

lappen Ranta

– tutkielma
yhdyssanoista, sanaliitoista ja niiden
poikkeamista
sosiaalisen median keskusteluissa

Heli Ristilä
Tampereen yliopisto
Kieli-, käännös- ja
kirjallisuustieteiden yksikkö
Suomen kieli
Pro gradu -tutkielma
Toukokuu 2013

Tampereen yliopisto

Suomen kieli

Kieli-, käännös- ja kirjallisuustieteiden yksikkö

RISTILÄ, HELI: *lappeen Ranta* – tutkielma yhdyssanoista, sanaliitoista ja niiden poikkeamista sosiaalisen median keskusteluissa

Pro gradu -tutkielma, 101 sivua + liitteet 7 sivua

Kevät 2013

Tässä tutkielmassa tarkastellaan yhdyssanoja ja sanaliittoja sosiaaliseen mediaan, Facebookiin, kirjoitetussa kielessä. Tutkielmassa selvitetään, paljonko yhdyssanoja on, millaisia niiden rakenteet ovat ja paljonko niissä tehdään poikkeamia suhteessa konventionaaliseen kirjoitusasuun. Sanaliittoja tarkastellaan suhteessa yhdyssanoihin ja niiden poikkeamamääriä sekä -tyyppjä esitellään.

Yhdyssanoja on 77 697 sanan aineistossa kaikkiaan 7634 kappaletta eli 9,8 % koko sanamäärästä. Konventiosta poikkeavia, osiinsa jakautuneita yhdyssanoja on 431, siten poikkeamaprosentiksi tulee 5,6 % koko yhdyssanamäärästä. Poikkeavasti yhteen kirjoitettuja sanaliittoja on 334. Yhdyssanoissa poikkeamia aiheuttaa eniten yhdysviivan käyttö: jos yhdyssanan osana on numero, lyhenne, merkki tai sanaliitto, joka konvention mukaisesti sanaan yhdistyessään vaatii ilmaisuun yhdysviivan, on poikkeaman syntyminen huomattavasti todennäköisempää kuin missään muussa yhdyssanatyypissä. Sanaliitoissa vaikeuksia tuottavat numeron ja sen määritettävän yhdistäminen (4 €) sekä pikkusanat, partikkelit ja konjunktiot, jotka ajautuvat yhteen konvention vastaisesti (*niinkuin*).

Substantiivien osuus on 85,2 % kaikista yhdyssanoista. Seuraavaksi eniten on adjektiiveja, 11,5 %. Aivan marginaalissa ovat yhdysverbit (0,7 %) ja muut sanaluokat (2,6 %). Yleisin yksittäinen yhdyssanatyyppejä on nominatiivisubstantiivimääritteiset yhdyssubstantiivit (NS+S), joita on 60,8 % kaikista yhdyssanoista. Määrittäjien muodosta yleisin on nominatiivi: nominatiivimääritteisiä yhdyssanoja tutkielman yhdyssanoista on 67 %.

Yhdyssanatyypeistä yleisin on määritysyhdyssanat (98,2 %), summayhdyssanoja ja muita yhdyssanatyyppejä on hyvin marginaalisesti.

Avainsanat: yhdyssana, sanaliitto, kielenhuolto, Facebook, sosiaalinen media, internet

Sisällysluettelo

1 JOHDANTO.....	1
1.1 Tutkielman kuvaus.....	1
1.2 Tutkielman taustaa.....	2
1.3 Tutkimuksen tavoitteita.....	4
2 YHDYSSANAT JA SANALIITOT.....	6
2.1 Yhdyssana per se.....	6
2.2 Yhdyssanan suhde sanaliittoon.....	9
2.3 Yhdyssanan kehyksiä: käsite, termi, yleisyys ja merkitys.....	11
2.4 Yhdyssanat viimeaikaisessa tutkimuksessa.....	14
3 AINEISTO.....	18
3.1 Aineiston keräys.....	18
3.2 Facebook tekstialustana ja aineistolähteenä.....	19
3.3 Aineistoa kuvaavia lukuja.....	22
3.4 Korpuksesta tutkimusaineistoksi.....	23
3.5 Yhdyssanan kriteerit – aineiston rajaus.....	27
3.5.1 Prefiksit ja muut epäitsenäiset määritteet.....	27
3.5.2 Johtimet -lainen, -moinen ja -mainen.....	28
3.5.3 Leksikaalistumat, vierassanat ja erisnimet.....	29
4 AINEISTON ANALYYSI.....	31
4.1 Yhdyssanat määritysosien määrän mukaan.....	31
4.2 Yhdysviiva ja erityisyhdyssanatyypit.....	33
4.3 Perusyhdyssanat.....	36
4.3.1 Yhdysosa-alkuiset yhdyssanat.....	40
4.3.2 Nominatiivimääritteiset yhdyssanat.....	46
a) Nominatiivimääritteiset substantiivit.....	47
b) Nominatiivimääritteiset adjektiivit.....	55
c) Nominatiivimääritteiset verbit ja muut marginaaliset esiintymät.....	57
d) Nominatiivimääritteisten yhdyssanojen yhdyssanatyypit.....	58
e) Nominatiivimääritteisten yhdyssanojen yhteenveto.....	60
4.3.3 Genetiivialkuiset yhdyssanat.....	62
a) Genetiivimääritteisten yhdyssanojen poikkeamat.....	64
b) Genetiivimääritteiset yhdyssanat sanaluokittain ja yhdyssanatyypeittäin.....	72
c) Genetiivimääritteiset yhdyssanat versus sanaliitot.....	75

4.3.4 Muut määritetyypit.....	78
4.4 Yhteenvedo kaikista yhdyssanoista.....	85
4.5 Yhteen kirjoitetut sanaliitot.....	88
4.5.1 Numeroilmaisut ja yhdysviiva.....	89
4.5.2 Liittosanat.....	91
4.5.3 Sulaumat.....	93
4.5.4 Liittoutuvat sanat.....	95
4.5.5 Kirjoitusvirheet.....	96
5 LOPUKSI.....	97
LÄHTEET.....	100
KUVAT	
Kuva 1. Yhdyssanan tyyppilyhenteen selite	40
TAULUKOT	
Taulukko 1. Yhdyssanojen rajaus.....	30
Taulukko 2. Moniosaiten määritysyhdyssanojen rakenne ja poikkeamat.....	32
Taulukko 3. Yhdysviiva.....	35
Taulukko 4. Erityisyhdyssanat määritysosan tyyppin mukaan.....	36
Taulukko 5. Yhdysosamääritteiset yhdyssanat määriteosan muodon mukaan.....	42
Taulukko 6. Yhdysosamääritteiset yhdyssanat sanaluokittain.....	45
Taulukko 7. Yhdyssanojen kombinaatiot yhdysosamääritteisten ryhmässä.....	46
Taulukko 8. Nominatiivimääritteiset yhdyssanat sanaluokan mukaan.....	60
Taulukko 9. Nominatiivimääritteiset yhdyssanat määritysosan sanaluokan mukaan.....	61
Taulukko 10. Nominatiivimääritteisten yhdyssanojen kombinaatiot.....	62
Taulukko 11. Genetiivimääritteiset yhdyssanat sanaluokan mukaan.....	73
Taulukko 12. Genetiivimääritteiset yhdyssanat määritysosan sanaluokan mukaan.....	74
Taulukko 13. Genetiivimääritteisten yhdyssanojen kombinaatiot.....	74
Taulukko 14. Genetiivimääritteisten sanaliittojen ja yhdyssanojen suhteet.....	78
Taulukko 15. Yhdyssanat, muut-ryhmä, sanaluokan mukaan.....	82
Taulukko 16. Yhdyssanat, muut-ryhmä, määritysosan sanaluokan mukaan.....	83
Taulukko 17. Yhdyssanojen muut-ryhmän kombinaatiot.....	85
Taulukko 18. Yhdyssanat sanaluokittain.....	87
Taulukko 19. Yhdyssanojen kombinaatiot.....	88

LIITTEET

LIITE 1 Internetviitteet koottuna.....	102
LIITE 2 Aineiston keruupaikat, Facebook-ryhmät.....	103
LIITE 3 Yhdysosa-alkuisten yhdyssanojen koontitaulukko.....	104
LIITE 4 Nominatiivimääritteisten yhdyssanojen koontitaulukko.....	105
LIITE 5 Genetiivimääritteisten yhdyssanojen koontitaulukko.....	106
LIITE 6 Yhdyssanojen koontitaulukko, muut määrittäjät.....	107
LIITE 7 Yhdyssanat koontitaulukko kaikki.....	108

1 JOHDANTO

1.1 Tutkielman kuvaus

Pro gradu -tutkielmani aiheena ovat yhdyssanat ja sanaliitot internetiin kirjoitetussa suomen kielessä. Tavoitteenani on esittää, paljonko ja millaisia yhdyssanoja tavalliset ihmiset¹ kirjoittavat, ja tutkimustulosten perusteella tarkastella, miten yhdyssanat ja sanaliitot hahmotetaan sekä millaisia poikkeamia niissä tehdään.

Suomalaiseen oppivelvollisuuteen kuuluu äidinkielen opetus (Perusopetuslaki § 11: ks. IV:FX). Koska äidinkielen oppiaineessa opetetaan kirjallisuuden ja kieliopin lisäksi oikeinkirjoitusta, voisi kuvitella, että kaikki suomalaiseen opetukseen osallistuneet hallitsisivat kirjoitetun suomen kielen konventiot, mutta näinhän ei ole. Yhteisessä kielessämme näkyy poikkeamia siitä, mitä meille on opetettu ja mitä kielenhuolto suosittaa. Kouluopetuksen taso saattaa vaihdella, samoin vaihdellee ihmisten kielellinen lahjakkuus ja ainakin näistä syistä voi syntyä poikkeamia. Lisäksi kielenhuollon ohjeita selatessa törmää hyvin nopeasti siihen tosiasiaan, että mutkikkaita, ulkoa muistettavia sääntöjä ja jopa sanalistoja on runsain mitoin ja toisaalta yksinkertaisia, kiistattomia sääntöjä kovin vähän. Sanapareilla voi myös esiintyä rinnakkaisia kirjoitusmuotoja, jotka ovat tai eivät ole säännönmukaisia siitä riippuen, mistä milloinkin puhutaan. Osa sanapareista varioi jopa ilman selitystä.

Milloin kieli on konventioiden mukaista, on siihen ainakin yksi selkeä syy: niin on opetettu ja kielenhuollossa suositettu. Korrekti kirjoitusasukin voi tietysti kertoa kielen ja ehkä jopa edelleen sen edustaman maailman hahmottamisesta, mutta emme kuitenkaan voi tietää, millä perusteella kirjoittaja on kielelliset valintansa tehnyt. Onko kirjoittaja valinnut korrektein kirjoitustavan vapaasti, oman kielellisen käsityksensä mukaan, vai vain ulkoa opittuna ja suositusten mukaisesti? Mielenkiintoisia tässä yhteydessä ovatkin erilaiset konventioista poikkeavat kirjoitusasut, jotka nähdäkseni kertovat ennen muuta yksilön henkilökohtaisesta valinnasta ja edelleen kielen yksiköiden hahmottamisesta. Mielenkiintoisia ovat myös ne yhdyssanat, joiden kirjoitusasuksi sanakirja tai kielenhuolto antaa sekä yhteen että erilleen kirjoitetut versiot: yhtäältä kumpi tapa silloin mieluummin valitaan ja toisaalta voiko sanaparien muodosta tai merkityksestä johtaa yleistyksiä ja tehdä ennakoiteja siitä, kumpi tapa on todennäköisempi?

¹ 'Tavallisella ihmisellä' tarkoitan tässä työssä henkilöä, joka kirjoittaa julkisuuteen (tässä Facebookiin: www.facebook.com) huvikseen ja harrastukseksi. Jo tässä yhteydessä on huomautettava, että aineistoni kirjoittajina on luultavasti myös työkseen kirjoittavia ja varmaan lingvistejäkin – tällainen seikka on mahdoton tarkistaa noin 2700 kirjoittajan osalta – mutta voimme turvallisesti olettaa, että heitä on tutkimustuloksen kannalta merkityksellisen pienen joukon koko kielenopasryhmästä.

Tämän tutkielman aineisto, Facebook-tekstit, jo kertonee lukijalle, että tavoitteenani on tutkia nimenomaan kieltä, joka pääsääntöisesti ei ole jyrkän muodollista vaan joka ennemminkin on kirjoitettua jutustelua kuin tarkkaan harkittua, viimeisteltyä ilmaisua, jollaista esimerkiksi sanomalehtiteksti on. Tutkielmassani tavoittelen kuvaa siitä, millaisia valintoja tehdään, kun ei luultavasti ehditä miettiä kielenhuollon sääntöjä saati tarkistaa kirjoitusasua sanakirjasta: miten kielenhuollon säännöt hallitaan intuitiivisella tasolla, miten yhdyssanojen ja sanaliittojen kirjoittaminen yleisesti osataan ja millaisissa kohdissa poikkeamia syntyy – tai ei synny.

1.2 Tutkielman taustaa

Kaupallisen internetin alkuaikoina, 1990-luvun puolivälin paikkeilla, suurin osa verkkoon kirjoittajista oli akateemista väkeä. Tilanne muuttui kuitenkin nopeasti, kun internetliittymiä alkoi saada korkeakoulujen, yliopistojen ja yritysten lisäksi tavallisiinkin talouksiin kohtuullisella hinnalla. Tilastokeskuksen mukaan vuonna 1996 internetyhteys oli 7 %:lla kotitalouksista. Yhdeksän vuotta myöhemmin vuonna 2005 kotitalouksista jo 58 % oli kytkettynä verkkoon, ja keväällä 2012 peräti 87 %:lla kotitalouksista oli vähintään yksi internetliittymä ja 45 %:lla jopa useampi (IV:SVT1). Käsi kädessä liittymien määrien kasvun kanssa kulkee internetin käytön monipuolistuminen. Aluksi kotitalouksissa käytettiin internetiä etupäässä sähköpostitteluun, tiedonhakuun sekä pankkiasioihin; vain 24 % käyttäjistä mainitsi käyttötarkoitukseksi keskustelupalstat ja chatit vielä vuonna 2005 (IV:SVT2). Vuonna 2012 käyttötarkoitusten kärjessä olivat edelleen sähköposti, tiedonhaku ja pankkiasiat, mutta keskustelupalstoja ja yhteisöpalveluja mainitsi käyttävänsä jo 49 % kaikista internetin käyttäjistä. Huomattavaa on, että mitä nuoremasta ikäryhmästä oli kyse, sitä yleisempää keskustelun käyminen ja yhteisöpalveluiden käyttö internetissä oli: 16–24-vuotiaista 88 % ilmoitti käyttäneensä internetin keskustelupalstoja ja yhteisöpalveluita (IV:SVT3). Nykyään verkossa hoidetaan muiden asioiden lisäksi sosiaalisia suhteita ja osallistutaan keskusteluihin. Internetistä on tullut luonteva kommunikointiväline, ja on perusteltua väittää, että yhä useampi ihminen viestii kirjoittamalla. Verkon välityksellä viestitään niin yksityisesti (sähköposti, erilaiset chat-ohjelmat kuten Messenger, Skype jne.) kuin julkisestikin esimerkiksi blogeissa, Twitterissä, Google+:-ssa tai Facebookissa. Julkisesta kirjoittamisesta on tullut arkipäiväistä jutustelua.

Väitän edellä esitetyn perusteella, että vaikka tavallisten ihmisten verkkoon kirjoittama teksti – erotuksena esimerkiksi kaunokirjallisuudesta ja sanomalehtitekstistä, olkoonpa niiden julkaisu-alusta mikä hyvänsä – ei tietenkään ole korvannut heidän puhettaan, on se kuitenkin astunut puheen rinnalle yleisenä kommunikointitapana. Voisi jopa ajatella, että nettitekstit ovat kirjoitettua puhetta ennemminkin kuin kirjoitettua tekstiä. Millä tavoin tämä näkyy kielenkäytössä? Kuinka sujuvaa kieli

on ja hallitsevatko ihmiset kielenhuoltosuositusten mukaisen korrektiin kielen – tai välittävätkö siitä edes? Asiasta on keskusteltu internetin yleistymisen aikana paikoin kiivaastikin, ja ilmiselvästi aihe koskettaa. Vähäisellä etsimisellä löytää helposti oikeakielisyyskeskusteluja, ja tässä yksi esimerkiksi: Kaksplus-lehden keskustelupalstalla (IV:KP) elokuussa 2008 esitettiin näkökulmia niin korrektiin kielen puolesta kuin vastaankin.

Onpa huonosti koulunsa käyneitä tällä palstalla, kun tuo yhdyssanojen kirjoitus takkuua NIIN MONELLA!! Ärsyttävää!!

kirjoittaa nimimerkki Älykkö. Hänelle vastaa nimimerkki Ärrimurri:

On se hyvä, jos joku voi ajatella olevansa älykkö jos osaa kirjoittaa yhdyssanoja.

Samaiselta palstalla avautuu myös nimimerkki ei oo välii kertomalla:

Tai minä en jaksa edes ajatella miten kuuluu kirjoittaa. Tää on täällä niin epävirallista et ihan sama.

Kielenhuollon asioista erityisesti yhdyssanavirheet tuntuvat joutuvan silmätikuksi ja kiinnostavan. Karkeasti ottaen voi sanoa, että henkilöä, joka ei hallitse yhdyssanojen oikeinkirjoitusta, pidetään lähtökohtaisesti hiukan tyhmänä ja ainakin kouluttamattomana siitä huolimatta, että yhdyssanavirheitä tehdään paljon ja yhdyssanojen myönnetään olevan vaikea kielen alue (esimerkiksi edellä mainittu IV:KP). Google-haulla (IV:GH) *yhdyssanavirhe* tulokseksi tulee 13 500 osumaa (31.1.2013, tulokset vaihtelevat päivittäin). Hakutuloksen alkuun sijoittuu parikin humoristista sivustoa, joiden kärkeen kiilaaminen kertoo osaltaan aiheen suosiosta ja edelleen siitä, että kirjoittamalla yhdyssanoja väärin voi helposti parodioida heppoista keskustelua. Suorastaan legendaksi on muodostunut tämän tutkimuksen pääotsikoksi hyvästä syystä valitsemani nettislangisana – tai paremminkin sanapari – *lappeen Ranta*, jolla ivallisesti huomautetaan yhdyssanavirheestä. Sanapari esitellään esimerkiksi Wikisanakirjassa nettislangia valottavassa luettelossa (IV:WS).

lappeen Ranta - ilmaus, jolla kommentoidaan yhdyssanavirheitä, etenkin jos niitä on paljon.

Yhdyssanat siis kiehtovat nettikirjoittajien mieltä. Näyttäisi siltä, että maallikoiden käsityksen mukaan yhdyssanojen osaaminen sijoittaa kirjoittajan johonkin kohtaan jatkumoa, jonka toisessa päässä on kouluttamaton tyhmä kirjoittaja ja toisessa koulutettu älykkö. Todellisuus tuskin on näin yksiselitteistä. Lingvistienkin mielestä yhdyssanojen hallintaa voidaan kuvata varsin laajalla skaalalla: Lantee ja Nieminen arvioivat, että yhdyssana on kirjoitetun kielen knoppi, jolla erotellaan sivistynyt kirjoittaja muista (Nieminen–Lantee b2010), kun taas Marja Ylikulju (Ylikulju 2003: 22) toteaa ykskantaan, että jos kirjoittaja ei hallitse nominatiivialkuisen yhdyssanan oikeinkirjoitusta, hän ei hahmota suomen kielen rakennetta – mikä mielestäni on melko jyrkästi sanottu, etenkin jos arvioidaan natiivisuomenkielistä, oman äidinkielenä intuitiivista asiantuntijaa. Kuvaavaa ja hiukan

ehkä huvittavaakin on se, että Ylikulju haksahdaa itsekin toteamuksensa yhteydessä kielioppisääntöä määrittäessään (mts. 22, alleviivaus omani):

Nykysuomessa ei voi olla nominatiivialkuista attribuuttia sanaliittojen alkuosana muuten kuin erisnimien ja niihin verrattavien elementtien yhteydessä (Vesikansa 1989: 224, Häkkinen 1990: 146). Erikseen kirjoittaminen on merkki siitä, että kirjoittaja ei hahmota suomen kielen rakennetta.

Nominatiivialkuinen attribuutti jo sinänsä on niin outo käsite, että voimme olettaa kirjoittajan tarkoittaneen nominatiivimuotoista attribuuttia. Sellainen toki voi nykysuomessa olla, nimittäin vaikka nominatiivisubstantiivi ei sanaliiton attribuuttina voikaan olla (kuin mainituissa erikoistapauksissa), adjektiivi (*nuori mies*) ja numeraali (*yksi talo*) ovat hyvinkin mahdollisia nominatiivimuotoisia attribuutteja sanaliitoissa. Tarkoitukseni ei tässä kuitenkaan ole osoitella sormella mainittua tutkielman tekijää vaan nostaa esille se seikka, että yhdyssanat ovat hankalahko kielen alue ja niiden määrittely voi olla melko mutkikasta. Palaan Lanteen, Niemisen ja Ylikuljun näkemyksiin vielä esitellessäni viimeaikaista yhdyssanatutkimusta (luvussa 2.4) sekä nominatiivisubstantiivimääritteisiin yhdyssanoihin useassa kohden tätä työtä (esim. luvuissa 2.3 Yhdyssanan kehyksiä jne. ja 4.3.2 Nominatiivimääritteiset yhdyssanat).

1.3 Tutkimuksen tavoitteita

Tutkimukseni on ennen muuta aineistolähtöinen – aineisto on ytimessä ja aineiston mukaan muotoutuu käsittely – ja se on laadultaan pikemminkin kvantitatiivinen kuin kvalitatiivinen, eli pyrin ryhmittelemään aineistoni mahdollisimman täsmällisesti. Selvitän lajittelun hankaluutta sekä aineistoa kuvailevassa luvussa että tarpeen mukaan yksityiskohtaisesti jaoteltujen tyyppien käsittelyn yhteydessä. Edelleen ryhmittelyn ja siitä johdettujen määrien mukaan pyrin tekemään yleistyksiä ja esittämään kvantitatiivista tietoa tavallisten kirjoittajien kielen yhdyssanoista, yhdyssanatyypeistä ja -poikkeamista. Yleistyksiä tehdessä on luonnollisesti mielenkiintoista katsoa jonkin verran laadullistakin kulmaa ja miettiä syitä sille, miksi tietyt ilmaisut ajautuvat yhteen ja tietyt erilleen.

Kvalitatiivista otetta tutkimukseeni tulee lisää, kun vien analyysia pidemmälle silloin, kun mielestäni on aihetta arvioida, mistä poikkeamatyyppi voisi syntyä ja voiko poikkeamien syntyä ennakoida esimerkiksi muodon tai merkityksen perusteella. Tarkastelun kohteena ovat myös ne ilmaisut, joille kielenhuolto ja sanakirja sallivat erilaiset kirjoitusasut; kumpi silloin mieluummin valitaan. Kaiken kaikkiaan tavoitteeni on tehdä katsaus nettikielen yhdyssanoihin, kuvailla niiden määriä, tyyppejä ja poikkeamia niin, että muodostuu käsitys siitä, missä asemassa yhdyssanat arkisessa kielessämme ovat niin määrällisesti kuin yleisyystyypeittäinkin. Tavoitteenani on myös tarkastella poikkeavia sanaliittoja, eli sanapareja jotka ovat ajautuneet yhteen konvention vastaisesti, ja erotella

niistä joitain keskeisiä tyyppejä. Toivoakseni työni palvelee jatkossa niitä tutkijoita, jotka tarvitsevat määrällistä tietoa siitä, paljonko kirjoitetussa arkisuomessa on yhdyssanoja, mitä tyyppejä ja millaisin poikkeamamäärin; työstäni voivat luonnollisesti hyötyä myös opettajat ja kielenhuoltajat.

Työni etenee niin, että ensin teen yleiskatsauksen suomen kielen yhdyssanoihin: tarkastelen, mikä yhdyssana oikeastaan on, pohdin yhdyssanan ja sanaliiton suhdetta ja yritän hahmottaa yhdyssanan kehäksiä käsitteen, termimäisyyden, yleisyyden ja merkityksen näkökulmista (luvut 2.1–2.3). Käyn läpi myös muutamia, lähinnä aivan viimeaikaisia, tämän työn kannalta relevantteja yhdyssanatutkimuksia ja oman työni yhteyksiä niihin (luvussa 2.4). Sitten esittelen aineistoni ja sen suhteen tekemäni ratkaisut niin keräyksen, tekstimassan rajauksen kuin muunkin käsittelyn suhteen (luku 3 alalukuineen). Huomautan jo tässä, että kaikki aineistoni esimerkit olen numeroinut ja muut, esimerkiksi tekstin joukossa olevat havainnollistamistarkoitusta varten keksimäni ilmaisut, on numeroimatta.

Varsinaisen työni ydinosa muodostaa yhdyssana-aineiston erittely (luku 4), jonka aloitan tarkastelemalla ensin yhdyssanan osien määrää ja määrän suhdetta poikkeamiin (luku 4.1). Oletukseni on, että mitä enemmän yhdyssanassa on osia, sitä useimmin sen oikeinkirjoitus tuottaa hankaluutta. Seuraavaksi esittelen yhdysviivan käyttöä ja samalla ne yhdyssanatyypit, jotka olen tässä työssä nimennyt erityisyhdyssanatyypeiksi ja jotka eivät ole aivan työni kaikkein suurimman mielenkiinnon kohteena juuri niiden erityisen luonteen vuoksi (luku 4.2). Erityisyhdyssanoista jatkan käsittelyä paneutumalla perusyhdyssanoiksi nimeämiini yhdyssanoihin (luku 4.3). Niiden käsittely etenee yksinkertaisimmista enemmän selitettäviin: ensin yhdysosa-alkuiset (luku 4.3.1), sitten nominatiivi- ja genetiivi-alkuiset (luvut 4.3.2 ja 4.3.3) ja luvun lopuksi ne yhdyssanat, joiden perusosaa määrittää jokin muu kuin edellä mainittu määriteosatyyppi (4.3.4). Teen vielä yhteenvedon kaikista yhdyssanoista (4.4) niiden piirteiden perusteella, jotka ovat kaikille yhdyssanoille yhteiset, ennen kuin siirryn käsittelemään konventiosta poikkeavasti yhteen kirjoitettuja sanaliittoja, jotka käsittelen tässä työssä siis eräänlaisina poikkeavina yhdyssanoina (luku 4.5).

Tutkielman lopuksi (luku 5) teen yhteenvedon tämän työn tuloksista.

2 YHDYSSANAT JA SANALIITOT

Suomen kielioppia lienee mahdoton kirjoittaa ilman, että ottaisi kantaa yhdyssanoihin, onhan sanojen yhdistäminen eli yhdyssanan muodostaminen – kuten niin usein todetaan – yleisin sananmuodostuskeino johtamisen ohella (esimerkiksi Häkkinen 1990: 144; Laaksonen–Lieko 2003: 109; Hakulinen et al. Iso suomen kielioppi, myöhemmin ISK § 146; Ikola 2001: 106). Yhdyssana on läsnä jo kielen lähtökohdassa, kieltä luomassa. Kaiken aiheesta kirjoitetun referointi ei kuitenkaan ole tämän työn kannalta tarpeellista eikä oleellistakaan, joten vaikka esittelenkin joitain näkemyksiä sanojen yhdistämisestä sananmuodostuskeinona sekä yhdyssanan ja sanaliiton suhteesta toisiinsa, teen sen lähinnä siinä määrin, kuin se on tarkoituksenmukaista tämän työn kannalta. Tässä luvussa käyn läpi nimenomaan yhdyssana- ja sanaliittokuvauksia. Yhdyssanan ja johtamisen suhdetta selvittelen aineiston rajaamista esittelevissä luvuissa (3.5.1 ja 3.5.2). Aineiston analyysiin paneutuessani käyn läpi kielenhuollon sääntöjä yhdyssanojen ja sanaliittojen konventionaalisesta kirjoittamisesta, sillä juuri niillä on merkitystä, kun aihetta käsitellään poikkeamien näkökulmasta. Tässä luvussa esittelen lopuksi muutamia yhdyssanatutkimuksia ja samalla sijoitan oman työni tähän tutkimuskenttään suhteessa aiempiin tutkimuksiin.

2.1 Yhdyssana per se

Kielen ominaispiirteisiin kuuluneen, että sen hämärärajaisuuden vuoksi kielikuvailut ja -oppisäännöt ovat mutkikkaita, täynnä täsmennyksiä ja poikkeuksia. Yhdyssanojen osalta hankaluus alkaa jo siinä vaiheessa, kun pyritään määrittämään, mikä oikeastaan on yhdyssana.

ISK määrittelee (ISK määritelmät: yhdyssana) yhdyssanan kielen yksikkönä yhdeksi sanaksi, joka muodostuu kahdesta tai useammasta sanasta ja joka kirjoitetaan normin mukaan yhteen ilman sananvälejä (esimerkki 1). Erikoistapauksissa voi olla yhdysviiva (esimerkki 2).

(1) *röyhelökangaskatto*

(2) *linja-autoasema*

Kieliopissa ISK (§ 401) kuitenkin toteaa, että yhdyssanan sisällä voi olla myös sananvälejä, esimerkiksi silloin kun määriteosa on sanaliitto (esimerkki 3).

(3) *Kaunotar ja Kulkuri -numerosta*

Yhdyssana on ulkoisilta, kirjoituksesta havaittavilta tuntomerkeiltään siis sana, joka koostuu vähintään kahdesta yhteen kirjoitetusta sanasta, tai se on sana, joka koostuu erilleen kirjoitetuista sanoista ja jonka sisällä voi olla yhdysviiva tai -viivoja. Esimerkin 3 yhdyssanan erottaa sanaliitosta kirjoituksessa vain pieni viiva pitkän ilmaisun viimeisen sanan *numerosta* edellä. Yhdyssanaa määrittäessä ilmiänsä tärkeämpää onkin ehkä miettiä, mikä on kielen yksikkö. Kielitoimiston verkkosanakirja (jatkossa MOT s.v. yksikkö) määrittelee yksikön (tässä mielessä) itsenäiseksi osaksi, joista muodostuu jaollinen kokonaisuus. Yhdyssana on siis kielen itsenäinen osa, joka on jaettavissa osiin.

Yhdyssanojen tyyppejä on useita. Määritys- eli determinatiivisten yhdyssanojen joukko on yhdyssanatyypeistä kaikkein yleisin (ISK § 408). Tämän tyyppin yhdyssanoissa toinen osa määrittää toista, osat ovat keskenään siis epäsymmetriset. Esimerkki 4, rakkaudentunnustus kahville, on yhtenäinen kommentti, jossa perusosa pysyy samana, mutta määrittäjä vaihtuu.

(4) *aamukahvi, taukokahvi, päiväkahvi, iltapäiväkahvi, töistäpäästyäkahvi, iltakahvi, yökahvi* <3

Vaikka määritysyhdyssanojen osien määrä voi olla hyvin lukuisa ja sen eri osat voivat olla itsekin yhdyssanoja ja siten jaollisia, on yhdyssana kuitenkin lopulta aina jaettavissa kahteen osaan (esimerkeissä 5 ja 6 näiden kahden osan välissä ||-merkki). Joskus kuitenkin on hankalaa päätellä, miten sana olisi jaettavissa kahtia: esimerkissä 6 on kaksi mahdollista tulkintaa.

(5) *[[[röyhelö][kangas]]||katto] : [[[linja-][auto]]||asema] : [vaki||[[käsi][kirjoittaja]]]*

(6) *mielenterveysongelma > [mielen||[[terveys][ongelma]]] vai [[[mielen][terveys]]||ongelma]*

Esimerkki 6 ei tosin ole kovin edustava: vaikka voidaan puhua sekä mielenterveydestä että terveysongelmasta, lienee selvää, että kyseessä on pikemminkin ongelma mielenterveydessä kuin terveysongelma mielessä. Tämä jaottelu kuitenkin selventänee, miten näkökulma asiaan muuttuu, jos sanajakoa katsotaan kahdelta (tai n) kannalta. Palaan yhdyssanan osien määrään luvussa 4.1 Yhdyssanat määritysosien määrän mukaan.

ISK kutsuu näitä yhdyssanan kahtiajaossa syntyneitä osia määriteosaksi (§ 403) ja edusosaksi (§ 404). Ilmaus on sikäli hankala, että edusosa on (käytännössä aina) sanan jälkimmäinen osa, se osa jota määritetään – *edus-* kuitenkin tuntuu viittaavan edustaan, *etu-* eli määriteosaan. (Vesikansa (Vesikansa 1989: 223) tosin esittelee yhdyssanatyypin, jossa hänen mukaansa jälkimmäinen osa määrittää ensimmäistä: *isäukko*, *äitimuuori*. Mielestäni tämän tyyppin voi kuitenkin hahmottaa appositioyhdyssanaksikin (ks. myös edempänä tässä luvussa), jossa osat ovat saman entiteetin tai asian kaksi puolta: samassa ilmaisussa hahmottuu siis entiteetti joka on sekä isä että ukko tai äiti että muori. Se kumpi osa määrittää kumpaa, ei ole päätettävissä.) Ongelma yhdyssanan osien nimeämisestä on tietenkin vähäinen, mutta sen verran häiritsevää, että otan tässä työssä käyttöön perinteiset ja

(ainakin) Häkkisen sekä Vesikansan käyttämät termit: **määriteosa** ja **perusosa** (Häkkinen 1990: 150; Vesikansa 1989: 214). Perusosa on siis määrittäjäyhdyssanan jälkimmäinen osa, se yleisviittaava sana ja perusta, jota edellä oleva määrittäjä täsmentää ja rajaa yleistarkoitteesta spesifimmäksi.

Yhdyssanat voivat olla myös rinnasteisia eli summayhdyssanoja (ISK § 398), jolloin osien välinen suhde on tasa-arvoinen eikä mikään osa määritä toista. Tällaisia yhdyssanoja ovat tyypillisesti esimerkiksi ammattinimikkeet (7) ja väreihin liittyvät ilmaisut (8).

(7) *Koomikko -taikuri*

(8) *punavihreä*

Erityistyyppejä ovat määrittäjäyhdyssanan alalajit: tarkoitetta kuvaileva bahuviiri (ISK § 424) (9),

(9) *rättipäät*

samasta asiasta kaksi (tai n) puolta esittelevä appositioyhdyssana (ISK § 422) (10),

(10) *mestari-kynäilijä*

toistoyhdyssanat (ISK § 425) ja

Hopihopi

(korpuksessa, käsiteltävästä aineistosta poistettu)

summayhdyssanojen tyyppiset mutta niistä selvästi eroavat yhdysnumeraalit (ISK § 780) (11).

(11) *kahtakymmentätuhatta*

Kaikkia edellä olevia tyyppejä esiintyy työssäni – joskin muutaman esiintyvän toistoyhdyssanan olen jättänyt huomioimatta yhdyssanaksi paitsi marginaalisuutensa myös epämääräisen muotonsa ja semantiikkansa vuoksi². Odotuksenmukaisesti määrittäjäyhdyssana on yleisin yhdyssanatyypin aineistossanikin. Tutkielmassani yhdyssanatyypin tarkastelu ei ole aivan olennaisin seikka, joten niiden rakenteeseen tai merkityskenttiin tämän syvemmin ei tässä kohden ole tarvetta paneutua. Palaan kuitenkin yhdyssanatyyppeihin vielä luvussa 4.3.2 Nominatiivimääritteiset yhdyssanat: d) Nominatiivimääritteisten yhdyssanojen yhdyssanatyypit (s. 58), lisäksi esittelen yhdyssanatyypin keskinäisiä määriä muissakin ryhmissä. Tässä kuitenkin jatkan sen tarkastelua, mikä erottaa yhdysosan erilleen kirjoitetuista ilmauksista; mikä erottaa (yhtenäisen) yksikön muista kielen osasista, erityisesti sanaliitoista tai liittosanoista.

² Nämä sanat ovat: *hahaha, huhhuh, hopihopi, höpöhöpö, jaajaa, kivakiva, makiatamakiata, nam-nam, namnamnam, oi oi, suomisuomisuomi, ymsyms.*

2.2 Yhdyssanan suhde sanaliittoon

ISK:n mukaan (ISK määritelmät: sanaliitto) sanaliitto on usean sanan muodostama rakenne, ja nimitystä sanaliitto käytetään yhtäältä erotukseksi yhdyssanasta ja toisaalta kuvaamaan kiteytyneitä ilmaisuja. Sanakirja (MOT s.v. rakenne) määrittelee rakenteen kokonaisuuden koostumistavaksi, osien suhteeksi tai olennaiseksi, koostetuksi osaksi. Rakenteen määrittely on ilmeisen hankalaa, ja sanakirjakin turvautuu sitä kuvatessaan vierassanoihin *strukturi* ja *konstruktio*. Edellä olevasta voi kuitenkin päätellä, että sanaliiton osat muodostavat yhdessä kokonaisuuden. Yhdyssanat ja sanaliitot ovat siis sikäli samanlaisia, että kumpikin on jonkinlainen sanojen kooste, yhtenäinen osa. Määrittelyjen sanavalinnoista voi edelleen päätellä – esimerkiksi siitä, että yhdyssana on *yksikkö* ja sanaliitto *kooste* – että yhdyssanan osien summa on kiinteämpi, sanamaisempi eli leksikaalisempi, kuin sanaliiton.

Vesikansa (Vesikansa 1989: 213) nostaa sanaliiton rinnalle termin liittosana, joka on hänen mukaansa kiinteän kokonaisuuden muodostanut sanaliitto. Häkkinen (Häkkinen 1990: 147) puhuu myös liittosanasta pyrkimyksenään tiivistää yhden termin alle ne sanaliitot, jotka muodostavat yhtä sanaa vastaavan kokonaisuuden (esimerkiksi *sen jälkeen ~ sitten*). Yhteen kuuluvissa sanoissa on siis jatkumo: löyhimmät = sanaliitot > kiinteämmät = liittosanat > kiinteimmät = yhdyssanat. Tämän mukaan näyttää siltä, että merkityksen rakentumisen raja on sanaliiton ja liittosanan välillä: liittosanaa vastaa yksi sana ja se on leksikaalistunut joskin jakautunut useaan osaan, kun taas sanaliitto on jollain tavalla merkitykseltäänkin useampisanainen ilmaus. Ortografinen eli kirjoitustavan raja sen sijaan näyttää olevan liittosanan ja yhdyssanan välillä: vaikka ne kumpikin muodostavat kiinteän, leksikaalistuneen ilmaisun, liittosana kirjoitetaan kahdeksi (tai useammaksi) erilliseksi osaksi, yhdyssanan osat kirjoitetaan yhteen. Tämän enempää en ota yhdyssanojen versus sanaliittojen ja liittosanojen osalta kantaa ilmaisujen tiiviyyteen tai löyhyyteen, ja siksi sanon jatkossa yhteen kuuluvia erilleen kirjoitettuja sanoja sanaliitoiksi silloin, kun viittaa yhteen ja erilleen kirjoitettujen sanaparien suhteisiin. Olkoon siis tässä tutkielmassa sanaliitto yleiskäsite sisältäen liittosanan, kun puhutaan sanaliittojen ja yhdyssanojen suhteista. Luvussa 4.5 (ja tarkemmin alaluvussa 4.5.3) kuitenkin otan liittosana-käsitteen uudelleen käyttöön, kun tarkastelen sanaliittojen merkityksen tiiviyt-
tä suhteessa sen osien yhteen kirjoittamiseen.

Kun sekä yhdyssana että sanaliitto voivat viitata kahden tai useamman sanan muodostamaan kokonaisuuteen, millä tehdä ero yhdyssanan ja sanaliiton välille? Tämän rajan vetäminen on ilmeisesti juuri se hankaluus, joka tuottaa horjuntaa kielenkäyttäjillä: milloin sanapari on tarpeeksi kiinteä kokonaisuus, jotta se kirjoitetaan yhteen, milloin yhteen kirjoittamisen perusteita ei ole. Miksi

kirjoitetaan yhteen *kivitalo*, kun tarkoitetaan 'kivistä tehtyä taloa', *kivistä taloa*, mutta erilleen *totta kai*, vaikka ei tarkoiteta totta 'luultavasti totta' vaan haetaan synonyymia sanalle *ehdottomasti*?

Yhdyssanan tuntomeriksi ja erotukseksi sanaliitosta esitetään monasti sanapainoa, mutta suositukset päättyvät usein huomioon, ettei sanapainon tarkkailusta välttämättä ole suurtakaan apua.

Tärkeä tuntomerkki, joka erottaa yhdyssanan sanaliitosta, on pääpaino. Yhdyssanassa se on alkuosan ensimmäisellä tavulla, -- poikkeustapauksissa yhdyssanan jälkiosa tosin lausutaan alkuosaa painokkaammin. -- yhdyssana voi painottua poikkeavasti affektipitoisessa puheessa -- moniosaisissa rinnaisissa yhdyssanoissa kuin suomalais-ruotsalais-saksalainen (sanakirja) kaikki osat ovat yhtä painokkaita.
(Vesikansa 1989: 213)

Näin kuvailee pääpainoa yhdyssanan tuntomerkkinä Vesikansa (Vesikansa 1989: 213, lihavoinnit tässä ja seuraavassa omani). Ongelma on ilmeinen – yhdyssanassahan voi tuon mukaan olla paino melkeinpä missä kohden hyvänsä – ja se syvenee, kun Vesikansa jatkaa:

On kuitenkin paljon yhdyssanoihin verrattavia sanaryhmiä, jotka totunnaisesti, jo kirjoitetun kieleemme keventämiseksi, suositetaan kirjoitettavaksi erilleen, siis sanaliitoina, vaikka pääpaino niissä on ensimmäisellä sanalla.
(Vesikansa 1989: 213)

Samoin näkemyksensä pääpainon suhteen joutuvat kiepsauttamaan niin Ikola (Ikola 2001: 107)

Yhdyssanassa on vain yksi pääpaino (1. tavulla), kun taas sanaliitossa on pääpaino kummankin (tai kunkin) sanan 1. tavulla -- Painotuksen tarkkaaminen ei useinkaan auta ratkaisemaan, onko kyseessä yhdyssana vai sanaliitto, sillä horjuntatapauksissa painotuskin horjuu.

kuin Häkkinenkin (Häkkinen 1990: 147).

Nyky Suomessa yhdyssanan selvimpänä tuntomerkkinä voidaan pitää painotusta --. Yhdyssanassa pääpaino on alkuosalla, kun taas sanaliitossa molemmat osat voivat saada pääpainon. Käytännössä painotuskin horjuu, ja esim. sellaiset syntagmat kuin kielenopettaja -- voidaan yhtä hyvin tulkita sekä sanaliitoiksi että yhdyssanoiksi.

Olen opiskellut suomen kieltä useita vuosia, enkä vieläkään osaa erottaa, missä kohden sanaa paino on, vaikka pyydetyn laisen painon osaankin tuottaa. Arvelen tästä syystä, että on myös muita kielenkäyttäjiä, joille painon tunnistaminen tuottaa vaikeuksia. Arvelen lisää, että vaikkakin väärässä kohdassa olevan painon luultavasti intuitiivisesti huomaisikin vähintään jonkinlaisena omituisuutena, painon osoittaminen sanajoukosta saati sen käyttäminen yhdyssanan ja sanaliiton erottamiseen toisistaan saattaa olla lähinnä teoreettinen mahdollisuus monelle tavalliselle kielenkäyttäjälle. Kun vielä pitkän linjan lingvistit – Vesikansa, Ikola ja Häkkinen – edellä toteavat, että vaikka painon tunnistaisi, ei siitä ihan aukottomasti voi päätellä, onko kyseessä yhdyssana vai sanaliitto, ei ole ihme, että **tavallista** kielenkäyttäjää yhdyssanan olemus voi hämmentää. Häkkinen toteaa (Häkkinen 1990: 147) ja samalla muotoilee tämän tutkimuksen ytimen:

Puhekielessä ei ole edes tarpeen erottaa yhdyssanoja ja sanaliittoja toisistaan. Ongelma konkretisoituu vasta kun on ratkaistava, k i r j o i t e t a n k o kyseiset elementit yhteen vai erikseen.

Tässä kohden voi tietysti pohtia, että jos puhekielessäkään ei ole merkitystä sillä, onko kyseessä yhdyssana vai sanaliitto, miten merkityksen tuottamisen tarve muuttuu, jos sama asia kirjoitetaan. Tähän asiaan törmäämme myös luvussa 2.4, jossa sivuan Niemisen ja Lanteen näkemyksiä yhdysosan painosta sen tunnusmerkkinä (ks. myös Nieminen–Lantee 2008).

2.3 Yhdysosan kehäksi: käsite, termi, yleisyys ja merkitys

Yhdysosan syntyä – sitä mitkä ehdot täyttyvät, jotta sanapari ajautuu yhteen – motivoidaan yleisesti sillä, että sanapari muodostaa yhdessä termin, käsitteen, erikoismerkityksen ja otetaanpa sanaparin yhdessä esiintymisen yleisyyskin usein esiin (ks. esimerkiksi Iisa et al. Kielenhuollon käsikirja jatkossa KK: 138, Kielitoimiston oikeinkirjoitusopas toim. Kankaanpää et al. jatkossa KO: 252, Häkkinen 1990: 153, Ikola 2001: 106, Saukkonen 1973: 332, Vesikansa 1989: 216). Termimäisyys, käsitteellisyys ja erikoismerkitys toteutuvatkin toki monissa yhdysosissa, esimerkiksi nimissä ja kuvainnollisissa ilmaisuissa kuten *Joensuu* ja *joensuu*, mutta on myös sanatyyppejä, joissa tämä peruste on mielestäni kestävä.

Kun suomessa on sovittu, että tietyn muotokriteerin täyttävät sanat, esimerkiksi nominatiivimuodossa oleva substantiivimäärite³ ja yhdysosamuotoinen määrite, kirjoitetaan **aina** yhteen pääsanaan kanssa, ei mielestäni ole oikein mahdollista väittää, että kyseessä olisi joka kerran termi (MOT s.v. termi: jnk erikoisalan käsitteen vakiintunut nimitys, ammattisana), käsite (MOT s.v. käsite: ajattelun luoma abstrakti hahmo; esineelle t. asialle ominaisten piirteiden kokonaisuus) tai erikoismerkitys, joka ei ole suoranaisesti sanojensa summa vaan esimerkiksi kuvainnollinen merkitys (*sinisilmäinen*) tai jokin erityinen lajinnimi (*sinisiipi*). Erityisesti tilapäiset sanat muodostavat käsitteellisyysongelman: jos yhdyssana syntyy hetken käyttötarkoitusta varten ja katoaa maailmasta sen jälkeen, missä vaiheessa siitä muodostuu käsite saati termi? Ongelmallinen on myös tässä tilapäiskontekstissa sanaparin vakiintuneisuudenkin vaikutus yhdysosan syntyyn. Saukkosen mukaan (Saukkonen 1973: 332)

Tällainen [determinatiivinen eli määritys-] yhdyssana on periaatteessa lyhenne suuremmasta kokonaisuudesta. Lyhenne syntyy silloin, kun sille on riittävän suuri frekvenssi ja kun sen merkitys on yksiselitteinen tyypistetyistä muodosta huolimatta.

Jos kyseessä on tilapäinen yhdyssana, sen osat eivät ole voineet esiintyä yhdessä kovin yleisesti, eikä tilapäisellä yhdysosalla ole tulevaisuutta. Häkkinen (Häkkinen 1990: 151) ottaakin selitykseksi muotin: kun sanaparit esiintyvät riittävän usein yhdessä, ne muodostavat **yhdysosan, josta saadun**

3 Paitsi erisnimiin liittyvissä appositiotyyppisissä titteli-ilmauksissa kuten *merirosvo Mustakorva* ja numeroihin liittyvissä määritteissä kuten *Kylmäkoskentie 22*. Esimerkit tämän työn aineistosta. Ks. myös 4.3.2 Nominatiivimääritteiset yhdysosat.

muotin avulla muistakin sanapareista voidaan muodostaa yhdyssanoja ilman, että niiden syntyä tarvitsee joka kerran selittää sanaparin yhdessä esiintymisen yleisyydellä. Samaan tapaan referoi Mäkisalo (Mäkisalo 2000: 64) Ryderin ideaa analogian merkityksestä: ennustettavuudessa on kyse mahdollisuuksista ja tilastosta, eli prototyypisimmistä tulee normi ja – kääntäen – mitä marginaalimpi ilmaus on, sitä paremmin se muodostetaan analogian avulla. Tilapäissanathan ovat luonnollisesti aivan äärimarginaalissa: ne syntyvät hetken tarpeeseen ja häviävät sen jälkeen. (Jos ne eivät häviä, ne eivät ole enää tilapäisiä.) Mielestäni muotti, tai analogia, onkin usein huomattavasti parempi selitys yhdyssanan synnylle yleisyyden, käsitteen tai terminäisyyden sijaan, tai vähintään niiden lisäksi. *Kivitalo*-muotin avulla voimme hyvin muodostaa *lakritsiradio*-sanan miettimättä, onko tai voiko se missään yhteydessä olla käsite, termi tai edes mahdollinen, olemassa oleva ja johonkin viittaava sana.

Lingvisteillä näyttää olevan hiukan erimielisyyttä siitä, miten yhdyssanan merkitys muotoutuu. Saukkonen (Saukkonen 1973: 332) pitää tunnettuna tosiasiana, että determinatiivisen yhdyssanan merkitys ei ole sen osien summa; Häkkisen (Häkkinen 1990: 153) mukaan yhdyssanan merkitystä ei useinkaan voi päätellä automaattisesti sen osista ja merkitysten vaihtoehtoja voi olla useitakin – Vesikansa (Vesikansa 1989: 216) taas ilmoittaa, että useimmiten yhdyssanan merkitys on sen osien perusteella selvä, mutta lieventää kuitenkin huomauttamalla, että nimenomaan yleisimmässä yhdyssanatyypissä, alisteisissa (eli determinatiivisissa eli määritys-) yhdyssubstantiiveissa määriteosan merkitys on usein yleinen (versus viittaa määräyksilöön) ja että yhdyssanan alkuperäinen merkitys on saattanut unohtua ja muuttua kuvainnolliseksi.

Tampereen yliopiston suomen kielen luennolla syksyllä 2011 professori Pajunen antoi pro gradu -tutkielman tekijöille oivan neuvon huomauttamalla: "Lingvisti ei tulkitse." Hänen sanojaan ei käy väänteleminen: kielentutkijan on syytä pitäytyä tiukasti aineistossaan eikä alkaa miettiä mahdollisia piilotettuja tarkoituksia. Kuitenkin yhdyssanojen kohdalla tämä on vaikeaa, tai ehkä jopa mahdotonta lingvistille siinä kuin tavalliselle kielenkäyttäjällekin. Sanakirjan (MOT s.v. tulkita) *tulkita*-sanalle antama ensimmäinen merkitys on tässä yhteydessä relevantti:

Tulkita: 1. ottaa selville; ymmärtää, käsittää; valita jk monista selitysmahdollisuuksista; selvittää, tehdä ymmärrettäväksi.

Yhdyssanan merkitys on toisinaan läpinäkyvä, eli sen viittaaman asian olemuksen voi nähdä suoraan sen osista. On esimerkiksi kohtuullista olettaa tarkoitetta näkemättä, että jo mainittu perhosen lajinimi *sinisiipi* tarkoittaa sinisiipistä perhosta. Läpinäkymättömien ilmaisujen kohdalla taas joutuu väistämättä tekemään jonkinlaisen tulkinnan, arvion tilanteesta, ja muodostamaan merkityksen niillä tiedoilla, joita jo on. Palataan nyt aiempaan kuvitteelliseen, mutta nyt jo maailmassa valmiiksi

olevaan sanaan: *lakritsiradio*. Sana on tietenkin tilapäinen – eivät sen paremmin Google kuin MOT-sanakirjakaan tunne sanaa (ainakaan 23.2.2013) eli voimme varsin huoletta arvioida sen universaalliseksi frekvenssiksi tässä kohden 2 – ja se tuntuu ensisilmäyksellä käsittämättömältä yhdistymältä. Jos kuitenkin tiedämme, että makeistehdas valmistaa *vadelmaveneitä*, *hedelmämailoja* ja *lakritsiradioita*, sana muuttuu käsitettäväksi. Ennen tätä mielessämme saattoi kehkeytyä maailmantietomme perusteella useitakin ideoita, jotka nyt on hylättävä. *Lakritsiradio* ei siis olekaan lakritsisten radioaaltojen vastaanotin (vrt. *ularadio*), lakritsia koskevaa ohjelmaa lähettävä radio (vrt. *rockradio*), lakritsia voimalähteenään käyttävä radio (vrt. *patteriradio*) eikä lakritsista rakennettu radio (vrt. *putkiradio*) vaan radion muotoinen lakritsista tehty makeinen (ei *radio* ensinkään). Jos sanojen välisiä suhteita haluaa ymmärtää, on niitä väistämättä tutkittava **ja** tulkittava: yhdyssanan osia on punta-roitava keskenään ja pohdittava, miten ne liittyvät toisiinsa paitsi yleensä myös **erityisesti kontekstissa**. Tavallisestihan tässä ei ole ongelmaa: yleinen maailmantieto ohjaa ymmärrystä mahdollisiin merkityksiin ja poimii kontekstissaan todennäköisimmän ilman, että asialla tavallisesti tarvitsee paljon vaivata päätään, mutta joskus kuitenkin vastaanottaja saattaa hämmentyäkin. Saman on huomannut Häkkinenkin, sillä hänen mukaansa (Häkkinen 1990: 144) tarkkoja sääntöjä yhdyssanan muodostamisesta tai tulkinnasta ei voi antaa eikä osien tuntemus ilman kontekstia (tai hakuteoksia) aina auta ymmärtämään, mikä sanan sisältö on.

Valitsin äskeiseksi esimerkiksi tarkoituksella nominatiivisubstantiivilla määritetyn substantiivin, *lakritsiradio*, koska sellainen on aina – riippumatta kiteymän asteesta, yleisyydestä, termimäisyydestä tai käsitteellisyydestä – yhdyssana, ja sellaisena kuvaa hyvin sanan olemassa olon vahvuutta tai tässä tapauksessa pikemminkin heikkoutta, häilyvyyttä. Jos määriteosa kuitenkin on geneetiivissä tai jossain muussa muodossa (kuin aiemmin mainituissa nominatiivimuodossa tai yhdysosamuodossa), sanat voidaan kirjoittaa yhteen tai erilleen. Joskus on mahdollista sekä yhteen että erilleen kirjoitus, eli kirjoittajan on pysyttävä valitsemaan, kumpi on asiayhteydessä oikea. Silloin kirjoittaja joutuu kielenhuollon mukaan pohtimaan, onko kyseessä termi, onko sanapari (tarpeeksi) vaikiintunut käsite, onko sanapaino vain ensitavulla tai tarkistamaan kuuluuko se johonkin erikseen laadittuun sanalistaan (ks. esim. KK: 138, 139, 141, 144, 146, 148; KO: 252, 253, 257–265), jotta sen oikeinkirjoitusasu selviäisi.

Huomautettava on tietenkin, että kirjoittamisen todellisuus ei liene yhtä hankalaa kuin edellä olevasta voisi päätellä. Käytännössä kirjoittaja harvoin joutuneen tarkistamaan, kuinka mikäkin sana kirjoitetaan, muutenhan tekstin tuottaminen kävisi mahdottoman työlääksi. Monien ilmaisujen kirjoitus opitaan ulkoa, monet säännöt ovat painuneet intuitiiviseen kielioppiin ja otetaan sieltä tarpeeseen erikseen miettimättä – ja monien ilmaisujen kirjoitus menee mönkään.

2.4 Yhdyssanat viimeaikaisessa tutkimuksessa

Yhdyssanoja tavallisten kirjoittajien internetissä julkaistuista teksteistä on tutkinut Marja Ylikulju vuonna 2003 Oulun yliopistossa julkaistussa pro gradu -työssään Internetin seuranhakuilmoitusten yhdyssanat. Aineistona Ylikuljun työssä on siis internetissä julkaistut seuranhakuilmoitukset, ja vaikka kirjoittaja kommentoi aineistoaan toteamalla, että sen "kielenkäyttö on ilmeisen spontaania" (Ylikulju 2003: 2), mielestäni tuollaista johtopäätöstä on hyvin hankala vetää nimenomaan tuosta aineistosta. Kielenkäyttö voi toki olla epävirallista tai leikittelevää, mutta ajatus siitä, että seuranhakuilmoitus – pyrkimys antaa itsestä erityisen hyvä kuva – kirjoitettaisiin spontaanisti eli harkitsematta, on minusta melko outo. Ylikulju perustelee käsitystään sillä, että kirjoittaja ei ole kirjoittaessaan tiennyt olevansa kielenopas lingvistiselle tutkimukselle. Uskallan väittää, että seuranhakuilmoitusta laatiessa tuo olisi tutkimuksesta tietoisellekin kirjoittajalle joka tapauksessa sivuseikka; tärkeämpää olisi varmaan tuossakin tapauksessa tehdä vaikutus ennemminkin potentiaaliseen kumppaniin kuin kielentutkijaan. Kun hyvän kuvan antaminen perustuu kirjoitukseen, kirjoitus silloin varmasti hiotaan mahdollisimman sujuvaksi ja virheettömäksi ennen julkaisua. Oman aineistoni sen sijaan katson suurimmalta osalta lähenevän spontaaniudessaan puhekieltä; palaan tähän myöhemmin aineistoa esitellessäni. Ylikuljun pro gradu -työ on perusteellinen esitys yhdyssanoista, jotka hän on luokitellut aiempien tutkimusten (mm. Saukkonen ja Vesikansa) mallien mukaan. Työssä käydään huolellisesti läpi myös eri-ikäisten ja kummankin sukupuolen kirjoittajien kielellinen kompetenssi yhdyssanojen osalta. Ylikuljun työstä on kuitenkin aikaa tätä kirjoittaessa kymmenen vuotta, ja katson hänen aineistonsa eroavan omastani sekä spontaaniudessaan että lajityypissään. Eroa Ylikuljun työhön on siinäkin, että kun Ylikulju määrittelee yhdyssanat (myös genetiivialkuiset) tiukasti oikeisiin ja väärin, minä pyrin hiukan pidemmälle menevään analyysiin eli tarkastelemaan, millaisin ehdoin (erityisesti genetiivialkuiset) yhdyssanat muodostavat tai eivät muodosta yhtenäisiä kombinaatioita tai voisiko joillain yhdyssanoilla olla perusteltuja syitä kirjoittaa ne joko sanaliitoiksi tai yhdyssanoiksi: aineistossani esiintyy jonkin verran sanoja, jotka jo sanakirjankin mukaan on mahdollista kirjoittaa yhdeksi tai kahdeksi sanaksi. On kuitenkin selvää, että työmme ovat jonkin verran samankaltaisia sekä tutkimuskohteeltaan että aineiston osalta, eli ne kummatkin käsittelevät yhdyssanoja julkisuuteen kirjoitetussa tavallisten ihmisten kielessä ja siksi Ylikuljun työn kanssa syntyy keskustelua tässä työssä myöhemminkin.

Anna Lantee on tutkinut pitkiä yhdyssubstantiiveja Helsingin Sanomien korpukselta vuonna 2010 Tampereen yliopistossa julkaistussa pro gradu -työssään (Lantee 2010). Tämä hänen työnsä ei juurikaan liity omaani muutoin kuin tietenkin yhdyssanojen rakenteellisen kuvauksen osalta, jotka ovat samat ylipäätään yhdyssanoille. Sen sijaan hänen muutamia yhdyssanatutkimuksensa yhdessä

Tommi Niemisen kanssa herättävät mielenkiintoa erityisesti tämän työn kannalta, sillä he oikeastaan kyseenalaistavat koko yhdyssanan olemassaolon. Esitelmässään (Nieminen–Lantee a2010) Lantee ja Nieminen huomauttavat aiempiin tutkimuksiinsa yhdyssanojen painosta (Nieminen–Lantee 2008) vedoten, että koska on epävarmaa voidaanko puheessa tuottaa yhdyssanoja (tai sanaliittoja) painon avulla eli onko puheessa lainkaan yhdyssanaa (tai sanaliittoa), voidaan kysyä missä määrin yhdyssanojen oikeinkirjoittamistakaan voidaan arvioida. Kielitieteen päivillä 2010 pidetyssä esitelmässään (Nieminen–Lantee b2010) he päätyvät väittämään, että sääntö yhteen kirjoittamisesta on "ortografinen knoppi", jolla erotetaan sivistynyt kirjoittaja muista. Oma aineistoani kerran toisensa jälkeen käydessäni läpi ja aihepiiriä pohtiessani, olen päätenyt hyvinkin taipuvaiseksi heidän kannalleen, sillä aineistossani en juurikaan ole törmännyt tapauksiin, joissa erilleen kirjoittaminen olisi tuonut sellaisen merkityksen tekstiin, jota ei vähintään kontekstissa voisi pitää täysin itsestäänselvänä tai ainakin hyvin helposti pääteltävissä – ja toisaalta, kuten jo edellisessä luvussa pohdin, yhdyssanan tulkinta lienee useimmiten kontekstisidonnaista, eli kirjoitetaanpa sanapari yhteen tai erilleen, tulkittavaksi se tulee vasta asiayhteydessään. (Ks. esimerkiksi luku 4.3.4 Muut määritetyypit: *ennenkuulumaton*.)

Yhdyssanojen suhdetta kieleen erityisesti oppimiskontekstissa on tutkinut Sari Oinonen Tampereen yliopistossa tehdyssä pro gradussaan Rakenteellinen kompleksistuminen peruskoulun seitsemäs-, kahdeksas- ja yhdeksäsluokkalaisten kirjoitelmissa (Oinonen 2008). Hänen mukaansa yhdyssanojen osuus kasvaa iän karttuessa (Oinonen 2008: 41). Usein kuulee arveltavan, että englannin kieli ohjaisi suomalaista yhdyssanakäsitystä niin, että koska englannissa sanat useimmiten kirjoitetaan erilleen, tapa on siirtynyt myös suomeen (esimerkiksi Heikkinen–Mantila 2011: 131, Leino 2005: 185, Ikola 2001: 108). Tähän liittyen palaamme vielä Ylikuljun pro graduun (Ylikulju 2003: 101), jossa hän huomauttaa mielestäni erittäin osuvasti:

"Äidinkieli on kirjoittajan vahvin kieli, ja sen rakenteen ja sääntöjen oppiminen on tie muiden kielten oppimiselle. Kun kirjoittaja pohtii äidinkielen yhdyssanan kirjoitusasua, on siinä usko, että siihen vaikuttaa englannin kielen ortografia. Luulisi vaikutuksen olevan toisinpäin paljon suurempi."

Mielenkiintoinen tulos Oinosen tutkimuksessa etenkin englannin kielen ohjailuun viitaten onkin se, että kaksikielisen luokan oppilaiden kirjoitelmissa oli yhdyssanoja huomattavasti enemmän kuin muilla, verrokkiryhmien yksikielisillä, luokilla (Oinonen 2008: 40). Kaksikielisen luokan opetuskielenä suomen lisäksi oli englanti, eli ne oppilaat, jotka käyttivät nimenomaan englantia toisena kielenään, kirjoittivat eniten yhdyssanoja. Vaikka luokalle oli valittu oppilaat pääsykokein ja siten he luultavasti olivat keskimäärin kielellisesti lahjakkaampia kuin ikätoverinsa, tämä tulos mielestäni tukee Ylikuljun arvelua englannin kielen vaikutuksesta suomenkielisten kielikäsitteeseen: englannin

kielen mallia ei ehkä käykään osoittelemisen silloin, kun halutaan tietää, millä perusteilla suomenkieliset jättävät yhteen kuuluvat sanat yhdistämättä.

Suvi Levonen esittää Tampereen yliopistossa tehdyssä pro gradussaan Verbaaliyhdyssanat HS 2000–2001 -korpuksessa (Levonen 2006) perusteellisesti, miten yhdyssanat ja etenkin verbaaliyhdyssanat syntyvät. Kun Levosen aineisto kuitenkin, kuten aiemmin mainitun Lanteenkin, on Helsingin Sanomien korpus, työmmme kohtaavat lähinnä teorian tasolla – Levosen gradun ydinaihe, verbaaliyhdyssanat, on vain yksi tyyppi tämän työn yhdyssanoista, ja sellaisenakin hyvin marginaalinen. Levonen arvioikin, että verbaaliyhdyssanat ovat oletettavasti erityisesti kirjoitetun kielen piirre (Levonen 2006: 9); oma aineistoni näyttää variaationsa perusteella asettuvan jonnekin kirjoitetun ja puhekielen välimaastoon (ks. luvut 3.2 Facebook tekstialustana ja aineistolähteenä sekä 4.4 Yhteenveito kaikista yhdyssanoista), joten verbaaliyhdyssanojen osalta tämän tutkimuksen aineisto on liian suppea. Esiintymiä ei yksinkertaisesti ole tarpeeksi, jotta niistä voisi johtaa minkäänlaisia näkemyksiä vertailtavaksi. Palaan kuitenkin Levosen tutkielman aihepiiriin verbaaliyhdyssanoihin vielä luvussa 4.3.4 Muut määritetyypit.

Jukka Mäkisaloon väitöskirjan *Grammar and Experimental Evidence in Finnish Compounds* (Mäkisalo 2000) viidestä artikkelista kiinnostavin tämän työn kannalta on kirjan toinen artikkeli, *Compounds in Finnish: on the Morpho-syntactic and Semantic Criteria*, jossa Mäkisalo jaottelee pienen aineiston yhdyssanat ja tekee sen mukaan selkoa suomen yhdyssanojen piirteistä. Artikkelissaan Mäkisalo huomauttaa (Mäkisalo 2000: 64), että Suomessa yhdyssanatutkimus on keskittynyt standardikieleen, joka on ortografisesti normitettua, ja vain vähän arki- tai puhekieleen. Tämä tutkimus pyrkii nimenomaan vastaamaan tähän: aineistoni, kuten jatkossa (luku 3.2) osoitan, on enemmän arkista jutustelua kuin standardoitua kieltä. Samoin poikkeamien nostaminen osaksi työn aihepiiriä tuo esiin sen puolen yhdyssanoista, johon normitus ei ole pystynyt vaikuttamaan. Työssään Mäkisalo keskittyy kaksiosaisiin yhdyssanoihin; tässä työssä aineistoon on laskettu myös useampi-osaiset yhdyssanat. Mäkitalon aineistona on 172 yhdyssanaa, jotka hänen kielenoppaansa ovat rajanneet viivoin radiossa puhutusta litteroidusta tekstistä; tässä työssä on 7634 yhdyssanaa, jotka kielenoppaani ovat itse tuottaneet (aineiston täsmällisiin määriin palaan myöhemmin). Vaikka aineistomme ja tutkimusmenetelmämme ovat erilaiset, Mäkisaloon tutkimuksessa tämän työn kannalta kiintoisaa on hänen aineistonsa suhteelliset yhdyssanamäärät. Palaan näihin tässä työssä vielä jatkossa luvussa 3.4, kun esittelen yhteenvedonomaaisesti tämän tutkielman aineistoa.

Tähän työhön vertailun vuoksi liittyy vielä hiukan vanhempi tutkimus: vuonna 1984 tehty tilastoesitys suomen kielen rakenteiden määrästä (Pajunen–Palomäki 1984). Aineistona tässä tutkimuksessa on Lauseopin arkisto ja se käsittelee sekä puhetta että kirjoitettua kieltä. Vertailen oman

aineistoni määriä tämän tutkimuksen tuloksiin, joskaan siinäkään ei käsitellä kieltä poikkeamien näkökulmasta. Tästäkin vertailusta lisää luvussa 3.4 Korpuksesta tutkimusaineistoksi.

3 AINEISTO

3.1 Aineiston keräys

Kun halutaan tarkastella jokapäiväistä, tavallisten kielenkäyttäjien käyttämää spontaania, arkista kieltä, on luontevaa kerätä aineisto sieltä, missä tavalliset ihmiset muutenkin, pyytämättä, kirjoittaisivat sekä tekstiä etukäteen suunnittelematta että sitä jälkikäteen muokkaamatta. Tähän tarkoitukseen sosiaalinen media tarjoaa runsaasti aineistoa. Valitsin aineistonkeruupaikaksi Facebookin (www.facebook.com) siitä syystä, että siellä on paljon eri-ikäisiä, eri koulutustaustaisia ja ylipäättään erilaisia ihmisiä keskustelemassa jopa saman otsikon alla, esimerkiksi ryhmässä. Jokin muukin keskustelufoorumi olisi tietysti käynyt myös, mutta Facebookin etuna on kuitenkin nimenomaan juuri se, että samassa palvelussa on hyvä käyttäjien kirjo edustettuna, kun taas yksittäisille, yhteen asiaan keskittyville keskustelufoorumeille tapaa kerääntyä melko homogeeninen joukko.

Kaikki aineistoni kommentit on kerätty Facebookin ryhmäsivuilta⁴, eli keskustelu on sikäli yleisluontoista, että se liikkuu – tai ainakin oletuksenmukaisesti liikkuu – ryhmän aihepiirin ympärillä; käyttäjien henkilökohtaiset sivut (ja edelleen henkilökohtaiset asiat) eivät siis kuulu tämän tutkimuksen aineistoon. Kaikki aineistoni ryhmät ovat avoimia, eli niiden kirjoituksia voi lukea kuka tahansa paitsi liittymättä ryhmään myös jopa kirjautumatta Facebookiin. Ryhmät valitsin sitä silmällä pitäen, että joukkoon sopisi mahdollisimman laaja kerros Facebookin käyttäjiä: kiusaamista vastustavassa ja poliittisissa ryhmissä sekä Yle-, kahvi- ja koiraryhmissä keskustelevat pääasiassa aikuiset, mutta jääkiekko-, Cheek- ja Aku Ankka -ryhmissä kirjoittajat ovat voittopuolisesti nuoria, jopa lapsia. Kaikissa ryhmissä on kumpikin sukupuoli edustettuna. Anonyymeja tai siltä vaikuttavia kirjoittajia on huomattava vähemmistö, eikä anonyymikirjoittelu poikkea tyypiltään oletettavasti⁵ omalla nimellään kirjoittavien teksteistä.

Keräsin tutkimusaineiston kahteen otteeseen: muutaman päivän aikana helmikuussa sekä uusintakierroksen touko-, kesä- ja heinäkuun aikana vuonna 2011. Keräsin aineiston käyden järjestelmällisesti läpi valikoimani Facebook-ryhmät, joista kopioin talteen kaikki ne seinäkirjoitukset ja kommentit, jotka sillä hetkellä olivat automaattisesti ilman erityistoimenpiteitä näkyvissä kaikille. Tekstimateriaali on kopioitu suoraan, eli teksti on samanlainen kuin se ryhmässä **keräyshetkellä** oli lukuun ottamatta kolmea seikkaa: 1) kaikki linkit, sähköpostiosoitteet ja puhelinnumerot olen pois-

⁴ Luettelo ryhmistä liitteessä 2.

⁵ Ei tietenkään ole mitään keinoa tutkia, ovatko kaikki kirjoittajat sitä, mitä väittävät olevansa – kun teksti on asiallista ja henkilöitymätöntä, ei sillä ole edes merkitystä.

tanut, 2) keskenään täsmälleen samanlaisista teksteistä olen jättänyt aineistoon vain yhden ja 3) kaikki sydän-merkit olen muuttanut muotoon <3⁶.

Keräämäni tyypisessä internetaineistossa on huomattava eräs tärkeä seikka: tämänkaltainen tekstien kokoelma voi myöhemmin muuttua julkaisualustallaan toisenlaiseksi, kuin mitä se on keräyshetkellä ollut. Tämä johtuu siitä, että tekstejä voivat lisätä ja poistaa jälkeinpäin niin sivun ylläpitäjä kuin kirjoittajakin, ja jos Facebook-profiilinsa poistaa kokonaan, poistuvat myös kaikki kyseisellä profiililla kirjoitetut kommentit. Muutoksia keräämäni aineiston ja internetissä olevan materiaalin välillä voi aiheuttaa myös se, että jo tätä kirjoittaessa kommenttien muokkaus Facebookissa on mahdollista – ominaisuus tuli palveluun kesän 2012 aikana. Tosin aineistonkeruuajana tällaista mahdollisuutta ei ollut, eli tuskin vanhoja kommentteja ainakaan merkittävässä määrin on käyty muuttelemassa. Vielä tätä kirjoittaessakin muiden käyttäjien on mahdollista tarkastella kirjoittajan tekemiä muutoksia: originaalikirjoitus siis on vieläkin lukijoille saatavilla, mikäli sitä ei ole poistettu kokonaan. Kirjoittajat ovat myös voineet poistaa ja kirjoittaa kommenttinsa kokonaan uudestaan, mutta tällaisen korvauksen jälkeen ne eivät enää ole olleet loogisissa paikoissa kommenttiketjuissa. Todennäköistä on siis, ettei tällaistaakaan olisi kovin usein tapahtunut – ainakaan niin usein, että sillä olisi merkitystä tämän työn kannalta.

Joka tapauksessa jos keräämieni tekstien alkuperäisiä julkaisupaikkoja tarkastellaan jälkeinpäin, siellä olevien tekstien koosteet ovat hyvin todennäköisesti jonkin verran erilaisia kuin aineistoni. Keräämäni aineistoon onkin syytä suhtautua samoin kuin mihin tahansa jähmeään, kehykseltään irrotettuun korpukseen: kerran kerätty on käsillä oleva aineisto, ja keräyksen jälkeen julkaisualustalle mahdollisesti tulleet muutokset elävät omaa elämäänsä toisaalla. Julkaisualustan ja korpuksen välillä ei ole yhteyttä eikä yhtäläisyyttä enää keräyshetken jälkeen. Tätä voisi verrata vaikka siihen, että tehdään tutkimus kirjasta, josta myöhemmin otetaan uusi, korjattu painos. Originaalipainoksesta tehty tutkimusta se ei kuitenkaan muuta.

3.2 Facebook tekstialustana ja aineistolähteenä

Sosiaalinen media ei enää tätä kirjoittaessa ole niin uusi asia, että sen perinpohjainen esittely olisi tässä yhteydessä tarpeen. Sanakirjamäärittely (Kielitoimiston sanakirja, myöhemmin MOT, s.v. sosiaalinen) kertoo aiheesta napakasti olennaisen:

Sosiaalinen media tietoverkkoja ja tietotekniikkaa hyödyntävä viestinnän muoto, jossa käsitellään vuorovaikutteisesti tuotettua sisältöä sekä luodaan ja ylläpidetään verkostoja, yhteisöllinen media.

⁶ Facebook konvertoi automaattisesti <3-merkkiyhdistelmästä ♥-merkin. Kirjoittajan näppäimistöltä lähtevä merkkiesitys on kuitenkin tässä yhteydessä riittävä ja asiallinen siksikin, että välttämättä ♥-merkki ei näy oikein jatkossa aineistoa käsitellessä.

Sosiaalinen media on siis netissä olevaa palvelua, johon – ainakin periaatteen tasolla – kuka tahansa voi liittyä ja jossa kuka tahansa voi osallistua keskusteluun. Mediumista riippuen kirjoitustyylit voivat vaihdella: siinä missä yhdessä kirjoittelu on virallisempaa ja asiapitoisempaa, voi se toisessa olla hyvinkin epämuodollista. Kuvailen tässä tutkimuslähteeni Facebookin kirjoitusten piirteitä siinä määrin, kuin aineistoni puitteissa olen niitä havainnoinut.

Facebookissa keskustelu on usein tunteikasta ja kiivastakin: tekstit vilisevät puhekielisyyksiä sekä erilaisia merkkejä, jotka ilmaisevat tunnetiloja: :-) :-D :-(<3 ja niin edelleen. Kirjoitukset seuraavat toisiaan nopeasti, mikä ilmenee keskustelun puhemaisuutena: virkkeet eivät useinkaan ole viimeisteltyjä ja puheenvuorot ovat samankaltaisia kuin internetin ulkopuolisessakin elämässä, eli jatkuvat elliptisinä vastauksina toisten teksteihin, kuten esimerkin 13 kommentti esimerkin 12 kommenttiin.

(12) *Vasabladet kysyy mitä puoluetta äänestät. Käykääpä äänestämässä!*

(13) *Käyty!*

Kyseessä on siis aineisto, joka on useimmiten ilmeisen spontaania, hiomatonta tekstiä, ja käy jatkuvaa sisäistä vuoropuhelua niin, että ainakin yleisesti ottaen eri kirjoittajien teksti on koherenttia keskenään ja sekä morfologisessa (näkyä esimerkiksi elliptisinä ilmauksina kuten esimerkeissä 12 ja 13) että asiayhteydessä toisiinsa (puhutaan yhdessä samasta aiheesta).

Facebookissa, ainakin näissä keräämissäni ryhmissä, käytettävä kieli varioi huomattavasti aivan konvention mukaisesta kirjakielisestä ja huolitellusta tekstistä (14) aina alatyylisiin ryöpsäykseen (15) ja varsinaista kirjoitusta sisältämättömiin kuvaaviin merkkijonoihin (16):

(14) *Vihreä ympäristövaliokunnan jäsen Timo Juurikkala kummastelee Kokoomuksen ympäristöohjelmaa. Esimerkiksi ilmastonmuutos sekä luonnon monimuotoisuuden suojeleminen ovat Kokoomukselta unohtuneet ihan kokonaan. "Kokoomus menee euro edellä", Juurikkala kommentoi.*

(15) *Vittu! Porvarit Nurin! Kakku käy liian pieneks Ahnaille! Sitäpaitti Olkiluodosta tuli helvetin kallis Skeitti-halli nuorisolle, ko, se ei tuu jahaan sähköä ninku, sit ikinä.... Ja Mä toistan... ikinä!*

(16) <3:)

Kommentointi keskittyy yleensä itse asiaan, ja vaikka kirjoittajat voivat saada hyvinkin kärkkäitä kommentteja, hyvin harvakseltaan ne koskevat kirjoitustapaa. Alatyylisiä ilmaisuja arvostelevia kommentteja en havainnut aineistossani ainuttakaan, ja vaikka yleensä kielioppivirheet – myös yhdyssanapoikkeamat – sallitaan huomauttelematta, panin merkille kahdesta syystä johtuvat kirjoitustavan kommentoinnit. Ensimmäinen on pyrkimys kohentaa tekstin ymmärrettävyyttä. Kahteen sa-

man kirjoittajan kirjoittamaan esimerkkiin (17 ja 18) jatkaa kolmannessa (19) toinen kirjoittaja neuvomalla ilmeisen hyväntahtoisesti hiukan selkeämpään ilmaisuun:

- (17) *Juolahtipa mieleeni; kuinka käy mielenterv. ong., joka toisinaan anoreks., toisinaan bulim.?*
- (18) *ei hyvin... Jos kerran ns. norm. ihmisenkin dieetti-pudotus ja taas painon nousu ovat epäterv., niin...*
- (19) *(Pahoittelen aiheeseen liittymätöntä kommenttia, mutta vinkkinä täytyy sanoa, kun tulin kiinnittäneeksi huomiota asiaan: Kannattaa yleensä selkeyden vuoksi kirjoittaa täällä kommentoidessa kokonaisina sanoja, ellei ole jotain erityistä tarvetta käyttää lyhenteitä. Muut ymmärtävät silloin paremmin, mitä tarkoitat ja on helpompi kommentoida ja keskustella. Toivottavasti tästä olisi apua keskustelun virittämiseen jatkossa, tarkoitin siis tosiaan ihan vaan hyvänä käytännön neuvona. Aihe on kuitenkin kiinnostava ja herättää varmasti ajatuksia monella)*

Vaikka varsinaisesti neuvomisesta ei ole esimerkissä 21 kyse, kriittinen palaute on kuitenkin sävyltään vielä kohtalaisen neutraali ja siitä näkyy pyrkimys ymmärtää – kieltämättä aika sekavaa – kirjoitusta (20).

- (20) *jep terveisiä hallintokoneistolle ja perustuslailliselle oikeudelle sananvapaudesta ja sosiaalisen median sensuroinnista ja SOINILLE JA PERUSSUOMALISILLE MINÄ JA MONI MUU HALUTAAN ITETÄÄ HYVISSÄ AJOIN AIOTTEKO PALAUTTAA LAIN JA JÄRJESTYKSEN MAAHAN TIEDÄTTEKÖ MITEN JA KUINKA ISOSTA TYÖSTÄ JO NYT ON KYSE JA ETTÄ SE RATKEAA KUTEN OLEN EDELLÄ ESITTÄNYT.*
- (21) *Täytyy antaa kyllä palautetta tuosta [x]:n⁷ kirjoitustaidosta, sori vaan. Eihän sitä jaksu lukea, kun ei siitä ymmärrä mitään! Liian paljon kirjoitusvirheitä. On se kumma kun puhutaan suomalaisuuden ja Suomen puolesta, mutta ei osata kirjoittaa edes yhdyssanoja, saatikka sitten sujuvaa kieltä, jotta siitä saisi jotain tolkkua.*

Huomattavaa on, että kommentoija (esimerkissä 21) pitää yhdyssanojen osaamista aivan kirjoittamisen perustasona ja vasta sen jälkeen hänen mielestään tulee yleisesti sujuvan kielen hallinta. Seuraava kommentoija (esimerkissä 22) on samaa mieltä palautteesta, mutta onnistuu jo itse rikkomaan pariakin kielenhuoltosääntöä, nimenomaan juuri yhdyssanojen osalta (alleiviivatut):

- (22) *[Edellisen kommentoijan] kanssa kyllä samaa mieltä tuosta [x]:n kirjoituksesta, alkukommentit aika vaikea selkoisia, loppuakohti sentään hieman ymmärrettävämpää :)*

Näissä ensimmäisissä esimerkeissä kyse on siis tekstin ymmärrettävyydestä ja ymmärrettävyyden kohentamisesta. Seuraavassakin esimerkissä (23) kyllä neuvotaan kirjoittajaa, mutta sävy on huomattavasti tiukempi: anteeksipyyntöä tai selittelyä kommentissa ei ole.

- (23) *Arvon [X Y]. Suomi, kun siihen viitataan valtiona, kirjoitetaan isolla kirjaimella, kuten kaltaisesi isäm maan ystävän luulisi tietävän. Myös lauseet aloitetaan isolla kirjaimella ja ne loppuvat pisteeseen, huutomerkkiin tai kysymysmerkkiin.*

7 Olen poistanut tästä tutkielmasta järjestelmällisesti kaikki kirjoittajiin liittyvät henkilötiedot ja milloin henkilöihin on tarpeen viitata, on viittaus merkitty hakasulkein [-]. Originaalikorpuksen teksteissä henkilötiedot näkyvät, ja luonnollisesti itse kirjoituspaikalla, julkisesti Facebookissa.

Koko asian tai asian esittäjän henkilöön kohdistuva vähättelevä vähättelemällä käytettyä kieltä onkin toinen huomaamani syy kieliasun kommentointiin. Kommentissa näkyy iva paradisisa ilmaisussa *kaltaisesi isäm maan ystävän*. Sama kirjoittaja jatkaa toisessa kommentissa (24):

(24) *EI ULKO MAALAISIA SUOMEEN KOSKA HE EIVÄT TAVANNEET TORILLA*

Tämän kommentin parodia avautuu ehkä vain niille, jotka ovat seuranneet erityisesti ulkomaalaisvastaisiin kohdistuvaa pilkkaa. Usein sen keinona on paitsi tekstin kirjoittaminen kokonaan versaalilla myös yhdyssanojen osien kirjoittaminen toisistaan erilleen, esimerkkinä vaikkapa Facebook-ryhmä ULKO MAALAISTEN TULI VUORTEN PÖLY PILVET TAKAISIN ULKO MAILLE! (IV:FBa), jonka ryhmäkuvaus selventää asiaa lisää:

TULI VUORTEN SAASTE PILVIEN PÄÄSY SUOMEN ILMA TILAA KIELLETTÄVÄ! MISSÄ ENÄÄ SEILAA KOTI MAISET POUTA PILVET KUN ILMA TILASSA ULKO MAALAISET LAAVA PÖLY PILVET VIE ELIN TILAN! JA VETERAANEILLE ILMAISET LENTO TUKI KOHDAT!

Yhteenvetona Facebookin kirjoittamattomista kirjoitussäännöistä voi siis sanoa, että vaikka alatyylisiä ilmaisia ja kielioppivirheitä yleisesti ottaen katsotaan läpi sormien, ainakin kahdesta syystä kirjoittajan kieleen kuitenkin voidaan puuttua: 1) halutaan ymmärtää kirjoittajaa ja siksi halutaan selkeämpää kieltä ja 2) halutaan vähätellä asiaa vähättelemällä siitä kirjoittavan kieltä. Jos siis kirjoittaa tekstiä, josta saa selvää eikä esitä kovin kärjekkäitä mielipiteitä, saa melko rauhassa kirjoittaa millaista kieltä haluaa tai osaa – suuria vaatimuksia kirjoitustaidolle tai paineita oikeinkirjoituksen tarkistamiseen ei Facebookissa siis nähdäkseni juuri ole.

3.3 Aineistoa kuvaavia lukuja

Korpuksen kertyi kaikkiaan yhteensä 5030 tekstikatkelmaa, eli Facebookin seinäkirjoitusta tai kommenttia⁸. Myöhemmin kutsun näitä teksteiksi tai kirjoituksiksi, koska tämän tutkimuksen kannalta ei ole merkitystä sillä, kumpaa tyyppiä ne ovat. Laskentatavasta riippuen koko korpuksen yhteenlaskettu sanamäärä on 80 617–80 970; tekstinkäsittelyohjelma (Libre Office 4.0.0.3) ilmoittaa sanamääräksi pienemmän kuin koko tekstimassa jaettuna sananvälien mukaan sanoiksi. Ero johtuu käsittelytavasta: tekstinkäsittelyohjelma laskee sanat paitsi välilyöntien myös joidenkin välimerkkien (esimerkiksi ?, ! ja ,) mukaan, mutta ei kuitenkaan huomioi kaikkia merkkijonoja sanoiksi (:-) ei ole ja <3 on sana tekstinkäsittelyn mukaan). Lopulliseen käsiteltävään aineistoon jää 77 697 sanaa (palaan aineiston rajaamiseen seuraavassa alaluvussa 3.4).

⁸ Seinäkirjoitus = Facebook-sivulla, eli seinällä, julkaistu teksti, kuva tai linkki. Tässä yhteydessä vain tekstit on huomioitu. Kommentti = Facebook-julkaisuun (seinäkirjoituksen) kirjoitettu kommentti, ts. ei itsenäinen julkaisu.

Yksittäisen tekstin, yhden kirjoittajan jättämän kommentin tai seinäkirjoituksen, keskimääräinen laskennallinen sanamäärä on 15,5. Sanamäärät vaihtelevat kuitenkin paljon, välillä 0–313, joten laskettu keskiarvo ei kuvaa hyvin (jos lainkaan) tyypillisen kirjoituksen mittaa. Mediaani kertoo hiukan paremmin yksittäisen tekstin sanamäärästä, ja se on 8, reilusti vähemmän kuin keskiarvo. Mediaania kuitenkin laskee 68 sellaista kommenttia, jossa ei ole sanoja lainkaan (ne ovat tyypillisesti hymiöitä tai sydämiä); kun ne jätetään laskuista, tekstipituuden mediaani nousee 9:ään.

Koko korpuksen uniikkien kirjoittajien, tämän tutkimuksen kielenoppaiden, määrä on 2693. Keskimäärin kielenoppaan kommentteja on aineistossa 1,9, vaihteluväli on 1–77 kommenttia. Mediaaniksi jää 1, eli huomattava osa kirjoittajista (1924) on jättänyt aineistoon vain yhden kommentin. Tämä on tutkimuksen kannalta erittäin hyvä asia, sillä näin aineistoon ei juurikaan pääse kertymään yksilönkielen piirteitä, vaan tekstimassa koskee aivan selkeästi suurinta osaa kirjoittajista: yksi ja kaksi kertaa kommentoineita on yhteensä 2392 eli 86,5 %, viisi tai sitä vähemmän kirjoittaneita on jo 2586 eli 96 % kaikista kielenoppaista. Lisäksi tyypillisesti ne nimet tai oikeammin nimimerkit, jotka ovat jättäneet eniten tekstejä, ovat sivustojen ylläpitäjiä (esimerkiksi eniten kirjoittanut YLE Areena, 77 kommenttia), joten on odotuksenmukaista, että niitä on kirjoittanut usean henkilön nettitiimi, ja tämäkin seikka siis vähentää yhden yksittäisen kirjoittajan tekstimäärää. En havaitse aineistossani korrelaatiota kommenttien lukumäärän ja sanamäärän välillä, toisin sanoen on mahdoton arvioida näissä puitteissa, tuottavatko usein kirjoittavat pidempiä kirjoituksia kuin harvoin kirjoittavat: kirjoittajan tuottamien kaikkien sanojen keskiarvo (sanamäärä / kommenttien määrä = keskiarvo) vaihtelee välillä 1 ja 253, eikä sillä näytä olevan riippuvuussuhdetta kommenttien määrän kanssa.

3.4 Korpuksesta tutkimusaineistoksi

Kun aloin työstää keräämääni korpusta tähän työhän sopivaksi aineistoksi, poistin ensin hymiöt ja sydämet sekä täysin vieraskieliset kommentit, eli ne joissa ei ole lainkaan suomalaisia sanoja. Poistin myös sellaiset irralliset merkit, jotka eivät suoranaisesti liity mihinkään sanaan, eli käytännössä ajatusviivat, ajatusviivan asemaiset yhdysviivat ja erilaiset hymiötyyppiset ilmaisut (esimerkiksi \o/ tai O_o). Aineistossa on tosin hymiöitä jäljelläkin: jos hymiö on kiinteästi kiinni sanassa, se on teknisistä syistä jäänyt aineistoon, koska se ei vaikuta sanamääriin eikä tilastointiin muutenkaan. Lopulta yhdistin eri sanoiksi jakautuneet yhdyssanat yhdeksi sanaksi. Tällaisia sanoja ovat esimerkiksi

ne yhdyssanat, joiden määrittäjänä on sanaliitto ja jotka sekä tekstinkäsittelyohjelma että oma laskentaohjelmani⁹ laskee erillisiksi sanoiksi niissä olevien sananvälien vuoksi (esimerkki 25).

(25) *Kaunotar ja kulkuri-lehdestä*

Samaksi sanaksi laskin myös ne yhdyssanat, jotka on kirjoitettu kielenhuollon sääntöjen vastaisesti erilleen toisistaan (esimerkki 26).

(26) *hame rahat.*

Elliptiset yhdyssanat – sellaiset yhdyssanaparit, joilla on sama määrite- tai perusosa, joka on jätetty toistamatta – laskin osiensa mukaan erillisiksi sanoiksi. Toisin sanoen tyypistä *eurooppa-ja maahanmuuttoministeriksi* tulee kaksi erillistä yhdyssanaa (esimerkit 27 ja 28). Toistamatta jääneen osan olen liittänyt mukaan sulkeissa (kuten esimerkissä 27 (*ministeriksi*)).

(27) *eurooppa-ja (ministeriksi)*

(28) *maahanmuuttoministeriksi.*

Huomattavaa on, että yhdyssanan sisällä saattaa olla yksi tai useampikin yhdyssana. Tässä työssä kuitenkin **yhdyssanaksi on laskettu ainoastaan viimeisin yhdistynyt muoto**, eli sanansisäisiä yhdyssanoja ei ole erikseen laskettu. Esimerkkien (29 ja 30) sanajoukoista jää laskuihin siis vain yhdet sanat.

(29) *Peruuttamalla ei pääse eteenpäin -äkkiryhmän*

(30) *valtiovarainministeripuolue*

Luvussa 4.1 Yhdyssanat määritysosien määrän mukaan käsittelen kuitenkin laajemmin näitä moniosaisia yhdyssanoja.

Kielenhuollon konventioiden vastaisesti yhteen kirjoitetut sanaliitot olen laskenut yhdeksi sanaksi (esimerkit 31 ja 32).

(31) *20-asteen*

(32) *niinkuin*

Näin järjesteltyyn aineistoon jäi 77 697 sanaa, numeroa tai niihin suoranaisesti liittyvää merkkiä (esimerkiksi %-merkki).

Perustelut sille, miksi jotkin ilmaisut ovat konvention mukaisia ja jotkin sen vastaisia, esitän kutakin sanantyyppiä käsittelevässä luvussa erikseen. Samoin selvitän myöhemmin, luvussa 3.5 Yh-

⁹ Tekstin analyysi on tehty käsityönä, mutta tekstimassan fyysisesti käsittelemiseksi ja laskemiseksi sain puolisoltoni Esa Ristilältä korvaamatonta apua. Ilman hänen koodaamiaaan skriptejä tekstin jaottelu ja laskeminen olisi ollut toivottoman työlästä ja kestänyt moninkertaisen ajan. Kiitos, Esa!

dyssanan kriteerit – aineiston rajausta, millaiset sanat ylipäättään olen laskenut yhdyssanaksi ja millä perusteilla. Tässä totean kokoavasti, että koko aineistossa:

- **Yhdyssanoja on 7634 kappaletta eli 9,8 % koko sanamäärästä.**
- **Poikkeavia yhdyssanoja on 431 kappaletta eli 5,6 % koko yhdyssanamäärästä.**
- **Poikkeavia sanaliittoja on 334 kappaletta.**

Huomattavaa on jo tässä se, että konvention mukaisten sanaliittojen laskeminen on mahdotonta, koska – kuten Häkkinenkin huomauttaa (Häkkinen 1990: 147) – sanaliitto voi olla mikä tahansa kahden tai useamman sanan syntaktinen kokonaisuus. Poikkeavien yhdyssanojen ja sanaliittojen suhde on noin 56 % : 44 %, kun kaikki poikkeamat yhteensä on 100 % (765 kappaletta). Yhdyssanapoikkeamia on siis hiukan enemmän, mutta mitenkään drastinen ero ei ole; sanat näyttävät ajautuvan konventiosta poiketen yhteen lähes yhtä usein kuin erilleenkin.

- **Sanoja, jotka voivat olla sekä yhdyssanoja että sanaliittoja, on 114 kappaletta eli 1,5 % kaikkien yhdyssanojen määrästä.**

Aineistossa on kahdeksan sellaista ilmaisua, jotka ovat sekä yhdyssana- että sanaliittopoikkeamia, joten kahdeksan esiintymää lasketaan päällekkäin sekä poikkeaviin yhdyssanoihin että poikkeaviin sanaliittoihin, kuten esimerkki 33.

(33) *Kansanmuisti sivuilta*

Poikkeama on sanaliitossa siksi, että *Kansan muisti* olisi pitänyt erottaa sanaliitoksi, ja yhdyssanas- sa siksi, että *sivuilta*-perusosa olisi pitänyt sitoa yhdysviivalla määräiteosaan; esimerkin 33 konventionaalinen muoto olisi *Kansan muisti -sivuilta*.

Nämä määrät tarkentuvat jatkossa pääluvussa 4 Aineiston käsittely ja sen alaluvuissa. Joka tapauksessa jo tässä vaiheessa voinee todeta, että melko marginaalisen kielen alueen kanssa ollaan tekemisissä, joskin toisaalta kielessä laskennallisesti varmaan yksi jos toinenkin ilmiönä tutkittava alue on kokonaisuudesta melko kapea.

ISK:n mukaan yhdyssanojen sanakirjafrekvenssi on noin 60–70 %, ja vaikka perussanat sen mukaan ovatkin käytössä kompleksisia sanoja yleisempiä (ISK § 146), on ero aineistoni yhdyssana- frekvenssiin (9,8 %) melkoinen. On tietenkin odotuksenmukaista, että mitä kompleksisempi sana on, sitä harvemmin se esiintyy. Esimerkiksi Suomen kielen taajuussanaston (Saukkonen et al. 1979: 41) mukaan kymmenen yleisintä sanaa ovat *olla, ja, se, ei, joka, että, tämä, hän, voida ja saada*; ensimmäinen kolmitavuinen sana on sijalla 37 oleva *kuitenkin*. Ensimmäinen yhdyssana –

maailma – on taajuussanaston yleisyysjärjestyksessä vasta sijalla 145 (Saukkonen et al. 1979: 43). Vertailun vuoksi laskin yhdyssanojen määrän CSC - Tieteen tietotekniikan keskus Oy:n julkaisemasta sanomalehtikielen taajuussanastosta (IV:SLTS). Käytin laskennassa samoja kriteereitä, joita omankin aineistoni suhteen olen noudattanut, ja tällä tavalla sain sanomalehtikielen taajuussanaston yhdyssanojen määräksi 7,18 %. Oletan aineistoni lähenevän puhekieltä ja olevan siksi vähemmän kompleksista kuin yhtenäiseksi, esimerkiksi sanomalehtijutuksi, kirjoitettu teksti. Tulos sanomalehtikielen yhdyssanafrekvenssistä olisikin kovin hämmäntävä, jos ei huomattaisi sitä seikkaa, että sanomalehtikielen **taajuussanastossa** on kyseessä 9996 **yleisintä** lemmaa, eli perusmuotoistettua (kanta)sanaa, joten joukosta ovat karsiutuneet kaikki tilapäissanat ja harvinaiset sanat. Taajuussanasto ei siis luonnollisestikaan sovi vertailuaineistoksi silloin, kun vertaillaan, paljonko missäkin tekstissä **ylipäättään** on yhdyssanoja – näin siitä huolimatta, että sanomalehtikielen taajuussanastossa on ilmoitettu sanojen frekvenssit. Luvussa 2.4 sivutun Mäkisalon väitöksen aineistossa yhdyssanojen määrä on 11,2 % (Mäkisalo 2000: 75). Jos yhdyssanojen määrästä vetää johtopäätöksen kielen kompleksisuudesta, on radiopuhe hiukan kompleksisempaa kuin tämän työn aineistona oleva internetteksti. Mäkisalo tosin koostaa aineistonsa hiukan eri tavalla, eli siitä puuttuvat sekä tilapäiset sanat että yhdyssanat, joissa on useampi kuin kaksi osaa – luultavasti hänen koko aineistonsa yhdyssanaprosentti on hiukan vinoutunut. Tilastotietoja suomen kielen rakenteesta 1 -tutkimuksen mukaan yhdyssanoja on puheessa 3,1 % ja kirjoitetussa kielessä 14,6 % (Pajunen–Palomäki 1984: 64). Tämän työn aineisto sijoittuu näiden välille: tässä on yhdyssanoja 6,7 %-yksikköä enemmän kuin puhutussa kielessä ja 4,8 %-yksikköä vähemmän kuin (painetussa) kirjoitetussa kielessä¹⁰.

Poikkeamien määrä aineistoni yhdyssanoissa on siis vain 5,6 %, mikä eroaa aika paljon Ylikuljun tuloksista yhdyssanavirheiden määrästä (Ylikulju 2003, tiivistelmä: "Aineistoni yhdyssanoista väärin erikseen on kirjoitettu 165 (10,7 %)".) Tähän eroon vaikuttanee ainakin osittain ne perusteet, joilla päätetään, mikä on poikkeama ja mikä ei. Ylikulju ilmoittaa laskeneensa genetiivimääritteisen sanan virheelliseksi, jos se on kirjoitettu toisin kuin sanakirjassa, paitsi jos lauseyhteys nimenomaan vaatii kirjoittamaan toisin. Hän ei kuitenkaan täsmennä, millainen sellainen lauseyhteys voisi olla. (Ylikulju 2003: 28). Minä olen lähtenyt tässä työssä siitä perusajatuksesta, että niiden yhdyssanojen kohdalla, joiden oikeinkirjoituksesta ei ole poikkeuksetonta sääntöä, en voi päättää kirjoittajan puolesta, onko kyseessä konventionaalinen muoto. Jos siis kirjoittaja on päättänyt jakaa sanan osiin, vaikka se on kirjoitettu sanakirjassa yhdyssanaksi, mutta mikään kielenhuollon sääntö ei jakamista suoranaisesti estä, en ole katsonut kyseessä olevan virheen. Poikkeamien määrän eroon tämän ja Ylikuljun työn välillä vaikuttaa varmasti myöskin se, että Ylikulju on jättänyt

¹⁰ Tilastotietoja suomen kielen rakenteesta 1 -tutkimuksen kirjakielinen aineisto koostuu sanoma- ja aikakauslehtien uutisista ja reportaaseista sekä tietokirjallisuudesta. (Pajunen–Palomäki 1984: 16)

työstään pois muutamia ilmeisen frekventtejä ja usein (tai yleensä) oikein kirjoitettuja yhdyssanoja. Palaan tähän vielä luvussa 4.3.2. Nominatiivimääritteiset yhdyssanat. On mahdollista, että myös laajempi aineistoni (tässä 7634 yhdyssanaa versus Ylikuljun 1542 yhdyssanaa) vaikuttaa asiaan.

Kielenoppaita tutkielmassani on siis 2693. Kirjoittajia, jotka kirjoittavat poikkeavasti yhdyssanat tai sanaliitot, on yhteensä 404. Heistä vain sanaliiton on kirjoittanut poikkeavasti 148, ainoastaan yhdyssanapoikkeaman on tehnyt 194 ja sekä sanaliitoissa että yhdyssanoissa poikkeamia tehneitä on 62. Yhdyssanan poikkeavasti kirjoittaneita on siis yhteensä 256, ja kun yhdyssanan ylipäättään kirjoittaneita on kaikkiaan 1358, voidaan laskea, että 18,9 % yhdyssanan kirjoittaneista on rikkonut konventiota – melkein viidennes. Kuten aiemmin on todettu, kaikkien sanaliittojen määrän laskeminen on mahdotonta, joten konventioista poikenneiden suhteen on tyydyttävä toteamaan vain yksinkertaisesti määrä, koska suhteellista näkökulmaa ei tässä voi tavoittaa.

3.5 Yhdyssanan kriteerit – aineiston rajaus

Kuvailin jo luvussa 2 Yhdyssanat ja sanaliitot niitä ongelmia, jotka liittyvät yhdyssanan ja sanaliiton välisiin suhteisiin. Rajankäyntiä on myös yhdyssanan ja johdetun sanan välillä (ks. esim. Koski 1982: 66–73). Vaikka kielenkäyttäjälle ei johtamisen ja yhdyssanan välisissä suhteissa ehkä olekaan juuri pohtimista – aineiston käsittelyn jälkeinen silmämääräinen arvioni on, että johtamisessa ei juuri tehdä poikkeamia kuin lähinnä johdosten liittämisisä lyhenteisiin¹¹ – tämän tutkielman vuoksi aihepiiriä on kuitenkin sivuttava, kun selvitän, millä perusteella olen jakanut sanat yhtäältä laskettaviin ja analysoitaviin yhdyssanoihin (sekä poikkeaviin sanaliittoihin) ja toisaalta analyysissa sivuutettaviin, tähän työhön kuulumattomiin johdettuihin sanoihin. Lisäksi osa perussanoista, etenkin lainasanoista, voi kuulostaa yhdyssanalta, vaikka ne eivät sitä ole. Näitä sanoja ei ole luonnollisesti laskettu yhdyssanoiksi; samoin on täysin vieraskielisten sanojen suhteen. Käyn seuraavassa läpi näitä ongelmallisia tyypejä ja selvitän, millaisia ratkaisuja niiden suhteen olen aineistossa tehnyt.

3.5.1 Prefiksit ja muut epäitsenäiset määritteet

Yhdyssanan määriteosana voi olla morfeemi, joka ei esiinny itsenäisenä (Häkkinen 1990: 154). Epäitsenäisiä morfeemeja voivat olla esimerkiksi prefiksit, sanaan liitettävät etuliitteet. Suomessa erityisen merkitysryhmän muodostavat kieltoprefiksit *epä-* ja *ei-*, joista *epä-* ei esiinny yksinään (ISK § 172). Sanakirja MOT (s.v. epä-) ilmoittaa *epä-*:n olevan yhdyssanan alkuosa ja ISK listaa *epä-*:n epäitsenäisten yhdysosien joukkoon (ISK § 406). Kun lisäksi *epä-* voi olla kantana (muille)

¹¹ Esimerkiksi *SDP-läiset*. Esimerkki on tämän tutkielman aineistosta.

johtimille ja on sillä tavalla luonteeltaan itsenäisen morfeemin kaltainen (Häkkinen 1990: 106), lasken tässä työssä kieltoprefiksillä *epä-* ja itsenäisyyden perusteella myös *ei-* alkavat sanat yhdyssanoiksi. *Epä-* liittyy perusosaan suoraan (paitsi jos perusosa alkaa ä-kirjaimella, esimerkiksi *epä-älyllinen*), mutta *ei-* vaatii aina yhdysviivan. Samamerkityksinen vierasperäinen *anti-*prefiksi liittyy perusosaan ilman viivaa (paitsi jos perusosa alkaa i-kirjaimella, esimerkiksi *anti-inflammatorinen*).

Aineistossani vierasperäisellä prefiksillä muodostetut yhdyssanat (käytännössä *anti-* ja *ex-*) olen laskenut yhdyssanaksi vain, jos perusosa on suomea. Tässä olen siis noudattanut periaatetta, jonka mukaan jonkin yhdyssanan osan on oltava suomea, jotta lasken sen tässä työssä yhdyssanaksi. Tämän mukaan suomalaisilla etuliitteillä ja etuliitteenomaisilla sanoilla (esimerkiksi: *esi-*, *edes-*, *yli-* ja *nyky-*) muodostetuista yhdyssanoista olen siis ottanut analyysiin mukaan kaikki.

3.5.2 Johtimet *-lainen*, *-moinen* ja *-mainen*

Johtimilla *-lainen* ja *-lainen* on samannäköisestä kirjoitusasusta huolimatta ero niin syntyhistoriasa, vokaaliharmoniassa kuin merkityksessäänkin. *-lainen* on alunperin – ja toisinaan vieläkin – itsenäinen lekseemi ja syntynyt kannasta *laji*; *-lainen* taas on edelleen johtunut lokaalijohtimesta *-lA* > *-lAinen* (Häkkinen 1990: 156) eikä esiinny yksin. *-lAinen* tarkoittaa viittauksensa kohteen kuuluvan kantasansa luokkaan, kun taas *-lainen*-johtimen merkitys on sama kuin *kaltainen*. *-lainen* liittyy kantasansa genetiivimuotoon *saman|lainen* (poikkeuksena taipumaton eri: *eri|lainen*), kun taas *-lAinen*-johdin voi liittyä useisiin erilaisiin vartaloihin *suoma|lainen*, *vepsä|lainen*, *kauniais|lainen*, *marxi|lainen* (mahdollisista vartaloista ks. esim. ISK § 190). Samalla tavalla on muotoutunut *-mo*-vartalosta *-moinen*, mutta *-moinen* osallistuu jo vokaaliharmoniaan (Häkkinen 1990: 156). Tosin *moinen*-sana myös voi olla yhdysosana mukautumatta vokaaliharmoniaan (esimerkiksi *minkä|moinen*) ja siksi sitä voidaan pitää yhdyssanan osanakin (Koski 1982: 67). *-lainen*-muodosta on kyse myös sanassa *täl|lainen*, jossa sulauma on muuttanut *nl*-yhtymän *ll*:ksi. *-moinen*-muodon kohdalla on sama kehitys: *tän|moinen* on muuttanut muotoon *tämmöinen*, siinä on siis tullut myös vokaaliharmonia jo mukaan. Tällaista sulautumista ja leksikaalistumista kielessä tapahtuu kaiken aikaa, ja alunperin erillään olleet sanat kiteytyvät niin tiiviiksi, että osien erotteleminen ei ole enää kovin selvää eikä tarkoituksenmukaistakaan: sanaparille on kehittynyt itsenäisen merkityksen lisäksi myös itsenäinen, omansa kaltainen kirjoitusasu.

-mainen-lopulla on sanassa (ainakin) kaksi mahdollista merkitystä. Se voi tarkoittaa viittauksensa kohteen olevan kantasansa kaltainen esimerkiksi *nynny|mäinen*, *aikuis|mainen* ja silloin sen johdin onkin oikeammin vokaaliharmoninen *-moinen* (ISK § 280). Sanan lopussa *-mainen*-kantana voi kuitenkin olla myös *maa*, silloin johdin on *-(i)nen*: esimerkiksi *kotimaa* > *kotimainen*, *ulkomaa*

> *ulkomainen* (ISK § 274). ISK:n mukaan (§ 280) -mAinen-johdin on yhdyssanamainen muun muassa siksi, että se voi osallistua ellipsiin, mutta toisaalta mainitussa luvussa ISK esittelee -mAinen-aineksen nimenomaan **johtimena**; -(i)nen johdin ei tietenkään itsessään muodosta yhdyssanaa, mutta *maa*-esimerkeissäni sen **kantana** on yhdyssana, ja yhdyssanoista johdetut sanat, niin johdoksia kuin ne ovatkin, ovat myös yhdyssanoja.

Olen katsonut vokaaliharmonian – eli sanan vokaalien mukautumisen samankaltaiseksi sanassa (ISK § 15) – olevan tärkeä erottava tekijä silloin, kun pohditaan, onko sana yhdyssana vai johdos: johdoksen vokaalit mukautuvat kannan vokaaliin (*päättäj*A**, *kaivaj*A**) kun taas yhdyssanan osien vokaalit pysyvät alkuperäisinä (*pää*AsiA**, *kaivuuty*Ö**). Samaan johtopäätökseen päätyi aineistonsa suhteen Mäkisalo, joskin hän huomauttaa, ettei vokaaliharmonia käy hänen aineistossaan yhdyssanoja johtimista erottavasta muuttujasta, koska tapauksia ei ole (Mäkisalo 2000: 72). Omassa aineistossani taas on runsaasti kansallisuussanoja, esimerkiksi *suomalainen*, jotka siis jo tällä määrittelyllä putoavat pois yhdyssanojen joukosta, ja edelleen -lAinen-sanojen epäitsenäisyys rajaa niitä painokkaammin. -mOinen-sanat jäävät myös pois vokaaliharmonian perusteella; aineistossani ei ole ainuttakaan yhdyssanaa, jossa -*moinen* olisi mukana (ilman vokaaliharmoniaa), joten rajausta ei sen pidemmälle tarvinnut tässä työssä pohtia. Koska -*lainen* ei osallistu vokaaliharmoniaan ja voi esiintyä itsenäisenäkin, on se näillä kriteereillä yhdyssanan osa ja siten myös tässä työssä katsotaan yhdyssanaksi.

3.5.3 Leksikaalistumat, vierassanat ja erisnimet

Suomessa on sanoja, jotka usean tavun vuoksi kuulostavat yhdyssanoilta, mutta eivät sitä ole. Tällaisia ovat esimerkiksi *hämähäkki* ja *artisokka* (ISK § 147). Tällaisia sanoja en luonnollisestikaan ole lukenut yhdyssanoihin; jos yhdyssanassa ei ole ollut selvästi erottuvia ja tunnistettavia osia, olen pyrkinyt tarkistamaan sanan etymologian – ja jos se ei ole onnistunut, olen jättänyt sanan laskuista kokonaan. Tällainen sana on esimerkiksi *eri|toten. Nimen|omaa* sen sijaan laskin yhdyssanaksi, koska, vaikka en sen etymologiaa saanutkaan selvitetyksi, katson sen olevan samaa kantaa kuin *ni|men|omainen*, joka taas on jo selvästi yhdyssana *omainen*-osansa vuoksi. *Omainen* ei osallistu vokaaliharmoniaan (esimerkiksi MOT s.v. ahaa-elämys: *välähdyksenomainen*) ja on omamerkityksinen 'kaltainen' tai 'tyyppinen', joten se on selvästi yhdyssanan osa. Voi tietysti pohtia, miten prototyyppinen perusosa se on: kumpi määrittää kumpaa esimerkissä (34)?

(34) *taian|omainen*

Onko kyseessä siis jokin 'omaisuus' eli tässä 'kaltaisuus', jota 'taika' luonnehtii vai onko ensin 'taika', jonka 'omainen' eli 'kaltainen' jokin on? Erityisesti adjektiivissa on runsaasti yhdyssanoja, joissa

osien välinen työnjako ei ole ihan selvä tai jaettavissa yksiselitteisesti määrite- ja perusosaan. Pohdin joidenkin yhdysosien suhteita vielä myöhemmin luvussa 4.3.2 Nominatiivimääritteiset yhdysosat: d) Nominatiivimääritteisten yhdysosojen yhdysosatyyppit (s. 58).

Yhdysosanamuotoiset erisnimet (*Arkadian|mäki, Heinä|luoma*) olen laskenut aineistoon, jos ne muuten täyttävät tämän työn yhdysosan kriteerit, eli edes toinen osa on suomalainen sana eikä kyseessä ole johdos (erisnimiin palaan vielä luvussa 4.3.2 Nominatiivimääritteiset yhdysosat: a) Nominatiivimääritteiset substantiivit). Kuukausien nimet (*kesä|kuu*) kuuluvat analysoitavaan aineistoon, samoin *keski|viikko*, mutta eivät luonnollisesti muut viikonpäivät – vaikka niissäkin ehkä loppuosan *tain* voisi *dag*-sanaksi vielä mieltääkin.

Edellä olevat mielessä pitäen esitän kokoavasti Taulukossa 1. Yhdysosojen rajausta millaiset ratkaisut olen tehnyt sen suhteen, mitkä näistä sanoista ovat ja eivät ole yhdysosia. Esimerkit aineistostani.

Yhdysana	Ei yhdysana
<i>epä selvä</i>	
<i>ei- sensuroitu</i>	
<i>anti piraatti järjestöt</i>	
<i>ex- presidentillä</i>	
<i>edes takas</i>	
<i>esi merkinä</i>	
<i>nyky piirtäjät</i>	<i>sukulainen</i>
<i>yli määräistä</i>	<i>sekalainen</i>
<i>jonkin lainen</i>	<i>tällainen</i>
<i>eri lainen</i>	<i>suomalainen</i>
	<i>aikamoinen</i>
	<i>semmoinen</i>
<i>koti mainen</i>	<i>nynnymäinen</i>
<i>perus suomalainen</i>	<i>persulainen</i>
<i>Heinä luoma</i>	<i>Urpilainen</i>
<i>keski viikko</i>	<i>torstai</i>
<i>nimen omaan</i>	<i>eritoten</i>
<i>taian omainen</i>	

Taulukko 1. Yhdysosojen rajausta

4 AINEISTON ANALYYSI

Tässä tutkielman ydinluvussa käyn läpi aineistoni yhdyssanoja eri näkökulmista. Aloitan yhdyssanojen ja yhdyssanapoikkeamien esittelyn tarkastelemalla yhdysosien määriä ja määrien suhdetta poikkeamiin. Sen jälkeen tarkastelen yhdysviivan käyttöä yhdyssanoissa. Olen nimennyt erityisyhdyssanatyypiksi joukon yhdyssanoja. Näihin kuuluvat sanaliitoista, lyhenteistä ja numeroista muodostuneet sanat. Tätä ryhmää tarkastelen ennen varsinaista syventymistä perusyhdyssanoihin, joita käsittelem niiden määriteosan muodon mukaan. Tämä perusyhdyssanojen osuus on työni painokkain osa. Sen jälkeen tarkastelen vielä yhdyssanoja yleisesti niiden piirteiden pohjalta, jotka ovat yhteisiä kaikille tyypeille. Aivan lopuksi käyn läpi niitä sanaliittoja, jotka on kirjoitettu konvention vastaisesti yhdeksi sanaksi.

4.1 Yhdyssanat määritysosien määrän mukaan

Sanojen yhdistäminen on rekursiivista, toisin sanoen sanoja voidaan liittää peräkkäin periaatteessa loputtomiin (ISK § 398 ja § 405, Mäkisalo 2000: 100). Yhdyssanan osina voivat siten olla yhdyssanat ja edelleen yhdistyneet yhdyssanat. Yhdyssanan osana voi olla myös sanaliitto, ja edelleen yhdyssanan osana olevan sanaliiton sisälläkin voi olla yhdyssanoja. Kuten aiemmin (luvussa 2.1 Yhdyssana per se) totesin, määritysyhdyssana on lopulta kuitenkin aina jaettavissa kahteen osaan; summayhdyssanassa sen sijaan osien määrä on jakamaton, eli lopullinen osien määrä on yhdyssanan osien tosiasiallinen määrä.

Aineistossani on kaikkiaan 885 sellaista yhdyssanaa, joissa on enemmän kuin kaksi osaa. Kutsun tällaisia sanoja jatkossa kompleksisiksi yhdyssanoiksi ISK:n mukaan. ISK katsoo kompleksisiin yhdyssanoihin kuuluvaksi vain ne yhdyssanat, joiden osana on yhdyssana (ISK § 405), mutta tässä työssä lasken kompleksisiksi yhdyssanoiksi myös ne yhdyssanat, joita joko määrittää tai jonka perusosana on sanaliitto. Kaikista yhdyssanoista kompleksisten yhdyssanojen määrä on 11,6 %. Tässä ryhmässä on sekä määritysyhdyssanoja että summayhdyssanoja: summayhdyssanoja on 5, niihin rinnastuvia erisnimiä 2 (*Alanko-Kahiluoto*), 1 epäselvä sana ("*tottakai-sapettaa*"), ja loput 877 on määritysyhdyssanoja.

Olen koostanut kompleksisia määritysyhdyssanoja kuvaavan taulukon (Taulukko 2. Moniosaisten yhdyssanojen rakenne ja poikkeamat), josta selviää tämän yhdyssanaryhmän määrite- ja perusosien suhteiden, määriteosan tyyppien, poikkeamien ja poikkeamaprosenttien määrät.

Määrite-osa	Perus-osa	Esimerkki	Määrä	Määriteosa perussana	Määriteosa sanaliitto	Määriteosa yhdyssana	Määriteosa muu	Poikkeamia	Poikkeama-prosentti
2	1	<i>elintarviketuotanto</i>	599	–	34	553	12	74	12,4
1	2	<i>Aalto-yliopisto</i>	164	153	–	–	11	25	15,2
3	1	<i>työväenlehtiperinne</i>	51	–	17	34	–	11	21,6
2	2	<i>10-vuotissyntymäpäivä</i>	35	–	7	26	2	6	17,1
4	1	<i>"Nyt tuli mutkia matkaan" -sivu</i>	7	–	5	2	–	2	28,6
3	2	<i>"Mikä siinä on" -musiikkivideo</i>	5	–	5	–	–	1	20
1	3	<i>Mikki-salapoliisisarja</i>	4	4	–	–	–	1	25
5	1	<i>"Maailman suurin Punainen Risti" -tapahtuma</i>	3	–	3	–	–	–	–
2	3	<i>viharikoslainsäädäntökäytännön</i>	2	–	1	1	–	–	0
2	4	<i>KOODINMURTO 2011 -viikkokysymyspäähkinä</i>	1	–	1	–	–	–	–
4	2	<i>Uusi suunta Rakennusliitolle vaaliliitto</i>	1	–	1	–	–	1	100
5	2	<i>Peruuttamalla ei pääse eteenpäin -äkkiryhmä</i>	1	–	1	–	–	–	–
6	1	<i>"minkä taakseen jättää sen edestään löytää"-tyylillä</i>	1	–	1	–	–	–	–
7	1	<i>"pelastakaa maailma ydinvoimalta, ydinvoimalla" komitea</i>	1	–	1	–	–	1	100
11	1	<i>"kaikki minulle/meille, tässä ja nyt! Kasvua! Enemmän ja enemmän minulle!" -ajattelu</i>	1	–	1	–	–	–	–
12	1	<i>"Kulttuuriteoksia on saatavilla jo niin paljon, ettei uusien tekemistä tarvitse tukea"-väittäjä</i>	1	–	1	–	–	–	–
Yhteensä			877	157	79	616	25	122	13,9

Taulukko 2. Moniosaisen määritysyhdyssanojen rakenne ja poikkeamat

Olen lajitellut sanatyypit yleisyyden mukaan: taulukon yläosassa ovat yleisimmät. Vaikka olen merkinnyt poikkeamaprocentin kunkin tyyppin kohdalle, on se harvinaisempien tyyppien vähälukuisuuden vuoksi relevantti ainoastaan koko aineiston osalta. Poikkeamia tehdään 13,9 %:ssa kaikista kompleksisista määritysyhdyssanoista, eli moniosaisissa yhdyssanoissa tehdään huomattavasti enemmän poikkeamia kuin yleensä yhdyssanoissa (koko aineiston yhdyssanojen poikkeamaprosentti on siis 5,6 %). Tämä on tietenkin odotuksenmukaista: kun sanoja tulisi kielenhuollon mukaan kirjoittaa useita peräkkäin ja kun sanapituus silloin kasvaa, alkaa kirjoittajan mielessä varmasti herätä epäily siitä, voiko niin pitkää sanaa ollakaan. Poikkeamaprosenttia nostaa myös sanaliitto määriteosassa: esimerkiksi niissä 2 + 1 -tyypin yhdyssanoissa, joissa määriteosa on sanaliitto, jää yhdysviiva pois 13 tapauksessa kaikkiaan 34:stä, kuten esimerkissä (35). (Yhdysviivan käytöstä perusteellemmin seuraavassa luvussa 4.2 Yhdysviiva ja erityisyhdyssanatyypit.)

(35) *HYVÄ IKÄ messuilla.*

ISK:n mukaan yhdysosien määrällä ei ole ylärajaa, mutta käytännössä niitä on harvemmin enemmän kuin viisi (§ 405). Näin on myös aineistossani: vaikka sanaliittomääritteiset yhdyssanat ovat pisimmät kaikista yhdyssanoista, ne pisimmät määritysyhdyssanat, joiden määrite on yhdyssana, ovat viisiosaiset (esimerkit 36 ja 37).

(36) *kukkahattu-sossutanttu -lauma*

(37) *ruotsi - suomi em-karsinta ottelua*

Pisin summayhdyssana on kymmenosainen (esimerkki 38, yhdyssana alleviivattu).

(38) *Tunturirundi starttas tänään! Karavaani kulkee seuraavasti:
Rovaniemi-Levi-Luosto-Ylläs-Levi-Ruka-Haukipudas-Kokkola...Siellä näkyy!*

Tämänkin sanakoosteen osina on yhdyssanoja *Rovaniemi* ja *Haukipudas*. On hiukan kyseenalaista, voiko tätä sanajonoa edes kutsua yhdyssanaksi vai onko kyseessä pikemminkin usean sanan ketjutettu ilmaus – ja toisaalta milloin summayhdyssana ei olisi sellainen? Summayhdyssanojen osuus kuitenkin sekä koko yhdyssanamäärästä (1,7 %) että kompleksisista yhdyssanoista (0,6 %) on sen verran marginaalinen, että ne voi käsitellä tässä työssä enemmän kurioositeettina kuin minään muuna (palaan yhdyssanatyypien jakaumaan vielä luvussa 4.4 Yhteenveto kaikista yhdyssanoista).

4.2 Yhdysviiva ja erityisyhdyssanatyypit

Tässä luvussa tarkastelen yhdysviivan käytön syitä ja poikkeamia aineistossani. Esiin nousevat myös yhdysviivan käyttöön kiinteästi liittyvät yhdyssanatyypit, jotka olen ryhmittänyt tässä tutkielmassa ja nimennyt erityisyhdyssanatyypiksi.

Yhdysviiva on sikäli olennainen osa tätä tutkielmaa, että milloin sen käyttö on kielenhuollon mukaan välttämätöntä ja se kuitenkin puuttuu, olen katsonut sanan poikkeamaksi, paitsi silloin kun yhdysviivan puuttuminen ei aiheuta koko ilmaisun hajoamista. Käytännössä kyseessä on silloin kahden saman vokaalin välistä puuttuva yhdysviiva (*alaaste*). Toisaalta kuitenkin, jos yhdysviivaa on käytetty asianmukaisesta syystä mutta konvention vastaisella tavalla – käytännössä yleensä silloin sen suhde sananväliin ei ole kielenhuollon suositusten mukainen – en ole laskenut koko sanaa poikkeavaksi. En myöskään ole käsitellyt koko sanaa poikkeavana, jos yhdysviiva on sanassa liikaa. Toisin sanoen: jos yhdyssanan osien yhdistämisestä on jotain jälkiä, kirjoittajan pyrkimys yhdistää ilmaisu kokonaiseksi on havaittavissa, olen katsonut sanan konvention mukaiseksi, vaikka yhdysviiva ei korrektilla kohdalla olisikaan. Tämän mukaan siis esimerkki 39 lasketaan poikkeaviin ja esimerkki 40 konvention mukaisiin yhdyssanoihin tässä työssä.

(39) *EU asioissa*

(40) *EU- asiassa*

Tällaiseen ratkaisuun päädyin siitä syystä, että tässä työssä on tarkoitus nimenomaan tarkastella sanojen yhdistämistä ja siinä tehtäviä sellaisia poikkeamia, jotka joko kokoavat tai hajottavat koko ilmaisun. Vaikka koko yhdyssanaa en yhdysviivan poikkeavan käytön vuoksi olekaan laskenut poikkeamaksi, olen kuitenkin eritellyt aineiston yhdysviivan käytön paitsi tarkoituksen myös konvention mukaan. Esitän seuraavaksi, millaisia variaatioita aineistossani on yhdysviivan suhteen sekä sen käytön syissä että kirjoittamistavassa.

Yhdysviivan käytön syitä analysoin aineistostani kaikkiaan kymmenen. Kielenhuollon mukaan yhdysviivaa käytetään yhdyssanassa, kun:

1. yhdysosana on erisnimi
2. yhdysosana on lyhenne
3. yhdyssanan osien rajalla on sama vokaali
4. halutaan jättää osa yhdyssanaa toistamatta eli elliptinen ilmaus
5. yhdysosana on sanaliitto
6. yhdysosana on numero
7. yhdysosana on sellainen vierassana, että sen hahmottaminen toisesta osasta on vaikeaa
8. yhdyssanassa on rinnasteisia osia.

(Kielenhuollon kannasta yhdysviivan käyttöön ks. esim. KO: 245–250, KK: 151–159.) Näiden lisäksi tässä tutkielmassa on vielä kaksi ryhmää. Yhdysviiva esiintyy aineistossani edellisten lisäksi, kun:

9. ilmauksessa on useita syitä käyttää yhdysviivaa
10. yhdysviivan käyttö on tarpeetonta.

Taulukossa 3. Yhdysviiva esitän koottuna, missä, minkä verran ja miten konvention mukaisesti yhdysviiva esiintyy aineistossani. Olen järjestänyt tyypit yleisyyden mukaan ja Käytön syy -sarakkeen selitteessä on edellä esittämään luetteloon numeroviittaus. Poikkeuksena edellä esittämäni jaotteleluun yhdysviivan käytön syistä olen aineistossani yhdistänyt ryhmään 7 myös muita marginaalisia käyttösyitä. Tällaisia ovat prefiksit *ex-* ja *ei-*, jotka esiintyvät aina yhdysviivan kanssa yhdyssanan alkuosana, sekä sovittu kirjoitusasu sanassa *dalai-lama*.

Käytön syy	Esimerkki poikkeava konvention mukainen	Määrä	Poikkeamia	Poikkeama- prosentti
Erisnimi (1)	<i>lines tarina</i> <i>Varsinais-Suomessa</i>	268	94	35,1
Lyhenne (2)	<i>spneutraali</i> <i>WHL-liigaa</i>	165	63	38,2
Vokaali (3)	<i>alaaste</i> <i>vuokra-asuntotuotannossa</i>	100	10	10,0
Ellipsi (4)	<i>lempi/liikanimiä</i> <i>olympia- ja hackamorekuolain</i>	90	23	25,6
Sanaliitto (5)	<i>nuija ja tosinuija pätkään</i> <i>Jumalia ja ihmisiä -elokuvan</i>	81	37	45,7
Numero (6)	<i>30 vuotiaan</i> <i>1900-luku</i>	61	21	34,4
Tarpeeton (10)	<i>pääministeri-puolue</i> –	37	37	100,0
Useita syitä (9)	<i>EX-SD -edustajaa</i> <i>1950–70-luvuilta</i>	33	23	69,7
Muu (7)	– <i>exit-strategia</i>	19	–	–
Rinnasteiset (8)	<i>Koomikko -taikuri</i> <i>oranssi-mustalta</i>	15	9	60,0
Yhteensä kaikki		869	317	36,5

Taulukko 3. Yhdysviiva

Kuten taulukosta (3) on nähtävissä, yhdysviivan käytössä tehdään poikkeamia erittäin usein ja poikkeamia tehdään kaikissa ryhmissä lukuun ottamatta muu-ryhmää (7). Ilmeisesti tämän ryhmän vierassanat ja prefiksit hahmottuvat erityisen painokkaasti yhdysviivalla erotettaviksi. Voi myös arvela, että vierassanoja ja prefiksejä käyttävät saattavat olla keskimääräistä valveutuneempia kirjoittajia.

Yhdysviivan kirjoittamisessa tehdään seuraavia poikkeamatyyppejä: 1) yhdysviiva puuttuu kokonaan, 2) yhdysviiva on väärässä paikassa ja 3) yhdysviivaa on käytetty tarpeettomasti. Kuten aiemmin totesin, tässä työssä yhdyssanapoikkeamaksi olen laskenut vain ne yhdysviivaan liittyvät poikkeamat, joissa tarpeellinen yhdysviiva puuttuu kokonaan ja siitä syystä ilmaus on jakaantunut erillisiksi sanoiksi, eli tyyppin 1 poikkeamat. Yhdysviivan puuttuminen on syynä 179 poikkeamaan 432:sta, eli peräti 41,4 %:ssa kaikista poikkeamista yhdysviivan puuttuminen hajottaa yhdyssanan kahteen (tai useampaan) osaan, kuten esimerkeissä 41, 42 ja 43.

(41) *pro-EU mielinen*

(42) *sunnuntai ilta*

(43) *Aku Anka viesteistä*

Yhdysviivan käyttöön liittyy usein aivan erityiset yhdyssanatyypit, joita tarkastelen seuraavaksi. Ryhmään lasken kuuluvaksi ne yhdyssanat, joiden osana on: 1) lyhenne, 2) numero tai merkki ja 3) sanaliitto. Tähän ryhmään eivät siis kuulu ne yhdysviivalliset yhdyssanat, joiden osien rajalla on vokaalipari (*tasa-arvo*), ne yhdyssanat joiden osana on vierassana (*extreme-laji*) tai prefiksimmäinen ai-

nes (*ex-presidentti*) eivätkä ne yhdyssanat, joiden osana on erisnimi (*Jyrkkä-Jyrki*) tai joiden osat ovat rinnasteiset (*filosofi-psykoanalyttikko*); nämä tyypit lasken kuuluvaksi seuraavan luvun 4.3 perusyhdysanoihin. Perusteena tähän on sanamaisuus: vaikka erityisyhdysanatyypeissä sanamaisuus toteutuukin ainakin osin, kun ilmaus esitetään puhuttuna (esimerkiksi *viisivuotias*), kirjoitettuna sanamaisuus katoaa (esimerkiksi *5-vuotias*). Juuri sanamaisuuden vuoksi laskin prefiksimäisten aineisten määrittämät yhdyssanat pois tästä ryhmästä, mutta syynä on myös niiden semanttinen samankaltaisuus sellaisten prefiksimäisten aineiden kanssa, jotka eivät vaadi yhdysviivaa sananrajalle (esimerkiksi *epä|ekologinen*). Sanaliiton osalta tilanne on taas melkein päinvastainen. Sanaliitto useasanaisena ilmaisuna lähenee lausetta, on ilmaisuvoimaltaan erittäin sanamainen, mutta käyttäytyy yhdysanassa yksittäisestä sanasta poikkeavasti ja samaan tapaan kuin lyhenne tai numero: se ei taivu eikä yhdisty toiseen yhdysosaan muutoin kuin yhdysmerkillä (ja sananvälillä) erotettuna.

Kaikkiaan tällaisia erityisyhdysanoja on aineistossani 319 eli 4,2 % kaikista yhdysanoista. Taulukossa 4. Erityisyhdysanat määritysosan tyypin mukaan esitän kokoavasti, miten tämän ryhmän tyypit jakautuvat alatyyppeihin ja minkä verran niissä on tehty poikkeamia.

Ryhmä	Esimerkki poikkeava konvention mukainen	Määrä	Poikkeamia	Poikkeama- prosentti
Lyhennesanat	<i>VT hullu ADHD-aikuisille</i>	169	50	29,6
Sanaliittosanat	<i>Älä Alistu kampanjan 45 minuuttia -ohjelmassa</i>	81	25	30,9
Numerosanat	<i>90 luvulla 10-kertaiset</i>	69	20	29,0
Yhteensä kaikki		319	95	29,8

Taulukko 4. Erityisyhdysanat määritysosan tyypin mukaan

Poikkeamamäärä on kaikkiaan huomattavan suuri: peräti 29,8 % tämän ryhmän sanoista on kirjoitettu konvention vastaisesti. Mikään määritysosatyyppi ei näytä olevan selvästi muita vaikeampi hallita vaan poikkeamia tehdään melko tasaisesti kaikissa tyypeissä. Sekä tästä erityisyhdysanojen ryhmätaulukosta (4) että yhdysviivan käyttötaulukosta (3) voi vetää johtopäätöksen, että yhdysviivan käyttö tuottaa kirjoittajille suuria hankaluuksia sekä silloin, kun pitää päättää, milloin yhdysviivaa ylipäätään tarvitaan, että silloin, kun pitää osata laittaa se oikeaan kohtaan.

4.3 Perusyhdysanat

Seuraavaksi käyn läpi perusyhdysanojen tyypit niiden määriteosan muodon mukaan jaoteltuna. Esitän kussakin alaluvussa yhdysanatyypin piirteet, määrän ja poikkeamat. Poikkeamat perustelen

kielenhuollon ohjeistuksin. Aivan ensiksi huomautan, että joskin yhdyssanan osana voi olla toinen yhdyssana (*ihmisoikeus|ministeri*), käsittelen yhdyssanan osana olevia yhdyssanoja tässä kuten perussanoja, sillä ne käyttäytyvät yhdyssanan osina samoin kuin muutkin sanat¹². Käytännössä siis tässä määrittelyyhdyssana käsitetään kaksiosaiseksi määriteosan ja perusosan kokonaisuudeksi, ja summayhdyssanaa pidetään rinnasteisten osiensa summana huolimatta siitä, onko jokin rinnasteinen osa jaettavissa vielä osiin.

Yhdyssanoja on jaettu merkitystensä mukaan monella tapaa ryhmiin. Tällaisia ryhmiä esittelee esimerkiksi Ylikulju läpi koko pro gradu -työnsä (Ylikulju 2003) viitaten Pauli Saukkosen (1973) ja Jouko Vesikansan (1989) lajitteluun. Ryhmiä esittelevät myös ISK § 409 (nominatiivimääritteisten osalta) ja § 411 (genetiivimääritteisten osalta) ja jopa Kielenhuollon käsikirja (Iisa et al. 2006: 183). Jotta voisi laskea, paljonko missäkin merkitysryhmässä on yhdyssanoja ja vetää johtopäätöksiä siitä, missä poikkeamia syntyy erityisesti – jos niitä ylipäätään erityisesti syntyy jossain – pitäisi koko aineiston yhdyssanat, 7634 kappaletta, lajitella merkityksiä erotteleviin ryhmiin. Tämän työn puitteissa se ei kuitenkaan ole mahdollista sanojen suuren määrän vuoksi, joten tässä luvussa keskityn kuvaamaan yhdyssanoja, niiden määriä ja poikkeamia, muotokriteerien näkökulmasta, käytännössä siis sanaluokkien ja jonkin verran johdosten tasolla. Sanan merkityksen käsite, milloin se tulee esiin tässä yhteydessä, on huomattavan laaja ja liikkuu siis lähinnä sanaluokan – ja edelleen johdosten – tasolla, ei niinkään yksittäisen sanan semantiikassa, joskin sitäkin paikoin sivutaan.

Yhdyssanan sanaluokka on sama kuin sen perusosan sanaluokka (ISK § 404). Olen jakanut yhdyssanat sanaluokkiin ja tehnyt jaon yhtäältä melko karkeasti – adjektiivit, substantiivit, verbit ja muut – mutta toisaalta olen jakanut adjektiivien ja substantiivien sanaluokkia välittömästi pienempiin yksikköihin. Adjektiivit olen jakanut edelleen -(i)nen-johdoksiin, karitiiveihin, partisiippeihin ja perusadjektiiveihin; substantiivit erisnimiin, verbikantaisiin johdoksiin ja partisiippeihin sekä perussubstantiiveihin. Olen ottanut sanaluokkajaossa perusteltuja vapauksia: en ole katsonut pelkääntään sanan sanakirja- ja taivutusmuotoon, kun olen ryhmitellyt sanoja, vaan olen katsonut samalla merkityksiä. Näin sekä adjektiivien että substantiivien joukkoon olen sijoittanut joitain nominaalisia verbimuotoja (partisiippeja) – ja olenpa yhden adjektiivisen numeraalin *viisikymppinen* määrännyt substantiiviksi siitä syystä, että sen merkitys yksinkertaisesti on 'seteli' ja siksi substantiivinen (esimerkki 44):

(44) *samalla kerään lahjoituksia matkalla lodelle, on se sitten euro
tai viisikymppinen ne lahjoitetaan LASTENKOTI LOUHELAN*

12 Jos esimerkiksi yhdyssana päättyy -nen-muotoon, muuttuu se yhdyssanan määriteosana s-muotoiseksi: *suomenhevonen* > *suomenhevos|jalostaja*; jos määriteosana olevan yhdyssanan viimeinen vokaali on sama kuin perusosan aloitettava vokaali, osien väliin tulee yhdysviiva: *yliopisto-opiskelu*, ja niin edelleen.

Verbikantaisia substantiiveja olen tarkistanut erityisesti Nykysuomen etymologisesta sanakirjasta (Häkkinen 2004, jatkossa NES). Johdoksissa olen huomionnut vain välittömästi verbistä johdetun substantiivin eli esimerkiksi adjektiivista johdetut -sUUs-substantiivit eli ominaisuuden nimet (*oikeudenmukaisuus, selväjärkisyys*) olen tässä työssä merkinnyt perussubstantiiveiksi, vaikka ne eivät sitä luonnollisesti yleisesti ottaen olekaan. Verbikantaiset johdokset olen edelleen lajitellut käyttäen pohjana ISK § 221:n lajittelua, mutta tiivistänyt sitä hiukan:

1. ryhmä: tekijä ja väline

johdokset: -jA, -(u)ri, -(i)lAs, -in, -e, -(i)ke

(45) *Aku-piirtäjä, perusduunari, korjauskertoimia*

2. ryhmä: teko ja tulos

johdokset: -minen, -U, -O, -Us, -ntA, -nti, -e, -UU, -nA, -Os, -ntO, -mUs, -mA

(46) *ammattitaidolla, arvopohdintaa, asiakasmaksuja, asiantuntemuksesta, Barks-äänestys, eduskuntatutkimuksen, eläinsuojelulainsäädännön, eläkejärjestelmä, hömppäpuhe*

3. teon paikka ja tilaisuus; paikka ja tilaisuus, ryhmä 3.

johdokset: -(i)mO, -jAiset, iAiset

(47) *ajatushautomon, vaalivalvojaisten*

Lisäksi olen merkinnyt verbimuodoista finiitti-, infinitiivi- ja partisiippimuodon (mutta en enää niiden alalajeja).

Määriteosan jaoin hiukan tarkemmin sanaluokkiin siitä yksinkertaisesta syystä, että määriteosassa sanaluokkia **on** enemmän: numeraalit ja pronominit. Toisaalta verbejä ei määriteosassa ole (paitsi joitain epäitsenäisiä nominaalisia yhdysosamuotoja: *uima-*). Määriteosa on siis jaettu seuraaviin ryhmiin: substantiivit, adjektiivit, numeraalit, pronominit ja muut. Esitän seuraavaksi, millaisia ongelmia sanaluokkajaossa oli, miten ne ratkaisin ja millä perustein.

Substantiiveilla ja adjektiiveilla, jotka ovat useimmin esiintyvät määriteosan sanaluokat (substantiivimäärittäjiä 5948 ja adjektiivimäärittäjiä 933), on jonkin verran ristikkäistä käyttöä: adjektiivia voidaan käyttää kuten substantiivia ja päinvastoin. Osa näistä rajalla olevista sanoista on yksinkertaisesti kaksimerkityksisiä, kuten vaikkapa yhdysosan alusta tuttu *elo*. Esimerkissä 48 *elo* viittaa elonkorjuun tuotokseen, 'satoon', ja esimerkissä 49 *elo* viittaa *elämään* ja paremmin vielä *elävään*.

(48) *Toivottavasti elokuun Turun suunnan keikoilla kuullaan uudestaan*

(49) *Torstai-illan elokuva tulee Uudesta-Seelannista*

Sanueen taustalla on toki sama elämään ja elämiseen liittyvä perusta, mutta nähdäkseni esimerkin 48 *elo* on kuitenkin substantiivisena viittaavampi – se voisi olla jopa referentiaalinen joissain yhteyksissä – kuin esimerkin 49 *elo*, jonka merkitys on selvästi kuvaileva, adjektiivinen. Sanojen välillä on linkki elämään liittyvien sanojen kautta, mutta aivan suoraa merkitysyhteyttä ei ole. Samantapainen sana on *huippu*, mutta sen merkitykset kuitenkin lähenevät toisiaan enemmän: *huippu* on 'korkein kohta', oli kyse sitten aineettomasta tai aineellisesta *huipusta*. Olen lähtenyt adjektiivien ja substantiivien erottelussa, milloin se ylipäätään on kyseenä, siitä että jos sana on kuvaileva, on se adjektiivi ja jos se taas selvästi viittaa johonkin entiteettiin, on se substantiivi. Tämän vuoksi *elo* jakautuu kahteen ryhmään, mutta *huippu* on substantiivi, vaikka se olisi adjektiivisessa käytössä; *luomu* on adjektiivi (*luonnonmukainen*), vaikka siitä puhuttaisiin ikään kuin tuotteena – tosin yhdyssanan alkuosana se helpommin mieltyy joka tapauksessa adjektiiviksi. Määrittämissä yhdyssanojen kohdalla huomasinkin mielestäni mielenkiintoisen ilmiön: vaikka määrittävä osa olisi hyvinkin substantiivinen, esimerkiksi *kivi- kivistalossa*, mieltyy se kuitenkin adjektiiviseksi täsmentäessään perusosaa. Olenkin useassa kohden tätä työtä tehdessä pohtinut, millä tavalla tämä adjektiivimaisuus vaikuttaa siihen, että nominatiivi- ja jopa yhdysosa-alkuiset yhdyssanat joskus jakautuvat osiinsa. Ehkä *kivi* mieltyy *kiviseksi*, *nyky-* *nykyiseksi* ja sen vuoksi ne koetaan erillisiksi sanoiksi: jos *kivinen talo* kirjoitetaan sanaliitoksi toisistaan erilleen, niin miksipä ei samamerkityksinen *kivi talo* myös, samoin *nykyinen aika* vertautuisi *nyky aikaan*. Palaan tähän ajatukseen ja näihin esimerkkeihin vielä seuraavan luvun (4.3.1 Yhdysosa-alkuiset yhdyssanat) loppupuolella.

Määriteosassa esiintyvien epäitsenäisten sanavartaloiden osalta tein sanaluokkajaon sen mukaan, mikä sanan kanta on. Jos kantaa oli mahdoton päättää (ns. nominaaliset epäitsenäiset vartalot ks. ISK § 416), laitoin määriteosan sanaluokaksi ryhmän muu, johon sijoittuivat tämän tyyppin lisäksi ne verbimuodot, joilla ei ole itsenäistä nominaalista käyttöä (*uima-*), adverbit (*ali-*, *läpi-* ja *yli*) ja sen kaltaiset muista sanaluokista eroavat ilmaisut (*erin-(omainen)*).

Kaiken edellä esitellyn lajittelun motiivina on saada tietoa siitä, millaisia yhdyssanoja käytetään missä määrin sekä millaisissa tyypeissä poikkeamia tehdään runsaammin, jos niitä ylipäätään tehdään enemmän joissain ryhmissä. Näin koko yhdyssanamäärä jakautuu ikään kuin matriisiin, jonka

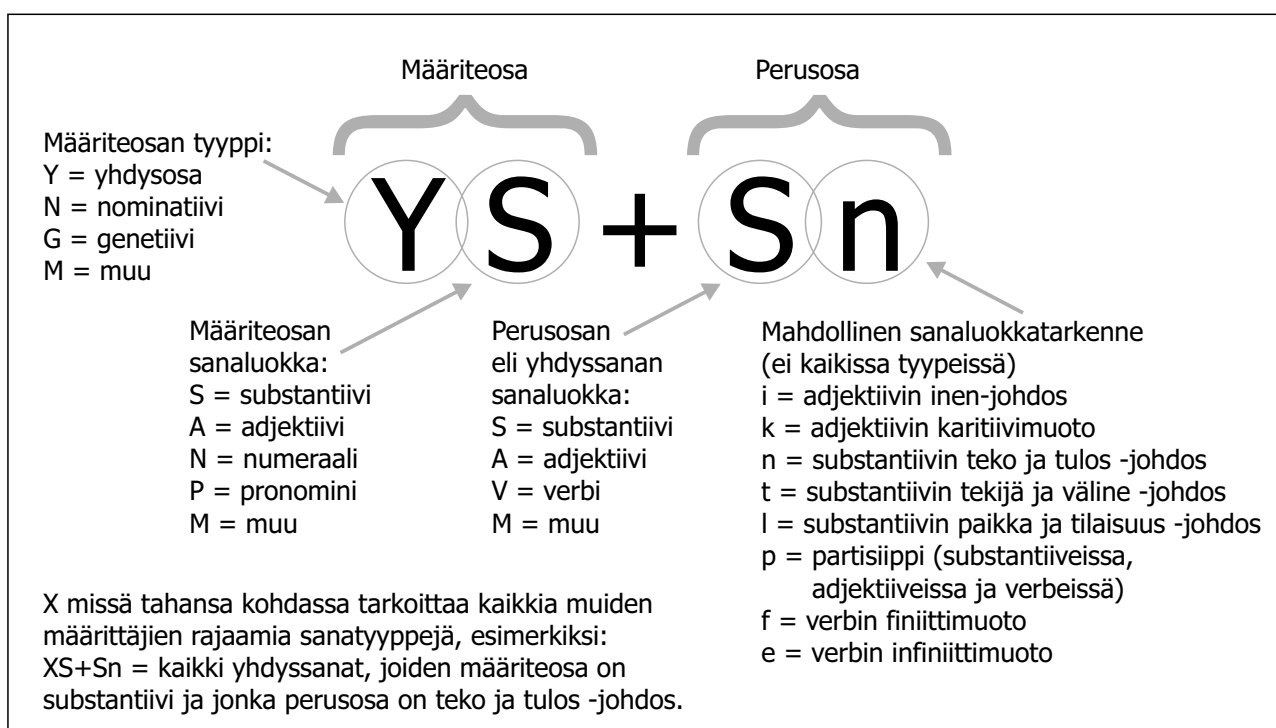
x-akselilla on määriteosan sanaluokka: substantiivit, adjektiivit, numeraalit, pronominit ja muut
 —> syvyysakselilla muoto: yhdysosa-, nominatiivi-, genetiivi- ja muu muoto

y-akselilla on yhdyssanan sanaluokka: substantiivi, adjektiivi, verbi ja muut.

Tämän matriisin perusteella esitän, millaisia määriä missäkin ryhmässä esiintyy ja missä tyypeissä tehdään poikkeamia enemmän, missä vähemmän. Käsittelyssä etenen **määriteosan muodon** mu-

kaan: esittelen ensin yhdysosa-alkuiset, sitten nominatiivialkuiset ja lopuksi genetiivialkuiset yhdysosat. Määriteosan muoto on ikään kuin kolmas ulottuvuus, se taivuttaa x-akselin sanaluokat vielä kerran ryhmiin. Yhdyssana muodostuu näin kolmiulotteiseksi: vaaka-akselilla (x) on määriteosan sanaluokka, syvyysakselilla sen muotoryhmä ja pystyakselilla (y) on perusosan sanaluokka.

Käytän paikoin yhdyssanoista selvyyden vuoksi lyhennettä, joka muodostuminen näkyy kuvasta 1. Yhdyssanan tyyppilyhenteen selite. Taulukot esitän niin, että se on samanlainen kaikissa ryhmissä huolimatta siitä, puuttuuko jokin esiintymätyyppi ryhmästä kokonaan; täten lukijalle lajittelun pohja pysyy näkyvillä ja samanlaisena tyyplistä toiseen ja vertailu eri ryhmien välillä – toivoakseni – helpottuu, kun pohja on aina sama.



Kuva 1. Yhdyssanan tyyppilyhenteen selite

4.3.1 Yhdysosa-alkuiset yhdyssanat

Joillain sanoilla on sellaisia erikoistuneita muotoja, jotka esiintyvät yhdyssanan määriteosina mutta eivät muuten kuulu sanan taivutusparadigmaan eivätkä koskaan esiinny yksinään. Näitä epäitsenäisiä, yksinään esiintymättömiä yhdyssanamuotoja kutsutaan casus componens -muodoiksi ja tässä tutkielmassa yhdysosamuodoiksi. Epäitsenäisyys koskee vain sanan muotoa; semanttisesti ne edustavat kokonaista sanaa merkityksineen. Ryhmän sanojen merkitykset ovat samanlaiset kuin näihin vertautuvissa nominatiivi- ja genetiivialkuisissakin yhdyssanoissa tai erilleen kirjoitetuissa vastaa-

vissa ilmaisuissa (*aikuisikä* : *aikuisen ikä*; *etäosallistua* : *osallistua etäältä*; *Pohjois-Suomi* : *pohjoinen Suomi* ja niin edelleen). Yhdysosamuodot voidaan jakaa seuraavalla tavalla:

1. partikkelien ja johdosten kantoina esiintyvät **vartalot** sekä **prefiksit**

(50) *ala|ikäisiä, epä|usko, keski|viikkona, tasa|vertaisesti, ulko|maalainen, ylä|puolelle*

2. (erityisesti -nen-loppuisten sanojen) **konsonanttivartalot**

(51) *aikuis|ikäni, ihmis|arvoa, Pohjois|-Suomen, yhteis|kunta*

3. loppuheittoiset ja muuten **lyhentyneet** sanat

(52) *etä|osallistua, pien|ryhmänä, puna|multa, sosiaali|tuet, uus|natsit, yht|äkkiä*

4. -mA-, -in-, U-, i- ja O-loppuiset **verbimuodosteet**

(53) *el|in|kaarensa, hak|o|teillä, katu|ma|päälle, ui|ma|-altaat*

5. i-, o-, io- ja Us-**johdokset**

(54) *huht|i|kuun, kolm|i|osainen, nel|i|vuotiaasta, Yhd|ys|valloissa*

6. **s-aineksen** sisältävät sanat

(55) *herra|s|mies, pari|s|kunta, tiili|s|kivejä*

(Häkkinen 1990: 154–155, ISK § 416–417.) Jo määriteosatyyppin nimitys *yhdysosamuoto* ja sen kuvailu *epäitsenäinen* kertovat, mistä yhdyssanassa kielenhuollon näkökulmasta on kyse: koska nämä määritemuodot eivät voi esiintyä yksin, ne kirjoitetaan aina yhteen määritettävänsä kanssa. Sääntö on yksinkertainen, eikä siitä ole poikkeuksia.

Rajasin tämän ryhmän aineistossani äärimmäisen tiukasti niin, että jos yhdyssanan määriteosalla on minkäänlaista itsenäistä käyttöä, olkoon miten marginaalista tai jopa hiukan erimerkityksistäkin, ja jos yhdyssanan määriteosa sanakirjankin mukaan voi olla itsenäinen, en ottanut yhdyssanaa tähän luokkaan. Tämän ryhmän ulkopuolelle jäivät esimerkiksi siis ne yhdyssanat, joiden määriteosana on *ei-*, *elo*, *perus*, *syys*, *lempi*, *teko*, *pakko*, *keino*, *tekstitys*, *vähä* ja *vara*. Joistain sanoista sanakirja (MOT) tosin antaa kaksi merkitystä, joista toinen on itsenäinen sijoissa taipuva sana ja toinen on rajattu yhdyssanakäyttöön ja ilmoitetaan taipumattomaksi. Mielestäni tämä jako on kuitenkin hiukan keinotekoinen, sillä harvoin yhdyssanan määriteosa muutenkaan taipuu (vaikkei se ihan tavatonta olekaan, esimerkiksi *omatunto* : *omaatuntoa*). Pelkästään sanakirjan (MOT) mukaan tehtävä jako olisi tässä sikälikin jonkin verran horjuva, että siellä esiintyy paitsi sellaisia sanoja, jotka ovat ainoastaan yhdyssanan alkuosina, mutta joista ei ole itsenäistä hakusanaa (esimerkiksi MOT s.v. yhdyssana, yhdyshenkilö, yhdysrakenteinen **mutta ei yhdys-**) että sanoja, joista on hakusanaa sekä erikseen yhdysosamuoto (esimerkiksi MOT s.v. super- tai nyky-) että yhdyssanamuotoja (MOT s.v. superego, nykyaika). Selkeyden vuoksi tässä ryhmässä on **ainoastaan** niitä yhdyssanoja,

joiden määriteosa on **yksiselitteisen epäitsenäinen** eikä esiinny koskaan missään muussa tehtäväsä kuin yhdyssanan määrittäjänä.

Yhdyssanoja, joiden määriteosa on yhdysosamuotoinen, on aineistossani 904 eli 11,8 % kaikista yhdyssanoista. Odotuksenmukaista on, että koska näiden yhdyssanojen määriteosat eivät esiinny yksinään, niiden oikeinkirjoittaminen yhdyssanan osanakaan ei horju. Ne voi siis olettaa kirjoitettavan poikkeuksetta yhteen perusosan kanssa – ja näin melkein onkin. Esitän taulukossa 5. Yhdysosamääritteiset yhdyssanat määriteosan muodon mukaan kokoavasti tämän ryhmän määrät ja poikkeamat yhdysosan tyyppin mukaan.

Ryhmä	Esimerkki poikkeava konvention mukainen	Määrä	Poikkeamia	Poikkeama- prosentti
Vartalo ja prefiksit (1)	<i>nyky tilanne avotalot</i>	403	10	2,5
Konsonanttivartalo (2)	<i>uutis kuvassa erityisavustajan</i>	321	4	1,2
Lyhentymät (3)	<i>täys järkinen pienihiukkaset</i>	116	8	6,9
Verbiuodosteet (4)	<i>asuin rakennuksia haukkumanimillä</i>	34	1	2,9
Johdokset (5)	<i>huhtikuun</i>	24	–	–
S-aineokset (6)	<i>herrasmies</i>	6	–	–
Yhteensä kaikki		904	23	2,5

Taulukko 5. Yhdysosamääritteiset yhdyssanat määriteosan muodon mukaan

Taulukossa (5) ryhmä-sarakkeen alla on yhdysosamuodon tyyppi kuten edellisen sivun luettelussa, johon myös sen numerot viittaavat. Taulukko on lajiteltu yhdysosatyyppin mukaan määräjärjestykseen: eniten on niitä yhdyssanoja, joiden määrittävinä yhdysosamuotoina on sanavartaloita ja prefiksejä ja vähiten niitä yhdyssanoja, joiden määriteosa sisältää s-aineuksen. Kahden harvinaisimman ryhmän (5 ja 6) sanat ovat usein vanhaa perua, kuten niiden esimerkkisanatkin, eivätkä nämä ryhmät liene kovin karttuvia, kun taas frekventteimmissä ryhmissä määriteosina on sanoja, joihin liittämällä on mahdollista muodostaa jatkuvasti uusia ilmaisuja.

Poikkeamia tässä yhdyssanatyypissä on ainoastaan kaikkiaan 2,5 %:ssa yhdyssanoja (vrt. koko aineiston keskimääräinen poikkeamaprosentti 5,6 %). Prosentuaalisesti eniten poikkeamia on tehty lyhentymäsanojen tyypissä (ryhmä 3): kahdeksassa poikkeamasta kahdessa on kyse ellipsistä, jolloin pois jäänyt yhdysviiva aiheuttaa poikkeaman (esimerkit 56 ja 57),

(56) *sosiaali ja terveys budjetti*

(57) *sosiaali ja terveys palveluihin*

kahdessa poikkeamassa kyseessä on tavanomaista useampiosainen, yhdyssanan määrittämä yhdys-
sana (esimerkit 58 ja 59)

(58) *sinipuna vaihtoehdon*

(59) *sosiaali raha hanat*

ja yhdessä genetiivinkaltainen määriteosa on saattanut vaikeuttaa yhdyssanan hahmottamista (esi-
merkki 60).

(60) *vihon viimeisintä*

Vihon on lyhentymä alkuperäisestä sanasta *vihdoin* (NES: 1484). Loput kolme ovat esimerkit 61, 62
ja 63, joiden poikkeamalle lienee sekä mahdoton että tarpeeton arvella syytä.

(61) *sosiaali rahat,*

(62) *täys järkinen*

(63) *yht äkkiä*

Määrällisesti eniten poikkeamia on ensimmäisessä ryhmässä (Vartalot ja prefiksit). Näissä
poikkeamissakin kahdessa on kyse ellipsistä ja yhdysviivan puuttumisesta (esimerkit 64 ja 65).

(64) *On ex-taistolaisia ja yhä enemmän myös ex-talous ja arvoliberaaleja.*

(65) *Miksi keski ja suurituloiset yleensä saavat lapsilisiä,*

Yhdysviivan puuttuminen on poikkeaman syynä myös tähän ensimmäiseen muotoryhmään kuulu-
van esimerkin 66 ilmaisussa, jossa yhdysviivaa olisi tarvittu yhdysosien rajalla olevan saman vo-
kaalin *a* vuoksi.

(66) *eikös se oo sitä tasa arvoa just?*

Kolmen poikkeaman syitä, esimerkit 67, 68 ja 69 on jälleen vaikea edes arvailla: niissä ei ole tarvit-
tu yhdysviivoja eikä niissä ole useita osia.

(67) *keski osalla,*

(68) *super hyvästä*

(69) *ULKO MAALAISIA*

Peräti neljän yhdyssanan määräiteosana oleva *nyky-* on irtautunut perusosasta, esimerkit 70–73.

(70) *nyky meno*

(71) *Nyky muslimit*

(72) *nyky politiikkaan*

(73) *nyky tilanne*

On tietenkin mahdoton sanoa, mistä tällainen poikkeama voisi johtua, mutta yksi selitys voisi olla se, että *nyky-* mieltyy lyhennykseksi sanasta *nykyinen*, joka on itsenäinen ja kirjoitetaan pääsanastaan erilleen. Muotti toimii siis ikään kuin toiseen suuntaan: sen mukaan ei nyt yhdistetäkään uutta sanaa vaan erotetaan yhdyssana osiinsa. Sama selitys voisi päteä paitsi joihinkin tämän ryhmän muihinkin poikkeamiin (esimerkit 61–63) myös joihinkin nominatiivimuotoisten yhdyssanojen poikkeamiin (*kivi talo ~ kivinen talo*); pohdinkin tätä jakoa näillä esimerkeillä mutta adjektiivimaisuuden näkökulmasta jo luvun alkukappaleessa.

Merkillepantavaa on, että määrällisesti suuressa ryhmässä, konsonanttivartaloisten ryhmässä, on tehty huomattavan vähän poikkeamia: esimerkit 74–77.

(74) *erityis koulutuksessa*

(75) *juuriedellispäivänä*

(76) *näennäis asioista*

(77) *uutis kuvassa*

Neljästä poikkeamasta yksi on vieläpä aiheutunut muussa kohdassa kuin konsonanttivartalon ja perusosan rajalla, eli esimerkin 75 *juuriedellispäivänä* ei oikeastaan, tiukasti ottaen, poikkeamana kuulu edes tähän ryhmään, vaikka muuten toki poikkeama onkin. Verbimuodoste-ryhmässä on yksi poikkeama (esimerkki 78).

(78) *asuin rakennuksia*

Kahdessa pienimmässä ryhmässä, johdokset ja s-ainekset, poikkeamia ei ole tehty lainkaan.

Kun yhdyssanasta usein sanotaan, että se muodostaa yhtenäisen käsitteen eikä sen osia hahmoteta erillisiksi (esimerkiksi Häkkinen, 1990: 153), ainakin tässä ryhmässä, missä puolet yhdyssanasta tosiaan on ihan konkreettisesti hankala hahmottaa erilliseksi ilmeisen epäitsenäisyytensä vuoksi, väite pitänee paikkansa.

Taulukossa 6. Yhdysosamääritteiset yhdyssanat sanaluokittain esitän kokoavasti, miten yhdysosamääritteiset yhdyssanat ovat jakautuneet sanaluokkiin.

Sanaluokka	Esimerkki poikkeava konvention mukainen	Määrä	Poikkeamia	Poikkeama- prosentti
Adjektiivit		153	6	3,9
– (i)nen-johdos	<i>vihon viimeisintä</i> <i>elinkautinen</i>	117	4	3,4
– karitiivi	– <i>pitkäaikaistyöttömillä</i>	7	–	–
– partisippi	<i>epämiellyttävä</i> –	5	–	–
– perusadjektiivi	<i>super hyvältä</i> <i>superhauskaa</i>	24	2	8,3
Substantiivit		725	16	2,2
– perussana	<i>näennäis asioita</i> <i>asumistuen</i>	564	12	2,1
– erisnimet	– <i>Pohjois-Koreassa</i>	67	–	–
– tekijä ja väline nimi (1)	– <i>pienyrittäjille</i>	28	–	–
– teko ja tulos (2)	<i>nyky meno</i> <i>kansallislaulu</i>	66	4	6,1
– paikka ja tilaisuus (3)	– –	–	–	–
– partisippi	– –	–	–	–
Verbit		4	–	–
– finiittimuoto	– <i>epäonnistui</i>	2	–	–
– infinitiivi	– <i>etäosallistua</i>	2	–	–
– partisippi	– –	–	–	–
Muut	<i>yht äkkiä</i> <i>yhteiskunnallisesti</i>	22	1	4,5
	Yhteensä kaikki	904	23	2,5

Taulukko 6. Yhdysosamääritteiset yhdyssanat sanaluokittain

Eniten tässä ryhmässä on yhdyssubstantiiveja, kaikkiaan 725 kappaletta eli 80,2 % kaikista yhdysosa-alkuisista yhdyssanoista. Seuraavaksi eniten on adjektiiveja, 153 – 16,9 %. Verbit ja muut-ryhmän yhdyssanoja on marginaalisesti: yhteensä 26 yhdyssanaa (2,9 %) tämän ryhmän sanoista. Poikkeamia on tehty niin vähän, että niiden suhteelliset määrät eivät ole relevantteja kuin koko ryhmän osalta, ja kaikkiaan ryhmän poikkeamaprosentti jääkin reilusti alle koko aineistoa koskevan poikkeamien keskiarvon (5,6 %).

Esitän vielä taulukossa 7. Yhdysosamääritteisten yhdyssanojen kombinaatiot, miten yhdyssanat ovat rakentuneet sanaluokittain, toisin sanoen millaisia määriteosan ja perusosan sanaluokkakombinaatioita tässä ryhmässä on tehty ja miten ne jakautuvat.

Sana- luokka	MÄÄRITEOSA											
	Substantiivi		Adjektiivi		Numeraali		Pronomini		Muut		Yht.	
	n	%	n	%	n	%	n	%	n	%	n	%
S	206	22,8	367	40,6	19	2,1			133	14,7	725	80,2
A	19	2,1	86	9,5	8	0,9	3	0,3	37	4,1	153	16,9
V			1	0,1					3	0,3	4	0,4
M			7	0,8					15	1,7	22	2,4
									Yhteensä	904	100,0	

Sanaluokka:

S = substantiivi, A = adjektiivi, V = verbi, M = muu

Taulukko 7. Yhdyssanojen kombinaatiot yhdysosamääritteisten ryhmässä

Liitteessä 3 on lisäksi koottuna koko yhdysosamääritteisen yhdyssanaryhmän tyyppien määrät kombinaatioineen ja poikkeamineen.

4.3.2 Nominatiivimääritteiset yhdyssanat

Useiden lähteiden mukaan kaikista määritesyhdyssanoista nominatiivimääritteisiä on eniten (esim. ISK § 409, Saukkonen 1973: 337; Häkkinen 1990: 151). Näin on aineistossanikin: nominatiivialkuisia yhdyssanoja on 5111, eli 67 % kaikista aineiston yhdyssanoista. Nominatiivialkuisuus liittyy yhdyssanoissa osien väliseen assosiaatioon ja yhteys on tapauskohtaista (ISK § 409). Käytännössä tämä tarkoittaa sitä, että yhdyssanan tulkintaa ohjaa konteksti ja yleinen maailmantieto siitä, millaisissa suhteissa asiat ylipäätään voivat keskenään olla: ymmärrämme esimerkiksi, etteivät *hiirihaukka-* ja *poliisikoira-*yhdyssanojen osat ole samanlaisissa suhteissa molemmissa sanoissa (hiiri ei työskentele sen paremmin haukkana kuin koirakaan metsästä poliiseja). Osien vapaasta ja tapauskohtaisesta assosiaatiosta johtuen myös tilapäiset ja uudet yhdyssanat – joita voi aineistossani odottaa olevan runsaasti – ovat useammin nominatiivi- kuin genetiivialkuisia (tai muilla määritemuodolla alkavia).

Nominatiivimäärittäjä on useimmiten substantiivi, ja silloin yhdyssanan kirjoittaminen onkin yksinkertaista: nominatiivisubstantiivi kirjoitetaan yhteen määritettävän sanan kanssa paitsi silloin, kun kyseessä on erisnimeen liittyvä tittelityyppinen ilmaus (esimerkki 79) tai numeraali-ilmauksen taipumaton määrite (esimerkki 80) (ISK § 400; KK: 252, 254).

(79) *merirosvo Mustakorva*

(80) *Kylmäkoskentie 22*

Jos nominatiivimuotoinen määrittävä sana on adjektiivi (tai muu), on tarkasteltava sanojen kokonaisuutta: jos se on kiteytynyt tai termimäinen, esimerkiksi jonkin lajin nimi, kirjoitetaan sanat yhteen (esimerkki 81), muuten erilleen. Osa nominatiivadjektiivialkuisista kiteytymistäkin kirjoitetaan erilleen (esimerkiksi *harmaa talous*, KK: 255).

(81) *kalpeanaama*

Myös aineistoni nominatiivimääritteisistä yhdyssanoista valtaosassa eli 91 %:ssa, määrittäjä on substantiivi. Lisäksi määrittäjinä on adjektiiveja 7,3 %, numeraaleja 0,8 %, pronomineja 0,5 % ja muita 0,3 %. Nominatiivimääritteiset yhdyssanat jakautuvat sanaluokkiin seuraavasti: substantiiveja 90,7 %, adjektiiveja 8,2 %, muut-ryhmään ja verbeihin kuuluvia yhteensä 1,1 %. Poikkeamia on tehty kaikkiaan 250, poikkeamaprosentti on siten 4,9 %. Esittelen seuraavaksi nominatiivimääritteiset yhdyssanat ensin sanaluokittain, sitten käyn läpi niitä yhdyssanoja, jotka voi kirjoittaa sekä yhteen että erilleen ja lopuksi teen yleiskatsauksen tämän ryhmän yhdyssanatyyppeihin.

a) Nominatiivimääritteiset substantiivit

Nominatiivimääritteisiä yhdyssubstantiiveja on yhteensä 4638 kappaletta, tämän ryhmän kaikista sanoista 90,7 %. Poikkeamia näissä substantiiveissa on tehty 229:ssä, joten 4,9 % nominatiivimääritteisistä yhdyssubstantiiveista on kirjoitettu poikkeavasti. Poikkeamien määrää selittää varmasti osaltaan se, että poikkeavista yhdyssanoista 69 (hiukan vajaa kolmannes) on kompleksisia, eli niissä on enemmän kuin kaksi osaa. Kompleksisten yhdyssubstantiivien poikkeamista 12 on vieläpä koostunut neljästä tai useammasta osasta. Yhdyssubstantiiveja määrittämässä on nominatiivimuotoisia substantiiveja (NS+S-sanat: 4350 – yhdyssubstantiivien määrittäjistä 93,8 %), adjektiiveja (NA+S-sanat: 247 – 5,3 %), numeraaleja (NN+S-sanat: 15 – 0,3 %), pronomineja (NP+S-sanat: 10 – 0,2%) ja muita (NM+S-sanat: 16 – 0,3 %). Esitän seuraavaksi määritetyypit edellä olevassa järjestyksessä esimerkkeineen ja selkeyden vuoksi käytän sulkeissa esitettyjä lyhenteitä kunkin sanaryhmän kohdalla (ks. myös kuva 1 sivulta 40).

Nominatiivimääritteisten yhdyssubstantiivien määrittäjinä on siis huomattavasti eniten substantiiveja: kaikkiaan 93,8 % eli 4350 kappaletta tämän tyyppin yhdyssanoista on substantiivi-substantiivipareja (NS+S). Poikkeamia NS+S-ryhmässä on tehty kaikkiaan 215 sanassa, eli 4,9 % on kirjoitettu konvention vastaisesti siitä huolimatta että, kuten aiemmin todettu, tämän pitäisi olla helpon yhdyssanaksi mieltyvä ryhmä niiden yhdyssanojen jälkeen, joita määrittää epäitsenäinen yhdysosamuoto.

NS+S-sanoista 3213 kappaletta eli 73,9 % kaikista NS+S-sanoista on johtamattomia perussanoja, joissa poikkeamia on tehty 176 sanassa eli 5,5 %:ssa. (Esimerkit 82 ja 83.)

Poikkeama	Konvention mukainen
(82) <i>jäsen porukka</i>	(83) <i>villatöppöset</i>

827 kappaletta eli 19 % on NS+Sn-sanoja: verbistä johdettuja 2. ryhmään kuuluvia teko ja tulos -sanoja. Poikkemia tässä joukossa on 22 kappaletta, poikkeamaprosentti on 2,7 %. (Esimerkit 84 ja 85.)

Poikkeama	Konvention mukainen
(84) <i>pakko maantieto</i>	(85) <i>eläkekorotus</i>

200 kappaletta eli 4,6 % on NS+Se-sanoja eli erisnimiä, joissa poikkeamia on tehty 12 kertaa, poikkeamaprosentti on 6 %. (Esimerkit 86 ja 87.)

Poikkeama	Konvention mukainen
(86) <i>räksytjä Urpiainen</i>	(87) <i>kanaali-koskinen</i>

NS+S-sanoista 98 kappaletta eli 2,3 % on verbistä johdettuja 1. ryhmään kuuluvia tekijä ja väline -sanoja, NS+St-tyyppiä. Poikkeamia niissä on 5 kappaletta, poikkeamaprosentti on 5,1 %. (Esimerkit 88 ja 89.)

Poikkeama	Konvention mukainen
(88) <i>mies laulaja</i>	(89) <i>turkistuottajat</i>

Marginaalissa on vielä 7 partisiippiä eli NS+Sp-sanaa (esimerkki 90)

(90) *luottamustehtävä*

ja 5 verbistä johdettua 3. ryhmään kuuluvia tekopaikkaan viittaavia NS+Sl-sanaa (esimerkki 91),

(91) *ajatushautomon*

joissa kummassakaan tyypissä ei ole poikkeamia tehty.

Hankalin substantiivimääritteinen ryhmä kirjoittajille on siis ollut erisnimiryhmä, jonka poikkeamaprosenttiin vaikuttavat jo aiemmin todettu ongelmallinen yhdysviiva sekä muutama perussuomalaisen puolunimeen liittyvä hankaluus. Pohtia voi tietysti, että jos puolueen nimi¹³ *PerusS* muuttuu muotoon *Perus S*, kummassa sanassa oikeastaan on poikkeama: alkuperäisessä, toistetussa

13 Aineiston keräyksen aikaan perussuomalaiset nostivat nimekseen *PerusS*, ja aiheesta keskusteltiin jonkin verran perussuomalaisten Facebook-ryhmässä. Nyt näyttää siltä, että nimi on eriytynyt ainoastaan naisten ja nuorten käyttöön puolueen sisällä ja puolueen virallisena nimenä on *Perussuomalaiset*. (IV:PS)

muodossa vai jopa molemmissa. Jo työtä aloittaessa aprikoin, otanko erisnimiä mukaan aineistoon lainkaan. Mietin erisnimien yhdyssanamaisuutta enkä keksinyt yhtään syytä, miksi ne pitäisi rajata työn ulkopuolelle (samoin toimin muidenkin kiteytyneiden sanojen kohdalla, tästä myöhemmin vielä tässä luvussa enemmän). Voisi kuvitella, että erisnimissä ei tehtäisi – niiden erityisyyden ja referentiaalisuuden vuoksi – poikkeamia lainkaan, mutta niinpä vain esiintyvät tässäkin aineistossa muun muassa tähän ryhmään kuuluva *kirkko nummi* ja genetiivialkuisissa *Arkadian mäki* – tosin aivan harvalukuisia tämäntyyppiset (yhdysviivattomat) poikkeamat kuitenkin ovat.

Ongelmallisia kielenkäyttäjille ovat varmasti sellaiset sanat, jotka jo alun perin ovat hankalassa muodossa (kuten *PerusS*) tai jopa kielenhuollon suositusten vastaisia. Tällaisista sanoista malliksi käy tässä ryhmässä poikkeava esimerkki 92.

(92) *perus juhla mokkaa*

Esimerkin 92 sanayhdistelmässä, hämmentävää kyllä, jo tuotteen alkuperäinen nimi on konvention vastaisesti muodossa *Juhla Mokka*. Esimerkin korrekti muoto pitäisi kirjoittaa siis *perus- Juhla Mokkaa*, jos haluaa säilyttää (konvention vastaisen) alkuperäisnimen. Hämmästyttävän usein lukuisista mahdollisista vaihtoehdoista yrityksen tai tuotteen nimeksi – keulakuvaksi – tuntuu valikoituvan sellaisia sanoja, jotka ovat oikeinkirjoitukseltaan ja taivutukseltaan mutkikkaita. Samoin hiukan oudolta tuntuu se, että yhdysviivaa vastustetaan niin pontevasti, että mieluummin hajotetaan yhdys-sanat osiinsa kuin käytetään yhdysviivaa. Esimerkiksi *Juhla Moka*n suhteen näyttää käyneen niin, että alkuperäisen kahvimerkki *Juhlasekoituksen* tie *Juhla Mokaksi* on käynyt pakkauksen etiketin vuoksi: ei ole haluttu laittaa pakkaukseen yhdysviivaa (tavuviivaa kun sanat on jaettu kahdelle riville etiketissä), niinpä ilmeisesti tuotenimi on päädytty kirjoittamaan konvention vastaisesti kaikissa yhteyksissä – mielestäni aika nolosti ottaen huomioon erityisesti, että kyseessä on aivan yksinkertaisin yhdyssanatyyppe eli nominatiivisubstantiivimääritteinen yhdysubstantiivi (NS+S-sana). (Nimen alkuperäisestä kirjoitusasusta ks. paahtimon internetsivuilta: IV:JM.) Palaan tällaisiin jo lähtökohdaisesti norminvastaisiin ilmaisuihin vielä sanaliittoja koskevassa luvussa 4.5.1 Numeroilmaisut ja yhdysviiva (sivulla 90).

Vaikka perussanojen, erisnimien sekä verbistä johdettujen tekijä ja väline -ryhmän (johdosryhmä 1) poikkeamaprosentti onkin jonkin verran nominatiivimääritteisten ryhmän keskimääräistä poikkeamaprosenttia (4,9 %) korkeampi, ei se kovin huomattavasti poikkea koko aineiston keskimääräisestä poikkeamaprosentista (5,6 %). Huomattava on kuitenkin runsaslukuisen verbistä johdettujen teko ja tulos -ryhmän (johdosryhmä 2) pieni virheosuus: vain 2,7 % sanoista on konvention vastaisia. Pientä prosenttia selittänee parhaiten ryhmään kuuluvien usein esiintyvien ja ilmeisen kiteytyneiden sanojen esiintymät. Tällaisia ovat esimerkiksi konvention mukaiset: *mielipide(kysely)*

41 kappaletta, *hallitus*-alkuiset sanat 72 kappaletta, *puohue*-alkuiset sanat 47 kappaletta, *vaali*-alkuiset sanat niin ikään 47 kappaletta, muutamia mainitakseni. On hyvinkin mahdollista, että mitä useammin sana esiintyy, sitä vähemmän siinä tehdään poikkeamia. Nettifoorumeilla, kuten Facebookin ryhmissä, keskustelu saattaa pitkäänkin käsitellä samaa asiaa, aihetta koskevat ja mielipiteen muodostukseen liittyvät sanat toistuvat ja sen vuoksi malleja on kirjoitusympäristössä niin paljon, että ne ohjaavat koko yhteisön kirjoitustapaa. Oletus tällöin tietysti on, että konvention mukaan kirjoitettu muoto esiintyy useimmin kuin poikkeavasti kirjoitettu. Pohdin samaa aihetta myöhemmin genetiivialkuisen yhdyssanojen yhteydessä (sivulla 66), silloin nimenomaan poikkeamien suhteen.

Nominatiivialkuisista yhdyssubstantiiveista 247:ää määrittää adjektiivi, eli nominatiivimääritteistä yhdyssanoista 5,3 % on adjektiivi-substantiivipareja (NA+S-sanat). Kielenhuollon mukaan nominatiivimääritteinen adjektiivi kirjoitetaan yhteen yleensä sen määrittämän substantiivin kanssa, jos ne yhdessä muodostavat erikoistuneen käsitteen nimityksen (KK: 254). Osa termimäisistäkin NA+S-sanoista kirjoitetaan vakiintuneesti kuitenkin erilleen, esimerkiksi *harmaa talous* ja *kevyt liikenne* (KK: 255). Kielenhuolto antaa ohjeita lähinnä kirjoittajalle, mutta jo olemassa olevan kielen tutkijalla on myös muita keinoja tarkastella yhdyssanan olemusta ja jopa jonkin verran kirjoittajan intentioita sen suhteen. Nimittäin yhdyssanapoikkeamia tarkastellessa on hyvä huomata tekstiyhteydestä seuraava seikka: jos määrittäjänä oleva adjektiivi **ei ole taipunut** pääsanansa kanssa, kyse on yhdyssanasta. Kongruoiva adjektiivikin voi toki olla yhdyssanan osana (esimerkiksi sanassa *isoaveljeä*), mutta kongruoimaton adjektiiviattribuutti (**iso veljeä*) on suomessa mahdottomuus, joten jos adjektiiviattribuutti on kongruoimaton, kyse on yhdyssanasta ja erilleen kirjoitettuna yhdyssanapoikkeamasta. Tilanne on eri taipumattomien adjektiivien osalta (*pikku poikaa*); taipumattomat adjektiivit eivät kuulukaan tässä tutkielmassa nominatiivimääritteisten yhdyssanojen ryhmään vaan ryhmään muut määritetyypit (ks. luku 4.3.4 Muut määritetyypit).

NA+S-sanoissa poikkeamia on kaikkiaan tehty 14, poikkeamaprosentti on siten 5,7 %, ja se on jonkin verran enemmän kuin koko nominatiivimääritteisten ryhmän poikkeamaprosentti (4,9 %), mutta likimain sama kuin koko aineiston poikkeamaprosentti (5,6 %). Eniten poikkeamia (esimerkiksi 93) on tehty NA+S-perussanojen ryhmässä (esimerkki 94), 10 poikkeamaa kaikkiaan 136 sanassa eli 7,4 %).

Poikkeama

(93) *lume demokratiaa*

Konvention mukainen

(94) *laajakaista*

NA+Se- eli erisnimiryhmässä on kaikkiaan 37 sanaa, esimerkit: poikkeamat 95 ja 97 sekä konvention mukainen 96. NA+Sn- eli teko ja tulos -johdosryhmässä on 52 sanaa, esimerkit poikkeavat 98

ja 100 sekä konvention mukainen 99. Kummassakin näissä ryhmässä on vain kaksi edelläkin mainittua poikkeavaa ilmausta, jotka kaikki on esitetty seuraavaksi.

Poikkeama	Konvention mukainen
(95) <i>kiva Jyrkiä</i>	(96) <i>Vapaavuori</i>
(97) <i>kokovartalo marissa</i>	
(98) <i>normaali verotuksessa</i>	(99) <i>normiuutisoitua</i>
(100) <i>pakko maantieto</i>	

NA+St-sanojen (tekijä ja väline) ryhmässä sanoja on 22, eikä niissä ole yhtään poikkeamaa, esimerkki 101.

(101) *mikroyrittäjät*

NA+Sl-sanojen (paikka ja tilaisuus) ryhmässä sekä partisiipeissa ei ole adjektiivimääritteisiä sanoja lainkaan.

NA+Sn- ja NA+Se-sanoissa tehtyjä virheitä on niin vähän, että niiden pohtiminen ei liene tarpeen, mutta perussanoissa (NA+S) tehtyjä kirjoitusvirheitä voi hiukan tarkastella. Kaikkiaan kymmenestä poikkeamasta viittä selittää se, että niissä on useampi kuin kaksi osaa ja näistä neljässä olisi tarvittu yhdysviivaa, joka on jäänyt pois (esimerkit 103–105). Esitän tässä poikkeuksellisesti selvyyden vuoksi myös konvention mukaiset variantit, jotka eivät siis esiinny aineistossani. Sulkeissa olevat sanat ovat jääneet elliptisen ilmaisuuden toiselle puolelle.

Poikkeama	Konvention mukainen
(102) <i>(lume) sananvapautta</i>	<i>-sananvapautta</i>
(103) <i>normi Aku tarinoihin</i>	<i>normi-Aku-tarinoihin</i>
(104) <i>ORIGINAL-JARE FANI</i>	<i>ORIGINAL-JARE-FANI</i>
(105) <i>pakko (työvoiman)</i>	<i>pakko-</i>

Kahdessa esimerkissä (102 ja 105) sulkeissa oleva sana on elliptisen ilmauksen toisto-osa, eli poikkeama on syntynyt siitä, ettei perusosaa esimerkissä 102 ja määriteosaa esimerkissä 105 ole sidottu yhdysviivalla yhtenäiseksi ilmaukseksi. Useaosaisissa yhdyssanoissa on jo yksi *pakko*-alkuinen ilmaus (105), ja niitä on kaksi lisää kaksiosaisien sanojen joukossa

(106) *pakko matematiikka*

(107) *pakko ruotsikin*

Voi tietysti ensinnä pohtia, onko *pakko*- adjektiivi lainkaan. Tässä yhteydessä, ja varmasti yhdyssanojen alkuosana muulloinkin, sen luokittelu adjektiiviksi on mielestäni perusteltua: se kuvailee, millaisesta perusosasta on kyse eli 'pakollisesta'. Toiseksi voi pohtia, miksi se pitää kirjoittaa yhteen määritettävänsä kanssa. Jos *pakko*- mielletään adjektiiviksi, ei se ole kongruoitunut (esimerkeissä 105–107, **pakkoa matematiikkaa* ei tunnu mahdolliselta muutenkaan); jos *pakko*- taas kuitenkin ajatellaan substantiiviksi, on se nominatiivimuotoisena joka tapauksessa kirjoitettava yhteen määritettävänsä kanssa. Analysoidaanpa se siis kummaksi hyvänsä, kuuluu se kiinteästi yhteen seuraavan sanan kanssa. Loput kolme poikkeavaa NA+S-sanaa ovat: *LEMPI ARTISTI*, *lume demokratiaa*, *populaari jutun*. Näistä *lume demokratiaa* liittyy edellä esitettyyn useaosaisten yhdyssanojen poikkeamaan, ja koko ilmaus kontekstissaan on:

(108) *Tuntuu että kiristetään, kiristetään vaan... Lume demokratiaa ja sananvapautta. Miten kauan Suomen kansa kestää???*

Lempi- on sanakirjan mukaan taipumaton yhdyssanan alkuosa (MOT s.v. lempi) ja *lume* sekä *populaari* ovat kongruoimattomia; nämä ovat siis perusteina poikeama-analyysille. *Lume*-sanalla on hyvä tilaisuus huomauttaa, että vaikka aineistoni onkin (yhden ihmisen käsiteltäväksi tällaisen tutkimuksen puitteissa) melko iso, marginaalisimmissa ilmaisuissa mukaan pyrkii väkisinkin yksilönkielen vaikutus. Jos poikkeamia on vain muutama ja jos niistä kaksikin on tehnyt sama kielenopas, ei sanan tai sanatyypin yleisestä taipumuksesta synnyttää poikkeamia voida oikein sanoa mitään yleispätevää. Tämä on tilanne paitsi *lume*-sanalla myös edellä käsitellyn *pakko*-sanalla suhteen: koko aineiston 24 *pakko*-alkuisesta sanasta viisi on kirjoitettu poikkeavasti. Poikkeamista neljä on tehnyt sama henkilö, joten on tarpeetonta pohtia tällä otannalla, onko *pakko*-sanassa itsessään tai sen käyttökontekstissa jotain sellaista, mikä irrottaa sen erilleen määritettävästään. Ilmiselvä tilanne yksinkertaisesti on se, että **eräs**, **yksi** tai **joku** kielenopas jostain syystä, jota emme mitenkään ainkaan luotettavasti voi päätellä, mieltää määritteen itsenäiseksi sanaksi ja kirjoittaa sen erilleen määritettävästään.

Loput nominatiivimääritteiset yhdyssubstantiivit ovat konvention mukaiset, eli kaikki ne yhdyssubstantiivit, joita määrittää nominatiivimuodossa oleva numeraali (NN+S), pronomini (NP+S) tai muun sanaluokan sana (NM+S), on kirjoitettu kielenhuollon sääntöjen mukaan konventionaalisesti. Näistä NN+S-sanoja on 15 kappaletta, ja ne jakautuvat niin, että 6 on perussubstantiivia (NN+S, esimerkki 109), 8 erisnimeä (NN+Se, esimerkki 110) ja 1 teko ja tulos -ryhmään kuuluva verbikantainen johdos (NN+Sn, esimerkki 111).

(109) *yksimielisyyteen*

(110) *Satakunnan*

(111) *kymppikevennysten*

NP+S-sanoja eli nominatiivipronominimääritteisiä yhdyssubstantiiveja on 10: perussanoja 8 (NP+S, esimerkki 112) ja teko ja tulos -ryhmän johdoksia 2 (NP+Sn, esimerkki 113).

(112) *monikulttuurisuutta*

(113) *itsetunto*

Muut-ryhmään eli NM+S-sanoihin analysoitui 16 sanaa: yksi NM+S-perussana (esimerkki 114), 14 NM+Se-sanaa (erisnimi, esimerkki 115) ja yksi NM+St (tekijä ja väline, esimerkki 116).

(114) *exit-strategia*

(115) *Iisalmeen*

(116) *polkupyöräilijöille*

Kaikkien nominatiivialkuisten yhdyssubstantiivien sanaryhmässä on yhteensä 35 sellaista yhdysanaa, jotka voi kirjoittaa joko yhteen tai erilleen (esimerkit 117–120).

(117) *normaali kilpailu*

(118) *nuorimies*

(119) *rikoksente kohetkellä*

(120) *Tosi leikkiä*

Verrannolliset ovat sanat *normaali* ja *tosi*, jotka voidaan käsittää joko adjektiiviksi tai substantiiviksi (MOT s.v. *normaali*; MOT s.v. *tosi*). Jos ne ajatellaan adjektiiveiksi eikä pidetä ilmauksia kiteytyneinä termeinä, ovat aineistossa olevat muodot *normaali kilpailu* ja *tosi leikkiä* konvention mukaiset; jos ne käsitetään substantiiveiksi, olisivat edellä todetun mukaan vain muodot *normaalikilpailu* ja *tosileikkiä* hyväksyttävät. *Normaalin* substantiivisuus ei oikeastaan voi tässä toteutua: kyseessä ei varmastikaan ole sanakirjan antama määritys *normaali*-substantiiville eli "toista suoraa t. tasoa vastaan kohtisuorassa oleva suora, kohtisuora", mutta se ei muuta tilannetta, sillä jo adjektiivitulkinalla pelkästään on molempien muotojen kirjoitusmahdollisuus, jos kirjoittaja päättää, että kyseessä on kiteymä. Koska määritettävä sanakin on nominatiivissa, kongruenssista ei voi tässä päätellä mitään.

Nuorimies viittaa yleisesti nuorukaiseen, *nuori mies* taas mieheen, joka on (erityisesti) nuori. MOT (s.v. *nuorimies*) ilmoittaa myös molemmat vaihtoehdot hyväksyttäväksi, vaikka ei teekään eroa merkitysten välille. Intuitiivisesti kuitenkin tuntuu, että voimme sanoa *17-vuotias nuorimies*

siinä kuin *17-vuotias nuorukainen* mutta *?17-vuotias nuori mies* tuntunee liikatoistolta. Esimerkin 121 kommentista on kuitenkin mahdoton päätellä, mikä on ollut kirjoittajan tarkoitus eikä liikatoistokaan tule vastaan, eli molemmat vaihtoehdot ovat hyväksyttävät.

(121) *Tainnu ottaa nuorimies Katainen letkautuksissaan mallia Herra Hakkaraiselta – –*

Rikoksentekehoketkellä on useampiosaisena yhdyssubstantiivina sikäli erityislaatuinen, että sen voi ajatella jakautuvan joko [*rikoksenteke*][*hetkellä*] tai [*rikoksen*][*tekehoketkellä*]. Jos päädytään jälkimmäiseen jakoon, kyseessä onkin genetiivimääritteinen yhdyssana, joka voitaisiin kirjoittaa myös *rikoksen tekehoketki*. Mielestäni tässä kontekstissa tuntuu loogisemmalta, että kyse on hetkestä, jossa on rikoksenteke (johon sijoittuu rikoksenteke, jossa tehdään rikos), ei tekehoketkestä jonka osana on rikos saati jonka omistaa rikos; *rikos* ei myöskään oikein tunnu sopivan *tekehoketken* objektiksi eikä subjektiksi (genetiivimäärittäjän tehtävästä ja tulkinnasta myöhemmin lisää luvussa 4.3.3 Genetiivi-alkuiset yhdyssanat). Kokonaisuudessaan lause on seuraava:

(122) *Mun mielestä rikoksen tehneet on luopununeet rikoksentekehoketkellä ihmisoikeuksistaan.*

On kuitenkin jätettävä sija toisellekin analyysille, vallankin kun se on täysin mahdollinen, joten siksi tämä sana sijoittuu ryhmään, jossa erilleen kirjoittaminen on mahdollista. Kaksi ilmaisua koskevat uutta vuotta tai uuttavuotta (esimerkit 123 ja 124):

(123) *mahtavaaa cheek uusvuas alkas jymeest jeeee....:)))*

(124) *Jee jee! Hyvää uuttavuotta :)*

Uusi vuosi -sanapari kirjoitetaan erilleen silloin, kun on kyse koko seuraavasta vuodesta, ja yhteen silloin, kun puhutaan vuodenvaihteen juhlapäivästä (SK: 192, MOT s.v. uusivuosi). Esimerkeissä voisi olla kyse koko seuraavasta vuodestakin, jolloin ilmaukset pitäisi kirjoittaa erilleen. Koska tarkoitusta ei voi tekstistä päätellä, ei näitä voi pitää poikkeamina. Nostan tässä esiin selvän poikkeaman (esimerkki 125), jotta kävisi ilmi, miten tämän tyyppin poikkeaman voi nähdä.

(125) *Onnee Biisistä ja Onnellista Uuttavuotta 2011!*

Kun *uuttavuotta* määrittää vuosilukua, koskee se silloin väistämättä koko vuotta, eli esimerkin 125 tapauksessa konvention mukainen ilmaus olisi – – *uutta vuotta 2011!* *Uuttavuotta* on myös hyvä esimerkki yhdyssanakongruenssista: siinä määriteosa on tosin partitiivissa, mutta koska se on kongruoitunut partitiivimuotoisen perusosan kanssa, on kyseessä kuitenkin nominatiiviadjektiivi. Tällaisia sanoja on aineistossani muitakin, esimerkiksi *jommankumman* ja *omastatunnosta*.

Loput 29 joko yhteen tai erilleen kirjoitettavista sanapareista ovat *tervetuloo*-toivotuksia (yksi tosin muodossa *teretuloo*), josta sanakirja (MOT s.v. tervetuloo) antaa vaihtoehdoksi myös erilleen

kirjoitetun version; aineistossani kaikki *tervetuloa*-ilmaisut on kuitenkin kirjoitettu yhdyssanaksi. Ilmaukset, jotka kielenhuollonkin mukaan voidaan kirjoittaa yhteen tai erilleen, ovat luonnollisesti konvention mukaisia, kirjoitetaan ne kummin hyvänsä, eli ne eivät kuulu poikkeamien joukkoon tässäkään. Jos kaikki esiintymät lasketaan, on erilleen ja yhteen kirjoitettujen sanaparien suhde 2 : 27, mutta jos vähennetään usein toistuvien sanojen vaikutusta ja kukin määritemuoto lasketaan vain kerran, on erilleen ja yhteen kirjoitettujen sanaparien suhde tässä ryhmässä 2 : 4.

b) Nominatiivimääritteiset adjektiivit

Nominatiivinen määriteosa on kaikkiaan 417 yhdysadjektiivissa. Kaikista nominatiivialkuisista yhdyssanoista adjektiivien määrä on 8,2 %. Tässä ryhmässä on tehty kaikkiaan 17 poikkeamaa ja poikkeamaprosentti on siten 4,1 % eli hiukan vähemmän kuin koko ryhmän poikkeamaprosentti (4,9 %) ja koko aineiston keskimääräinen poikkeamaprosentti (5,6 %).

Nominatiivialkuisia yhdysadjektiivejakin määrittää useimmiten substantiivi. Substantiiviadjektiivipareja (NS+A) on 269, ja niiden osuus nominatiivimääritteisistä yhdysadjektiiveista on 64,5 %. Poikkeamia NS+A-sanoissa on kymmenen, poikkeamaprosentti on melko alhainen 3,7 %. NS+A-sanat jakautuvat tyypeiksi seuraavasti: 17 näistä on perusadjektiivia, niistä poikkeama on neljässä. (NS+A-esimerkit: 126 ja 127).

Poikkeava

(126) *sika hieno*

Konvention mukainen

(127) *arvoliberaaleja*

NS+Ai-sanoja (inen-johdoksia) on 240, poikkeama on kuudessa. (NS+Ai-esimerkit: 128 ja 129).

Poikkeava

(128) *karja tilalliset*

Konvention mukainen

(129) *alkuperäinen*

Karitiiveja eli NS+Ak-ryhmän sanoja on kahdeksan (esimerkki 130) ja partisiippeja eli NS+Ap-sanoja neljä (esimerkki 131), kummassakaan tyypissä ei ole poikkeamia.

(130) *ammattitaidottomia*

(131) *koulukiusatun*

Nominatiivialkuisia adjektiiviadjektiivipareja (NA+A) on 116 eli 27,8 % nominatiivialkuisista yhdysadjektiiveista on NA+A-sanoja. Poikkeamia näissä on tehty kuusi, ja ne kaikki on tehty NA+Ai-sanoissa (inen-johdoksissa), joita onkin valtaosa (110 – 94,8 %) NA+A-sanoista (esimerkit 132 ja 133).

Poikkeama(132) *pieni palkkaiset***Konvention mukainen**(133) *hyvätuloinen*

Kolme NA+A-sanaa on perusadjektiiveja (esimerkki 134),

(134) *MUSTAHOPEA*

ja partisiippeja on kolme (NA+Ap, esimerkki 135),

(135) *tervetulleiksi*

eikä kummassakaan ryhmässä ole siis tehty poikkeamia. NA+Ak-sanoja (karitiiveja) ei ole.

Nominatiivialkuisista yhdysadjektiiveista 20 on numeraalimääritteistä eli NN+A-sanaa (4,8 %). Näistä kaksi on konvention mukaista NN+A-perussanaa (esimerkki 136) ja 18 NN+Ai-sanaa, joissa yksi poikkeama (esimerkit 137 ja 138).

Poikkeama(137) *yksi selitteinen***Konvention mukainen**(136) *parivuotias*(138) *kaksipiippuinen*

Karitiiveja ja partisiippeja ei NN+A-sanoissa ole.

12 nominatiivialkuista yhdysadjektiivia alkaa pronominilla (NP+A-sanat). Näistä yhdeksän on NP+Ai-sanoja (esimerkki 139) ja kolme NP+Ap-sanoja (esimerkki 140); poikkeamia ei NP+A-sanoissa ole lainkaan.

(139) *moniarvoinen*(140) *itsejauhettuna*

Muut määrittäjät -ryhmässä (NM+A) ei ole sanoja.

Kolme nominatiivimääritteisistä yhdysadjektiiveista on sellaisia, että niiden osat voisi kielenhuollon mukaan kirjoittaa joko yhteen tai erilleen. Ne kaikki ovat *itse*-pronominalkuisia partisiippeja (NP+Ap): *itseaiheuttamansa*, *itsejauhettuna* ja *itsemaksavat*. Kielenhuollon mukaan partisiippi voidaan kirjoittaa yhteen edeltävän sanan kanssa, jos kokonaisuuden merkitys on kiteytynyt tai muuttunut kuvalliseksi (KO: 261). Ilmeisesti näin on kirjoittajan mielestä näissä sanoissa käynyt, eikä nähdäkseni ole mitään syytä väittää toisinkaan – myös oma intuitioni sanoo, että nämä voivat kaikki olla yhdyssanoja.

Nominatiivimääritteisiä yhdysadjektiiveja kaikkiaan ja niissä tehtyjä poikkeamia on sen verran vähän, että on mahdoton vetää johtopäätöksiä siitä, mitkä nominatiivialkuiset yhdysadjektiivi-

tyypit olisivat kirjoittajille muita vaikeampia – oikeastaan voisi todeta melkeinpä päinvastaisesta näkökulmasta: poikkeamaprosentti on sen verran pieni (4,1 % poikkeamia nominatiivialkuisista yhdysadjektiiveista), että yleisesti ottaen voi sanoa nominatiivialkuiset yhdysadjektiivit hallittavan varsin hyvin. Prosentuaalisesti eniten poikkeamia on tehty johtamattomien perusadjektiivien ryhmässä, mutta koska siinäkin poikkeama : kaikki -suhde muodostuu vähäisistä esiintymistä 4 : 22, ei johtopäätöksiä voi vetää. Toisaalta taas suurilukuisin nominatiivialkuisen yhdysadjektiivien ryhmä on selkeästi inen-johdokset (NX+Ai), joiden osuus nominatiivialkuisista yhdysadjektiiveista on 90,4 % (377 kappaletta) ja poikkeamaprosentti vain 3,4 %. Se kertonee paitsi, että inen-adjektiivi hakeutuu usein yhdyssanaksi myös että sellaisena sen oikeinkirjoitus hallitaan hyvin.

c) Nominatiivimääritteiset verbit ja muut marginaaliset esiintymät

Verbejä nominatiivimääritteisten yhdyssanojen ryhmässä on ainoastaan kolme. Ne ovat kaikki infinitiivisiä: kaksi infinitiiviä ja yksi partisiippi. Infinitiivit (esimerkit 141 ja 142) osallistuvat verbiketjuun, joten ne ovat predikaatin tehtävässä ja siten selkeästi verbejä. Verbejä määrittävät substantiivit (NS+Ve-esimerkki 142 ja NS+Vp-esimerkki 143) sekä adjektiivi (NA+Ve, esimerkki 141).

(141) *Miksi täälläkin joutuu pakkolukemaan Perssujen absurdin sairasta propagandaa?*

(142) *Lupaani rullaluistella helsingistä Äänekoskelle(lodelle) nyt tulevana kesänä.*

(143) *Käsikirjoittama.*

Viimeinen esimerkki (143) on tosiaan yksinäinen sana kommentissa, ja se on vastaus edelliseen kommenttiin: – – *siinä on kuullemma suomalaisen piirtämä tarina* – –. Kontekstissaan sen merkitys siis on 'on suomalaisen käsikirjoittama tarina'. Agenttipartisiippina se on niin verbimäinen (vaatii subjektin ja objektin), että katson sen kuuluvaksi verbien joukkoon, vaikka se onkin adjektiivin tavoin (ellipsin) attribuuttina. Nominatiivimääritteisten adjektiivien joukkoon sen sijaan olen laskenut refleksiivisen agenttipartisiipin *itseaiheuttamansa*, joka sisältää tekijän jo itsessään ja mieltyy siksi itsenäisempänä adjektiivimaiseksi ja, kuten jo aiemmin adjektiivien yhteydessä totesin, verbeistä adjektiivien ryhmään olen analysoinut muutaman muunkin partisiipin (*itsejauhettuna, itsemaksavat, koulukiusattu, tervetullutta*), samoin substantiiveihin kaksi (*ristisanatehtävä, verkkopalveluvastava*). Kriteerinä on ollut niiden nominaalinen käyttö lauseenjäsenenä ja etenkin substantiivien ryhmässä jo aivan selvä leksikaalistuma, mitkä syyt erottavat ne verbien tehtävästä.

Jäännösryhmässä muut on 53 yhdyssanaa. Niiden määrittäjänä on 31 substantiivina (NS+M-sanat), 8 adjektiivina (NA+M-sanat), 8 numeraalia (NN+M-sanat) ja 6 pronominia (NP+M-sanat). Aivan viimeisin mahdollinen ryhmä nominatiivimääritteisille sekalaisille esiintymille, eli muut nominatiiviset määrittäjät määrittävät muu-sanaluokan yhdyssanoja (NM+M-sanat) on tyhjä, toisin sa-

noen kaikki nominatiivialkuiset yhdyssanat oli mahdollista sijoittaa jommankumman osansa (määrittäjän tai perusosan) mukaan johonkin sanaluokkaryhmään.

Nominatiivialkuisten muut yhdyssanat -ryhmässä on 36 adverbia, 14 numeraalia ja 2 pronominia; pronomineina ryhmässä ovat *jompikumpi* ja sen poikkeava muoto *jomman kumman*. -*sti*-johdoksella voisi muodostaa adverbien myös substantiivista (ISK § 373), mutta aineistoni kaikki 24 -*sti*-johdoksella muodostettua yhdysadverbia on johdettu inen-adjektiivista – ja ne ovat konvention mukaisia (esimerkit 144–147):

(144) *henkilökohtaisesti*

(145) *lähtökohtaisesti*

(146) *tiukkapipoisesti*

(147) *vaihtoehtoisesti*

Muut-ryhmässä on kaikkiaan neljä poikkeamaa: *vuosi kymmeniä*, *sata nolla*, jo mainittu *jomman kumman* sekä ryhmän ainoa useampi kuin kaksiosainen yhdyssana, jonka perusosana on sanaliitto ja jonka sanaluokkaa on siksi mahdoton sanoa (esimerkki 148).

(148) *pakko sitä sun tätä*

Poikkeamien joukko on niin pieni ja sekalainen, että sen perusteella niiden syistä ei voi päätellä mitään.

d) Nominatiivimääritteisten yhdyssanojen yhdyssanatyypit

Nominatiivimääritteisistä yhdyssanoista: 4 on yhdysnumeraaleja, 87 on summayhdyssanoja, 65 on epäselviä (joista selostus seuraavassa kappaleessa) ja loput 4955 on määritisyhdyssanoja. Summayhdyssanojen joukkoa kasvattaa erityisesti 69 *maailma*-sanaa. Kaikki yhdysnumeraalit on kirjoitettu konvention mukaisesti, eikä summayhdyssanoissakaan ole kuin kaksi poikkeamaa: *vankila hotellien* ja *perushoitaja yksinhuoltaja*.

Kun ollaan luonnollisen kielen ryhmittelyn kanssa tekemisissä, tulee väistämättä eteen epäselviä esiintymiä, joista on hankalaa sanoa, mihin ryhmään ne oikeastaan kuuluvat. Nominatiivialkuisissa yhdyssanoissa on pienehkö joukko (kaikkiaan 65, eri ilmaisuja 20) sanoja, jotka eivät tunnu oikein sujuvasti sopivan sen paremmin summa- kuin määritisyhdyssanojenkaan joukkoon. Otan tässä muutaman perusteluineen esimerkiksi. Runsaslukuisin yksittäinen sana (42 esiintymää), joka aiheuttaa hankaluutta tässä suhteessa, on *vaihtoehto*. Se ei oikein voi olla määritisyhdyssana: *vaihto* ei määritä *ehto*a eikä toisinpäin; se ei myöskään liene summayhdyssana: sekä *vaihto* että *ehto*. Molemmien osan merkitykset kuitenkin sanassa ovat jollain lailla mukana: *ehto* niin että se vaatii

jonkinlaiset perusteet (vaihtoehdon pitää sopia aiempaan kontekstiin, esimerkiksi kultakalan tilalle on saatava 'vedessä elävä eläin', eikä siten esimerkiksi 'kirjallisuusteoriaa' voida edes harkita tähän kohtaan) ja *vaihto* niin että jotain voi vaihtaa (edelleen samassa kontekstissa, samoin ehdoin johonkin toiseen: vaihdetaan kultakala lehtikalaan). Samalla tavalla merkitystä kantavat useaosaiset nimet: *Aalto-Setälä*¹⁴, *Annastiina*, *ViikkoVapaus*. Molempia osia tarvitaan, että entiteettiä kuvattaisiin kokonaan, mutta ei silti voine sanoa, että *Setälä* määrittäisi *Aaltoa* tai toisinpäin, tai että henkilö on sekä *Aalto* että *Setälä* samaan tapaan kuin hän voi summatiivisesti olla *parturi* ja *kampaaja*. Nimiä tässä joukossa on 19, mutta huomattavaa on, että kaikki nimet eivät kuulu tähän joukkoon. Esimerkiksi *Kiviniemi* tai *Halla-aho* sopivat sujuvasti määrittämissanojen ryhmään: *kivi* kuvaa millaisesta *niemestä* on kyse; *aholla* lienee usein *halla* – vähintään teoriassa tämän tyyppin nimet voisivat olla myös yleiskäyttöisiä substantiiveja. Loput yhdyssanatyypiltään erityisryhmään sijoittamani nominatiivialkuiset yhdyssanat ovat *meikämanne*, *kuusnolla* ja *sata-nolla* (2 esiintymää). *Meikämanne* lie monen kautta muodostunut leikillisehkö ilmaus; *mei-käläinen* ilmaisuun viitataan Nykysuomen etymologisessa sanakirjassakin (NES: 1260) vain lyhyesti johtimen kantana, enkä muualtakaan sen alkuperästä saanut tietoa, joten sanan alkuperä jää tässä hämärän peittoon. *Kuusnolla* ja *sata-nolla* saattavat äkkiä katsoen vaikuttaa olevan numeraalisia ääripää- tai tulosilmausjua, mutta näin ei kuitenkaan (tässä) ole. *Kuusnolla* tarkoittaa jokseenkin samaa kuin *ylivoimainen*, *ylivoimaisesti* (ei sanakirjasana; oma, kokemukseen pohjautuva tieto). Tässä yhteydessä *sata-nolla* on vastaus *kuusnolla*-ilmaisuun ja tarkoituksena on korostaa ylivoimaisuutta entisestään. Esimerkit 149–151 ovat peräkkäin ja liittyvät toisiinsa.

(149) *Cheek voittaa kaikki kuusnolla! : D <3*

(150) *oikeestaa sata nolla :) <3*

(151) *<3 <3 eiku sata-nolla ! <3*

Näiden ilmaisujen alkuperänä on tietysti tulosilmaus: jos esimerkiksi jokin ottelu päättyy kuusi–nolla (6–0), voi voittajan arvioida olleen melko ylivoimainen. Mielestäni ilmaus *kuusnolla* on jo niin vakiintunut, ettei sitä voida enää käsittää muuna kuin adverbileksikaalistumana. Tässä on myös elävä esimerkki analogiasta tai muotista, josta oli puhe jo luvussa 2.3 Yhdyssanan kehyksiä jne.: *kuusnolla*-muotin mukaan voidaan ainakin samassa kontekstissa tehdä siitä itsestään komparatiivi *sata-nolla*.

Nostin nämä erityiset nominatiivialkuiset yhdyssanat esille yhtäältä sen vuoksi, että kävisi ilmi, millaisia variaatioita yhdyssanatyypeissä voi olla ja toisaalta esimerkiksi siitä, millaisia hankaluuksia hyvin tarkkarajaiset kielenkuvausmallit voivat aiheuttaa, etenkin jos elävä kieli halutaan ai-

14 Kyseessä on nimi, joka ei muodostu aiemmasta ja avioliiton kautta saadusta nimestä, jollainen voisikin mieltä summayhdyssanaksi: henkilö voi olla sekä yhtä että toista sukua (samaan tapaan kuin olla *parturi-kampaaja*). Tässä tietääkseni *Aalto-Setälä* on yhden suvun nimi.

van jäännöksettömästi lajitella ryhmiin. ISK:n mukaan määrittelyyhdyssanat ovat endosentrisiä (eli koko yhdyssana tarkoittaa samaa kuin sen perusosa), mutta bahuviiri, joka kuitenkin on määrittelyyhdyssanan alatyyppejä, on sen mukaan eksosentrisen, eli muodostuu molemmista osistaan eikä toinen edusta toista (ISK § 408). Samaan tapaan *vaihtoehto*, *Aalto-Setälä* ja *satanolla* edustavat kokonaisuutena tarkoitetaan, mutta eivät silti rakenna sitä ikään kuin irtonaisista osista, kuten summayhdyssana rakentuu (*maa + ilma = maailma*, *parturi+kampaaja = parturi-kampaaja*, kahden eri ammattin taitaja). Esitänkin, että määrittelyyhdyssanoissa on myös muita eksosentrisiä tyyppisiä kuin mainittu bahuviiri ja edellä käsittelemäni sanat edustavat tätä tyyppiä.

Tässä yhteydessä ei kuitenkaan ole tarpeen tämän enempää lähteä analysoimaan yhdyssanatyyppejä, kuten alussa (luvussa 2.1 Yhdyssana per se) jo totesinkin. Nämä tyyppiltään epäselvinä pitämäni yhdyssanat lopulta kuitenkin lajittelin perinteisiin ryhmiin, tästä lisää luvussa 4.4 Yhteenve-to kaikista yhdyssanoista.

e) Nominatiivimääritteisten yhdyssanojen yhteenve-to

Esitän seuraavaksi taulukossa 8. Nominatiivimääritteiset yhdyssanat koottuna koko ryhmän esiintymien määrät ja poikkeamat sanaluokittain.

Sanaluokka	Esimerkki poikkeava konvention mukainen	Määrä	Poikkeamia	Poikkeama- prosentti
Adjektiivit		417	17	4,1
– (i)nen-johdos	<i>nopea tahtinen</i> <i>ahdasmielisiä</i>	377	13	3,4
– karitiivi	<i>ikärajattomia</i> –	8	–	–
– partisiippi	<i>käsikirjoitettuja</i> –	10	–	–
– perusadjektiivi	<i>ankka hullu</i> <i>sanavalmis</i>	22	4	18,2
Substantiivit		4638	229	4,9
– perussana	<i>äänestys sydeemejä</i> <i>aamukahvi</i>	3364	186	5,5
– erisnimet	<i>rähinä cheek</i> <i>Kaurismäen</i>	259	14	5,4
– tekijä ja väline nimi (1)	<i>hallitus neuvottelija</i> <i>asianajaja</i>	121	5	4,1
– teko ja tulos (2)	<i>Hallitus neuvottelut</i> <i>kesäkokous</i>	882	24	2,7
– paikka ja tilaisuus (3)	<i>Jokela-ravintola</i> –	5	–	–
– partisiippi	<i>luottamustehtävä</i> –	7	–	–
Verbit		3	–	–
– finiittimuoto	– –	–	–	–
– infinitiivi	<i>pakkolukemaan</i> –	3	–	–
– partisiippi	– –	–	–	–
Muut	<i>laajamittaisesti</i> <i>vuosi kymmeniä</i>	53	4	7,5
	Yhteensä kaikki	5111	250	4,9

Taulukko 8. Nominatiivimääritteiset yhdyssanat sanaluokan mukaan

Tulokset poikkeavat jonkin verran Ylikuljun tuloksista, etenkin adjektiivien osalta. Hänen aineiston-
sa nominatiivialkuisissa adjektiiveissa on virheellisiä muotoja (virheellisiä : norminmukaisia
13 : 136 >) 9,6 % ja saman ryhmän substantiiveissa (41 : 763 >) 5,4 % (Ylikulju 2003: 98). Eroa se-
littänee jonkin verran erilainen analyysitapamme: Ylikulju on esimerkiksi laskenut adjektiivista *-sti-*
johdoksella johdetut sanat adjektiiveiksi, kun tässä työssä ne ovat adverbien joukossa (ryhmässä
muut), sekä numeroalkuiset yhdysadjektiivit nominatiivialkuisten ryhmään, kun tässä työssä ne ovat
erityisyhdyssanojen joukossa erillisessä numeroalkuisten ryhmässä. Nominatiivialkuisten yhdys-
substantiivien osalta tässä työssä olevaa hiukan pienempää virheprosenttia selittänee parhaiten se,
että Ylikulju on raakannut aineistostaan pois sellaiset sanat, jotka hänen mukaansa "nykyisin yleis-
esti hahmotetaan yhden sanan ilmauksiksi. Tällaisia ovat esimerkiksi *maailma*, *elokuva* ja kaikki
kuukaudennimet. Kaikki nämä tapaukset on kirjoitettu oikein yhdyssanaksi." (Ylikulju, 2003: 3).
Omassa aineistossani kaikki mainitut sanat ovat mukana, samoin kaikki muutkin yhdyssanat, jotka
voivat – jonkun mielestä, jossain mielessä, jossain yhteydessä – hahmottua yhdeksi jakamattomaksi
sanaksi (ehkä tällainen voisi olla jo aiemmin pohdittu *vaihtoehto*). Mielestäni Ylikuljun raja-
us on tilastoinnin kannalta virheellinen: miten voi antaa luotettavaa tietoa yhdyssanojen oikeinkirjoitukse-
sta, jos rajaa aineiston ulkopuolelle jo lähtökohtaisesti ne sanat, jotka kirjoitetaan oikein? Toisaalta
0,5 %-yksikön ero ei kuitenkaan ole kovin merkittävä, joten suurta osaa ne eivät kuitenkaan näytte-
le; on itse asiassa hiukan hämmästyttävää, etteivät ne vaikuta tuon enempää. Mahdollisesti kyseessä
voisi olla tekstilajin ero, mutta ehkä pitävin selitys on se, että tekstimassa tässä työssä on kaikkiaan
niin suuri, että vaikka nämä sanat kohtalaisen yleisiä ovatkin, eivät ne vaikuta isossa joukossa juuri
suuntaan eivätkä toiseen.

Taulukossa 9. Nominatiivimääritteiset yhdyssanat näkyy määriteosan sanaluokan mukaan ko-
koavasti, miten nominatiivialkuisten määriteosat ovat jakautuneet sanaluokkiin.

Määriteosan sanaluokka	Esimerkki poikkeava konvention mukainen	Määrä	Prosenttia kaikista	Poikkeamia	Poikkeama- prosentti
Substantiivi	<i>prätkä jengin</i> <i>kielikeskustelussa</i>	4652	91,0	226	4,9
Adjektiivi	<i>nopea tahtinen</i> <i>hyväpalkkaisten</i>	372	7,3	21	5,6
Numeraali	<i>yksi selitteinen</i> <i>tuhatpäinen</i>	43	0,8	2	4,7
Pronomini	– <i>monikulttuurisuutta</i>	28	0,5	1	3,6
Muut	– <i>salassapitolaki</i>	16	0,3	–	–
Yhteensä kaikki		5111	100	250	4,9

Taulukko 9. Nominatiivimääritteiset yhdyssanat määriteosan sanaluokan mukaan

Substantiivi on siis ylivoimaisesti yleisin määrittäjä nominatiivialkuisissa yhdyssanoissa (91 %). Seuraavana yleisyydessä tulevia adjektiiveja on alle kahdeksan prosenttia. Muut sanaluokat jäävät yhteensäkin laskettuna marginaaliin: alle kahden prosentin osuuteen (1,7 %). Adjektiivimääritteissä yhdyssanoissa poikkeamia tehtiin prosentuaalisesti hiukan enemmän kuin substantiivimääritteissä, mutta ero on niin pieni, että etenkin kun poikkeamaan osallistuneiden adjektiivimäärittäjien ryhmä ei kaikkiaan ole kovin suuri (21 esiintymää), ei määrittäjän sanaluokasta liene löydettävissä varsinaisesti mitään poikkeamia selittävää tekijää; samoin on tulos niin numeraali- kuin pronominiyhmänkin osalta. Poikkeamia ei tehty lainkaan muut määrittäjät -ryhmän sanoissa.

Nominatiivimääritteisten sanojen kombinaatiot jakautuvat taulukon 10. Nominatiivimääritteisten yhdyssanojen kombinaatiot mukaan.

Sana- luokka	MÄÄRITEOSA											
	Substantiivi		Adjektiivi		Numeraali		Pronomini		Muut		Yht.	
	n	%	n	%	n	%	n	%	n	%	n	%
S	4350	85,1	247	4,8	15	0,3	10	0,2	16	0,3	4638	90,7
A	269	5,3	116	2,3	20	0,4	12	0,2			417	8,2
V	2	0,0	1	0,0							3	0,1
M	31	0,6	8	0,2	8	0,2	6	0,1			53	1,0
											Yhteensä	5111 100,0

Sanaluokka:

S = substantiivi, A = adjektiivi, V = verbi, M = muu

Taulukko 10. Nominatiivimääritteisten yhdyssanojen kombinaatiot

Taulukosta näkee, että ylivoimaisesti suurin ryhmä on NS+S-yhdyssanat; NS+A- ja NA+S-sanoja on lähes yhtä paljon ja neljänneksi eniten on NA+A-sanoja. Neljän suurimman ryhmän joukko, joka siis koostuu substantiiveista ja adjektiiveista, kattaa 97,5 % nominatiivialkuisista yhdyssanoista – kaikki muut kombinaatiot ovat 2,5 %:n osuudella hyvin marginaalisia.

Lisäksi liitteessä 4 on koottuna koko nominatiivimääritteisten yhdyssanojen ryhmän kombinaatiot, niiden määrät ja poikkeamat yhteen taulukkoon.

4.3.3 Genetiivialkuiset yhdyssanat

Kun sanaparin määrittäjän muotona on genetiivi, on aina tapauskohtaista, milloin sanat kirjoitetaan yhteen, milloin erilleen. Samoissakin sanapareissa voi olla vaihtelua: joskus kyseessä on myös merkityksen muuttuminen, joskus edes merkityksen muuttuminen ei tee eroa.

Kielenhuollon mukaan genetiivimuotoinen määrittäjä kirjoitetaan yhteen pääsanansa kanssa, kun kokonaisuudella on **kiteytynyt merkitys** tai se on esimerkiksi (erikoisalan) **termi** tai ammattinimike (KO: 257). Kielenhuolto huomauttaa, että **erilleen kirjoittaminen voi korostaa alkuosan viittaavan johonkin tiettyyn asiaan** (KO: 257) ja samaan tapaan sanoo Vesikansa viitatessaan erityisesti yhdyssubstantiiviin (huomattavaa on tietysti, että vain substantiivilla voi ylipäätään olla sellainen referentti, joka ei riipu kontekstista): "Jos – – määriteosana on genetiivimuotoinen substantiivi tai adjektiivi – – yhdyssubstantiivi on lähinnä käytöltään termin luonteinen tai ainakin – vastaavaan sanaliittoon verrattuna – tavalla tai toisella erikoismerkityksinen." (Vesikansa 1989: 216.) Vaikka painotus voi siis vaihdella akselilla erityinen–yleinen, itse sisällön merkitys ei välttämättä varsinaisesti muutu kirjoitetaanpa sanapari yhteen tai erilleen.

Jotkin genetiivialkuiset yhdyssanat ovat siis kiteytyneet tarkoittamaan erityismerkitystä. Esimerkiksi *talon poika* versus yhteen kirjoitettu sanaparinsa *talonpoika* nostetaan usein esiin, kun tätä piirrettä halutaan havainnollistaa. *Talonpoika* viittaa maanviljelijään (joskin termi tuntuu auttamattoman vanhentuneelta), kun taas *talon poika* voi viitata minkä hyvänsä talon poikaan. Jos tällaista erityistä viitettä haettaisiin, kuitenkin tuntuisi, että talo pitäisi motivoida jotenkin tekstiyhteydessä: jollain tavalla talo siis on läsnä ja selitettävänä viittaussuhteena *talon pojassa*, kun taas *talonpojasta* voidaan puhua esimerkiksi ammattina ilman minkäänlaista kytköstä itse taloon.

Jotkin sanaparit on sovittu kirjoitettavaksi erilleen toisistaan. Tällaisia ovat esimerkiksi kielten nimet. On vain yksi suomen kieli, eli se tuntuu hyvinkin erityiseltä ja jopa referentiaaliselta (on osoitettavissa, joskin työläästi, mikä ja millainen on suomen kieli), silti sanaparin konventionaalinen muoto kirjoitetaan aina erilleen, sanaliitoksi: *suomen kieli*. Ei ole ihme, jos kielenkäyttäjää hämmentää se, että sanaliitosta johdettu -inen-adjektiivi (eli GS+Ai-sana) kuitenkin kirjoitetaan konventionaalisesti yhdyssanaksi: *suomenkielinen*. (KO: 257.)

Kun nominatiivimääritteisten yhdyssanojen osien väliset suhteet ovat varsin vapaat ja assosiaatiiviset (ks. luku 4.3.2 Nominatiivimääritteiset yhdyssanat), eli osien välisiä erilaisia merkityssuhteita on lukuisia, liikkuu genetiivialkuisten yhdyssanojen osien suhde pienemmällä alueella eikä ole samalla tavalla monitulkintainen. Tärkeimmät merkityssuhteet ovat omistus- tai siihen läheisesti liittyvä osa-kokonaisuussuhde (*isänmaa, leivänpala*) sekä verbaalikantaisissa yhdyssanoissa objektia tai subjektia vastaava merkitysosa (*sairaanhoido, kansanäänestys*). Tämä luonnollisesti rajaa genetiivimääritteen käyttöä yhdyssanoissa, ja genetiivialkuisia yhdyssanoja onkin vähemmän kuin nominatiivialkuisia. (ISK § 411.)

Kielenhuollossa on likimain koko suomen kielen tietoisien ohjailun ajan (toisin sanoen niin kauan kuin kielenhuoltoa on ylipäätään ollut) väännetty kättä siitä, kumpi muoto – nominatiivi vai

genetiivi – milloinkin tulisi valita yhdyssanan määriteosaksi. Suomen yhdyssanoja on yritetty kii-vaastikin ohjata siihen suuntaan, että määriteosa olisi mieluummin genetiivissä kuin nominatiivissa, koska nominatiivialkua on pidetty svetisminä, ja etenkin 1900-luvun alkupuoliskolla vieraista vaikutteista (vallankin ruotsin vaikutuksesta) haluttiin päästä eroon (ks. esimerkiksi edustava kuvaus aiheesta: Vesikansa 1989: 227–229, myös Ikola 2001: 111.) Nykyään kielenhuolto ei tunnu kanta-van tästä kovin suurta huolta: Kielitoimiston oikeinkirjoitusopas (painos vuodelta 2010) ei ota kan-taa aiheeseen lainkaan, ja Kielenhuollon käsikirja (painos vuodelta 2006) antaa melko yleisiä ohjei-ta motivoiden valintaa nominatiivin ja genetiivin välillä lähinnä niin, että nominatiivi viittaa ylei-sempään ja genetiivi erityisempään (KK: 134–138). Ehkä keskustelu aiheen ympärillä on laimentu-nut sen vuoksi, että kielenhuolto on onnistunut nominatiivi-genetiiviohjailussa varovastikin arvioi-den heikosti. Aineistoni yhdyssanat ovat nominatiivi : genetiivi -suhteessa 5111 : 882. Kaikkiaan ai-neistoni yhdyssanojen määrittäjät jakautuvat seuraavasti:

• nominatiivi	5111	67 %
• yhdysosa	904	11,8 %
• genetiivi	882	11,6 %
• erityisyhdyssanat	317	4,2 %
• muut	420	5,5 %

Genetiivejä on määrittäjinä aineistossani kolmanneksi eniten, eikä genetiivi määrittäjänä ei ole suhteellisesti sen dominoivampi muoto kuin yhdysosamuotokaan. Koska genetiivi kuitenkin on kolmanneksi yleisin määrittäjämuoto ja yleisesti käsitelty yhdyssanan alkuosana kielenhuollon oppais-sakin nimenomaan yhteen ja erilleen kirjoittamisen näkökulmasta (esimerkiksi KO: 257–260, KK: 144–145, SK: 187–190), on sen tarkastelu tässä erikseen varsin perusteltua.

a) Genetiivimääritteisten yhdyssanojen poikkeamat

Hankaluus genetiivialkuisten yhdyssanojen tutkimisessa poikkeamien näkökulmasta on siinä, mitä katsoo konvention mukaiseksi ja mitä sen vastaiseksi. Kun kielenhuolto yleisesti motivoi genetiivi-alkuisia yhdyssanoja niiden kiteymällä ja yleiseen viittaavalla merkityksellä, liikutaan alueella, joka on ennemminkin hämäräkö ja mielipiteeseen perustuva kuin aivan selvärajainen ja poikkeamatto-min säännöin määritelty. Tässä työssä lähdenkin siitä, että en voi päättää kielenoppaan puolesta, on-ko genetiivimääritteinen sanapari hänen mielestään yhdyssana vai sanaliitto eli täytyvätkö genetiiv-ialkuisen sanaliiton kohdalla ne kriteerit, jotka muodostaisivat sen yhdyssanaksi. Korkeintaan voin ajatella, että molemmat muodot ovat mahdolliset. Ajatus on tietysti hiukan kyseenalainen, koska vaikka teoriassa ehkä **mikä tahansa** genetiiviattribuutilla täydennetty sana voisi jossain kontekstis-

sa olla yhdyssana, ei näin varmasti käytännössä ole. Joitain poikkeuksia edellä esitetystä rajauksesta olenkin tehnyt lähinnä kiteymän, merkityksen muuttumisen tai referentiaalisuuden perusteella, ja ne esitän seuraavaksi.

Genetiivimääritteisistä yhdyssanoista rajasin yhteensä 39 poikkeamaa, jotka käyn läpi perusteluineen. Poikkeuksellisesti laitan esimerkkiin koko sen lauseen (tai lausetta vastaavan osan), jossa ne esiintyvät. Tavoittelen tällä ratkaisulla – sekä kaikkien poikkeamien käsittelemisellä sekä niiden esittämisellä (suppeassa) kontekstissaan – sitä, että lukijalle muodostuisi käsitys paitsi siitä, millaisia genetiivimääritteiset yhdyssanat ovat ja miten ne muodostuvat, myös siitä millaisissa tapauksissa poikkeamia syntyy, mitä ne ovat ja millaisin perustein ne ylipäättään ovat poikkeamia. Nämä seuraavaksi esittämäni sanaparit ovat siis nähdäkseni konventiosta poikkeavia. Aineistossani on kuitenkin useita sanapareja, joista sanakirja (MOT) esittää vain yhteen kirjoitetun version, mutta joiden toisistaan erilleenkin kirjoittaminen nähdäkseni voi olla perusteltua, enkä ole sellaisia katsonut poikkeamiksi (tällaisia ovat esimerkiksi: *kaupungin johdon*, *lausunnon antajalta*, *murhan himo*, *rahan käyttöä*, *suunnan muutoksia*, *tiedon murusesta* ja *tulon "tasaus"*).

Yksinkertaisimmat poikkeamiksi määriteltävät sanaparit ovat erisnimiä, sillä niillä on oma erityinen kirjoitustapansa, jota ei käytännössä milloinkaan varioida. Niitä on tässä poikkeamien joukossa kaksi ja kummatkin tarkoittavat samaa (152 ja 153).

(152) *KAKSI akuuttia tavoitetta: Nr 1 päästä ensi lauaintaina Finlandiahiihdossa maaliin. Nr2 päästä Arkadian mäelle.*

(153) *Lupailleet paljon Arkadian mäellä, mutta toivoneet syvästi hiljaa että nopesati luonnollisella tavalla poistuisivat heidän rasiitteesta.*

Erisnimeen vertautuvia ovat instanssien nimet, sillä niillä on yleensä sovittu kirjoitusasu. Aineistossa on kaksi tätä tyyppiä (154 ja 155):

(154) *Valtakunnan oikeusaloittaa valkopesu ohjelmansa ja huomataan että maahan on saatava riippumaton perustuslaki tuomioistuini.*

(155) *Yhtä johtajista syytetään yleisesti viraomaisrikollisuudesta sisäasian ministeriö tutkinut rikospoliisi tutkinut käynyt oikeutta suomessa jne.*

Kuten esimerkeistä huomaa, kirjoittajilla on muunkin laisia ongelmia kirjoittamisessa ja enemmän kuin mitään muuta nämä vaikuttavat marginaalisilta, jostain syystä tehdyiltä lipsahduksilta kuin kovin frekventeiltä poikkeamatyypeiltä.

Usein esiintyvä erisnimiin liittyvä poikkeama sen sijaan on seuraavaksi esiteltävä tyyppi. Yhdyssanaksi laskettavia ovat sellaiset sanaparit, joissa toinen osa on genetiivimuotoinen erisnimi ja toinen tarkoittaa jotain siihen liittyvää, mutta ei kuitenkaan itse erisnimen tarkoitteelle kuuluvaa ei-

kä sen osaakaan. Nämä sanaparit vaativat väliinsä yhdysviivan, ja sen puuttuminen aiheuttaa siis poikkeaman. (KK: 153.) Tällaisia esiintymiä on 13 (esimerkit 156–168).

- (156) *saako tonne Espoon keikalle lippuja hankittua vielä jostaan?*
- (157) *oli suht asiallinen espoon keikka! thänks!*
- (158) *Kiitos eilisestä Inkeröisten keikasta.*
- (159) *kiitti Inksan keikast!(;*
- (160) *Inksan keikka oli ihana!:(*
- (161) *Olin joensuun keikalla ensi kerran katsoo sinua jare.*
- (162) *Vuonna 2005 Kiteen keikalla oli ehkä 7 ihmistä yleisössä.*
- (163) *Olin yks niistä, joka oli silloin Kiteen keikalla, heh =D*
- (164) *koska ne aikasempien keikkojen liput tulee myyntiin? niinku esim lahden keikan 7.1*
- (165) *Beninin prossiveistokset British Museumissa ovat riittävä syy Lontoon matkaan*
- (166) *oisko millo mäntyharjun keikka tiedos??*
- (167) *mikse täällä oo kuvia seinäjoen keikast joka oli 5.3 ku olin siel oot muuten tosi söpö*
- (168) *varmaan ainoon tumman sedän on nähny afrikan lomamatkalla...*

Kahta lukuun ottamatta (*Lontoon matkaan* ja *afrikan lomamatkalla*, esimerkit 165 ja 168) kaikki tämän tyyppin sanat koskevat siis keikkoja, ja ne on kirjoitettu Cheek-ryhmään. Tässä sopii pohtia jo aiemmin (nominatiivimääritteisten substantiivien kohdalla, sivulla 50) esittämäni arviota siitä, miten ympäristö ohjailee kirjoitustavassa: jos joukossa olisi muoto *Inksan-keikka*, vaikuttaisiko sen esimerkki niin, että ryhmässä yleisemminkin alettaisiin käyttää yhdysviivaa tässä yhdyssanatyypissä? Toisaalta silmämääräinen arvioni tästä tyyppistä muissakin yhteyksissä on, että hyvin yleisesti yhdysviiva jätetään pois ja sanat kirjoitetaan erilleen. Painotan nyt kuitenkin, että kyseessä on arvio ja asiasta jotain varmaa sanoakseen pitäisi se tutkia erikseen, sillä tämän aineiston perusteella on ehkä liioiteltua sanoa, että yhdysviiva jätettäisiin yleisestikin pois siitä huolimatta, että yhtään ainutta konvention mukaan kirjoitettua *Erisnimen-keikka-* tai *Erisnimen-matka-*tyyppistä yhdyssanaa aineistossani ei ole.

Seuraavaksi nostan esiin sellaiset sanat, jotka nähdäkseni ovat kiteymiä ja joilla on siksi sovitettu kirjoitusasu, ja ne esiintyvätkin melko usein (nimenomaan kiteyminä).

- (169) *Mielen kiinnolla odotan puolustusministerin esiintuloa asian johdosta, --*
- (170) *Kun kuulin uutisen, että turkistuottajat aikovat antaa vaalirahaa puolueille, alan perua aiempaa kannan ottoani tässä keskustelussa.*
- (171) *..lopettaminen järkevin toimen pide*
- (172) *Nimen omaan tärkeä, ja muutenkin olisi mielestäni hyvin tärkeää luoda yhteishenkeä ja on aivan varma,*

(Viimeisin esimerkki 172 on vastaus edelliseen kommenttiin "Tärkeä etappi !") Kaikista näistä on sanakirjassa (MOT s.v. mielenkiinto, kannanotto, toimenpide, nimenomaan) ainoastaan yhdyssanaksi kirjoitettu hakusana. Usein yhdyssanamaisuutta aprikoivia neuvotaan kokeilemaan, josko sanaparin väliin sopisi jokin liitepartikkeli, ja jos ei sovi, kyseessä on yhdyssana (KO: 253, KK: 139). Esimerkiksi liitepartikkeli *-kin* sopii tähän tarkoitukseen hyvin. Ei tunnu löytyvän kontekstia, jossa olisi luontevaa sanoa **mielenkin kiinnolla*, **kannankin ottoa*, **toimenkin pide* saati **nimenkin omaan*. Kiteymänä voi pitää myös ihmistarkoitteista *viranomaista*, joka on sillä tavalla läpinäkymätönkin, että perusosa *omainen* ei tarkoita 'kaltaista' eikä *omaista* merkityksessä 'perheenjäsen' vaan se liittyy niin kiinteästi yhteen *viran*-sanana kanssa, että ne ovat merkitykseltäänkin erottamattomat (esimerkki 173).

- (173) *maassa kidutetaan viranomaisten toimesta ihmisä viranomaisten toimesta sekä eristetään kaikessa hiljaisuudessa joka on sotatiloissa käytettävä poliittinen pakkokeino rauhan ajan suomessa.2 myyjätärtä kertoi heillä olevan mies puolinen sukulainen joka viran omaisten toimesta on joutunut henkisisistä ja fyysisistä syistä työkyvyttömyyseläkkeelle.*

Näitä esiintymiä on kaksi, ja ne kumpikin on kirjoittanut sama kielenopas, joka ehkä hiukan hämmentävästi on kuitenkin onnistunut tuottamaan poikkeaman kanssa samaan (esimerkin 173) kommenttiin myös kaksi konvention mukaista muotoa. Muitakin ihmisiä tarkoittavia sanoja on poikkeamien joukossa. Aivan selvä poikkeama on seuraavassa esimerkissä (174), jossa ellipsin toinen puoli on kirjoitettu yhteen, mutta toinen jätetty merkitsemättä yhdysviivalla.

- (174) *Intel sponsaa kouluille matikan ja bilsanmaikkoja, samalla ku täällä luokkakoot kasvaa, kouluja lakkautetaan, ja rahaa menee taxikyteihin*

Poikkeaman selvyys johtuu siis siitä, että osa ellipsisistä on kirjoitettu yhteen ja osa jäänyt erilleen; periaatteessa mahdollinen olisi myös muoto *matikan ja bilsan maikkoja*. Seuraavat ihmistarkoitteiset sanat olen merkinnyt poikkeamiksi, mutta ne alkavat ehkä olla jo hiukan rajatapauksia.

- (175) *Nyt tehdään oikeaa puoluepolitiikkaa, jossa takinkääntäjät ja p:n nuolijat huomataan.*
- (176) *Kymmeniä miljardeja vuosittain veron maksajien rahoja säästävä yhden luokun kelan maksettava parempi perusturva pitäisi jonkun saada hallituksessa vaikka hallituksen yli kävellen aikaan.*

Kummastakin sanasta – *p:n nuolijan* lyhenteen tulkinta lienee itsestäänselvyys – on sanakirjassa hakusana (MOT s.v. perseennuolija, veronmaksaja) nimenomaan yhdyssanaksi kirjoitettuna. Ongelmallisia poikkeaman kannalta nämä ovat siksi, että on tilanteita, joissa määriteosa voisi saada attribuutin (*pomon p:n nuolijat, x:stä johtuneen veron maksaja*) ja silloin sanapari ilmenisi sanaliittona. Näiden sanojen kanssa mennään siis sille alueelle, jolla on tarkasteltava kontekstia ja sen mukaan päätettävä, onko kyseessä yhdyssana vai ei, kun aiemmissa sanoissa oli kontekstista riippumatta kyseessä niin kiteytyneet sanaparit, että yhdyssana toteutuu aina. Kummassakin kommentissa nähdäkseen on selvästi nähtävillä, että kyseessä on yleiset entiteetit, ei erityiset, referentiaaliset viittauskohdet. Ensimmäisessä lyhenne on voinut vaikuttaa erilleen kirjoitetun muodon valintaan (*p:nnuolija* ei ehkä tunnu oikealta), koska se rinnastuu *takinkääntäjään*, jonka kielenopas on tuottanut konvention mukaan. Kuvainnollinen *takinkääntäminen* esiintyy myös poikkeamana tässä ryhmässä (esimerkit 177 ja 178).

(177) *mitäs se Laaninen sanoikaan noista hallitusneuvotteluista ja takin käännöistä,,,,,,*

(178) *Ei halitukseen kun keran on sanottu että uusien vaalien kautta vasta hallitukseen ei mitään takin kääntöjä !!!!!!!!!!!!!!!*

Sanakirjan mukaan (MOT s.v. takinkääntö) ilmaus on yhdyssana. Sanaparia voi testata myös liitepartikkelilla *-kin* eli *takinkin kääntö* versus *takinkääntökin*, jolloin yhteen kirjoitetun ilmaisun eriytyminen kuvainnolliseen käyttöön käynee selvästi ilmi: *takinkin kääntö* johdattaa pohtimaan, mitä muuta on käännetty – kuvainnollisessa ilmaisussa ei liene mahdollista löytää muuta kontekstissa *käännettävää* toisin kuin konkreettisessa ilmaisussa, jolloin *käännettävänä* voisi olla (esimerkiksi pyykinpesun jälkeen) vaikkapa *housut*. Kuvainnollinen ilmaus on myös seuraavan esimerkin (179) poikkeama *päivän selvä*.

(179) *Tämä on ollut päivän selvä asia jo vuosikymmeniä.*

Ilmaus on paitsi kuvainnollinen myös jokseenkin läpinäkymätön: vaikka se on usein toistuva ja sillä tavalla tuttu, ei sen merkitys käy suoraan ilmi sanojen suhteesta. Ei ole selvää, miten *päivä* on *selvä*, kun se voi samaan aikaan olla esimerkiksi sateinen, sumuinen tai synkkä; ilmaus on eriytynyt omaan käyttöönsä ('ilmeinen') sovitussa kirjoitusasussa ja siksi sen osien erilleen kirjoittaminen on poikkeama.

Ihmisten toimiin liittyvät seuraavat poikkeamat, joissa genetiivimäärite on subjektin asemassa (esimerkit 180 ja 181).

(180) *Eikös olisi aika suunnitella syksyksi kansain vaellusta eduskunnan eteen kertomaan heille miten asioita pitää hoitaa*

(181) *Kansan äänestys suoraan kelan kautta maksettavasta paremmasta perusturvasta verorahoilla*

Ensimmäisessä esimerkissä (180) on määriteosa vanhassa monikkomuodossa (*kansain*), joka harvoin nykyään esiintyy yksin ja jo sinällään ohjaa yhdyssanatulkintaan. Kommentissa ilmaus on synonyymi *mielenosoitukselle*, kyseessä ei ole siis minkään tietyn kansan (saati tiettyjen kansain) vaeltaminen, johon sanakirjan (MOT s.v. kansainvaellus: kansojen siirtyminen uusille asuma-alueille; vars. germaanien liikehtiminen Euroopassa v. 400–600) antama merkitys viittaa. *Kansanäänestys* on myös sikäli erityinen tapaus, että kuten sanakirja (MOT s.v. kansanäänestys) sanoo, kyseessä on: "äänestys jossa äänioikeutetut ilmaisevat suoraan mielipiteensä jstak asiasta." *Kansanäänestys* käsitteenä tarkoittaa siis ainakin arkimerkityksessään muuta kuin äänestystä säännöllisesti järjestetyssä vaalissa, esimerkiksi eduskuntavaalissa. Ei voida nähdäkseni sanoa, että eduskunta- tai kunnallisvaalit olisivat kansanäänestys, vaan kansanäänestys on lisä-äänestys, joka järjestetään, kun halutaan saada kansan mielipide jostain erityisesti esille nousseesta asiasta. Kyseessä on siis termi ja siten yhdyssana. Inhimilliseen toimintaan liittyy myös seuraavan esimerkin (182) termi.

(182) *Ja suomi on maailman mestari!! <3*

Maailmanmestari on maailmanmestaruuskilpailun voittaja (MOT s.v. maailmanmestari); titteli vertautuu ammattinimikkeeseen siinä mielessä, että sen voi ansaita vain yhdellä tavalla, maailmanmestaruuskilpailuun osallistumalla. Tämä kuuluu myös niihin yhdyssanoihin (kuten edellä esitellyt *p:n nuolijat* ja *veron maksajat*), joiden pelkälle määriteosalle voi kuvitella joissain erityisyhteyksissä attribuutin ja silloin sanapari hajoaisi (esimerkiksi vaikkapa scifi-kirjallisuudessa: *hän oli tunnetun maailman mestari*). Tällaisesta käytöstä ei kuitenkaan tässä yhteydessä ole kyse: kommentti on jääkiekko-ryhmästä ja se on julkaistu välittömästi vuoden 2011 jääkiekon maailmanmestaruusvoiton jälkeen.

Aikaan liittyviä poikkeamia on neljä (esimerkit 183–186):

(183) *Turha kyllä olisi antaa mitään armon aikaa enää Kataiselle.*

(184) *Nyt on kohtalon hetket Suomessa ja jokainen huokaus, rukous Isälle tiettäväksi on arvokkaampaa kuin esim. 10 x Nokian globaalinen arvo kullassa...*

(185) *Hyvää joulun aikaa sinne liigalle! :p*

(186) *Jos miettisitte muuta kellon aikaa, useampi voisi osallistua.*

Armon aika (183) sanaliittona merkinnee 'armollista aikaa', kun taas *armonaika* yhdyssananä sanakirjan mukaan (MOT s.v. armonaika) tarkoittaa erityisesti jonkin tehtävän suorittamiseen varattua aikaa, jäljellä tai käytettävissä olevaa aikaa. Esimerkin keskustelussa kyse on eduskuntavaalien jälkeisestä hallituksen muodostamisesta, joka oli kommentissa mainitun Kataisen tehtävä. *Armonaika* viittaa tässä siihen aikaan, joka Kataiselle tulisi (tai ei tulisi) myöntää tehtävään. Yhdyssanan ja sa-

naliiton välillä voidaan nähdä siis merkitysero. Samalla tavalla merkitysero on *kohtalon hetket* ja *kohtalonhetket*-ilmauksilla (184). Sanakirjan mukaan (MOT s.v. kohtalonhetket) kyseessä on jonkin olemassaolon tai tulevaisuuden kannalta erityisen tärkeä hetki, eli ikään kuin käännekohta. On ehkä hiukan hankalaa nähdä, mikä merkitysero sanaparilla on, ennen kuin nostaa esiin referentiaalisuuden ja kääntää sanajärjestystä hiukan: *Suomen kohtalon hetket* -ilmaus viittaisi suoraan Suomen valtion tilanteeseen, Suomen kohtaloon, kun taas *kohtalonhetket Suomessa* viittaa yleisemmin Suomen kansalaisiin, valtioon, luontoon, teollisuuteen tai mitä nyt ylipäätään kontekstissa käsitellään ja halutaan käsitellä. Kun referentiaalisuus on tullut todennetuksi, *kohtalon hetket Suomessa* -ilmaus käy käsittämättömäksi (minkä tai kenen kohtalosta on kyse?) ja on siten poikkeama. *Joulunaika* (185) taas on konkreettisempi ilmaus kuin edelliset aikaan liittyvät sanaparit. Se on tarkkarajainen, selvästi ilmaistava aika, joka alkaa jouluna ja päättyy loppiaiseen. Niin sanakirja (MOT s.v. joulunaika) kuin Kotimaisten kielten keskuskin (jatkossa Kotus, IV:KO) ottavat kantaa *joulunajan* yhdyssanamaisuuteen. Sanakirjan mukaan hiukan hämmäntävästi *joulunaika* on sama kuin *joulun seutu*, eli jollain tavalla jälkimmäinen sanapari on nähtävästi löyhempi ilmaus, vaikka tarkoite on sama. Kotuksen kannanotto taas liittyy nimenomaan yhdyssanamaisuuteen ja on vastaus kysymykseen "Mitkä joulu-alkuiset sanat ovat yhdyssanoja? Joulunodotus vai joulun odotus, jouluntulo vai joulun tulo, joulunaika vai joulun aika?" Vastaus kuuluu seuraavasti: "Joulunaika on vakiintunut käsite. Sana viittaa tiettyyn tarkkarajaiseen aikaan (yleensä jouluaatosta loppiaiseen), ja kirjoitetaan tavallisesti yhdyssanaksi." Vastauksen "tavallisesti" ehkä antaisi luvan arvella, että aivan aukoton tuokaan ohje ei ole vaan että sanaparin voisi joissain yhteyksissä kirjoittaa erilleenkin. *Kellon aika* -ilmaisussa (186) nousee taas esiin referentiaalisuus: voisi tietysti ajatella, että on kyseessä jonkin tietyn kellon aika, mutta koska tässä tarkoitetaan kellonaikaa yleisessä merkityksessä (vertaa esimerkiksi *päivämäärään*), on kyseessä poikkeama. Kaikki nämä neljä aikaa koskevaa poikkeavaa ilmaisua ovat kuitenkin sellaisia, ettei niiden merkitys mielestäni muutu sanaliiton ja yhdyssanan variaatioissa niin drastisesti, etteikö asia kävisi selväksi, vaan kyse on ennemminkin tekstin hienosäädöstä kuin siitä, että lukija joutuisi aivan ymmälle näiden ilmaisujen äärellä siksi, että ne on kirjoitettu sanaliitoksi eikä yhdyssanoiksi (tai toisinpäin).

Jäljellä on kolme poikkeamaa, joissa kaikissa konvention vastaisuus on itsestään selvä. Ensimmäisessä poikkeaman (esimerkki 187) ilmaisee kongruenssin puute:

(187) *Suvaitsevaisuus joissakin asioissa on muodostunut mitä törkeimmäksi suvaitsemattomuudeksi kristillisperinteellisiä elämän arvoja kohtaan.*

Sanakirjan (MOT s.v. elämänarvo) mukaan *elämänarvo* esiintyy usein monikossa, kuten esimerkiksi 187:kin, ja tarkoittaa elämässä arvokkaita asioita. Hakusanan perusteella se on yhdyssana. Jos il-

mauksen haluaisi kirjoittaa sanaliitoksi, tarkoittamaan tiettyä elämää, kongruoisi sen määre, eli lauseke olisi: *kristillisperinteellisen elämän arvoja*. Kun attribuutti *kristillisperinteellisiä* kuitenkin kongruoi jälkimmäisen sanan eli *arvoja* kanssa, liittyy se nimenomaan siihen, ei *elämään*. Yhdessä *elämänarvoja* voisi saada attribuutikseen partitiivimuotoisen *kristillisperinteellisiä*, erilleen kirjoitetun *elämän arvoja* ei.

Aineiston rajausta selvittäessäni (luvussa 3.5.2 Johtimet *-lainen*, *-moinen* ja *-mainen*) käsitteelin jo *lainen*-sanana olemusta ja totesin siinä, että *lainen* voi esiintyä yksinkin. Sanakirja (MOT s.v. samanlainen) antaa *samanlainen*-sanasta vain yhdyssanamuodon, ja Kielitoimiston oikeinkirjoitusopas nostaa *samanlainen*-sanana esimerkiksi pronominin ja *lainen*-sanana yhteen kirjoittamisesta (KO: 260). Tosin KO huomauttaa siinä yhteydessä, että pronomini ja *lainen*-sana kirjoitetaan **yleensä** yhteen, eli vaikka *samanlainen* on kohdassa esimerkkiyhdyssanana, voi lukija jäädä hienoiseen epätietoisuuteen siitä, onko nimenomaan tuossa esimerkkisanassakin mahdollisuus sanaliittomuotoon (kuten esimerkissä 188).

(188) *tää on kyl kova jätkä ja hyvät biisit ja saman lainen kampaus ku mulla*

Viimeinen tämän ryhmän poikkeama (esimerkki 189) on tehty tarkoituksella ja olikin jo työn alkupuolella esillä, kun kuvailin aineistoa (luvussa 3.2 Facebook tekstialustana ja aineistolähteenä).

(189) *Suomi, kun siihen viitataan valtiona, kirjoitetaan isolla kirjaimella, kuten kaltaisesi isäm maan ystävän luulisi tietävän.*

Kyseessä on siis ilmiselvä parodia kommentissa, jossa kuitenkin annetaan kielioppineuvoja. Poikkeama syntyy siitä, että isänmaa yhdyssanana tarkoittaa valtiota (ei kenenkään tietyn isän maata) ja parodia ilmenee muuten korrektissa tekstissä puhekielisen muodon vuoksi: sananrajalla on assimiloitunut ***m isäm maa*** – assimilaatiosta huolimatta sanapari on kirjoitettu erilleen.

Tässä olivat kaikki aineiston genetiivimääritteisissä yhdyssanoissa tehdyt sellaiset poikkeamat, joissa erilleen kirjoittaminen on nähdäkseni selvästi poikkeama. Näistä laskettuna poikkeamaprosentiksi tulee 4,4 %. Ylikulju (Ylikulju 2003: 98) on laskenut omassa työssään vain adjektiivit ja substantiivit, joten vertaan tuloksia näiden kesken. Genetiivialkuisia adjektiiveja ja substantiiveja on aineistossani yhteensä 840 kappaletta. Niissä on tehty poikkeamia (adjektiivit 2, substantiivit 36) yhteensä 38 ilmauksessa. Poikkeamaprosentti on silloin 4,5 %. Ylikuljun aineistossa on genetiivialkuisia yhdysadjektiiveja ja -substantiiveja yhteensä (138 ja 153) 274 kappaletta ja niissä virheitä (20 ja 31) yhteensä 51 eli virheprosentti on 18,6 %. Ero on huikea, mutta selittyy varmasti eri tavalla arvioida poikkeama (tai virhe, kuten Ylikulju niitä työssään kutsuu). Esimerkiksi nostan kaksi ilmausta, jotka tässä työssä olisivat olleet norminmukaisia, mutta jotka Ylikulju on ensin analysoinut vir-

heiksi ja sitten esittänyt selityksenä sille, miksi on hankalaa erottaa sanaliitto ja yhdyssana toisistaan. Sanaparit ovat *läheisyyden kaipuu* ja *kauneuden kaipuu* (Ylikulju 2003: 53). Ylikulju ilmoittaa, että verbien teonnimijohdokset kirjoitetaan **usein** yhteen edeltävän määritteen kanssa, mutta huomauttaa, että lauseyhteydestä riippuen myös erikseen kirjoittaminen on mahdollista (Ylikulju 2003: 51). Työssä ei kuitenkaan käy ilmi, mikä tai millainen olisi se lauseyhteys, jossa erilleen kirjoittaminen voisi toteutua konvention mukaisesti. *Kauneudenkaipuu*-sanalle Ylikulju esittää ohjailevaksi vastineeksi sanakirjasanan, mutta *läheisyyden kaipuu* -ilmauksen yhdyssanamaisuus jää roikkumaan ilmaan (Ylikulju 2003: 53). Mielestäni ei ole perustetta sille, että kyseinen sanapari pitäisi ehdottomasti kirjoittaa yhdyssanaksi eikä siten ole perustetta myöskään väittää, että niiden sanaliittomuoto olisi virheellinen. Itse asiassa kielenhuoltajana valitsisin mieluummin erilleen kirjoitetut versiot **kummastakin** suositeltavammaksi selkeytensä vuoksi.

b) Genetiivimääritteiset yhdyssanat sanaluokittain ja yhdyssanatyypeittäin

Seuraavaksi esitän kokoavasti, miten genetiivimääritteiset yhdyssanat jakautuvat sanaluokkiin ja edelleen – samoin jaoin kuin nominatiivialkuiset yhdyssanat – sanaluokkien alaluokkiin. Kuten edeltävässä luvussa kävi ilmi, poikkeamia tässä ryhmässä on melko vähän lukumäärällisesti (ja luonnollisesti silloin suhteellisestikin), mutta esitän ne tässä samalla omissa sanaluokissaan. Tässä luvussa tarkastelen myös, miten genetiivimääritteisten yhdyssanojen määriteosa jakautuu sanaluokkiin ja millaisia kombinaatioita tässä ryhmässä on. Luvun lopuksi käyn vielä lyhyesti läpi genetiivimääritteisten yhdyssanojen jakautumisen yhdyssanatyyppiin.

Esitän taulukossa 11. Genetiivimääritteiset yhdyssanat sanaluokan mukaan koottuna genetiivimääritteisten yhdyssanojen ja niissä tehtyjen poikkeamien jakauman sanaluokan ja sanaluokkien alaryhmän mukaan.

Sanaluokka	Esimerkki poikkeava konvention mukainen	Määrä	Poikkeamia	Poikkeama- prosentti
Adjektiivit		167	2	1,2
– (i)nen-johdos	<i>saman lainen</i> <i>ajankohtaista</i>	157	1	0,6
– karitiivi	<i>isänmaatonta</i> –	4	–	–
– partisiippi	<i>asiantuntevia</i> –	5	–	–
– perusadjektiivi	– <i>päivän selvä</i>	1	1	100,0
Substantiivit		673	36	5,3
– perussana	<i>sisäasiain ministeriö</i> –	293	27	9,2
– erisnimet	<i>Arkadian mäelle</i> <i>Kankaanniemi</i>	78	2	2,6
– tekijä ja väline (1)	<i>veron maksajien</i> <i>työntekijä</i>	128	2	1,6
– teko ja tulos (2)	<i>takin kääntöjä</i> <i>lihantuotannossa</i>	171	5	2,9
– paikka ja tilaisuus (3)	<i>ääntenlaskemosta</i> –	1	–	–
– partisiippi	<i>kaupunginvaltuutettua</i> –	2	–	–
Verbit		7	–	–
– finiittimuoto	<i>kyseenalaistan</i> –	3	–	–
– infinitiivi	<i>kyseenalaistaa</i> –	1	–	–
– partisiippi	<i>kyseenalaistanut</i> –	3	–	–
Muut	<i>Nimen omaan</i> <i>todennäköisesti</i>	35	1	2,9
	Yhteensä kaikki	882	39	4,4

Taulukko 11. Genetiivimääritteiset yhdyssanat sanaluokan mukaan

Yhdyssubstantiivit ovat tässäkin ryhmässä suurin sanaluokka. Niitä on kaikkiaan 673, eli 76,3 % kaikista genetiivimääritteisistä yhdyssanoista. Adjektiivien ryhmä on seuraavaksi suurin, 167 yhdysadjektiivia on yhteensä määritetty genetiivillä, ja koko genetiivimääritteisten ryhmästä se on 18,9 %. Genetiivimääritteisissä yhdyssanoissa tehdyt poikkeamat jakautuvat niin, että kaksi niistä on tehty adjektiiveissa (*saman lainen*, *päivän selvä*), yksi muut-ryhmän sanoissa (*Nimen omaan*) ja kaikki loput yhdyssubstantiiveissa. Yhdyssubstantiiveissa poikkeamat jakautuvat niin, että 27 on tehty perussanoissa (GS+S), 2 erisnimissä (GS+Se), 2 tekijä ja väline -johdosryhmässä (GS+St) sekä 5 teko ja tulos -johdosryhmässä (GS+Sn). Partisiipeissa (GS+Sp) sekä paikka ja tilaisuus -johdosryhmissä (GS+Sl) ei ole tehty poikkeamia lainkaan. Koko genetiivimääritteisten yhdyssanojen ryhmän poikkeamaprosentti on 4,4 %, ja eniten poikkeamia sisältävän alaryhmän, perussubstantiivien (GS+S), poikkeamaprosentti on 9,2 %.

Seuraavaksi tarkastellaan genetiivimääritteisten yhdyssanojen määrittäjien sanaluokkia, ja koottuna ne ovat nähtävänä taulukossa 12. Genetiivialkuiset yhdyssanat määriteosan sanaluokan mukaan jaoteltuna. (Mukana on tyhjä muut-ryhmä muiden yhdyssanojen mallin mukaisesti ja ryhmien keskinäisen vertailun vuoksi.)

Määriteosan sanaluokka	Esimerkki poikkeava konvention mukainen	Määrä	Prosenttia kaikista	Poikkeamia	Poikkeama- prosentti
Substantiivi	<i>kansan äänestys hampaankoloon</i>	805	91,3	38	4,7
Adjektiivi	<i>saman lainen todennäköiseltä</i>	18	2,0	–	–
Numeraali	– <i>yhdenvertaisuus</i>	2	0,2	–	–
Pronomini	<i>saman lainen jonkinlainen</i>	57	6,5	1	1,8
Muut	– –	–	–	–	–
Yhteensä kaikki		882	100	39	4,4

Taulukko 12. Genetiivimääritteiset yhdyssanat määritysosan sanaluokan mukaan

Genetiivimääritteisiäkin yhdyssanoja määrittää kaikkein useimmin substantiivi (805 kappaletta – 91,3 %). Seuraavaksi suurin määrittäjäryhmä ovat pronominit, joita on 57, eli 6,5 %. Määrittäjinä marginaalissa ovat adjektiivit (18 – 2 %) ja numeraalit (3 – 0,3 %). Muita sanaluokkia määrittäjinä ei ole. Määriteosina on perussanojen lisäksi yhdyssanoja: yksi yhdyssana on neliosainen (*homovastaistenvastaistenvastaisilla*), 57 kolmiosaista (esimerkiksi *pääkallonpaikalta*) ja loput 824 ovat kaksiosaisia. Poikkeamista 7 on tehty kolmiosaisissa yhdyssanoissa, loput 32 poikkeamaa on kaksiosaisissa yhdyssanoissa.

Genetiivialkuisien kombinaatiot ja niiden määrät käyvät ilmi koottuna taulukosta 13. Genetiivimääritteisten yhdyssanojen kombinaatiot.

Sana- luokka	MÄÄRITEOSA											
	Substantiivi		Adjektiivi		Numeraali		Pronomini		Muut		Yht.	
	n	%	n	%	n	%	n	%	n	%	n	%
S	664	75,3	7	0,8	2	0,2					673	76,3
A	99	11,2	11	1,2			57	6,5			167	18,9
V	7	0,8									7	0,8
M	35	4,0									35	4,0
Yhteensä											882	100,0

Sanaluokka:

S = substantiivi, A = adjektiivi, V = verbi, M = muu

Taulukko 13. Genetiivimääritteisten yhdyssanojen kombinaatiot

Substantiivimääritteiset substantiivit, GS+S-sanat, ovat tässäkin ryhmässä kaikkein yleisin kombinaatio (662 – 75,1 %). Seuraavaksi eniten on GS+A-sanoja (99 – 11,2 %) ja hiukan ehkä yllättävästikin kolmanneksi suurin ryhmä on GP+A-sanat (57 – 6,5 %). Pronominimäärittäjien ryhmän kokoa kasvattavat sellaiset kaltaisuuteen liittyvät adjektiivi-ilmaisut kuin *eräänlainen, jonkinlainen, min-käläinen, monenlainen, muunlainen ja samanlainen*.

Genetiivimääritteiset yhdyssanat ovat yhtä lukuun ottamatta määritysyhdyssanoja: esimerkin 190 sanan olen sen epämääräisyydestä huolimatta katsonut summayhdyssanaksi.

(190) *Suomessa on kaksi uusi vihertävänkellertävänkukista puoluetta ko nykyinen vasemmisto ja vihertävät aikaemmin.*

Määritysyhdyssanoissa on kuitenkin nähdäkseni samalla tavalla eksosentrisiksi laskettavia, eli ilmaistakseen yhden merkityksen molemmat osat vaativia, yhdyssanoja kuin nominatiivimääritteisissäkin (ks. luku 4.3.2 Nominatiivimääritteiset yhdyssanat: d) Nominatiivimääritteisten yhdyssanojen yhdyssanatyypit). Tällaisia ovat esimerkiksi *viran-* ja *asianomainen, nimenomaan* ja *toimenpide*, joista analysoimalla ehkä on löydettävissä genetiiviyhdyssanalle tyypillinen osa-kokonaisuussuhde tai subjektin tai objektin tarkoite, mutta jotka pintatasolta vaikuttavat pitkälle kiteytyneiltä, erottamattomilta yhdistymiltä.

Liitteessä 5 on taulukoituna nähtävillä koko genetiivimääritteisten yhdyssanojen kooste.

c) *Genetiivimääritteiset yhdyssanat versus sanaliitot*

Genetiivimääritteistä yhdyssanaa motivoi usein sen yleisviittauksellisuus. Verrattuna genetiiviattribuutilla muodostettuun sanaliittoon yhdyssanan tarkoite on yleisempi kuin sanaliiton. Tätä ajatusta ilmentää esimerkiksi osa-kokonaisuus- versus omistussuhteen vertailu sanaliiton ja genetiivimääritteisen yhdyssanan välillä: kun sanaliitossa genetiiviattribuutti tulkitaan omistussuhteeksi, yhdyssanassa genetiivimäärite ohjaa tulkintaa ennemminkin kuuluvuussuhteeseen, esimerkkinä vaikkapa *lasten kengät* ('joidenkin tiettyjen lasten kengät') ja *lastenkengät* ('lasten käyttöön tarkoitettut kengät') (Jaakola 2004: 146). On kuitenkin huomattava, että vaikka sanaliitto voi viitata erityiseen entiteettiin, se voi myös viitata yleiseen – merkille pantavaa on myös, että aina yleisen ja erityisen välille ei voi tehdä selvää eroa. Tarkastelen seuraavaksi, mikä genetiivimääritteisten yhdyssanojen ja genetiiviattribuutilla määritettyjen sanaliittojen suhde on.

Rajasin aineistostani ryhmän, joka koostuu genetiiviattribuutista ja sen pääsanasta. Genetiiviattribuutti voi esiintyä peräkkäinkin, ja sellaisissa tapauksissa aineistoon tuli samasta lausekkeesta niin monta osaa kuin siinä on genetiiviattribuutteja (esimerkki 191).

(191) *Pruukin kennelin omistajalla*

> *Pruukin kennelin*

> *kennelin omistajalla*

Koska tarkoitus on tarkastella, miten ilmaisut jakautuvat sanaliittoihin ja yhdyssanoihin, jätin ryhmään vain ne sanaparit, joista ylipäätään on mahdollista muodostua yhdyssana. Pois rajautuivat silloin sellaiset genetiiviattribuutit (ja niiden pääsanat), joissa oli jokin liitepartikkeli, possessiivisuffiksi tai muu liite, joka estäisi yhdyssanan muodostumisen (esimerkiksi *toimituksen tarkoin vartioima* jää pois kiilasanan *tarkoin* vuoksi). Pois jäivät myös kongruenssissa muodostuneet genetiivit, jotka eivät oikeastaan ole genetiiviattribuutteja. (Esimerkiksi *mustavalkoisen elokuvan ohjaajan* -lausekkeesta jäisi tähän ryhmään vain *elokuvan ohjaaja*, koska *mustavalkoisen* on kongruoitunut *elokuvan* kanssa. Taivuttamaton muoto ensimmäisestä sanaparista on *mustavalkoinen elokuva*, kun taas jälkimmäisen sanaparin *elokuvan* pysyy genetiivinä, vaikka *ohjaaja*-sana taipuisi toisessa yhteydessä, esimerkiksi *elokuvan ohjaajaa*.)

Kuten on jo todettu, yhdyssana viittaa mieluummin lajiin ja yleiseen kuin yksilöön tai erityiseen. Tätä ajatusta vasten on hankala perustella, miksi on olemassa *Lontoon-matka*-tyyppisiä yhdysanoja, joissa määrittäjänä on erisnimi ja jota käytetään silloinkin, kun puhutaan tietystä, erityisestä matkasta (*Se viimekesäinen Lontoon-matkamme oli kyllä erityisen hauska*). Samoin yleinen–erityinen-jaottelussa ongelmia aiheuttavat suoraan kontekstista viitteensä saavat eli deiktiset genetiiviattribuutit, pronominit, jotka nähdäkseni muutamia lukuun ottamatta (esimerkiksi kvanttoripronominit: *moni, muutama, kaikki, mikään*) ovat tiettyyn, kontekstisidonnaiseen tarkoitteeseen viittaavia ja siten erityisiä, eivät yleisiä¹⁵. Sekä erisnimien että pronominiensa rajaaminen ei genetiiviattribuuteista tästä syystä voi olla perusteltua, koska ne muodostavat yhdyssanojakin (pronominit genetiivimääritteisten ryhmässä jopa frekventisti, ks. edellinen luku).

Edellä esitetyin rajauksin ryhmään jäi 3692 sanaliittoa, jossa on genetiiviattribuutti ja sen määrittämä pääsana, jatkossa puhun näistä genetiivimääritteisinä sanaliittoina (versus genetiivimääritteiset yhdyssanat). Tämän työn puitteissa ei ole mahdollista tehdä kattavaa vertailua koko aineiston genetiivimääritteisten sanaliittojen ja yhdyssanojen kesken, mutta joitain sanaryhmiä voin esimerkinomaisesti nostaa esiin.

Todetaan ensin, että yhdyssanoista olen katsonut 26:n olevan sellaisia, jotka voisi kirjoittaa erilleenkin, esimerkiksi: *kansantahtoa, lumenauraus, politiikantekoon, rasistinleiman ja turkisten vastustajat*. Siihen nähden, että genetiivimääritteisen yhdyssanan muodostumista määräävät säännöt

¹⁵ Pronominin referentti itsessään voi kyllä olla yleismerkityksinen, mutta siihen viittaava pronominihan on väistämättä erityinen, eli sillä on spesifi viittauskohde kontekstissaan.

ovat melko väljät, määrä tuntuu pieneltä (se on noin 3 % kaikista genetiivialkuisista yhdyssanoista) ja kertonee, että genetiivimääritteiksi yhdyssanoiksi valikoituvat useimmin ne, jotka ovat pitkälle kiteytyneitä ja jo olemassa olevina sanapareina frekventtejä. Tällaisia ovat aineistossani esimerkiksi *äidinkieli, viranomainen, viikonloppu, vaaleanpunainen, todennäköinen, puheenjohtaja* ja niin edelleen. Toisesta suunnasta katsoen taas genetiiviattribuutilla rakentuneiden sanaliittojen suuri määrä kertoo samasta seikasta: vähemmän yhteen kiteytyneet, harvemmin esiintyvät – ja varmasti merkitykseltään erityisimmät – ilmaukset kirjoitetaan mieluummin erilleen toisistaan, sanaliitoksi kuin yhdyssanoiksi. Joukossa on sellaisiakin sanapareja, joista sanakirja (MOT s.v. omantunnonkysymys, pakkasenpurema, rumannäköinen, sodanjälkeinen) antaa sekä erikseen että yhteen kirjoitetut muodot, esimerkiksi *omantunnon kysymys, pakkasen purema, ruman näköinen ja sodan jälkeinen*.

Valitsin muutaman kohtalaisen frekventin sanaryhmän tässä tarkasteltavaksi. Kaikki ovat perussubstantiiveja (sulkeissa lopullinen lukumäärä, yhteensä 201): *aika* (20), *asia* (30), *eläin* (15), *elämä* (8), *kansa* (116) ja *luonto* (12). Näistä sanoista yhdistetyt yhdyssanat rajasin pois, eli kunkin ryhmän yhdyssanan määrittäjänä tai sanaliiton genetiiviattribuuttina on kyseisen sanan yksikön tai monikon genetiivimuoto yksinään (*ajan, aikain, asia, asioiden, asiain* ja niin edelleen). Yhdyssanojen ja sanaliittojen perusosat lajittelin sanaluokan ja edelleen johdostyyppin mukaan samoin, kuin olen rajannut tässä työssä yhdyssanatkin. Tavoitteeni tällä otteella oli tarkastella, miten sanaluokat ja niiden johdostyyppit vaikuttavat sanaparien yhdistymisiin tai sanaliitoina pysymisiin. Tyypiltään hän sanat ovat – paitsi siis perussubstantiiveja – elämään liittyviä, yhtäältä hyvinkin konkreettisia (vertaa esimerkiksi *rakkaus-, teoria-* ja *politiikka-*sanoihin) ja toisaalta kuitenkin jonkin verran abstrakteja (vertaa esimerkiksi niiden alakäsitteisiin *tunti, tarina, kissa* ja *suomalaiset*). On syytä huomauttaa tässä, että sanojen semantiikka saattaa vaikuttaa niiden yhdistymiskäyttäytymiseen ja siksi-kin tulosta voidaan pitää vain esimerkkinä ja suuntaa antavana.

Otannasta paljastui seuraavaa: ainoastaan substantiivien ryhmässä oli variaatiota, muut ryhmät jakoivat sanat joko yhdyssanoihin tai sanaliittoihin. Adjektiivien johdostyypeissä inen-adjektiivit ja partisiipit ovat kaikki yhdyssanoja, perusadjektiivit sanaliittoja. Verbit ja muut-ryhmän sanat ovat kaikki sanaliittoja. Substantiiveista, joita tässäkin ryhmässä oli selvästi eniten (153 – 76,1 %), kaikkiaan 93 (60,8 %) on yhdyssanoja ja 60 (39,2 %) on sanaliittoja. Substantiivit jakautuvat niin, että perussubstantiiveja sekä teko ja tulos -ryhmän sanoja on noin puolet ja puolet yhdyssanoissa, mutta tekijä ja väline -ryhmän sanoissa yhdyssanat ovat selvästi enemmistönä: 33 tekijännimeä yhdyssanoina ja vain 2 sanaliittona. Luultavasti tekijä ja väline -ryhmässä tulosta ohjailee usein esiintyvä *kansanedustaja*, mutta mielenkiintoista olisi tarkastella, ajautuvatko yleisestikin ihmistä tarkoittavat sanat (muita useammin) yhdyssanoiksi, niin kuin voisi arvella.

Sanaluokka	Esimerkki yhdyssana sanaliitto	Yhdyssana		Sanaliitto		Yhteensä
		k	%	k	%	
Adjektiivit		26	89,7	3	10,3	29
– (i)nen-johdos	<i>ajankohtaista</i> –	25	100,0	–	–	25
– partisippi	<i>asiantuntevia</i> –	1	100,0	–	–	1
– perusadjektiivi	– <i>aikojen surkein</i>	–	–	3	100,0	3
Substantiivit		93	60,8	60	39,2	153
– perussana	<i>kansanrintama</i> <i>elämän alku</i>	27	45,0	33	55,0	60
– tekijä ja väline (1)	<i>kansanedustaja</i> <i>eläinten pitäjä</i>	33	94,3	2	5,7	35
– teko ja tulos (2)	<i>luonnonsuojelua</i> <i>eläinten kohtelu</i>	33	56,9	25	43,1	58
– paikka ja tilaisuus (3)	– –	–	–	–	–	–
Verbit		–	–	3	100,0	3
– finiittimuoto	– –	–	–	–	–	–
– infinitiivi	– –	–	–	–	–	–
– partisippi	– <i>luonnon sanelema</i>	–	–	3	100,0	3
Muut	– <i>asian puolesta</i>	–	–	16	100,0	16
Yhteensä kaikki		119	59,2	82	40,8	201

Taulukko 14. Genetiivimääritteisten sanaliittojen ja yhdyssanojen suhteet

Yllä taulukossa 14. Genetiivimääritteisten sanaliittojen ja yhdyssanojen suhteet on kokoavasti edellä esitetyt tämän testin tulokset. Huomautan silti vielä kerran, että tulos on enemmän suuntaa antava ja lähinnä hypoteesin kaltainen kuin edustava otos koko aineistosta.

4.3.4 Muut määritetyypit

Edellä olen käynyt läpi aineistostani ne yhdyssanat, joita on määrittänyt numero, lyhenne, sanaliitto tai yhdysosa-, nominatiivi- tai genetiivimuotoinen (yhdyss- tai perus)sana. Jäljellä on ryhmä muut määritetyypit, jonka esiintymiä esittelen seuraavaksi.

Tämän ryhmän määrittäjinä ovat muissa kuin nominatiivi- tai genetiivisijoissa taipuneet nominatit (*maassa|muutto*), adverbit (*läsnä, ali, auki, hyvin, pahoin, ohi, ilmi*), taipumattomat adjektiivit (*pikku, eri, ensi, viime*) ja sellaiset itsenäiset joskin prefiksimäiset ainekset, jotka eivät sulaudu perusosaan vaan vaativat yhdysviivan (*ei-*)¹⁶. Lisäksi joukossa on joitain sellaisia sanoja, joiden määrittäjät ovat niin epäselviä, ettei niiden muodosta voi saada selvää tai joita ei voi luontevasti sijoittaa mihinkään muuhun ryhmään (*Maria!-, PeruSS-*). Kaikkiaan ryhmässä on 420 yhdyssanaa, ja niissä on tehty 25 poikkeamaa. Poikkeamaprosentti on siten 6 % – melko lähellä koko aineiston keskimääräistä poikkeamaprosenttia (5,6 %).

¹⁶ Huomattava on, että *ex-*prefiksi on epäitsenäinen ja siksi kuuluu ryhmään yhdysosa-alkuiset.

Tämän ryhmän poikkeamien määrittely kuitenkin on, jos mahdollista, vielä hankalampaa kuin se on genetiivimääritteisten yhdyssanojen ryhmässä. Poikkeavista sanoista osasta on sanakirjassa (MOT) hakusana, johon on helppo tehdä vertailu suoraan – toisaalta näissäkin sanoissa on joitain, joille kielenhuollollisia, ehdottomia perusteita kirjoittaa yhteen tai erilleen ei ole. Tällaisia ovat:

Aineistossa (määrä jos enemmän kuin 1)	s.v. MOT-sanakirjassa, esimerkki tai selitys
(192) <i>anteeksi antava</i>	<i>anteeksiantava, anteeksiantava Jumala</i>
(193) <i>itsestään selvyyksiä (2)</i>	<i>itsestäänselvyys, itsestään selvä asia, selviö, truismi</i>
(194) <i>mukana olo</i>	<i>mukanaolo, Vasemmiston mukanaolo hallituksessa.</i>
(195) <i>olemassa olostaan</i>	<i>olemassaolo, se että on olemassa, oleminen, eksistenssi</i>
(196) <i>puolesta puhuja (2)</i>	<i>puolestapuhuja, puohustaja, puoltaja</i>
(197) <i>uskoon tulostani</i>	<i>uskoontulo, uskovaiseksi tuleminen</i>

Olen merkinnyt nämä poikkeamiksi sekä sanakirjaesiintymän perusteella että oman kieli-intuitioni vuoksi mutta melko vastentahtoisesti, sillä mielestäni näissä sanapareissa ei ole kyseessä mitenkään erityisen kiteytynyt tai terminäinen merkitys, eikä merkitys nähdäkseni muutukaan kirjoittaa sanaparin yhteen tai erilleen, sanoopa oma intuitioni tai sanakirja mitä hyvänsä. Kumpikaan yhdyssanan osa ei ohjaa muodollaan tai merkityksellään erityisesti kirjoittamaan sanapareja ehdottomasti yhteen (kuten esimerkiksi yhdysosa-alkuisissa yhdyssanoissa tapahtuu). Merkityksen muuttumisesta kirjoitustavan mukaan nostan esiin tämän ryhmän poikkeavan sanaparin (esimerkki 198), jossa yhdyssanaksi kirjoitettuna merkitys on kuvainnollinen, joskin sekin kantaa ainakin osittain konkreettisen, erilleen kirjoitetun sanaparin, merkitystä ('jotain sellaista niin erikoista, että siitä ei ole edes kuultu aiemmin').

Aineistossa	s.v. MOT-sanakirjassa, selitys
(198) <i>ennen kuulumatonta</i>	<i>ennenkuulumaton, vertaansa vailla oleva, tavaton</i>

Poikkeamissa on neljä sanaparia, joissa määrittäjänä on taipumaton adjektiivi, ja joista kielenhuolto antaa selvät ohjeet. Taipumattomat adjektiivit kirjoitetaan pääsanastaan erilleen, paitsi kiteymissä ja -inen-adjektiivin kanssa aina yhteen. *Ensi*-adjektiivi kirjoitetaan pääsanastaan erilleen, jos sen merkitys on 'seuraava' mutta yhteen, jos sen merkitys on 'ensimmäinen'. (KO: 256)

(199) *ensi asuntoon*

(200) *ensi sijaisesti*

(201) *eri arvoisuuden*

(202) *pikku hiljaa*

Ensiasunto (esimerkki 199) tarkoittaa erityisesti, termimäisesti ensimmäistä asuntoa (vanhempien luota muuttamisen jälkeen tai ensimmäistä omistusasuntoa), *ensisijaisesti* (200) on *ensisijainen*-adjektiivista muodostettu adverbijohdos ja *eriarvoisuus* (201) on käsite, jonka merkitys on joko 'epätasa-arvoisuus' tai 'arvoltaan erilaista' (MOT s.v. eriarvoisuus). *Pikkuhiljaa* (202) on vakiintunut adverbi, jonka merkitys on jotakuinkin 'hitaasti' (MOT s.v. pikkuhiljaa).

Kaksi poikkeavista sanoista on *päin*-sanoja, joiden suhteen kielenhuolto on ottanut uuden kannan viime vuosina. Aiemmin ohjeena oli, että *päin* kirjoitetaan yhteen partikkelien kanssa (*eteenpäin, tännepäin*) ja erilleen nomineista (*kotiin päin, lännestä päin*) (esimerkiksi SK: 186, huom. painovuosi 2005, tarkistettu versio vuodelta 1996). Nyt voidaan *päin*-sana kielenhuollon mukaan kirjoittaa pääsanastaan yhteen tai erilleen, vaikka alkuosana olisikin adverbi tai partikkeli (KO: 263). Tässä ryhmässä onkin runsaasti *päin*-sanoja, mutta näissä poikkeavissa sanoissa kuitenkin kyseessä on hiukan erilaiset tapaukset: esimerkissä 203 yhteen kirjoittamista ohjaa alkuosa *edes-* ja esimerkissä 204 *päin*-sana onkin alkuosana; sanapari *päinvastoin* esiintyy yhdyssanana myös kielenhuollon antamassa sanalistassa (KO: 265) sekä sanakirjassa (MOT s.v. päinvastoin).

(203) *edes päin*

(204) *päin vastoin*

Kolmella poikkeamalla on sovittu kirjoitusasu: esimerkki 205, joita on kaksi poikkeamaa tässä ryhmässä, on *päinvastoin*-sanan tapaan sanalistassa (KO: 264) ja sanakirjassa (MOT s.v. jotenkuten); esimerkin 206 sanan konventionaalinen kirjoitusasu vaatii yhdysviivan: *dalai-lama* (MOT s.v. dalai-lama). Vierasperäinen titteli ehkä houkuttelee ajattelemaan, että kyseessä olisi erisnimen tapainen sana tai sanapari, ehkä peräti erisnimi, ja siksi osat ovat ajautuneet erilleen ja onpa ne kirjoitettu isoilla alkukirjaimillakin. *Dalai-laman* alkuosa on taipumaton ja vierasperäinen ja siksi sana on tässä ryhmässä; kokonainen yhdyssana kuitenkin on perussubstantiivi, jonka taipuu loppuosastaan nominin tavoin.

(205) *joten kuten (2)*

(206) *Tuomiojan mukaan Kiina ei painostanut Dalai Laman tapaamisessa, --*

Loput seitsemän poikkeamaa ovat useaosaisia, yhdyssanamäritteisiä yhdyssanoja. Esitän niistä sekä aineistoesimerkin, että konventionaalisen kirjoitusasun ja esimerkin alla poikkeaman syyn.

Poikkeama aineistossa

konventionaalinen muoto

(207) *hälläväliä politiikkaan*

hälläväliäpolitiikkaan (tai hälläväliä-politiikkaan)

Edellä olevassa poikkeamassa (207) määrittäjän ja määritettävän kongruenssi puuttuu. Hälläväliä-yhdyssana ei oikeastaan voi taipua (enempää, sen perusosan on *väli*-sanana partitiivi), mutta jos sitä yrittäisi kongruoida niin että sanat voisi kirjoittaa toisistaan erilleen, olisi tulos jotakuinkin: *hälläväli(äi)seen politiikkaan*. Kaksi poikkeamista (208 ja 209) on tehty numeroilmauksissa:

(208) *kuuskut luvulla*

kuuskyluvulla

(209) *puol toista vuotiaankin*

puoltoistavuotiaankin

Vaikka tässä työssä olen lajitellut numeron, merkin ja lyhenteen yhdyssanan alkuosana erityisyhdys-sanatyypin määrittäjäksi, ovat nämä tyypit kuitenkin muodoltaan nominatiivisia (KK: 140) ja sillä perusteella ne kirjoitetaan yhteen seuraavan substantiivin kanssa. *-vuotias*-sana on lisäksi epäitse-näinen yhdyssanan perusosa, joka vaatii **aina** yhteen kirjoitetun määrittäjän, tässä se on siis *puoltoista-*. Seuraavissa poikkeamissa (210–212) määrittäjänä on nimi.

(210) *Orjantappuraa biisi*

Orjantappuraa-biisi

(211) *PeruSS meikäläiset*

PeruSS-meikäläiset

(212) *Suoli 24 kuluttajia*

Suoli24-kuluttajia

Orjantappuraa, vaikka onkin partitiivimuoto, on sitä seuraavan substantiivin *biisi* määrittäjä, koska se on *biisin* nimi. Samalla tavalla *PeruSS-* määrittää nimenomaan (ryhmän) nimenä *meikäläisiä*. Nimeä tarkoittaa myös *Suoli 24*, joka on parodia nettiportaalin nimestä, *Suomi24*-sanasta (IV:S24): *Suoli24-kuluttajat* ovat siis nettiportaaliin kirjoittajia. Viimeinen esimerkki (213) on aineistossa suoraan linkin ([]-merkintä) perässä, joten voidaan arvella sen olevan kommentti linkin sisältöön.

(213) *[] - ”tottakai-sapettaa”*

totta kai -? (totta kai – sapettaa?)

Yhdysviiva voi tarkoittaa ajatusviivaakin, ja siinä tapauksessa sen ympärillä pitäisi olla sanavälit. Tästä esimerkistä on vaikeaa sanoa mitään varmaa, eli se jää väistämättä epäselväksi.

Seuraavaksi tarkastelen, miten tämän ryhmän yhdyssanat ovat jakautuneet sanaluokkiin. Kuten muissakin ryhmissä, yhdyssubstantiiveja on eniten, 198 kappaletta eli koko ryhmästä 47,1 %. Seuraavaksi eniten on adjektiiveja, 100 – 23,8 %. Muut-ryhmän sanoja on 80 – 19 % ja yhdysverbejä 42 – 10 %. Huomattava on yhdysverbien suuri määrä tässä ryhmässä, mikä onkin helposti selitettävissä: adverbi on luonteeltaan verbin määrite ja kun tämän ryhmän määritteistä iso osa on tosiaan taipumattomia adverbeja, ei ole yllätys, että verbit hakeutuvat nimenomaan tähän ryhmään. Kun lasketaan yhteen vielä verbikantaiset adjektiivit, eli partisiipit ja karitiivit, sekä substantiivien deverbaalit johdokset – teko ja tulos, tekijä ja väline sekä paikka ja tilaisuus -ryhmät – ja partisiipit, määrä nousee huomattavasti: 207 kappaletta eli liki puolet (49,2 %) tämän ryhmän sanoista on verbejä tai verbikantaisia nomineja. Esitän vielä taulukossa 15. Yhdyssanat, muut-ryhmä koottuna, miten tämän ryhmän yhdyssanat poikkeamineen jakautuvat sanaluokkiin.

Sanaluokka	Esimerkki poikkeava konvention mukainen	Määrä	Poikkeamia	Poikkeama- prosentti
Adjektiivit		100	2	2,0
– (i)nen-johdos	<i>viimeaikaisten</i> –	69	–	–
– karitiivi	<i>ennen kuulumatonta</i> <i>yhteensopimatonta</i>	3	1	33,3
– partisiippi	<i>ei-sensuroitu</i> –	26	–	–
– perusadjektiivi	<i>puol toista vuotiaanki</i> <i>ylileveää</i>	2	1	50,0
Substantiivit		198	15	7,6
– perussana	<i>ensi asuntoon</i> <i>ensi-illoissa</i>	60	9	15,0
– erisnimet	<i>Pikku cheek</i> –	2	–	–
– tekijä ja väline nimi (1)	<i>puolesta puhuja</i> <i>yksinyrittäjiä</i>	28	3	10,7
– teko ja tulos (2)	<i>uskoon tulostani</i> <i>esiintuloa</i>	107	3	2,8
– paikka ja tilaisuus (3)	– –	–	–	–
– partisiippi	<i>valtaapitäviä</i> –	1	–	–
Verbit		42	1	2,4
– finiittimuoto	<i>Allekirjoitin.</i> –	17	–	–
– infinitiivi	<i>ylläpitämään</i> –	16	–	–
– partisiippi	<i>anteeksi antavan</i> <i>ilmiantanut</i>	9	1	
Muut	<i>taaksepäin</i> <i>joten kuten</i>	80	7	8,8
	Yhteensä kaikki	420	25	6,0

Taulukko 15. Yhdyssanat, muut-ryhmä, sanaluokan mukaan

Taulukossa huomattavaa on edellä esitetyn lisäksi se, että että poikkeamia on verbeissä vain yksi ja adjektiiveissakin vain kaksi. Keskimäärin ryhmän poikkeamaprosentti (6 %) on lähellä koko aineiston keskiarvoa (5,6 %).

Tämän ryhmän määrittäjänä on ylivoimaisesti eniten ryhmään muut sijoittunut kielenaines, 262 – 62,4 %. Ne ovat taipumattomia adverbeja (*alle, ohi, yli, päin*) tai leksikaalistuneita nomini-kantaisia sanoja (*edellä, toisin, väärin, perin*) – sanan alkuperää en ole lähtenyt selvittämään kaukaa, vaan muut-ryhmälle kriteerinä on ollut sanakirja (MOT): jos sanakirjassa on määriteosan sana selityksellä "Sana on taipumaton tai vaillinaisesti taipuva" on se sijoittunut tähän ryhmään (esimerkiksi MOT s.v. puolesta). Seuraavaksi eniten on adjektiiveja, 73 – 17,4 %, ja vasta kolmanneksi suurimpana ryhmänä ovat substantiivit, jotka muissa ryhmissä ovat ylivoimaisesti isoin määrittäjäjoukko. Tässä ryhmässä niiden määrä on vain 62, eli 14,8 % kaikista määrittäjistä. Numeraaleja ja pronomineja on melko marginaalisesti, yhteensä 23 – 5,5 %. Suhteellisesti eniten poikkeamia on tehty numeraalimääritteisissä yhdyssanoissa, mutta koska ryhmä on niin marginaalinen, ei sen prosenttilukuun kannata takertua, eikä tässä mikään määrittäjäryhmä nouse erityisen selvästi esiin. Esitän vielä taulukossa 16. Yhdyssanat, muu-ryhmä määritysosan sanaluokan mukaan koottuna, miten tämän ryhmän yhdyssanat poikkeamiseen jakautuvat määritysosiensa perusteella.

Määriteosan sanaluokka	Esimerkki poikkeava konvention mukainen	Määrä	Prosenttia kaikista	Poikkeamia	Poikkeama-prosentti
Substantiivi	<i>uskoon tulostani kantaaottavin</i>	62	14,8	4	6,5
Adjektiivi	<i>pikku hiljaa ensisijaisesti</i>	73	17,4	4	5,5
Numeraali	<i>kuuskut luvulla kolminkertaistaa</i>	8	1,9	2	25,0
Pronomini	<i>moninkertaistuu –</i>	15	3,6	0	0,0
Muut	<i>puolesta puhujia yksinkertaisesti</i>	262	62,4	15	5,7
Yhteensä kaikki		420	100	25	6,0

Taulukko 16. Yhdyssanat, muut-ryhmä, määritysosan sanaluokan mukaan

Yhdyssanojen kombinaatioista (ks. myös taulukko 17) suurin joukko on muu-määritteiset substantiivit eli MM+S-sanat, joissa perusosan substantiivia määrittää adverbi (esimerkki 214) tai muu aines (esimerkki 215). Näitä on vajaa kolmannes koko ryhmästä, 29,3 %.

(214) *yksinoikeuden*

(215) *ei-ideologiaa*

Seuraavaksi eniten on MM+M-sanoja, joissa sekä määrittäjä että perusosa ovat adverbeja (esimerkki 216). Näitä on 60, eli 14,3 %.

(216) *päinvastoin*

Kolmanneksi suurimmassa, MS+S-ryhmässä, substantiivi määrittää substantiivivia (217):

(217) *käyttöönottoa*

Tässä ryhmässä on 47 sanaa, eli 11,2 % tämän ryhmän sanoista. Kolme suurinta ryhmää ovat yhteensä 54,7 % koko ryhmästä, ja loput tyypit jakautuvat melko marginaalisesti. Koska ryhmä kuitenkin on hiukan sekalainen ja siksi ehkä jopa vaikeasti hahmotettava, esitän kaikista lopuistakin olemassa olevista sanoista esimerkit:

Sanatyyppi	esimerkki	määrä – %-osuus
<i>MM+A</i>	(218) <i>päinvastainen</i>	45 – 10,7 %
<i>MM+V</i>	(219) <i>irtisanoutua</i>	34 – 8,1 %
<i>MA+A</i>	(220) <i>ainutlaatuisten</i>	33 – 7,9 %
<i>MA+S</i>	(221) <i>pikkumiehen</i>	27 – 6,4 %
<i>MA+M</i>	(222) <i>ensisijaisesti</i>	13 – 3,1 %
<i>MS+A</i>	(223) <i>neuvoa-antavan</i>	13 – 3,1 %
<i>MP+A</i>	(224) <i>minkäänlaista</i>	8 – 1,9 %
<i>MP+M</i>	(225) <i>täälläpäin</i>	4 – 1 %
<i>MP+V</i>	(226) <i>moninkertaistuu</i>	3 – 0,7 %
<i>MN+V</i>	(227) <i>kaksinkertaistua</i>	3 – 0,7 %
<i>MN+M</i>	(228) <i>kahtakymmentätuhatta</i>	3 – 0,7 %
<i>MN+S</i>	(229) <i>kuuskut luvulla</i>	1 – 0,2 %
<i>MN+A</i>	(230) <i>puol toista vuotiaankin</i>	1 – 0,2 %
<i>MS+V</i>	(231) <i>lukuunottamatta</i>	1 – 0,2 %
<i>MS+M</i>	(232) <i>jumalauta!</i>	1 – 0,2 % (tämä voisi olla verbikin)

Määriteosien ja perusosien kombinaatiot selviävät koostettuna vielä taulukosta 17. Yhdyssanojen muut-ryhmän kombinaatiot.

Sana- luokka	MÄÄRITEOSA											
	Substantiivi		Adjektiivi		Numeraali		Pronomini		Muut		Yht.	
	n	%	n	%	n	%	n	%	n	%	n	%
S	47	11,2	27	6,4	1	0,2			123	29,3	198	47,1
A	13	3,1	33	7,9	1	0,2	8	1,9	45	10,7	100	23,8
V	1	0,2			3	0,7	3	0,7	34	8,1	41	9,8
M	1	0,2	13	3,1	3	0,7	4	1,0	60	14,3	81	19,3
									Yhteensä		420	100,0

Sanaluokka:

S = substantiivi, A = adjektiivi, V = verbi, M = muu

Taulukko 17. Yhdyssanojen muut-ryhmän kombinaatiot

Liitteessä 6 on lisäksi kaikkien muut-ryhmän yhdyssanojen koontitaulukko, johon on koottuna kaikki tämän ryhmän yhdyssanojen määrät poikkeamiseen.

Näin kaikki aineistoni yhdyssanat on tullut käytyä läpi. Seuraavaksi esitän vielä tiivistelmäomaisen yhteenvedon, jossa kokoan yhteen kaikki esitetyt yhdyssanat.

4.4 Yhteenvedo kaikista yhdyssanoista

Tässä koosteluvussa tarkastellaan lyhyehkösti kaikkien aineistoni yhdyssanojen yhdyssanatyyppejä, poikkeamia, sanaluokkia sekä määriteosan ja perusosan kombinaatioiden suhteita.

Kaikkiaan aineistossa on siis 7634 yhdyssanaa, joissa on tehty kaikkiaan 431 poikkeamaa, poikkeamaprosentiksi tulee siten 5,6 %. Yhdyssanatyypiltään aineisto jakautuu siten, että siinä on:

- 4 yhdysnumeraalia 0,05 % *kakskyt* ei poikkeamia 0 %
- 129 summayhdyssanaa 1,69 % *maailman* 10 poikkeamaa 7,8 %
- 7500 määritysyhdyssanaa 98,24 % *poikaleirillä* 416 poikkeamaa 5,4 %
- 1 epäselvä 0,01 % *tottakai-sapettaa* 1 poikkeama 100 %

Nominatiivimääritteisten yhdyssanojen kohdalla pohtimani (ks. sivu 58) epäselvät yhdyssanatyypit sijoitin yllä olevaan listaan seuraavasti: osan kaksiosaisista nimistä (*Suomi24, Esa-Pekka, Annastiina, Ruuhonen-Lerner, ViikkoVapaus*) sijoitin summayhdyssanojen joukkoon, samoin *meikämanne-, kuus-nolla- ja satanolla-ilmaisut*; määritysyhdyssanoihin laskin *vaihtoehdon* ja ne nimet, joista voi erottaa määritysosan, esimerkiksi *Ikävalko, Seinäjoki ja Savonlinna*. Yksi jäi vielä jaottelun jälkeenkin epäselväksi. Tämän aineiston valossa näyttää siltä, että summayhdyssanoihin kasautuu poikkeaa-

mia enemmän kuin määritysyhdyssanoihin – tosin kovin pitkälle meneviä johtopäätöksiä ei näin pienen aineiston valossa tietenkään kannata tehdä.

Tämän tutkielman yhdyssanat ovat jakautuneet sanaluokkiin seuraavasti:

- substantiiveja 6506 85,2 %
- adjektiiveja 876 11,5 %
- verbejä 55 0,7 %
- muita 197 2,6 %

ISK:n mukaan kaunokirjallisuudessa yhdyssanoja on tekstisanoista: substantiiveja 82,1 %, adjektiiveja 8,85 %, verbejä 0,05 % ja muita yhteensä 8,8 % (ISK § 407, muiden yhteenlaskenta omani). Substantiivit ja adjektiivit noudattelevat samansuuntaisesti lukuja, mutta verbien ja muiden osalla ero on kohtalaisen iso. ISK:n tietolähteenä on Eino Mikkolan saksankielinen teos vuodelta 1971, eikä minulla valitettavasti ole mahdollisuutta tutkia alkuperäisteosta, joten tässä jää valistuneen arvauksen varaan arvio, että kaunokirjallinen teksti on koherentimpaa kuin nettikirjoittelun tapainen kohtalaisen lyhyt ja yhteen asiaan keskittyvä jutustelu ja siksi se vaatii myös enemmän muuta kielenainesta (konjunktioita, partikkeleita ja adverbjeja) siteeksi kuin aineistoni kaltainen teksti. Saukkosen (Saukkonen 1973: 338) mukaan Nykysuomen sanakirjan yhdyssanat ovat jakautuneet suunnilleen seuraavasti: substantiiveja 89 %, adjektiiveja 10 %, verbejä 0,3 % ja partikkeleita 0,7 %. Hiukan näissäkkin on eroa oman aineistoni lukuihin, mutta kaiken kaikkiaan varsin lähellä ollaan. Saukkonen on tarkastellut vielä yhdysadjektiiveja ja -substantiiveja lähemmin ja osittain vertailua voi jatkaa tämän työn aineistoon. Hänen mukaansa yhdysadjektiiveista nominatiivialkuisia on noin 75 %, genetiivialkuisia noin 22 % ja muulla alkavia noin 3 %. Tämän tutkielman aineiston yhdysadjektiiveista: 47,6 % on nominatiivialkuisia, 19,1 % genetiivialkuisia, 17,5 % yhdysosamääritteisiä, 4,5 % erityisyhdyssanoja ja muulla muodolla alkavia 11,4 %. Saukkosen yhdyssanat ovat siis sanakirjasanoja, eli niissä ei luultavasti ole juurikaan numero- tai lyhennemääritteisiä adjektiiveja (*24-vuotiaalle*, *32-sivuinen*, *EU-kritiikittömät*), joita tässä aineistossa sitä vastoin on runsaasti. On myös mahdollista, että erilaiset analyysitavat ovat siirtäneet tässä työssä nominatiivialkuisista yhdysadjektiiveista runsaan joukon sanoja erityisyhdyssanojen ryhmään (esimerkiksi juuri mainitut numero- ja lyhennealkuiset sanat). Genetiivimäärite ei mitä luultavammin kuitenkaan ole kovin herkkä tällaisille vaihteluille, ja tämän työn ja Saukkosen laskelmien määrät ovatkin melko lähellä toisiaan. Saukkonen nostaa sanakirjan yhdyssubstantiiveista esiin kaksi erityistyyppiä. Hänen mukaansa yhdysubstantiiveissa on prefiksi- ja casus componens -alkuisia noin 7 %. Tässä työssä en kuitenkaan ole eritellyt prefiksejä erikseen, vaan ne ovat itsenäisyytensä tai epäitsenäisyytensä mukaan jakautuneet

joko yhdysosa-alkuisiin tai muulla määritteellä alkaviin yhdyssanoihin, joten vertailua ei voi tehdä. Saukkonen ilmoittaa myös verbikantaisia yhdyssubstantiiveja olevan 12 % (Saukkonen, 1973: 338). Omassa aineistossani verbikantaisia substantiiveja on 24,2 % kaikista yhdyssubstantiiveista, eli ero on melkoinen. On kuitenkin mahdollista, että jälleen erilaiset analyysitavat ja sanojen frekvenssi vaikuttavat niin paljon, ettei deverbaalien yhdyssubstantiivien vertailukaan kannata. Hämmentävän lähellä kuitenkin kaikkiaan sanakirjasanojen prosentuaaliset osuudet ovat tämän aineiston määriä, etenkin ottaen huomioon esiintymien vaikutuksen tähän aineistoon (tekstiaineisto, eli toistoa ja frekventtejä ilmaisuja on) ja Saukkosen aineistoon (sanakirja-aineisto, ei toistuvia ilmaisuja). Taulukossa 18. Yhdyssanat sanaluokittain esitän tämän koko työn aineiston jakautumisen sanaluokkiin poikkeamiseen.

Sanaluokka	Esimerkki poikkeava konvention mukainen	Määrä	Poikkeamia	Poikkeama- prosentti
Adjektiivit		876	42	4,8
– (i)nen-johdos	<i>ULKO MAALAISIA</i> <i>keski-aikainen</i>	729	21	2,9
– karitiivi	<i>ennen kuulumatonta</i> <i>pitkäaikaistyötön</i>	23	1	4,3
– partisippi	– <i>valkopesty</i>	46	–	–
– perusadjektiivi	<i>puol toista vuotiaanki</i> <i>ylileveää</i>	78	20	25,6
Substantiivit		6506	373	5,7
– perussana	<i>vero rahoilla</i> <i>aamukahvi</i>	4517	300	6,6
– erisnimet	<i>Etelä Savoona</i> <i>Linnasilta</i>	413	18	4,4
– tekijä ja väline nimi (1)	<i>nato komentajan</i> <i>asianajaja</i>	309	11	3,6
– teko ja tulos (2)	<i>terveys palveluihin</i> <i>puurakentamisen</i>	1250	43	3,4
– paikka ja tilaisuus (3)	– <i>ajatushautomon</i>	6	–	–
– partisippi	<i>2 tehtävä</i> <i>valtaapitäviä</i>	11	1	9,1
Verbit		55	1	1,8
– finiittimuoto	<i>ulosmitataan</i> –	22	–	–
– infinitiivi	<i>allekirjoittamassa</i> –	20	–	–
– partisippi	<i>anteeksi antavan</i> <i>kyseenalaistanut</i>	13	1	7,7
Muut	<i>joten kuten</i> <i>tiukkapiipoisesti</i>	197	15	7,6
Yhteensä kaikki		7634	431	5,6

Taulukko 18. Yhdyssanat sanaluokittain

Taulukosta on nähtävillä, että ryhmiä, joissa ei ole ainuttakaan poikkeamaa ovat: adjektiivipartisiipit (XX+Ap), substantiivien paikka ja tilaisuus -johdosryhmän sanat (XX+Sl), verbien finiittiset muodot (XX+Vf) ja verbien infinitiivimuodot (XX+Ve) – verbeissä on tehty suhteellisesti vähiten poikkeamia ylipäättään ja toisaalta taas perusadjektiiveissa niitä on tehty suhteellisesti eniten.

Taulukossa 19. Yhdyssanojen kombinaatiot esitän vielä koottuna, miten yhdyssanat ovat yhdistyneet sanaluokittain. Ylivoimaisesti yleisin tyyppi on S+S-sanat, joita on 71,52 %. Seuraavaksi suurin ryhmä ovat A+S-sanat, joita on 8,59 % ja kolmantena tulevat S+A-sanat, 5,34 %.

Sana- luokka	MÄÄRITEOSA											
	Substantiivi		Adjektiivi		Numeraali		Pronomini		Muut		Yht.	
	n	%	n	%	n	%	n	%	n	%	n	%
S	5460	71,52	656	8,59	78	1,02	10	0,13	302	3,96	6506	85,22
A	408	5,34	247	3,24	58	0,76	80	1,05	83	1,09	876	11,47
V	10	0,13	2	0,03	3	0,04	3	0,04	37	0,48	55	0,72
M	69	0,90	28	0,37	12	0,16	10	0,13	78	1,02	197	2,58
											Yhteensä	7634 100,0

Sanaluokka:

S = substantiivi, A = adjektiivi, V = verbi, M = muu

Taulukko 19. Yhdyssanojen kombinaatiot

Kun substantiivi on sanaluokka, joka koskee entiteettejä, konkreettisia ja abstrakteja asioita, näyttää siltä, että hahmotamme maailman etupäässä asioina ja niiden väliset suhteet (muut), määrät (numeraalit), toiminta (verbit) ja kuvailut (adjektiivit) jäävät huomattavan taka-alalle, ainakin silloin, kun kyse on yhdyssanoista.

Liitteessä 7 esitän koottuna koko aineiston yhdyssanat alaryhmineen ja poikkeamineen. Seuraavaksi esittelen aineistoni sellaiset sanaliitot, jotka on kirjoitettu poikkeavasti yhteen; ne ovat siis mieltyneet kirjoittajalle yhdyssanoina – niitä ei kuitenkaan ole tässä tutkielmassa laskettu edellä käsiteltyyn yhdyssanojen joukkoon.

4.5 Yhteen kirjoitetut sanaliitot

Konvention vastaisesti yhteen kirjoitettuja sanaliittoja on aineistossani kaikkiaan 334. Niitä olisi ollut pari enemmän, mutta jätin tästä ryhmästä pois sellaiset aivan ilmeiset kirjoitustavasta johtuvat yhteen ajautuneet ilmaisut kuin esimerkkien 233 ja 234 tapauksessa olevat alleviivatut sanaparit.

(233) *käytin samaa ostosta vertailussa kuin Kataja(kok) eduskunnassa.*

(234) *maahan pitäisi hommata pikapuolin hallitus-perusteluna että ulkolaiset valtiovieraat joutuvat siirtämään vierailujaan*

Esimerkissä 233 on jäänyt sananväli pois sulkeiden edeltä, mutta sanat kyllä nähdäkseni erottuvat sulkeella tarpeeksi, jotta niitä voidaan pitää erillisinä tässä yhteydessä, vaikka kirjoittajalle ilmiselvää

lipsahdus tietysti onkin tapahtunut. Samantapainen lipsahdus on esimerkin 234 kohdalla: yhdysviiva on käytetty ajatusviivan asemasta ja jätetty sen molemmiin puolin kuuluvat sananvälit pois. Näitä tapauksia on todella aivan kourallinen ja mielestäni ne ovat ilmeisyytensä vuoksi niin eri tyyppiä kuin muut tässä käsiteltävät sanat, joten niiden sivuuttaminen on perusteltua; palaan tähän vielä loppuksi luvussa 4.5.5 Kirjoitusvirheet.

Konvention vastaisesti yhteen ajautuneet sanaparit muodostavat melko sekalaisen joukon, mutta jonkinlaisia merkitysryhmiä ja tapaustyyppejä siitä voi erottaa. Yhdyssanojen tapainen, sanaluokkiin keskittyvä, käsittely ei ole tässä kuitenkaan järkevä, joskin yksi ryhmä (luvussa 4.5.3 Sulautumat) tätäkin puolta sivuaa. Valotan yhteen liittyneitä sanaliittoja eri tyyppien – jotka ovat keskenään kovin erilaiset – näkökulmasta. Kielenhuollon säännöt yhdyssanoista liittyvät myös sanaliittoihin, sillä ne määrittelevät, milloin sanaparit kirjoitetaan yhteen ja milloin erilleen. Käsittelin näitä sääntöjä yhdyssanojen kohdalla niin kattavasti, ettei niihin ole syytä enää sanaliittojen yhteydessä palata. Osa sanaliitoiksi ajautuneista sanapareista esiintyy yleisesti sanakirjoissa (etenkin liittosanat, luku 4.5.2) ja kielenhuollonkin antamissa sanalistoissa (ks. esim. KO: 264–265). Esitän aineistoni konvention vastaiset sanaliitot esimerkkeinä kattavasti mutta mahdollisimman tiiviisti. Koska ryhmissä on päällekkäisyyksiä eivätkä ne siten ole toisiaan poissulkevia, täsmällisiä tilastoja en tästä ryhmästä esitä, mutta yksittäisten ilmaisujen ja ilmaistyyppien esiintymämääriä kuitenkin nostan esiin harvinaisten tai yleisten tapausten toteamiseksi.

4.5.1 Numeroilmaisut ja yhdysviiva

Suurin yksittäinen ryhmä, jonka voi poikkeavista sanaliitoista erottaa, ovat erilaiset numeroilmaisut, jotka kuvailevat esimerkiksi prosentti- tai rahamääriä, ajanjaksoja ja ylipäättään numeerisesti ilmaistavia asioita. Numero on osallisena peräti 107 poikkeavassa sanaliitossa, eli noin kolmanneksessa tästä ryhmästä. Tyypillisesti numero kirjoitetaan suoraan yhteen määritettävänsä kanssa (esimerkit 235 ja 236); joskus yhteen kasautuu pidempiäkin ilmaisuja (esimerkit 237 ja 238).

(235) *2.60€*

(236) *4vuotta*

(237) *max.2000e/kk*

(238) *500milj.euroa*

Koko aineistossa on numero + perusosa -pareja 367. Poikkeavasti niistä on kirjoitettu siis 107 eli 29,2 %. Poikkeamamäärä on melko merkittävä ja seurailee – päinvastaiseen tapaan eli yhdistyen kun ei pitäisi yhdistyä – erityisyhdyssanojen poikkeamaprosenttia (29,8 %). On ilmeisen selvää, että

numeron ja muun merkin sekä lyhenteen liittäminen sanaan tuottaa ongelmia kirjoittajille. Joko tarpeellinen yhdysviiva jätetään pois (ks. luku 4.2 Yhdysviiva ja erityisyhdys-sanatyypit) tai numero kiilautuu kiinni määritettäväänsä.

Yhdysviiva on aiheuttanut hankaluutta tässäkin ryhmässä, ja luonnollisesti niin, että sitä on käytetty konvention vastaisesti sitomaan ilmaisuja, joiden kuuluisi olla erillään. Esitän seuraavaksi selvyuden vuoksi esimerkkien lisäksi myös konventionaalisen muodon niistä sanoista, joissa yhdysviiva on liikaa. Luettelossa on kaikki tämän ryhmän yhdysviivapoikkeamat.

Aineistoesimerkki	konvention mukainen tai alkuperäinen
(239) <i>5-tähden</i>	<i>5 tähden</i>
(240) <i>20-pennin</i>	<i>20 pennin</i>
(241) <i>20-asteen</i>	<i>20 asteen</i>
(242) <i>7-Päivää-lehti</i>	<i>7 Päivää -lehti</i>
(243) <i>HS-Politiikka</i>	<i>HS Politiikka</i> (alkuperäinen)
(244) <i>kiva-koulu</i>	<i>KiVa Koulu</i> (alkuperäinen)
(245) <i>Juhla-Mokkaa</i>	<i>Juhla Mokkaa</i> (alkuperäinen)
(246) <i>"Jutan"-poppoo</i>	<i>"Jutan" poppoo</i>
(247) <i>Keskustan-kannattajat</i>	<i>Keskustan kannattajat</i>
(248) <i>Mikki-Hiiren</i>	<i>Mikki Hiiren</i>
(249) <i>silloin-tällöin</i>	<i>silloin tällöin</i>
(250) <i>Käteisen -kelkkaan</i>	<i>Käteisen kelkkaan</i>
(251) <i>OY-Suomi-AB</i>	<i>Oy Suomi Ab</i>

Ensimmäiset esimerkit (239–242) ovat tyypiltään samankaltaisia kuin muutkin tämän ryhmän poikkeavat yhteen kirjoitetut numeroilmaisut, mutta niissä määrite on sidottu yhdysviivalla perusosaan. *7-Päivää-lehti*-ilmaisussa loppuosan *-lehti*-sanon liittyminen perusosaan (*7 Päivää*) on konvention mukainen, joskin sen kirjoitustavassa on horjuntaa viimeisen yhdysviivan kohdalla (ks. konvention mukainen asu); sanaliittopoikkeama on siis syntynyt ensimmäisen yhdysviivan kohdalla, numeron *7* ja *Päivää*-sanon välissä.

Jo aiemmin tässä työssä (ks. Nominatiivimääritteiset substantiivit, s. 49) nostin esiin sellaiset ilmaisut, jotka ovat jo lähtökohtaisesti normin vastaiset. Esimerkit 243 ja 244 ovat tätä tyyppiä (ks.

kirjoitusasu: IV:HS¹⁷ ja IV:KK) sekä esimerkki 245 *Juhla-Mokka* (kirjoitusasu: IV:JM) jopa sama, joka oli jo aiemmin esillä (s. 49). Oikeastaan poikkeama on siis paradoksaalisesti syntynyt siitä, että kirjoittaja ei ole toistanut poikkeamaa vaan on kirjoittanut – tai pyrkinyt kirjoittamaan – konvention mukaisesti. Esimerkit 246–248 ovat ihmistarkoitteisia – Mikki Hiirikin on siinä määrin inhimillinen, että hänet voidaan tähän joukkoon laskea – mutta kaksi ensimmäistä kuitenkin ryhmämerkityksessä. On mahdotonta arvailla näiden poikkeamien syitä, mutta ehkä ihmistarkoitteisissa sanaliitoissa voisi olla jonkinlainen tendenssi ajautua yhteen, vallankin kun esimerkeissä 246 ja 247 on kyseessä tarkkarajainen joukko. Toisaalta määrittäjänä on genetiivialkuinen erisnimi, samoin kuin esimerkissä 250, ja siten ne lähenevät tyypiltään GS(e)+S-tyypin eli *Lontoon-matka*-tyypin yhdyssanaa. Vaikka yhtään konvention mukaan kirjoitettua tämän tyypin yhdyssanaa ei aineistossani olekaan, ei se tarkoita, etteikö niitä olisi ollenkaan ja etteikö samalla muotilla voitaisi kirjoittaa sanoja konvention vastaisesti: *Inksan-keikka* > *Jutan-poppoo*. Esimerkki 249 on hyvin liittosanamainen merkityksessään 'joskus' ja tulee vastaan seuraavassa luvussa (4.5.2 Liittosanat). Viimeinen esimerkki 251 kuvaa Suomea yrityksenä ja on siten melko kuvainnollinen. Kuvainnollisuus onkin ehkä syy siihen, miksi kirjoittaja on halunnut sitoa osakemuodot Suomi-sanaan. Palaan kuvainnollisiin ilmaisuihin vielä luvussa 4.5.3 Sulaumat.

4.5.2 Liittosanat

Liittosana vastaa muodoltaan sanaliittoa: se koostuu kahdesta tai useammasta osasta, jotka kuuluvat yhteen. Häkkisen mukaan liittosana eroaa sanaliitosta siten, että sanaliitto voi tarkoittaa mitä tahansa kahden sanan yhteenliittymää, kun taas liittosanan merkitys on ilmaistavissa yhdellä sanalla tai se on eriytynyt osiensa merkityksestä (Häkkinen 1990: 147–148). Tässä työssä sanaliitto-käsite on ollut tähän asti riittävä siksi, että näkökulma sanaliittoon on muodostunut yhdyssanan kautta: sanaliitto on tarkoittanut yksinkertaisesti kahta sanaa, jotka on kirjoitettu joko konvention mukaan yhteen tai sen vastaisesti. Kun nyt tarkastellaan sanaliiton sisäisen rakenteen tiivyyttä, on liittosana-käsitteelle tilaa. Tässä työssä rajaan kuitenkin liittosanojen ryhmään ainoastaan ne sanat, joiden merkitys on selvästi ilmaistavissa yhdellä sanalla; kuvainnolliset merkitykset käsittelem seuraavassa luvussa (4.5.3 Sulaumat).

Liittosanoiksi katsottavia yhteen kirjoitettuja sanaliittoja on 84. Tyypeiltään ne ovat etupäässä adverbeja, kuvaavat tapaa tai intensiteettiä, aikaa ja puhujan suhtautumista. Esitän tässä ensin yksisanaisen variantin ja sitten sitä tarkoittavat aineistoesimerkit luettelona, esiintymämäärä sanan perässä sulkeissa.

17 *HS Poliitiikkaan* on aineiston keräyksen jälkeen tullut lisää *.fi*-liite, eli nyt sen muoto on *HS.fi Poliitiikka*. Vaikka muoto on muuttunut, on se edelleen konvention vastainen.

Aikaan viittaavat:**vastaava sana**

aina, jatkuvasti
nyt
heti
lopussa, lopulla, lopulta

äskettäin, taannoin
joskus
kauan

aineistoesimerkit

(252) *kokoajan (6), kaiken aikaa (1), iätajat (1)*
 (253) *tällä hetkellä (3)*
 (254) *samantien (1), samantein (1), ensitilassa (1)*
 (255) *viimemetreillä (1), viimekädessä (1),
 ennenpitkään (1)*
 (256) *viimeaikoina (1)*
 (257) *silloin-tällöin (1)*
 (258) *aikapäiviä (1)*

Tapaan, paikkaan, määrään ja intensiteettiin viittaavat:**vastaava sana**

kuten
toisin
kovasti, paljon, tuntuvasti
siten
sieltä
voimalla, tehokkaasti
ankarasti, kiireellä, ponnistellen
paljon
vähän
jotenkin
kaikkea, kaikki
paikalta, tapahtumasta

aineistoesimerkit

(259) *niinkuin (13), niinkun (4), ikäänkuin (1)*
 (260) *eritavalla (4)*
 (261) *aikalaila (3)*
 (262) *sitämyöten (1), sillätavalla (1), sillälailakin (1)*
 (263) *sieltäkäs (2)*
 (264) *olantakaa (1)*
 (265) *hikihatussa (1)*
 (266) *moninverroin (1)*
 (267) *jossain määrin (1)*
 (268) *jollainlailla (1)*
 (269) *kaiken maailman (1)*
 (270) *paikan päältä (1)*

Puhujan suhtautuminen, näkökulman painotus, kysymys**vastaava sana**

ehdottomasti
edelleen, lisäksi, toiseen, vielä
kuitenkin
edelleen, myös, samoin
miksi
kuitenkin, toisaalta
toivottavasti
oikeastaan, todellisuudessa

aineistoesimerkit

(271) *tottakai (11)*
 (272) *sitäpaitsi (4), sitäpaitti (1)*
 (273) *jokatapauksessa (2)*
 (274) *niinikään (1)*
 (275) *minkävuoksi (1), minkätakia (1)*
 (276) *sensijaan (1)*
 (277) *toivomukseen (1)*
 (278) *itseasiassa (1)*

Adjektiivinen**vastaava sana**

hyvä, kiinnostava, tavoiteltava
huolimaton

aineistoesimerkit

(279) *"tavoittelemisen" arvoinen (1)*
 (280) *juostenkustulta (1)*

Muu**vastaava sana**

itse

aineistoesimerkit

(281) *omasuu (1)*

Määristä näkyy, ettei kovin frekventtejä ilmaisuja ole kuin *niinkuin* ja *tottakai*. *Niinkuin*-sanon frekvenssiä selittää vanha tapa: ennen se oli yhdyssana ja edelleen varmasti ainakin osittain kirjallisuuden ja muiden vanhojen tekstien vuoksi sellaiseksi mieltyy. Esimerkkilause e1 on vuonna 1873 julkaistusta ja sen jälkeen varsin laajasti luetusta klassikosta, Kiven Seitsemästä veljeksestä (Kivi 2000

[1873]: 18); esimerkki e2 Sillanpään vuonna 1934 julkaistusta Ihmisiä suviyössä -romaanista (Sillanpää 1948 [1934]: 151) ja esimerkki e3 Shakespearen Romeosta ja Juliasta, suomennos vuodelta 1955 on Yrjö Jylhän käsialaa (Shakespeare 1955: 125).

- (e1) *Minä en kiellä, en käske; tee niinkuin sydämesi parhaaksi katsoo.*
- (e2) *– hän tiesi myös, että emäntä oli jo ehtoolla mennyt Syrjämäkeen, niinkuin pitikin mennä sellaisen ihmisen, joka nyt vähän enemmän ymmärsi kuin joku tavallinen akka.*
- (e3) *Ikäänkuin se nimiois tuliputken ammus, joka hältä veis hengen, niinkuin serkun hältä tappoi kirottu käsi tää.*

Niinkuin ei juuri vaihda merkitystään kirjoitetaan se sitten yhteen tai erilleen, mutta *tottakai*-sanan osien erilleen kirjoittamisen suositusta on vaikea ymmärtää. Sanaparinhan voi helpostikin ymmärtää tarkoittavan yhdessä melkein päinvastaista kuin erilleen kirjoitettu (*tottakai* = 'ehdottomasti'; totta kai = 'mahdollisesti totta'); sen yhteen kirjoitettu variantti on kolmitavuinen, kahdeksan merkin mittainen sana, eli varsin kohtuumittainen ja helposti hahmotettava. Erilleen kirjoittaminen ei siis mitenkään merkittävästi selkeytä erityisen pitkää sanaa, sen merkityskin on aivan eri – joten miksi ihmeessä kielenhuollossa on päädytty muotoon *totta kai*?

Muut ilmaisut eivät siis toistu kovin usein, mutta toisaalta ovat niin samankaltainen joukko, että niiden esittäminen mielestäni kannattaa jo sen vuoksi, että käy ilmi, miten – aika usein jopa lyhyemmät – konvention mukaiset ilmaisut varioivat poikkeavien ilmaisujen kanssa. Poikkeamat saattavat rakenneosiansa merkityksiltä hahmottua yksinkertaisempina, konkreettisempina, kuin niiden yksisanainen konvention mukainen vastineensa ja olla puhekielisempinä tutumpiakkin. Koska sanaparien merkitys on kuitenkin kovin kiinteä, sanamainen, ajautuvat ne yhteen ja seurauksena on, jokseenkin paradoksaalisesti, kaikkiaan pidempi ilmaus kuin konvention mukainen yhden sanan ilmaus olisi ollut.

Tässä ryhmässä on joitain sellaisia sanoja, jotka näyttävät esiintyvän muita useammin yhdysosan osina, ja palaan niihin luvussa 4.5.4 Liittoutuvat sanat.

4.5.3 Sulaumat

Osa sanoista on sulautunut yhteen niin, että sanapari muodostaa jonkinlaisen yhtenäisen, konkreettisen merkityksen. Jaottelin niitä sanaluokkiin sen mukaan, mikä niiden merkitys on. Näin lajiteltuja sanapareja kertyi 79. Merkitykset eivät ole samalla tavalla selviä kuin yhdyssanoissa, mutta joitain samoja piirteitä on. Esimerkiksi objektin ja verbin sulaumia on muutamia, esimerkit 282:

- (282) *veditherneet, perheidenyhdistämisiin, maailmanmarkkinahonnaksilasketaan, huumoriakinpitää, elokuvantekemisestä, kuullasitä, laskipäätänsä*

Verbi ja adverbi ovat ajautuneet yhteen esimerkeissä 283:

- (283) *yhdessätekemisen, vaihtumisenjälkeen, poisjääminen, poisjättämistä, mitenmene, kuriinlaittamiseksi (2), kiusaamisenpolulle*

Verbin ja subjektin liittoja lienevät esimerkit 284:

- (284) *seon, mäensaa, mitääntuli, meilläöolisi*

Yksi verbiketjukiin on, esimerkki 285:

- (285) *pääsiskattoo*

Konkreettisia substantiiveja ovat esimerkin 286 sanat:

- (286) *valkoisemiehen, Uuttavuotta (3), Turunsinappia, Tummapaahto, T:kokoomuksenjäsen, suomenpoikaa, suomenkansan, Suomenkansalaiset, Suomenarmeijan, rek.hieroja, PetenKoiratarvike, Perushumalaistensivu, paraslaulaja, Natonjäsenenä, linnanjuhlissa (2), kuvasusta, Kepunkuntia, kaisaniemenpuistossa, jehovantodistajikaan, HyväJyrki!, hiivasyndroomanoireet, HIENOKUVA, eduskunnanpuhemiesheltä*

Kaikki ilmaisut tarkoittavat yhtä asiaa, osalla on referenttikiin (esimerkiksi *kaisaniemenpuistossa, HyväJyrki!, PetenKoiratarvike, Perushumalaistensivu*). Kuvailuvia, adjektiivisia substantiiveja joukossa ovat esimerkiksi *valkoisenmiehen* ('länsimaalainen'), *suomenpoika* ('suomalainen'). Edellisiä esimerkkejä (286) abstraktimpia ovat seuraavat substantiivit, esimerkki 287:

- (287) *työnperässä, toissavuonna, suorapuhe, rankkapaikka, pieniosa, parinviikon, parihengen*

Näistä osan voi hyvin mieltää adverbiksiin, esimerkiksi *pieniosa* tarkoittaa 'vähän', *rankkapaikka* 'vaikea'. Adverbityyppisiä ilmaisuja rakennetaankin substantiivien kautta, merkitys on silloin kuvainnollinen, kuten esimerkeissä 288–290:

- (288) *niminimeltä*

- (289) *rivienvälistä*

- (290) *samassaveneessä*

Kuvainnollisuus jollain tavalla saattaisi selittää, miksi sanat on haluttu kirjoittaa yhteen; yhtenä yhdyssanan kriteerinähän on eriytynyt merkitys, joka usein on kuvainnollinen. Tässä tyypissä toimisi siis jälleen muotti sanojen muodostajana (ks. 2.3 Yhdyssanan kehyksiä: jne.). Seuraavissakin esimerkkisanoissa (291) merkitys on jokseenkin adverbiaalinen ja siten abstrakti, mutta samankaltaista kuvainnollisuutta kuin esimerkeissä 288–290 niissä ei ole.

- (291) *juuriedellispäivänä, kauhojemmekaa, loppuakohti, pitkäaikaan, työnperässä*

Johonkin näiden välimaille asettuu esimerkin 292 ilmaus, jonka tarkoitus on esittää, että jokin asia menee liian pitkälle ja sen vuoksi tulee hylätyksi, tarpeettomaksi tai huonoksi (kyseessä on ravitermi *pitkä laukka*: jos hevonen laukkaa ravikilpailussa liian pitkään, se hylätään kisasta).

(292) *pitkälaukka*

Adjektiivisissa ilmaisuissa perusosana on usein infiniittinen verbimuoto, esimerkit 293:

(293) *ennaltaehkäisevää, ennaltasovittuna, koroillaelävälle, poliittisestisitoutumaton, ulospäinsuuntautuvassa, uudelleen katsottava, ylläoleva*

Nämä esimerkin 293 sanat ovat sellaisia, että aivan ehdottomasti ne eivät edes ole poikkeamia: on mahdollista nähdä ne kiinteinä, erityismerkityksisinä ja vakiintuneina käsitteinä. Myös muunlaisia adjektiivismerkityksisiä yhteensulautuneita sanaliittoja on, esimerkit 294:

(294) *hyväntahdon, parasmahdollinen, pitkäntähtäimen, tosiharvoista*

On tietenkin ihan mahdoton – ainakin tällä otannalla – sanoa, miksi tämän ryhmän ilmaisut ovat ajautuneet yhteen. Sanat ovat kiistatta merkitystekemisissä toistensa kanssa, mutta se on yksinään kestämätön peruste niiden yhteen kirjoittamiselle, sillä samalla perusteella pitäisi kaikki sanat samassa kontekstissa kirjoittaa yhteen. Osalla kuitenkin on selvästi yksi referentti ja se varmasti selittää yhteensulaumaa. Myös samakantaisten sanojen vaikutuksia joukosta on havaittavissa, esimerkiksi *hyväntahdonlähettiläs* > *hyvätahto*. Erityisesti sellaiset yhteensulautuneet sanat, joiden osana on pronomi, ovat kuitenkin niin hankalia ymmärtää, että tyydyn tässä vain lyhyesti esittämään ne, esimerkit 295:

(295) *monellamuutoslaisella, munmielestä, määmikkää, seuus*

4.5.4 Liittoutuvat sanat

Jotkin sanat näyttävät ajautuvan yhdyssanoiksi muita useammin. Tällaisia ovat esimerkiksi jotkin taipumattomat adjektiivit (esimerkit 296), adverbit (esimerkit 297), ja konjunktiot (esimerkit 298).

(296) *viimevuosi (5), viimehallituskeikallakin, viimemetreillä, viimeaikoina, viimekädessä, eripuolilla, eriasiaa*

(297) *yhtäpaljon (2), yhtäaikaa (2), ainakun*

(298) *ennenkuin (7), niinkuin (17)*

Hankalimmat kirjoittajalle näistä erillisistä pikkusanoista näyttävät olevan *kuin-* ja *kun-*sanat, jotka ovat kaikkiaan 34 kertaa liittyneet konvention vastaisesti toiseen sanaan. Hiukan ehkä yllättäviä yhteen liittyviä sanoja aineistossani ovat verbit *ei* ja *olla*, kummallakin 7 konvention vastaista esiintymää (esimerkit 299 ja 300).

(299) *eise, eikun (4), eikai (2)*

(300) *onse (2), onkai, onjo, ollaja, olittekste, olehyvä*

Nämä pikkusanat (kaikki lukuun ottamatta *viime*-sanaa ovat nelikirjaimisia tai pienempiä) ovat ilmeisen epäitsenäisiä merkitykseltään ja ehkä siksi ajautuvat kiinni toisiin sanoihin. Osaan vaikuttaa luultavasti puhekieliset variantitkin: erityisesti puhepartikkeleina esiintyvä *niinku* saattaa toimia muottina niin *niinkuin*-liittymälle kuin muillekin *kuin-* ja *kun-*sanoille (*mutku, muutaku, eiku* jne.). Puhekielisiä ovat myös *no*-keskustelupartikkeliin sitoutuneet *Nojoo...* ja *Noniin. Niin, näin ja vaan* ovat myös tällaisia sanoja ja esiintyvätkin kukin kerran muissa yhtymissä kuin jo mainituissa: *sitä-vaan, näinse ja niinmitä*.

4.5.5 Kirjoitusvirheet

Osa tämän ryhmän sanoista on suoraan arvioitava virheiksi, koska niille ei voi kuvitella minkäänlaista selitystä. Tätä tyyppiä on jokseenkin vähän, ja selityksenä virheellisyyksille voi pohtia silkan lyöntivirheen lisäksi sitä, että Facebook saattaa näyttää rivitetyn tekstin jollain tavalla väärin: sanat ovat esimerkiksi eri rivillä mutta julkaisu- tai enter-napin painalluksen jälkeen ne ovatkin yhdessä. Facebookissa ei ole mahdollista tehdä sitovia välilyöntejä, eli osa yhteen ajautuneista numeroilmaisista saattaa olla tietoisia: on haluttu pitää numero samalla rivillä perusosan kanssa. Joitain yli- ja varsinkin epämääräisiä yhdysviivoja saattaa selittää tavutuksen tarve, eli kirjoittaja on halunnut jakaa sanat eri riveille, ja ne ovat saattaneet Facebookissa vielä ollakin niin, mutta tässä keräysaineistossa ne kuitenkin ovat ajautuneet taas peräkkäin ja edelleen poikkeamaksi. Koska näitä on mahdoton arvioida tai tietää, ei mitään poikkeamia ole syytä käsitellä vahinkoina – lisäksi on huomattava, että kun esimerkiksi välilyönti usein joko osuu liikaa juuri yhdyssanan rajalle tai jää juuri siitä pois, missä yhdyssanamaisuuden voi hahmottaa, ei suuri tällaisten poikkeamien joukko voi olla sattumaa. Sattuman pitäisi tilastollisesti osua yhtä usein muihinkin sananosiin, ja näinhän ei ole (ne olisivat silloin mukana aineistossa). Esitän seuraavaksi luettelona sellaiset poikkeamat, kaikkiaan 11, joita en niiden epäselvyyden vuoksi alkanut edes analysoida ja jotka voinee sijoittaa juuri sattuman kautta syntyneisiin virheisiin.

(301) *hetkelläJos, jailman, jaKes., jooaika, Nimmaristaja, sivuillaniVaali tilaisuuksistani, työtäjä, vastavaalien, vastaavaneduskunnan, valtakunnan oikeusaloittaa, silläSuomessa*

5 LOPUKSI

Kaiken kaikkiaan tutkimusaineistosta jää minulle käsitys, että vaikka yhdyssanavirheistä puhutaan toisinaan paljonkin ja usein varsin kärtyiseen sävyyn, konventioista tehdään poikkeamia lopulta varsin vähän. Mielestäni kuva internetiin kirjoittavien kirjoitustaidosta, ainakin yhdyssanojen osalta, on varsin myönteinen. Poikkeavat yhdyssanat saattavat toki osua tekstistä silmään konvention mukaisesti kirjoitettuja herkemmin ja siksi voi tuntua, että niitä on paljon, mutta tämän aineiston valossa näyttää pikemminkin siltä, että niitä on melko vähän.

"Vähän" on tietenkin suhteellinen käsite ja toki voidaan sanoa, että 5,6 % on poikkeamien määräksi aivan liikaa. Vastaani ei tätä tutkielmaa kirjoittaessa kuitenkaan tullut ainuttakaan sellaista poikkeamaa, jossa sanan alkuperäinen tai tarkoitettu merkitys muuttuisi niin paljon, että poikkeavasta muodosta seuraisi suorastaan sekaantumisen vaara – ja täydellisyyteen kielen kanssa tuskin päästään koskaan. Jos ja kun yhdyssana- ja sanaliittopoikkeamia, kielenhuollon kannalta ja maallikoiden käsityksen mukaan kirjoitusvirheitä, kuitenkin halutaan entisestäänkin vähentää, se käy luultavasti tuloksellisimmin puuttamalla hankalimpiin kohtiin: numeroiden, merkkien ja lyhenteiden liittämistä sanaan ja yhdysviivan käyttöä yleisestikin opettamalla saataneen kohennettua tilastoja jo huomattavasti. Osviittaa yhdyssanojen ja sanaliittojen muihinkin hankaliin kohtiin saa tämän tutkielman tilastoista, vaikka työni ei luonnollisesti reseptinä näihin kielivaivoihin voikaan toimia.

Toisesta suunnasta katsoen – vaikka tämän tutkielman poikkeamien määrä on mielestäni pikemminkin positiivisen vähäinen kuin masentavan runsas – täytyy huomata, että jo yhdyssanojen ja sanaliittojen kohdalla poikkeamia tehneitä kirjoittajia on kielenopasryhmästä melkein viidennes. Jos tarkasteltaisiin nettikirjoittelua muiden kielenhuollon alueiden näkökulmasta, arvelen, että poikkeamia tekevien joukko kasvaisi huomattavasti ja virheiden vaikutuksetkin alkaisivat näkyä selvemmin. Esimerkiksi pilkun käyttöön liittyvät virheet saattavat muuttaa asiasisältöjäkin jo merkittävästi. Äidinkielen opetus ja kielenhuolto ovatkin yhä tärkeämpi osa kansalaistaidon rakentumista, kun enenevässä määrin viestitään kirjoittamalla verkon yli. Sinänsä on ilahduttavaa, että ihmiset kirjoittavat rohkeasti ja paljon huolimatta siitä, tekevätkö he virheitä. Joskus internetin kielikeskusteluja seurattaessa on tullut mieleen, että ehkä äidinkielen opetuksessa ja kielenhuollossa olisi syytä opettaa myös kieleen liittyvää hyvää käytöstä ja tehdä selväksi, että kirjoittajan kielitavoille naureskelu ja kielivirheiden pilkkaaminen on yksinkertaisesti huonoa käytöstä, vaikka itse konventionaalisen kielen hallitsisikin.

Tätä tutkielmaa voi toivoakseni lukea myös muusta kuin opetuksen ja kielenhuollon suunnasta. Kun suunnittelin tätä työtä, minulla oli kunnianhimoinen tavoite tehdä yleisyystutkimuksen lisäksi laadullista analyysiä ja paneutua poikkeamien semantiikkaan perusteellisesti. Varsin pian, jo suunnitteluvaiheessa, kävi ilmi, että tavoite oli työmäärältään pro gradu -työksi mahdoton: jo tässä työssä käytetyn aineiston luokittelu ja tilastointi osoittautui varsin mittavaksi urakaksi. Yhden työhön kirjatun aineistoa kuvaavan luvun vuoksi olen saattanut tehdä tunteja työtä. Olen kuitenkin lopputulokseen tyytyväinen. Aineiston määrä, sen käsittelyn työläydestä huolimatta, on varsin perusteltu, sillä kvantitatiivista tutkimusta ei voi tehdä kovin pienillä määrillä. Toisaalta jos aikoo tehdä laadullista tutkimusta, pitää ensin tietää, mitä tutkii. Tilastotieto kielestä on tärkeää silloin, kun halutaan tutkia kielen yleisiä piirteitä eikä takertua marginaaliin. Jos taas halutaan tutkia marginaalia, on tiedettävä, mikä se on. Murteita on tutkittu, koululaisten kirjoitelmia ja sanomalehtikieltä on analysoitu, mutta tietääkseni tavallisten ihmisten internetkirjoituksia ei ole kovinkaan paljon vielä tutkittu. Vaikka ei juuri tämän aineiston tapaista kieltä – yhdyssanoja, sanaliittoja tai ylipäättään tavallisten ihmisten kirjoittamaa (internet)kieltä – tutkisikaan, voi tästä tutkimuksesta olla apua vertailevana aineistona. Toivon siis, että tutkimuksestani hyötyvät paitsi kielenopettajat ja -huoltajat myös ne, jotka tutkivat yhdyssanoja tai sanaliittoja ja tarvitsevat tietoa niiden yleisyydestä, tyypeistä, poikkeamista ja niin edelleen.

Olen tarkoituksella jättänyt tähän tutkielmaan solmimattomia langanpätkiä, ratkaisemattomia ongelmia, jotka ehkä innostavat jatkopohdintaan tai jopa -tutkimuksiin. Sanaliittojen ja yhdyssanojen suhdetta raapaisin pinnasta juuri tässä tarkoituksessa: niiden suhteen tutkiminen olisi paitsi mielenkiintoista myös varsin relevanttia sanojen yhdistymiseen liittyvää selvitystyötä. Ihmistarkoitteisia sanoja pohdin käsitteellisyyden näkökulmasta: ihminen tarkastelee asioita omalla inhimillisellä tavallaan ja siksi ihmisiin liittyvä saattaa kielessäkin näkyä painokkaammin, ehkä sanojen yhdistämisessä nimenomaan ihmistarkoitteisten yhdyssanojen runsautena. Toisaalta voi löytyä myös muita merkitystyyppisiä, jotka kiteytyvät mieluusti yhdyssanoiksi. Erottaviakin tyyppisiä pohdin hiukan. Erityisen kiinnostavaa olisi tutkia sitä, millaisin ehdoin nominatiivimuotoinen substantiivimäärite irtoaa perusosasta, silloin kun se irtoaa, kuten *kivi talosta*. Onko olemassa joitain erotettavia merkitysryhmiä, joissa muita useammin yhdyssanan osat jakautuvat sanaliitoksi? Ylipäättään yhdyssanojen jako merkitysryhmiin olisi mielenkiintoista; vanhastaanhan sitä on jokin verran harrastettu, mutta mahdollisesti aiempiin jakoihin löytyisi uusia näkökulmia ja ehkä jopa ryhmiä tutkielmani aineiston kielestä. Adjektiivimaisuuteen liittyvät yhdyssanojen osat – substantiivien adjektiivimaisuus määriteosana sekä partisiippien adjektiivimaisuus perusosana – olisivat kiinnostavia tarkasteltavia nimenomaan sanaluokkien rajankäynnin näkökulmasta: millaisia määritteitä partisiipit voivat yh-

dyssanan perusosina saada ja miten se muuttaa niiden verbimäisyyttä adjektiivien suuntaan; miten substantiivi taipuu kuvailevaksi, kun se saa määritettävän? Yhdyssanojen konventioita olisi mielenkiintoista tarkastella ihmislähtöisesti: onko kirjoittajassa itsessään, hänen koti- tai koulutaustassaan tai asenteessaan kirjoittamiseen syitä sille, miten yhdyssanat mieltävät? Kiinnostavaa olisi tarkastella myös, miten kirjoitusympäristö ohjaa yhdyssanojen esiintymiseen: johtaako esimerkiksi ympäristön paine kirjoittamaan sanat konvention vastaisesti tai sen mukaan? Tämän aihepiirin ympärillä pyörivistä kysymyksistä on suorastaan runsauden pulaa ja mielestäni ne ovat kiehtovia jo sananmuodostuksenkin, kielen lähtökohdan, alkupisteen, kannalta.

LÄHTEET

- HAKULINEN, AULI – VILKUNA, MARIA – KORHONEN RIITTA – KOIVISTO, VESA – HEINONEN, TARJA RIITTA – ALHO, IRJA 2004: *Iso suomen kielioppi*.
Verkkoversio: <http://scripta.kotus.fi/visk/etusivu.php> 18.12.2011
- HEIKKINEN, VESA – MANTILA, HARRI 2011: *Kielemme kohtalo*. Gaudeamus Helsinki University Press, Helsinki.
- HÄKKINEN, KAISA 2004: *Nyky-suomen etymologinen sanakirja*. WS BOOKWELL OY. JUVA.
- HÄKKINEN, KAISA 1990: *Mistä sanat tulevat*. Sisälähetysseuran kirjapaino Raamattutalo. Pieksämäki.
- IISA, KATARIINA – OITTINEN, HANNU – PIEHL, AINO 2006: *Kielenhuollon käsikirja*. Yrityskirjat Oy. Jyväskylä: Gummerus Kirjapaino Oy.
- IKOLA, OSMO 2001: *Nyky-suomen opas*. Åbo Akademis tryckeri. Turku.
- ISK = HAKULINEN, AULI – VILKUNA, MARIA – KORHONEN RIITTA – KOIVISTO, VESA – HEINONEN, TARJA RIITTA – ALHO, IRJA 2004: *Iso suomen kielioppi*.
Verkkoversio: <http://scripta.kotus.fi/visk/etusivu.php> 18.12.2011
- IV = KS. INTERNETVIITTEET KOOTTUNA, LIITE 1
- JAAKOLA, MINNA 2004: *Suomen genetiivi*. Suomalaisen Kirjallisuuden Seuran Toimituksia. Helsinki.
- KANKAANPÄÄ, SALLI (TOIM.) – HEIKKILÄ, ELINA – KORHONEN, RIITTA – MAAMIES, SARI – PIEHL, AINO 2010: *Kielitoimiston oikeinkirjoitusopas*. Kotimaisten kielten tutkimuskeskuksen julkaisuja 147. Tampere: Juvenes Print.
- KIELITOIMISTON SANAKIRJA 2.0. VERKKOVERSIO:
<http://helios.uta.fi:2070/mot/uta/netmot.exe?motportal=80> [viitattu 8.12.2011]
- KIVI, ALEKSIS 2000: *Seitsemän veljestä*. Otavan Kirjapaino Oy. Keuruu.
- KK = IISA, KATARIINA – OITTINEN, HANNU – PIEHL, AINO 2006: *Kielenhuollon käsikirja*. Yrityskirjat Oy. Jyväskylä: Gummerus Kirjapaino Oy.
- KO = TOIM. KANKAANPÄÄ, SALLI – HEIKKILÄ, ELINA – KORHONEN, RIITTA – MAAMIES, SARI – PIEHL, AINO 2010: *Kielitoimiston oikeinkirjoitusopas*. Kotimaisten kielten tutkimuskeskuksen julkaisuja 147. Tampere: Juvenes Print.
- KOSKI, MAUNO 1982: *Suomen johto-opin morofologiaa*. Åbo Akademis kopieringscentral. Turku.
- LAAKSONEN, KAINO – LIEKO, ANNELI 2003: *Suomen kielen äänne- ja muoto-oppi*. Oy Finn Lectura Ab. Helsinki: Hakapaino Oy.
- LANTEE, ANNA 2010: *Pitkät yhdyssubstantiivit Helsingin Sanomissa*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.
- LEINO, PIRKKO 2005: *Suomen kielioppi*. Kustannusyhtiö Otava. Otavan Kirjapaino Oy. Keuruu.

- LEVONEN, SUVI 2006: *Verbaaliyhdyssanat HS 2000–2001 -korpuksessa*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.
- MOT = Kielitoimiston sanakirja 2.0. Verkkoversio:
<http://helios.uta.fi:2070/mot/uta/netmot.exe?motportal=80> [viitattu 8.12.2011]
- MÄKISALO, JUKKA, 2000: *Grammar and Experimental Evidence in Finnish Compounds*. University of Joensuu. Faculty of Humanities. Joensuun yliopistopaino. Joensuu.
- NES = HÄKKINEN, KAISA 2004: *Nykysuomen etymologinen sanakirja*. WS Bookwell Oy. Juva.
- NIEMINEN, TOMMI – LANTEE, ANNA 2008: *Monitulkintaisten yhdyssanojen sivupaino*, saatavilla 30.11.2011: <http://papers.legisign.org/posters/lantee-nieminen2008.pdf>
- NIEMINEN, TOMMI – LANTEE, ANNA a2010: *Yhdyssana suomen kielessä ja puheessa*, saatavilla 30.11.2011: <http://papers.legisign.org/posters/nieminen-lantee2010.pdf>
- NIEMINEN, TOMMI – LANTEE, ANNA b2010: *Yhdyssana suomen kielessä ja puheessa*, saatavilla 30.11.2011: <http://papers.legisign.org/presentations/nieminen-lantee2010.pdf>
- NIEMINEN, TOMMI – LANTEE, ANNA 2011: *Yhdyssanasta suomen kielessä ja puheessa*, saatavilla 31.1.2013: http://epublications.uef.fi/pub/urn_isbn_978-952-61-0391-4/urn_isbn_978-952-61-0391-4.pdf
- OINONEN, SARI 2008: *Rakenteellinen kompleksistuminen peruskoulun seitsemäs-, kahdeksas- ja yhdeksäsluokkalaisten kirjoitelmissa*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.
- PAJUNEN, ANNELI – PALOMÄKI, ULLA 1984: *Tilastotietoja suomen kielen rakenteesta 1*. Kotimaisten kielten tutkimuskeskus. Valtion painatuskeskus. Helsinki.
- SAUKKONEN, PAULI 1973: *Suomen kielen yhdyssanojen rakenne*. Teoksessa Suomalais-ugrilaisen seuran toimituksia 150. Suomalais-ugrilainen seura. Suomalaisen Kirjallisuuden Kirjapaino Oy. Helsinki.
- SAUKKONEN, PAULI – HAIPUS, MARJATTA – NIEMIKORPI, ANTERO – SULKALA, HELENA 1979: *Suomen kielen taajuussanasto*. Porvoo: WSOY.
- SHAKESPEARE, WILLIAM 1955: *Romeo ja Julia*. Teoksessa *William Shakespearen suuret draamat I*. Kustannusyhtiö Otavan kirjapaino. Keuruu.
- SILLANPÄÄ, F. E 1948: *Ihmisiä suviyössä*. Kustannusyhtiö Otavan kirjapaino. Helsinki.
- SK = LEINO, PIRKKO 2005: *Suomen kielioppi*. Keuruu: Otavan Kirjapaino Oy.
- VESIKANSA, JOUKO 1989: *Nykysuomen sanavarat*. WSOY:n graafiset laitokset. Juva.
- YLIKULJU, MARJA 2003: *Internetin seuranhakuilmoitusten yhdyssanat*. Suomen kielen pro gradu -tutkielma. Oulun yliopisto.

LIITE 1**Internetviitteet koottuna:**

IV:FX = <http://www.finlex.fi/fi/laki/ajantasa/1998/19980628#a628-1998> [viitattu 3.5.2013]

IV:GH = <http://www.google.fi/search?q=yhdyssanavirhe> [viitattu 31.1.2013]

IV:KP = <http://kaksplus.fi/keskustelu/plussalaiset/mitas-nyt/1138634-arsyttavat-yhdyssanavirheet-palstalaisilla/> [viitattu 31.1.2013]

IV:FB = Facebookin internetviitaukset tekstissä aakkostettu [viitattu 14.2.2013]
a = <http://www.facebook.com/groups/109345692436185/>

IV:SVT1

Suomen virallinen tilasto (SVT): Väestön tieto- ja viestintätekniiikan käyttö [verkkajulkaisu].
ISSN=1799-3504. 2012, 1. Internetin käytön muutokset . Helsinki: Tilastokeskus
[viitattu: 21.4.2013].

Saantitapa: http://www.stat.fi/til/sutivi/2012/sutivi_2012_2012-11-07_kat_001_fi.html

IV:SVT2

Tilasto: Väestön tieto- ja viestintätekniiikan käyttö [verkkajulkaisu].

ISSN=1799-3504. 2005, Internetin käyttötarkoitukset keväällä 2005, prosenttia internetin käyttäjis-
tä. Helsinki: Tilastokeskus [viitattu: 21.4.2013].

Saantitapa: http://www.stat.fi/til/sutivi/2005/sutivi_2005_2006-06-12_tau_001.html

IV:SVT3

Suomen virallinen tilasto (SVT): Väestön tieto- ja viestintätekniiikan käyttö [verkkajulkaisu].

ISSN=1799-3504. 2012, Liitetaulukko 12. Internetin käyttötarkoitukset 3 kuukauden aikana iän ja
sukupuolen mukaan 2012, %-osuus väestöstä . Helsinki: Tilastokeskus [viitattu: 21.4.2013].

Saantitapa: http://www.stat.fi/til/sutivi/2012/sutivi_2012_2012-11-07_tau_012_fi.html

IV:WS = <http://fi.wiktionary.org/wiki/Liite:Nettislangi#L> [viitattu 31.1.2013]

IV:SLTS = *Suomen sanomalehtikielen taajuussanasto*. 2004 Kielipankki, CSC - Tieteen tietotekniikan
keskus Oy [viitattu 23.2.2013]

<http://www.csc.fi/tutkimus/alat/kielitiede/taajuussanasto-B9996/view>

Taajuussanaston lisenssiehdot: <http://creativecommons.org/licenses/by-nd-nc/1.0/fi/>
[viitattu 26.2.2013]

IV:JM = <http://juhlamokka.fi/tarina/> [viitattu 11.4.2013]

IV:KO = http://www.kotus.fi/?i=453&s=2607#faq_453 [viitattu 16.4.2013]

IV:PS = <http://www.perussuomalaiset.fi/> [viitattu 24.4.2013]

IV:S24 = <http://www.suomi24.fi/> [viitattu 20.4.2013]

IV:HS = <https://www.facebook.com/HSfiPolitiikka>

IV:KK = <http://www.kivakoulu.fi/> [viitattu 26.4.2013]

LIITE 2**Aineiston keruupaikat, Facebook-ryhmät:**

Kahvi

<http://www.facebook.com/pages/Kahvi/97103637697>

YLE Teema

<http://www.facebook.com/yleteema>

Aku Anka

<http://www.facebook.com/ankkalinna>

Yksikin kiusattu on liikaa

<http://www.facebook.com/group.php?gid=47995964015>

Koirat

<http://www.facebook.com/pages/Koirat/188817855914>

Perussuomalaiset

<http://www.facebook.com/perussuomalaiset>

Kokoomus

<http://www.facebook.com/kansallinenkokoomus>

Suomen Sosialidemokraatit SDP

<http://www.facebook.com/sosialidemokraatit>

Vasemmistoliitto

<http://www.facebook.com/vasemmisto>

Vihreät

<http://www.facebook.com/vihreat>

Jääkiekko

<http://www.facebook.com/pages/JÄÄKIEKKO/313279022148>

KD - Kristillisdemokraatit

<http://www.facebook.com/pages/Kristillisdemokraatit/157614802920>

Suomen Kommunistinen Puolue

<http://www.facebook.com/group.php?gid=17771783816>

Suomen Senioripuolue

<http://www.facebook.com/group.php?gid=109478305749058>

Itsenäisyyspuolue

<http://www.facebook.com/pages/Itsen%C3%A4isyyspuolue/135338196509910>

Köyhien asialla

<http://www.facebook.com/group.php?gid=148510568534>

Piraattipuolue

<http://www.facebook.com/Piraattipuolue>

Muutos 2011 - Förändring 2011

<http://www.facebook.com/Muutos2011>

Vapauspuolue

<http://www.facebook.com/pages/Vapauspuolue/130607886992574>

Cheek

<http://www.facebook.com/cheekfb>

LIITE 3

Yhdysosamääritteisten yhdyssanojen koontitaulukko

Kaikki yhdysosa-alkuiset yhdyssanat yhteensä 904

	Määriteosa																							
	Substantiivi			Adjektiivi			Numeraali			Pronomini			Muut			Yhteensä								
	k	k-%	p	k	k-%	p	k	k-%	p	k	k-%	p	k	k-%	p	k	k-%	p						
Substantiivit	206	28,4	2	1,0	0,7	1	367	50,6	12	3,3	3,1	9	19	2,6	17	89,5	133	18,3	2	1,5	725	80,2	16	2,2
perussana	143	69,4	1	0,7	0,7	1	291	79,3	9	3,1	3,1	9	17	89,5	17	89,5	113	85,0	2	1,8	564	77,8	12	2,1
erisnimi	39	18,9					18	4,9									10	7,5			67	9,2		
tekijä ja väline (1)	2	1,0					19	5,2									7	5,3			28	3,9		
teko ja tulos (2)	22	10,7	1	4,5	4,5	1	39	10,6	3	7,7	7,7	3	2	10,5	2	10,5	3	2,3			66	9,1	4	6,1
paikka ja tilaisuus (3)																								
partisiippi																								
Adjektiivit	19	12,4					86	56,2	3	3,5	3,5	3	8	5,2	8	5,2	37	24,2	3	8,1	153	16,9	6	3,9
perussana	4	21,1					8	5,2	1	12,5	12,5	2	2	25,0	2	25,0	10	27,0	1	10,0	24	15,7	2	8,3
inen-johdos	15	78,9					68	79,1	2	2,9	2,9	6	6	75,0	6	75,0	25	67,6	2	8,0	92	60,1	4	4,3
karitiivi							7	8,1													7	4,6		
partisiippi							3	2,0									2				5	3,3		
Verbit							1	25,0									3	75,0			4	0,4		
finiitti							1	100,0									1	33,3						
infiniitti																	2	66,7						
partisiippi																								
Muut							7	31,8									15	68,2	1	6,7	22	2,4	1	4,5
Yhteensä	225	24,9	2	0,9	0,9	2	461	51,0	15	3,3	3,3	27	3,0	3,0	3	0,3	188	20,8	6	3,2	904	100	23	2,5

k-% = sanaluokan kohdalla prosentti kaikista määriteosatyypeistä, alaryhmän kohdalla pääryhmän prosentista (esimerkiksi: substantiivien perussanan prosentti koko substantiivi-substantiivi-sanojen (GS+S-sanojen) joukosta)

LIITE 4

Nominatiivimääritteisten yhdyssanojen koontitaulukko

Kaikki nominatiivialkuiset yhdyssanat yhteensä 5111

	Määriteosa																		
	Substantiivi			Adjektiivi			Numeraali			Pronomini			Muut			Yhteensä			
	k	k-%	p	k	k-%	p	k	k-%	p	k	k-%	p	k	k-%	p	k	k-%	p	
Substantiivit	4350	93,8	215	4,9	247	5,3	14	5,7	15	0,3	0,0	10	0,2	16	0,3	4638	90,7	229	4,9
perussana	3213	73,9	176	5,5	136	55,1	10	7,4	6	40,0		8	80,0	1	6,3	3364	72,5	186	5,5
erisnimi	200	4,6	12	6,0	37	15,0	2	5,4	8	53,3				14	87,5	259	5,6	14	5,4
tekijä ja väline (1)	98	2,3	5	5,1	22	8,9								1	6,3	121	2,6	5	4,1
teko ja tulos (2)	827	19,0	22	2,7	52	21,1	2	3,8	1	6,7		2	20,0			882	19,0	24	2,7
paikka ja tilaisuus (3)	5	0,1														5	0,1		
partisiippi	7	0,2														7	0,2		
Adjektiivit	269	64,5	10	3,7	116	27,8	6	5,2	20	4,8	1	5,0	12	2,9		417	8,2	17	4,1
perussana	17	6,3	4	23,5	3	0,7			2	10,0						22	5,3	4	18,2
inen-johdos	240	89,2	6	2,5	110	94,8	6	5,5	18	90,0	1	5,6	9	75,0		377	90,4	13	3,4
karitiivi	8	3,0														8	1,9		
partisiippi	4	1,5			3	0,7						3	25,0			10	2,4		
Verbit	2	66,7			1	33,3		0,0		0,0			0,0			3	0,1	0	0,0
finiittinen																			
infiniittinen	1	50,0			1	100,0										2	66,7		
partisiippi	1	50,0														1	33,3		
Muut	31	58,5	1	3,2	8	15,1	1	12,5	8	15,1	1	12,5	6	11,3	1	53	1,0	4	7,5
Yhteensä	4652	91,0	226	4,9	372	7,3	21	5,6	43	0,8	2	4,7	28	0,5	1	5111	100	250	4,9

k-% = sanaluokan kohdalla prosentti kaikista määriteosatyypeistä, alaryhmän kohdalla pääryhmän prosentista (esimerkiksi perussanan prosentti koko substantiivi-substantiivi-sanojen (NS+S-sanojen) joukosta)

LIITE 5

Genetiivimääritteisten yhdyssanojen koontitaulukko

Kaikki genetiivialkuiset yhdyssanat yhteensä 882

	Määriteosa																			
	Substantiivi			Adjektiivi			Numeraali			Pronomini			Muut			Yhteensä				
	k	k-%	p	k	k-%	p	k	k-%	p	k	k-%	p	k	k-%	p	k	k-%	p		
Substantiivit	664	98,7	36	5,4	7	1,0	2	0,3							673	76,3	36	5,3		
perussana	284	42,8	27	9,5	7	100,0	2	100,0							293	43,5	27	9,2		
erisnimi	78	11,7	2	2,6											78	11,6	2	2,6		
tekijä ja väline (1)	128	19,3	2	1,6											128	19,0	2	1,6		
teko ja tulos (2)	171	25,8	5	2,9											171	25,4	5	2,9		
paikka ja tilaisuus (3)	1	0,2													1	0,1	0	0,0		
partisiippi	2	0,3													2	0,3	0	0,0		
Adjektiivit	99	59,3	1	1,0	11	6,6						57	34,1	1	1,8	167	18,9	2	1,2	
perussana	1	1,0	1	100,0		0,0									1	0,6	1	100,0		
inen-johdos	89	89,9			11	100,0						57	100,0	1	0,6	157	94,0	1	0,6	
karitiivi	4	4,0													4	2,4	0	0,0		
partisiippi	5	5,1				0,0									5	3,0	0	0,0		
Verbit	7	100,0	0	0,0		0,0									7	0,8	0	0,0		
finiitti	3																			
infiniitti	1																			
partisiippi	3																			
Muut	35	100,0	1	2,9		0,0									35	4,0	1	2,9		
Yhteensä	805	91,3	38	4,7	18	2,0	0	0,0	2	0,2	0	0,0	57	6,5	1	1,8	882	100	39	4,4

k-% = sanaluokan kohdalla prosentti kaikista määriteosatyypeistä, alaryhmän kohdalla pääryhmän prosentista
(esimerkiksi: substantiivien perussanan prosentti koko substantiivi-substantiivi-sanojen (GS+S-sanojen) joukosta)

LIITE 6

Yhdyssanojen
koontitaulukko,
muut määrittäjät

Kaikki ryhmän muut yhdyssanat yhteensä 420

	Määriteosa												Yhteensä							
	Substantiivi			Adjektiivi			Numeraali			Pronomini			Muut			k	k-%	p	p-%	
	k	k-%	p	p-%	k	k-%	p	p-%	k	k-%	p	p-%	k	k-%	p	p-%	k	k-%	p	p-%
Substantiivit	47	23,7	4	8,5	27	13,6	2	7,4	1	0,5	1	100,0	123	62,1	8	6,5	198	47,1	15	7,6
perussana	10	21,3	2	20,0	23	85,2	2	8,7	1	100,0	1	100,0	26	21,1	4	15,4	60	30,3	9	15,0
erisnimi																				
tekijä ja väline (1)	17	36,2	1	5,9	2	7,4														
teko ja tulos (2)	19	40,4	1	5,3	2	7,4														
paikka ja tilaisuus (3)																				
partisiippi	1	2,1																		
Adjektiivit	13	13,0			33	33,0			1	1,0	1	100,0	45	45,0	1	2,2	100	23,8	2	2,0
perussana																				
inen-johdos					33	100,0			1	100,0	1	100,0	1	2,2			2	2,0	1	50,0
karitiivi																				
partisiippi	13	100,0																		
Verbit	1	2,4							3	7,3			34	82,9	1	2,9	41	9,8	1	2,4
finiitti																				
infiniitti	1	100,0							3	100,0			15	44,1			17	41,5		
partisiippi																				
Muut	1	1,2			13	16,0	2	15,4	3	3,7			60	74,1	5	8,3	81	19,3	7	8,6
Yhteensä	62	14,8	4	6,5	73	17,4	4	5,5	8	1,9	2	25,0	262	62,4	15	5,7	420	100	25	6,0

k-% = sanaluokan kohdalla prosentti kaikista määriteosatyypeistä, alaryhmän kohdalla pääryhmän prosentista
(esimerkiksi: substantiivien perussanan prosentti koko substantiivi-substantiivi-sanojen (GS+S-sanojen) joukosta)

LIITE 7

Yhdyssanat
koontitaulukko
kaikki

Kaikki yhdyssanat yhteensä 7634

	Määriteosa																													
	Substantiivi						Adjektiivi						Numeraaali						Pronomini						Muut					
	k	k-%	p	p-%	k	p-%	k	k-%	p	p-%	k	k-%	p	p-%	k	k-%	p	p-%	k	k-%	p	p-%	k	k-%	p	p-%				
Substantiivit	5460	83,9	315	5,8	656	10,1	29	4,4	78	1,2	12	15,4	10	0,2					302	4,6	17	5,6	6506	85,2	373	5,7				
perussana	3818	69,9	255	6,7	463	70,6	22	4,8	65	83,3	11	16,9	8	80,0					163	54,0	12	7,4	4517	69,4	300	6,6				
erisnimi	321	5,9	16	5,0	57	8,7	2	3,5	9	11,5									26	8,6			413	6,3	18	4,4				
tekijä ja väline (1)	249	4,6	9	3,6	42	6,4													18	6,0	2	11,1	309	4,7	11	3,6				
teko ja tulos (2)	1056	19,3	35	3,3	94	14,3	5	5,3	3	3,8			2	20,0					95	31,5	3	3,2	1250	19,2	43	3,4				
paikka ja tilaisuus (3)	6	0,1																					6	0,1						
partisiippi	10	0,2							1	1,3	1	100,0											11	0,2	1	9,1				
Adjektiivit	408	46,6	15	3,7	247	28,2	9	3,6	58	6,6	12	20,7	80	9,1	1	1,3			83	9,5	5	6,0	876	11,5	42	4,8				
perussana	26	6,4	7	26,9	11	1,3	1	9,1	30	51,7	11	36,7							11	13,3	1	9,1	78	8,9	20	25,6				
inen-johdos	347	85,0	8	2,3	223	90,3	8	3,6	28	48,3	1	3,6	75	93,8	1	1,3			56	67,5	3	5,4	729	83,2	21	2,9				
karitiivi	13	3,2			7	2,8							1	1,3					2	2,4	1	50,0	23	2,6	1	4,3				
partisiippi	22	5,4			6	0,7							4	5,0					14	16,9			46	5,3						
Verbit	10	18,2			2	3,6			3	5,5			3	5,5					37	67,3	1	2,7	55	0,7	1	1,8				
finiittinen	3	30,0			1	50,0							2	66,7					16	43,2			22	40,0						
infiniittinen	3	30,0			1	50,0			3	100,0									13	35,1			20	36,4						
partisiippi	4	40,0											1	33,3					8	21,6	1	12,5	13	23,6	1	7,7				
Muut	69	35,0	4	5,8	28	14,2	3	10,7	12	6,1	1	8,3	10	5,1	1	10,0			78	39,6	6	7,7	197	2,6	15	7,6				
Yhteensä	5947	77,9	334	5,6	933	12,2	41	4,4	151	2,0	25	16,6	103	1,3	2	1,9			500	6,5	29	5,8	7634	100	431	5,6				

k-% = sanaluokan kohdalla prosentti kaikista määriteosatyypeistä, alaryhmän kohdalla pääryhmän prosentista (esimerkiksi perussanan prosentti koko substantiivi-substantiivi-sanojen (NS+S-sanojen) joukosta)