ja vastaanottoa Kansalliskirjaston digitaalisten lehtikokoelmien pohjalta. Kolmanneksi on tarkasteltu digitaalisissa ympäristöissä tapahtuvaa kirjoittamista ja lukemista.[6] Korpustutkimuksen alueella kansainvälisesti tunnettuja ovat jo mainitut CLiC-yhteistyöprojekti sekä The Stanford Literary Lab. Esimerkiksi Diane Santos (2018) tutkii portugalilaista kirjallisuutta elektronisessa muodossa, ja Anna Čermáková (ks. esim. Čermáková & Chlumská 2017; Čermáková & Mahlberg 2018) hyödyntää lastenkirjallisuuden digitaalisia korpuksia. Vuonna 2018 julkaistiin Frontiers-alustalla ”Computational Linguistics and Literature” -teemanumerossa Anna Kazantsevan toimittamana yhteensä viisi kirjallisuudentutkimuksellista artikkelia.[7] Charles Dickensin tuotannosta tehtyä korpusta tutkitaan tyylin näkökulmasta CLiC-Dickens-projektissa, jonka tavoitteena on osoittaa tietokoneavusteisten menetelmien käytettävyys ja hyöty kirjallisuuden ja esimerkiksi teosten henkilöhahmojen tutkimuksessa.

Korpuslingvistiikka tarkastelee usein suurikokoisia digitaalisia aineistoja laskennallisten menetelmien avulla. Erityisesti kirjallisuudentutkimuksessa puhutaan *korpusstilistiikka*-nimisestä (*corpus stylistics*) tutkimussuuntauksesta, jossa kaunokirjallisten teosten tyyliä tutkitaan korpuslingvistiikan menetelmillä (esim. Mahlberg ym. 2016; Stockwell 2017; McIntyre & Walker 2019). Franco Morettin (2013 [2000]) lanseeraama määrällinen kaukoluvun (*distant reading*) metodi sopii luonnehtimaan tämänkaltaista tutkimusotetta, jossa kaunokirjallisia (tai muita, ks. esim. Kannisto 2016) aineistoja ei lähilueta vaan lähestytään tietokoneavusteisesti laajana kokonaisuutena,

https://doi.org/10.21435/tl.271

jolloin niistä voidaan eritellä esimerkiksi romaanilajin kehitystä, henkilöhahmojen verkostoja, teemoja ja tyylipiirteitä. Katherine Bode (2018) on kirjoittanut Australian kansalliskirjaston Trove-tietokannan pohjalta datarikasteista, kontekstualisoitua ja ylirajaista kirjallisuuden historiaa 1800-luvun Australiasta. Hän hyödyntää bibliometriikan, verkostoanalyysin ja topiikkimallinnuksen kaltaisia digitaalisia metodeja.

Korpuslingvistiikan laskennallisten menetelmien avulla voidaan tutkia tiettyjen sanojen tai sanaryhmien käyttöä ja käytön vaihtelua esimerkiksi romaanikokoelman eri osissa tai yhden romaanin lukujen välillä. *Iijoki*-korpuksen kohdalla tämä voisi tarkoittaa romaaneissa tiiviisti toistuvien ja helpohkosti esiin louhittavien voimasanojen käytön ja niiden vaihtelujen seuraamista. Korpuksen avulla on jo saatu selville (ks. Saviniemi tässä teoksessa), että Päätalon kielessä *hitot* ja *saamarit* esiintyvät melko harvoin, mutta niiden sijaan etenkin nykylukijalle vieraammat voimasanat, kuten *hiikkari*, *horpotti* ja *lempo*, tulevat esiin. Voimasanojen käytössä on eroja myös miesten ja naisten sekä uskovaisten ja ei-uskovaisten henkilöhahmojen välillä. Edelleen voi kysyä, käyttelikö *Iijoki*-sarjan aloittanut 52-vuotias Päätalo – joka sarjan alussa kuvasi itseään lapsena ja nuorena – rehvakkaammin voimasanoja kuin sarjan loppupuolen iäkkäämpi, muuttuneessa elinympäristössä aikuista itseään kuvaava kirjailija-kertoja. Jo nämä esimerkinomaiset tarkastelut osoittavat, että pelkästään yksittäisten sanojen, esimerkiksi voimasanojen, digitaalisen tekstilouhinnan avulla voidaan tutkia konteksteja, joihin teokset sijoittuvat, kuten niiden sosiokulttuurista ympäristöä. Lisäksi voidaan tarkastella tekstin tekijyyden, henkilöhahmojen, kerronnallisten keinojen ja tyylin kysymyksiä.

Iijoki-korpuksesta voisi tarkastella muutaman tyypillisen sanan toistumista louhimalla tekstimassasta esiin yksityiskohtia. Näin tekstistä olisi mahdollista tunnistaa luotettavammin teoksen tyylin kannalta merkittävää toistoa, ja tätä tietoa voitaisiin soveltaa paitsi kirjailijan tyylin analysoinnissa myös tarkoituksenmukaisten käännösratkaisujen tekemisessä (Čermáková 2015: 375). Kerronnan kannalta kiinnostavan dialogin käytön yleisyyttä voi tarkastella etsimällä sitä merkitsevät repliikkiviivat. Sanaston tasolla tietyt murresanat on mahdollista seuloa esiin: ”Mää räyryn rasvaksi, tiämmä!” ärähtelee

https://doi.org/10.21435/tl.271

Hyvästi, Iijoki -romaanin minäkertojan kipeä, äreä ja ”emäkrapulaa” sairastava kirvesmieskaveri junassa matkalla Tampereelta Kyröskoskelle (Päätalo 1995: 5).

Edellä mainittua Päätalon romaania lukiessa huomio kiinnittyy toiseenkin tekstin erityispiirteeseen: Päätalo käyttää runsaasti paikannimiä ja erisnimiä. Erilaiset erisnimet, kuten paikkakunnat, on mahdollista tunnistaa automaattisesti nimientunnistusjärjestelmällä (Luoma ym. 2020). Näin voidaan tarkastella erilaisten paikkojen määrää ja merkitystä teksteissä, kuten sitä, onko niitä runsaasti kautta koko sarjan vai vaihteleeko nimien määrä. Pelkästään edellä siteeratun *Hyvästi, Iijoki* -romaanin yhdellä sivulla (Päätalo 1995: 5) nimiä on runsain mitoin: *Tampere*, *Kyröskoski*, *Ikaalinen*, *Juhtimäen Jauli*, *Sasinmäki*. Nimien runsas käyttö sitoo romaanit vahvasti paikkaan, mikä taas on omiaan lisäämään Päätalon tekstin dokumentoivana pidettyä kronikoinnin (ks. Saviniemi & Mantila tässä teoksessa) tuntua, tekstin realistisuutta ja mimeettisyyttä: tämä on totta ja tapahtunut, oikeissa paikoissa oikeille ihmisille.

Päätalon suosima realistinen, todentuntua ja arkipäiväisyyttä korostava tyyli on ollut suomalaisen lukijakunnan suosiossa pitkään, viimeistään 1970-luvulta lähtien (Eskola 1979, 1990; Eskola & Linko 1986; Jokinen 1997). Nämä piirteet viehättävät edelleen lukijoita, mistä kertoo esimerkiksi Enni Mustosen (oik. Kirsti Manninen) *Syrjästäkatsojan tarinoita* -sarjan suuri menestys niin kirjastoissa kuin kirjakaupoissakin (Launis ym. 2018). Nykylukijoiden kirjallista makua luonnehtivat sarjamaisuus ja kotimaisuus (ks. Launis ym. 2018) eli piirteet, jotka kuvaavat Päätalon tuotantoa. Lisäksi työn kuvaus, jolla Päätalon ja Väinö Linnan lukijasuosiota on selitetty (Eskola 1990: 96–102), on yksi suomalaisen romaanin pisimmistä jänteistä. Tämä jänne jatkuu edelleen nykykirjallisuudessa (Ojajärvi 2013: 154–156; Launis ym. 2018: 15). Realistinen tyyli siis liittynee Päätalon ilmiömäiseen lukijasuosioon (ks. Ylönen 2013; Kuivasmäki 1974; Janatuinen 2005), ja juuri tähän tekstin tyylin tutkiminen toisi lisäarvoa. Korpuksen avulla päästään siis käsiksi siihen, mistä päätalomainen, omaelämäkerrallinen todentuntuisuus tyylin, ilmaisujen, kerronnan ja sanaston tasolla yksityiskohtaisemmin rakentuu.

Edellä esitetyn yksittäisten sanaryhmien tarkastelun lisäksi korpusta voi käyttää romaanien ominaispiirteiden analysointiin koko

https://doi.org/10.21435/tl.271

tekstin tai ainakin kokonaisten lukujen tasolla. Näin on mahdollista tutkia romaaneista nousevia teemoja ja teemojen vaihtelua. Kuvassa 9 esitetyn, romaanisarjan yleisimpiin sanoihin perustuvan analyysin lisäksi tähän sopii topiikkimallinnus (ks. esim. Kannisto 2016; Bode 2018; Navarro-Colorado 2018; Santos 2018; Kyröläinen & Laippala 2020), jota käytetään laajalti digitaalisessa kirjallisuudentutkimuksessa, kun taas verkostoanalyysi (*network analysis*) on esimerkiksi kirjahistorioitsijoiden suosiossa. Topiikkimallinnus on ohjaamattomaan koneoppimiseen perustuva menetelmä, jonka avulla pystytään aineistolähtöisesti, ilman etukäteen esitettyjä hypoteeseja tarkastelemaan aineistossa toistuvia aihepiirejä. Käytännössä aihepiirit määräytyvät tekstissä yhdessä esiintyvien sanojen kautta, ja analyysin pohjan luovat jokaista topiikkia kuvaavat avainsanat, jotka menetelmä määrittää. Näin pystytään tarkastelemaan suuresta aineistosta nousevia topiikkeja ilman lähilukua. Eri topiikkien vaihtelua voi vertailla korpuksen eri osien välillä.

Kuva 9. Koko *Iijoki*-sarjan yleisimmät saneet eli sanaesiintymät. Mitä yleisempi sana on, sitä suuremmalla pistekoolla se on kuvassa.

https://doi.org/10.21435/tl.271

Kuvan 9 sanapilvi on laadittu Pythonin Wordcloud-paketilla normalisoimalla isot kirjaimet pieniksi, poistamalla välimerkit ja filtteröimällä pois yleisimmät konjunktiot, apuverbit ja osa pronomineista. Tyypillisesti tekstiaineistojen yleisistä sanoista voi päätellä aineiston aihepiirejä sekä tyylillisiä ominaisuuksia. Sanapilven motivoimassa *Iijoki*-sarjan laadullisessa analyysissä voisi keskittyä tarkastelemaan muun muassa sarjan aihepiirejä ja henkilöitä: mitä *Iijoki*-sarjan aihepiireistä tai henkilöistä voidaan analysoida ja tulkita sen perusteella, että sen yleisimpiin sanoihin kuuluvat esimerkiksi maskuliiniset *kalle*, *veikko*, *janne*, *ville*, *isä*, *mies* ja *poika* mutta myös sanat *äiti* ja *vaimo* nousevat silti yleisimpien sanojen joukkoon? Kertovatko *minä*-pronominin monet muodot siitä, että *Iijoki*-sarja asemoituu omaelämäkerrallisiksi ”minän kertomuksiksi” (ks. Rojola 2002: 70), ja mistä kielivät raportoivat verbit *sanoi* ja *jatkoi*? Entä mitä kertovatkaan *Iijoki*-sarjasta sanan *ei* korostunut yleisyys sekä se, että sen yleisempiin sanoihin lukeutuu myös *perkele*?

Päätaloa on pidetty maskuliinisena, miesten maailmaa kuvaavana ja mieslukijoiden suosiossa olevana kirjailijana (Janatuinen 2005). Päätalon teoksissa esiintyvä minäkertoja on tiedetysti hyvin tunteellinen mies. Teoksissa kuvaillaan monenlaisia tunteita ja affekteja, joita koskevia sanoja ja ilmauksia korpuksesta on mahdollista louhia esiin. *Iijoki*-korpuksesta voitaisiin tarkastella pojasta mieheksi kasvun kuvaamisen tunne- ja affektikirjoja: Kuvastavatko erilaiset tunnekirjot eri ikävaiheita? Vaihtuvatko tunnerekisterit aikakausien, iän tai paikkojen mukaan? Ovatko jotkin tunteet ja affektit, kuten Päätalon tuotantoon tiiviisti liitetty häpeä, läsnä aina ja niin mittavasti kuin on oletettu? Tuloksia olisi mahdollista vertailla joihinkin muihin isoihin aineistoihin, esimerkiksi mediateksteihin, mikä tarjoaisi mahdollisuuksia moniaineistoisen suomalaisten miesten tunteita eri vuosikymmeninä tarkastelevaan tutkimukseen. Muutenkin olisi kiinnostavaa miettiä korpuksen tarjoamia tutkimusmahdollisuuksia yhdistettynä joihinkin muihin laajoihin, louhittavissa oleviin aineistoihin. Yksi mielenkiintoinen aihe voisi olla moniaineistoinen massadatatyyppinen suomalaisten metsä- tai luontosuhteita tarkasteleva tutkimus.

https://doi.org/10.21435/tl.271

Ylipäänsä Päätalolle ominainen toisto on erityisen mielenkiintoinen näkökulma korpustutkimukseen. Liisi Huhtala (1972: 93) näki Päätalon käyttämässä murresanastossa mekaanista toistuvuutta. Toisto onkin olennainen osa sittemmin julkaistua *Iijoki*-sarjaa. Ylönen (2017: 65, 294) pitää toistamista ensinnäkin Päätalon tapana korostaa tapahtumien merkitystä, minkä lisäksi se vahvistaa lukukokemuksen terapeuttisuutta. Toki toistolla on merkitystä myös niiden lukijoiden kannalta, jotka eivät ole seuranneet sarjaa alusta saakka (Ylönen 2013: 276). Lisäksi Päätalolla on selvästi ollut pyrkimyksenä murresanojen opettaminen lukijoille, ja toisto liittynee siihen tavoitteeseen. Runsas murresanaston käyttö onkin aineiston vahvuus: Räisänen (1999: 96) arvioi kansankielen sanojen nousevan Päätalon teoksissa määrältään tuhansiin.

Korpuksen tekstejä voisi tulevaisuudessa verrata geneettisen kritiikin näkökulmasta (ks. Mantila, Keskimaa & Saviniemi tässä teoksessa) Taivalkosken Päätalo-instituutissa säilytteillä oleviin käsikirjoitusversioihin, etenkin jos viimeksi mainitut digitoitaisiin sähköisesti muokattavaan muotoon. Päätalon kirjojen toimitusprosessille oli leimallista käsikirjoituksen merkittävä tiivistäminen ja lyhentäminen. ”Joka kerran olen joutunut lyhentämään käsikirjoitusta, en koskaan lisäämään”, kirjailija on todennut (Sainio 1973: 15). Koska Päätalon käsikirjoituksista pisimmät olivat peräti yli tuhatsivuisia, *Iijoki*-sarjan teosten tekstiä jouduttiin yleensä lyhentämään enimmillään 20–25 prosenttia. Kustantajan edustajat ja kirjailija itse lyhensivät käsikirjoituksia yhteistyönä, ja tiivistämistä jouduttiin tekemään käsikirjoituksissa kauttaaltaan. (Marttinen 1995: 176, 1999: 165, tiivistämisestä ks. myös Lappalainen 1999: 173.) Käsikirjoitusten rakenteen vuoksi tekstiä ei voinut lyhentää pelkästään alusta, lopusta tai keskeltä (Marttinen 1999: 165). Kustannustoimittaja Juhani Syrjä piti Päätaloa ”hyvin kaukaa virittävänä” kirjailijana, joka saattaa viitata tekstissään johonkin vasta 200 sivun päässä tulevaan asiaan. Toisaalta Syrjän kärjistyksen mukaan Päätalon teoksissa saatetaan kertoa sama asia ainakin kolmeen kertaan. (Marttinen 1999: 165.) Tällaisten rakenteellisten ratkaisujen vertaaminen painetussa kirjassa ja käsikirjoitusversioissa voisi sekin olla kiinnostavaa uutta tietoa tuottava tutkimuskohde. *Iijoki*-korpuksen tutkimukselliset käyttömahdollisuudet ovat todellakin lukuisat, ja tässä olemme todennäköisesti sivunneet niistä vain murto-osaa.

https://doi.org/10.21435/tl.271

Lopuksi: Kadonnutta aikaa tutkimassa

Palaamme näin päätäntölukumme lopuksi vielä sen alussa kuvailemaamme haastattelutilanteeseen ja anekdoottiin Päätalon erityislaatuisesta muistista sekä pyrkimyksestä rehellisyyteen: Päätalo olisi toki voinut jo kunnioitettavassa lähes 80 vuoden iässään ja kyseisessä haastattelutilanteessa vain vastata toimittajan kysymykseen, että tottahan talosta täydellinen tuli, kun hän sen itse omin käsin rakensi. Siihen ei umpirehellisestä Päätalosta anekdootin mukaan kuitenkaan ollut. Tarina jatkuu vielä siten, että vasta myöhemmin toimittaja-Majanderille valkeni kollegansa Malmbergin itsensä esittäneen samaisen kysymyksen Päätalolle jo aiemmin. Silloinkin Päätalo oli kuljettanut toimittajaa kellariinsa katsomaan ja todistamaan tuota samaa, kohtalokasta valuvikaa. Tapahtuman kerronnallinen toisto sinetöi Majanderin mukaan todisteen Päätalosta Suomen uskottavimpana kirjailijana. Jos rakennusmestari on niin rehellinen, että on valmis toistamiseen kertomaan ventovieraalle eteläsuomalaiselle sanomalehtitoimittajalle rakennustöissä tekemästään virheestä, hänen on oltava omaelämäkerrallisten teostensa tarinoiden minäkertojanakin umpirehellinen, tai ainakin niin umpirehellinen kuin mahdollista. Tarinan pitäisi siis näin vahvistaa uskoa, että Päätalo pyrki kykyjensä mukaan kertomaan teoksissaan kokemuksistaan totuuden ja vain totuuden, minkä lisäksi hänellä oli pettämättömän tarkka, jopa – tai ennen kaikkea – itselleen armoton muisti.

Sen toteamiseen, että Päätalo pyrki teostensa tekijänä mahdollisimman totuudenmukaiseen, vaikkakin toki aina subjektiiviseen kuvaukseen oman elämänsä kulusta, ei tietenkään tarvita helsinkiläisen sanomalehden miestoimittajia, vaikka anekdootti talonsa kellarin valuvikaa esittelevästä kirjailijasta kuriositeettina sinänsä sympaattinen onkin. Totuudellisuuden tavoittelu näyttäytyy paitsi Päätalon teosten omaelämäkerrallisen kerronnan eetoksena myös niille hyvin lajityypillisenä piirteenä, jota kirjallisuudentutkija Lea Rojola (2002) on kuvannut suoranaiseksi elämäkerrallisten teosten ja ”minän kertomusten” lajilaiksi. Tämä lajilaki tuottaa lukijoille lupauksen totuudesta ja kertojan vilpittömästä rehellisyydestä, ja sen noudattamista Päätalon tapauksessa tämän toistuvasti hyvin pikkutarkaksi kuvattu muisti vielä tukee. Omaelämäkerrallisten teosten lajilakia noudatellen Pää-

https://doi.org/10.21435/tl.271

talo lupaa ja vakuuttaa lukijoilleen kuin oikeuden edessä kunniansa ja omantuntonsa kautta, että todistaa ja kertoo elämästään kaiken totuuden mitään siitä salaamatta, mitään siihen lisäämättä tai mitään siitä muuttamatta.

Varmaan juuri nämä seikat todella ovat yksi keskeinen syy Päätalon suosioon: hänen teoksensa välittävät lukijoilleen tunnetta tekijän vilpittömästä rehellisyydestä ja tämän tekemän muistelutyön pedanttisesta huolellisuudesta ja tarkkuudesta – ja tekevät siis kirjailijasta uskottavan. Tämä uskottavuus on kannatellut jo Päätalon ikätoverien sukupolvesta seuraavaan ja sitä seuraavaan lukevaan sukupolveen. Päätaloa luetaan edelleen, kuten myös tutkitaan, ja tämä käsillä oleva teos osoittaa mitä erinomaisimmin, miten moneen Päätalon tuotanto taipuu ja miten monitahoisen kulttuurisen varannon hänen teoksensa juuri uskottavuudessaan tutkijoillekin tarjoavat. Päätalon teosten arvo ei olekaan siis pelkästään siinä, miten uskottavasti niissä kuvataan tekijän omaa elämää. Se on myös siinä, mitä ja miten niissä kerrotaan kaikesta muusta tekijän ympärillä näyttäytyneestä, jo menneestä elämästä ja ihmisistä sekä näitä ympäröineistä valta- ja alakulttuureista ja koko yhteiskunnasta.

Silti ei sovi unohtaa sitäkään, ettei Päätalon teosten lumo voi mitenkään selittyä pelkästään niiden uskottavuudella. Ovathan esimerkiksi lakitekstit, erilaisten tuotteiden ja kulutushyödykkeiden käyttöohjeet tai vaikkapa lääkeaineiden tuoteselosteetkin lajityypeilleen uskollisesti totuudenmukaisia ja uskottavia. Siitä huolimatta niitä ei lueta viihtymis-, eläytymis- tai varsinkaan lumoutumistarkoituksessa kuin korkeintaan niitä hyödyntävässä kokeellisessa kirjallisuudessa. Päätalon teoksia kuitenkin luetaan, ja niitä luetaan sukupolvesta toiseen. Päätalon tavassa kertoa tarinoita on oma veto- ja tenhovoimansa, eikä häntä ole perusteettomasti tituleerattu selkosten Proustiksi. Vaikka Päätalon tuotanto tarjoaa monitahoisen kulttuurisen varannon, jonka avulla voimme etsiä kadonnutta aikaa ja jota voidaan tutkia niin kielitieteen, murretutkimuksen, historian kuin elokuvatutkimuksenkin näkökulmista, kuten tässä teoksessa on tehty, myös teosten kaunokirjalliselle laadulle on annettava se painoarvo, joka sille kuuluu.

Etsikäämme siis selkosten Proustimme teoksissaan kuvaamaa kadonnutta aikaa, kuten tässäkin teoksessa on tehty, ja nauttikaamme samalla siitä, *miten* tuota kadonnutta aikaa erinäisin kaunokirjallisin

https://doi.org/10.21435/tl.271

ja kerronnallisin keinoin kuvataan. Näiden tavoitteiden saavuttamisessa voimme uutena kvantitatiivisena, laajojen tutkimusaineistojen hankinnan ja kokoamisen menetelmänä hyödyntää digitaalista tekstilouhintaa sen eri muodoissa. Tekstilouhinta voi johtaa uusin silmin tehtyjen havaintojen myötä kokonaisiin uusiin tutkimuskysymyksiin ja -näkökulmiin. Näiden aineistomassojen analyysiin ja tulkintaan ja varsinaisten johtopäätösten tekoon tarvitsemme silti edelleen – ja todennäköisesti vielä entistä kriittisemmin – laadullista tutkimusotetta.

Viitteet

1 Päätalon tuotannosta on tehty jonkin verran opinnäytteitä, mutta esimerkiksi väitöstasoista tutkimusta siitä on toistaiseksi tehty kirjallisuudentutkimuksen alalla hyvin vähän. Kuitenkin Päätalo on tuottanut teoksissaan poikkeuksellisen runsaasti aineistoa, jossa riittää tutkittavaa yli tieteenalarajojen. Onkohan niin, että Päätaloa lukevat rakastavat ja kunnioittavat kohdettaan niin paljon, etteivät tohdi lähteä tutkimaan aihetta? Ja lieneekö toisaalta niin, että Päätaloa lukemattomat eivät rakasta kohdetta tarpeeksi lähteäkseen sitä tutkimaan, vaan tuotannon laajuus pikemminkin kavahduttaa heitä kauemmaksi? Vaikka Päätalon tuotantoon kohdentunut vähäinen tutkimuksellinen kiinnostus ja sen kaunokirjallista laatua koskeva vähättely (ks. Ylönen 2013: 323–326) lienevät vähentyneet, voi kuitenkin olla, että etenkin tuotannon laajuus on estänyt tutkijoita tarttumasta Päätalon teoksiin.

2 *Iijoki*-sarjan genreä on pohdittu enemmänkin. Karonen (1999: 203–204, 207) näkee siinä monta todellisuutta ja päätyy pitämään sitä luotettavana muistelmadokumenttina, jossa omaelämäkerta ja romaani lomittuvat. Vahtola taas luonnehtii sarjaa omaelämäkerralliseksi muistelmateokseksi, joka on toteutettu kaunokirjallisin menetelmin ja joka perustuu muun muassa omiin kokemuksiin, haastatteluihin ja kirjallisiin lähteisiin (Vahtola 1999: 151, ks. myös Ylönen 2013: 284; Keskimaa 2018: 24). *Iijoki*-sarjan teoksia on määritelty omaelämäkerrallisiksi romaaneiksi sarjan ilmestymisvuosikymmenten ajan, joskin kronikkaromaaneista ja muistelmista puhuminen alkoi yleistyä 1990-luvulla (Ylönen 2013: 294). Ylönen (2013: 307 ja siinä mainitut lähteet) on havainnut, että sarjan tultua valmiiksi sen luokittelu omaelämäkerraksi tai autobiografiaksi vaikuttaisi lisääntyneen. Roos (1999: 30) pohtii, että sarja on omaelämäkerta ja kaunokirjallisuutta eikä siis ”täysiverinen” omaelämäkerta. Hän heittää ehdolle autofiktion (ranskaksi ja englanniksi *autofiction*) (vrt. Lauri & Ojajärvi 2017: 122) ja muistelmaproosan (ruotsiksi *memoarprosa*) käsitteet, kunnes päätyy *omaelämäkertaepiikkaan*. *Iijoki*-sarjaa voi luonnehtia hyvällä syyllä myös *tunnustuskirjallisuudeksi* (ks. esim. Karkulehto 2007: 143–144).

3 CLiC (bham.ac.uk).

https://doi.org/10.21435/tl.271

4 Ks. https://www.finlit.fi/fi/tutkimus/edith/aleksis-kivi-korpus-sks.
5 Ranskassa on käynnissä korpusprojekti Proustin kirjeistä, ks. https://cahier.hypotheses.org/corr-proust ja http://proust.elan-numerique.fr/.
6 Ks. esim. projektit "LibDat: Towards a More Advanced Loaning and Reading Culture. A Study, based on Digital Material, of Contemporary Finnish Loaning and Reading Habits" (Suomen Akatemia, http://virtual.vtt.fi/virtual/libdat/), COST-verkosto "Women Writers in History. Toward a New Understanding of European Literary Culture" (http://www.womenwriters.nl/index.php/COST_Action),"Tekstit liikkeessä: Naiskirjailijoiden tuotannon vastaanotto Suomessa ja Venäjällä 1840–2020" (Emil Aaltosen säätiö, https://tekstitliikkeessa.com/tekstit-liikkeessa/) sekä *Avaimen* (2/2021) teemanumero "Kirjallisuus digitaalisissa ympäristöissä", toim. Hanna-Riikka Roine ja Laura Piippo, https://journal.fi/avain/issue/view/7763.
7 Ks. https://www.birmingham.ac.uk/schools/edacs/departments/englishlanguage/research/projects/clic/index.aspx.

Lähteet

Allison, Sarah – Heuser, Ryan – Jockers, Matthew – Moretti, Franco – Witmore, Michael 2011: Quantitative formalism. An experiment. – *Pamphlets of the Stanford Literary Lab, Pamphlet 1*, January 15, 2011. https://litlab.stanford.edu/LiteraryLabPamphlet1.pdf (11.4.2021).

Avain. Kirjallisuudentutkimuksen aikakauslehti. Kirjallisuus digitaalisissa ympäristöissä. Vol 18 Nro 2 (2021) https://journal.fi/avain/issue/view/7763 (4.1.2022).

Bode, Katherine 2018: *A world of fiction. Digital collections and the future of literary history*. Ann Arbor: University of Michigan Press. https://doi.org/10.3998/mpub.8784777.

Čermáková, Anna 2015: Repetition in John Irving's novel A widow for one year. – *International Journal of Corpus Linguistics* 20(3) s. 355–377.

Čermáková, Anna – Chlumská, Lucie 2017: Expressing place in children's literature. – T. Egan & H. Dirdal (toim.), *Cross-linguistic correspondenses. From lexis to genre* s. 75–96. Amsterdam: John Benjamins Publishing Company.

Čermáková, Anna – Mahlberg, Michaela 2018: Translating fictional characters. *Alice* and *the Queen* from the Wonderland in English and Czech. – Anna Čermáková & Michaela Mahlberg (toim.), *The corpus linguistics discourse. In honour of Wolfgang Teubert* s. 223–253. Amsterdam: John Benjamins Publishing Company.

Eskola, Katarina 1979: *Suomalaiset kirjanlukijoina*. Helsinki: Tammi.

Eskola, Katarina 1990: *Lukijoiden kirjallisuus Sinuhesta Sonja O:hon*. Helsinki: Tammi.

Eskola, Katarina – Linko, Maaria 1986: *Lukijan onni. Poliitikkojen, kulttuurieliitin ja kirjastonkäyttäjien kirjallisista mieltymyksistä.* Helsinki: Tammi.

https://doi.org/10.21435/tl.271

Fingerroos, Outi – Haanpää, Riina 2006: Muistitietotutkimuksen ydinkysymyksiä. – Outi Fingerroos, Riina Haanpää, Anne Heimo & Ulla-Maija Peltonen (toim.), *Muistitietotutkimus. Metodologisia kysymyksiä*. Tietolipas 214. Helsinki: Suomalaisen Kirjallisuuden Seura.

Heikkinen, Antti 2019: *Kallio-poika*. *Peilikuvassa Kalle Päätalo*. Helsinki: Gummerus.

Huhtala, Liisi 1972: *Kalle Päätalon Koillismaa-sarjan tarkastelua*. Eripainossarja 6. Turku: Turun yliopiston kotimaisen kirjallisuuden laitos.

Janatuinen, Jenni 2005: *Miehenkuva*. Helsinki: Johnny Kniga.

Jokinen, Kimmo 1997: *Suomalaisen lukemisen maisemaihanteet.* Jyväskylä: Jyväskylän yliopisto.

Kannisto, Maiju 2016: Uusi, ehompi, paras? Digitaaliset tekstilouhinnan työkalut televisiotutkimuksessa. – *Lähikuva* 1/2016 s. 62–70.

Karkulehto, Sanna 2007: *Kaapista kaanoniin ja takaisin. Johanna Sinisalon, Pirkko Saision ja Helena Sinervon teosten queer-poliittisia luentoja*. Acta Universitatis Ouluensis, B, Humaniora 81. Oulu: Oulun yliopisto. http://urn.fi/urn:isbn:9789514286117 (13.4.2021).

Karonen, Vesa 1999: Juuret Iijoen törmässä. Missä latva? – Tero Liukkonen (toim.), *Päätalon juurella. Juhlakirja 11.11.1999 Kalle Päätalon täyttäessä 80 vuotta* s. 199–208. Helsinki: Gummerus.

Keskimaa, Sari 2018: *Iijoki-sarja kielielämäkertana*. Acta Universitatis Ouluensis, B, Humaniora 165. Oulu: Oulun yliopisto. http://urn.fi/urn:isbn:9789526220536.

Kirstinä, Leena 2006: Päätalo, Kalle (1919–2000). – *Suomen kansallisbiografia*, osa 7, Negri–Pöysti s. 850–852. Studia Biographica 3. Helsinki: Suomalaisen Kirjallisuuden Seura. https://kansallisbiografia.fi/kansallisbiografia/henkilo/4834 (10.12.2021).

Kirves, Jenni 2005: Kalle Päätalo ja rakkauden nälkä. Selkosten Proust oli myös selkosten Miller. – *Parnasso* 1/2005 s. 44–49.

Koivisto, Päivi 2017: *Pienten lajien maailmat. Lyhyt suomalaisen autofiktion historia 1/3*. https://www.helmet.fi/fi-FI/Tapahtumat_ja_vinkit/Vinkit/Pienten_lajien_maailmat_Lyhyt_suomalaise(129798) (29.5.2020).

Kuivasmäki, Riitta 1974: *Kolme suomalaista menestyskirjailijaa ja heidän lukijansa. Väinö Linnan, Hannu Salaman ja Kalle Päätalon tuotannon vastaanoton tarkastelua*. Yleisen kirjallisuustieteen lisensiaatintyö. Tampereen yliopisto.

Kyröläinen, Aki-Juhani – Laippala, Veronika 2020: Määrällinen korpuslingvistiikka. – Milla Luodonpää-Manni, Markus Hamunen, Reetta Konstenius, Matti Miestamo, Urpo Nikanne & Kaius Sinnemäki (toim.), *Kielentutkimuksen menetelmiä I–IV* s. 487–524. Suomalaisen Kirjallisuuden Seuran Toimituksia 1457. Helsinki: Suomalaisen Kirjallisuuden Seura. https://doi.org/10.21435/skst.1457.

Lappalainen, Jussi T. 1999: Tutkimuksen ja fiktion yhtäläinen synty. – Tero Liukkonen (toim.), *Päätalon juurella. Juhlakirja 11.11.1999 Kalle Päätalon täyttäessä 80 vuotta* s. 165–174. Helsinki: Gummerus.

https://doi.org/10.21435/tl.271

Launis, Kati – Cherny, Eugene – Neovius, Mats – Nurmi , Olli – Vainio, Mikko 2018: Mitä naiset lukevat? Vantaan kaupunginkirjaston digitaalisen lainausdatan tulkintaa. – *Avain* 4/2018 s. 4–21. https://journal.fi/avain/article/view/76461 (14.4.2021).

Lauri, Jarkko – Ojajärvi, Jussi 2017: 1950-luku. Taloa rakentamassa. – Jussi Ojajärvi & Nina Työlahti (toim.), *Maamme romaani. Esseitä kirjallisuuden vuosikymmenistä* s. 113–147. Nykykulttuurin tutkimuskeskuksen julkaisuja 121. Jyväskylä: Nykykulttuuri.

Luoma, Jouni – Oinonen, Miika – Pyykönen, Maria – Laippala, Veronika – Pyysalo, Sampo 2020: A broad-coverage corpus for Finnish named entity recognition. – *Proceedings of the 12th Language resources and evaluation conference* s. 4615–4624. https://www.aclweb.org/anthology/2020.lrec-1.567 (14.4.2021).

Mahlberg, Michaela – Stockwell, Peter – de Joode, Johan – Smith, Catherine – O'Donnell, Matthew Brook 2016: CLiC Dickens: Novel uses of concordances for the integration of corpus stylistics and cognitive poetics. – *Corpora* 11(3), 433–463. https://doi.org/10.3366/cor.2016.0102.

Majander, Antti 2019: Kirjallisuuskriitikon kauhunhetki. Kalle Päätalo polvistui eteeni. Marraskuussa täyttyy 100 vuotta omaelämäkerrallisuuden maailmanennätysmiehen syntymästä. – *Helsingin Sanomat* 3.8.2019. https://www.hs.fi/blogi/majander/art-2000006193321.html (3.6.2020).

Marttinen, Eero 1995: *Mestarin eväät. Kalle Päätalon taival kuvin ja sanoin*. Jyväskylä: Gummerus.

Marttinen, Eero 1996: *Mestarin selkonen. Kalle Päätalon Koillismaa sanoin ja kuvin*. Jyväskylä: Gummerus.

Marttinen, Eero 1999: *Mestari elämän telineillä. Kalle Päätalo ystävien silmin*. Jyväskylä: Gummerus.

McIntyre, Dan & Walker, Brian 2019: *Corpus stylistics. Theory and practice*. Edinburgh: Edinburgh University Press.

Moretti, Franco 2013 (2000): *Distant reading*. Lontoo: Verso Books.

Navarro-Colorado, Borja 2018: On poetic topic modeling. Extracting themes and motifs from a corpus of Spanish poetry. – *Frontiers in Digital Humanities*. https://doi.org/10.3389/fdigh.2018.00015.

Niemi, Juhani 1999: Kansan suosikki ja kaanon. – Tero Liukkonen (toim.), *Päätalon juurella. Juhlakirja 11.11.1999 Kalle Päätalon täyttäessä 80 vuotta* s. 35–41. Helsinki: Gummerus.

Ojajärvi, Jussi 2013: Työn syrjässä. – Mika Hallila, Yrjö Hosiaisluoma, Sanna Karkulehto, Leena Kirstinä ja Jussi Ojajärvi (toim.), *Suomen nykykirjallisuus 2. Kirjallinen elämä ja yhteiskunta* s. 154–156. Suomalaisen Kirjallisuuden Seuran Toimituksia 1390:2. Helsinki: Suomalaisen Kirjallisuuden Seura.

Parente-Čapková, Viola 2017: Naisten kirjoittaman kirjallisuuden ylirajainen vastaanotto. Kohti kirjallisten toimijuuksien historiaa. – *Avain* 4/2017 s. 36–49. https://journal.fi/avain/article/view/69305/30757 (14.4.2021).

Pietilä, Arto 1980: Veri veti kohti selkosia. – *Pääsiäis-Pellervo* 2/1980 s. 8–11.

Päätalo, Kalle 1995: *Hyvästi, Iijoki*. Jyväskylä: Gummerus.

https://doi.org/10.21435/tl.271

Rajala, Panu 1999: Kuinka Päätalo otti lukijan kyytiinsä. – Tero Liukkonen (toim.), *Päätalon juurella. Juhlakirja 11.11.1999 Kalle Päätalon täyttäessä 80 vuotta* s. 175–185. Helsinki: Gummerus.

Rojola, Lea 2002: Läheisyyden löyhkä käy kaupaksi. – Markku Soikkeli (toim.), *Kurittomat kuvitelmat. Johdatus 1990-luvun kotimaiseen kirjallisuuteen* s. 69–99. Turun yliopisto, Taiteiden tutkimuksen laitos, Sarja A 50. Turku: Turun yliopisto.

Roos, J. P. 1999: Mutta onko tämä omaelämäkerta? – Tero Liukkonen (toim.), *Päätalon juurella. Juhlakirja 11.11.1999 Kalle Päätalon täyttäessä 80 vuotta* s. 25–30. Helsinki: Gummerus.

Räisänen, Alpo 1981: Kalle Päätalo syntymäseutunsa kielen ja kulttuurin välittäjänä. Professori Alpo Räisäsen esitelmä Päätalo-Savotassa 12.–13.7.1980. – *Taivalkosken Sanomat* 7 s. 5–7.

Räisänen, Alpo 1999: Kestävämpi kuin Kollajan poola. – Tero Liukkonen (toim.), *Päätalon juurella. Juhlakirja 11.11.1999 Kalle Päätalon täyttäessä 80 vuotta* s. 91–98. Helsinki: Gummerus.

Sainio, Mauri 1973: Kalle Päätalo. Pitkän tekstin mies. – *YV* 1/1973 s. 14–17.

Santos, Diane 2018: Literature studies in Literateca. Between digital humanities and corpus linguistics. https://www.linguateca.pt/Diana/download/DianaSantosFSCEO.pdf (14.4.2021).

Schöch, Christof 2017: Topic modeling genre. An exploration of French classical and enlightenment drama. – *Digital Humanities Quarterly* 11 (2).

Sinnemäki, Anssi 1998: Kuka olikaan ensimmäinen? https://www.hs.fi/kulttuuri/art-2000003733229.html (10.12.2021).

Stockwell, Peter 2017: Poe's Gothic ambience. – Linda Barone (toim.), *Over many a quaint and curious volume of (not) forgotten lore. Edgar Allan Poe across disciplines, genres and languages* s. 7–23. Newcastle: Cambridge Scholars Publishing.

Teräs, Kari 2018: Mitä on muistitieto, ja miten se poikkeaa muusta historiallisesta tiedosta? – *Alusta!* 6.7.2018. https://alusta.uta.fi/2018/07/06/mita-on-muistitieto-ja-miten-se-poikkeaa-muusta-historiallisesta-tiedosta/ (12.4.2021).

Vahtola, Jouko 1999: Pelastettu kulttuuri. – Tero Liukkonen (toim.), *Päätalon juurella. Juhlakirja 11.11.1999 Kalle Päätalon täyttäessä 80 vuotta* s. 151–163. Helsinki: Gummerus.

Vahtola, Jouko 2001: Kalle Päätalo 1919–2000. – *Päätalo-sanomat* s. 27–30.

Vehviläinen, Matti 1978: *Kalle Päätalo*. Jyväskylä: Gummerus.

Ylitalo, Riikka 2010: Kalle Päätalo, Pohjois-Suomen suuri kansankirjailija. – Sinikka Carlsson, Liisi Huhtala, Sanna Karkulehto, Ilmari Leppihalme & Jaana Märsynaho (toim.), *Pohjois-Suomen kirjallisuushistoria* s. 119–121. Suomalaisen Kirjallisuuden Seuran Toimituksia, Tiede 1261. Helsinki: Suomalaisen Kirjallisuuden Seura.

Ylönen, Ritva 2013: *Tervaksinen toteemi. Kalle Päätalon Iijoki-sarjan vastaanotto ja vaikutus*. Acta Electronica Universitatis Tamperensis 1335. Tampere: Tampere University Press. https://urn.fi/URN:ISBN:978-951-44-9214-3.

Ylönen, Ritva 2017: *Kalle Päätalo. Kirjailijan elämä*. Helsinki: Suomalaisen Kirjallisuuden Seura.

https://doi.org/10.21435/tl.271