

Puhekielisyys verkkokeskustelussa

Suomen kieli ja viestintä
Maisterin tutkinnon tutkielma
Jouni Karppinen
2011

PUHEKIELISYYS VERKKOKESKUSTELUSSA

Tutkimuksen tavoitteet

Tässä tutkielmassa luotiin katsaus kahteen hyvin erilaiseen aihepiiriin, suomalaiseen sosiolingvistiseen variaationtutkimukseen sekä ulkomaiseen verkkokeskustelun tutkimukseen. Perinteistä sosiolingvististä puhekielen variaationtutkimusta sovellettiin uudelle alueelle, kirjoitettuun suomenkieliseen verkkokeskusteluun. Kyse oli siis eräänlaisesta soveltavasta pilottitutkimuksesta. Verkkokeskustelusta tutkittiin sen puhekielisyyttä sekä keskustelijoiden kielenkäyttötapoja, ja eri keskustelutyyppejä verrattiin toisiinsa.

Tutkimusaineisto ja -menetelmät

Tutkimusmenetelmä oli kvantitatiivinen ja perustui siihen, että pääkaupunkiseudun puhekielessä tyypillisesti esiintyvien kielenpiirteiden suhteellisia määriä selvitettiin arkisista verkkokeskusteluista. Internetissä yleisistä keskustelutyypeistä valittiin kolme oletettavasti puhekielisyyttä vahvasti ilmentävää: reaaliaikainen IRC-keskustelu, Facebookin tilapäiviytykset sekä City.fi-portaalin keskustelupalsta.

Tulokset

Verkkokeskustelun puhekielisyyssaste jäi odotusten mukaisesti puhetta matalammaksi. Keskustelijat voitiin jakaa melko selvärajaisiin ryhmiin sen mukaan, ovatko heidän viestinsä vahvasti puhekielisen vaikutelman antavia, pääosin yleiskielisiä vai siltä väliltä. Aivan samat puhekielisyyttä ilmentävät piirrevariantit eivät olleet yleisiä verkkokeskustelussa kuin puheessa, vaan useimmin käytettyjä olivat koko Suomen alueelle levinneet yleissuomalaiset piirteet. Näiden piirteiden yleiskieliset variantit antaisivat mahdollisesti liian kirjakielisen leiman, jota keskustelijat vaikuttivat välttelevän. Kolmesta tutkitusta keskustelutyypistä keskustelupalsta oli selvästi vähiten puhekielinen. IRC-keskustelu ei näyttänyt merkittävästi Facebookia puhekielisempänä, vaikka sen reaaliaikainen luonne olisi antanut näin olettaa. Tämä voi johtua Facebookin käyttäjien suuresta keskinäisestä hajonnasta.

Variaationtutkimus menetelmänä näytti soveltuvan suhteellisen hyvin myös verkkokeskustelun tutkimukseen, mutta sen rajallisuus erityisesti mahdollisten kysymyksenasettelujen suhteen tuli selvästi esiin. Sillä on samat tunnetut heikkoudet kuin puheen tutkimuksessa, kuten kapea-alaisuus ilman muiden tutkimusmenetelmien tukea. Myös aineiston riittävään kokoon on kiinnitettävä huomiota.

Avainsanat

Puhekieli, sosiolingvistiikka, variaationtutkimus, verkkokeskustelu

Sisällysluettelo

1 Johdanto	1
1.1 Tutkimuksen tavoite.....	3
1.2 Tutkielman rakenne	7
2 Puhekieli.....	8
2.1 Käsitteitä.....	8
2.2 Puhekielen elementit	10
2.3 Pääkaupunkiseudun puhekieli	12
2.3.1 Kehitysvaiheet	13
2.3.2 Kielellisiä piirteitä	17
2.4 Sosiolingvistinen puhekielentutkimus Suomessa	26
3 Verkkokeskustelu	30
3.1 Teknisiä ominaisuuksia.....	31
3.2 Kielenkäyttö	33
3.3 Sosiaalisia ominaisuuksia.....	39
3.4 Keskustelutyypit.....	43
3.5 Verkkokeskustelun variaationtutkimus.....	47
4 Analyysi.....	51
4.1 Aineisto.....	51
4.2 Tutkimusmenetelmä	57
4.3 Tulokset.....	60
5 Päätelmät	73
Lähteet	78
Liite 1: Keskustelutyypien puhekielisyyshyvitys ja -asteet piirteittäin	85
Liite 2: Informanttien puhekielisyyshyvitys ja -asteet piirteittäin	86

1 Johdanto

Verkkokeskustelua on tutkittu koko sen reilun parinkymmenen vuoden ajan, jonka se on ollut tavallisten ihmisten ulottuvilla. Tutkimustahti on ollut kiihtyvää, ja verkkoviestinnän tutkimusta on tehty toisaalta muun muassa pedagogisista sekä psykologisista lähtökohdista (Cho 2010) ja toisaalta kielitieteellisesti esimerkiksi diskurssianalyysin, sosiolingvistiikan, semiotiikan sekä etnografian näkökulmista (Herring 2010a). Kiinnostus on ymmärrettävää, koska yhä suurempi osa kaikesta ihmisten välisestä viestinnästä tapahtuu internetin ja muiden viestintäverkkojen välityksellä. Tunnetusti tärkeä sanaton viestintä jää verkkoviestinnässä pitkälti pois, jolloin kielen merkitys korostuu, ja monet käyttävät tietomassoista erottautumiseen juuri kielellisiä tehokeinoja.

Internetissä kieli voi olla ilmiänsä kirjoitetun lisäksi puhuttua. Monilta osin puheen voi olettaa verkossa jäljittelevän tilanteita verkon ulkopuolella (vertaa esimerkiksi Skype-puheluita matkapuhelimeen). Kirjoitetussa kielessä eroja on sen sijaan enemmän (vertaa esimerkiksi sähköpostia kirjeeseen). Siinä on enemmän vaihtelua erilaisten käyttötapojen määrän vuoksi, ja se on säilynyt paljon puhetta ja videota yleisempänä kanavana käyttää kieltä verkossa (Herring 2010a). Myös vahvojen kielellisten normien puuttuminen internetissä johtaa suurempaan vaihteluun (Cho 2010). Näistä syistä juuri kirjoitettu verkkokeskustelu on kiinnostava tutkimusaihe.

Monenlaisten kirjoittamalla tapahtuvien verkkokeskustelujen tiedetään ilmiänsä huolimatta jäljittelevän puhekieltä (ks. esim. Mann & Stewart 2000, 181–184). Tämä tulee esiin erityisesti arkisessa ja vapaamuotoisessa keskustelussa, kuten chateissa. Toisaalta verkkokeskustelun väitetään olevan aivan oma tapansa käyttää kieltä, erillään sekä yleis- että puhekielestä (Crystal 2002, 47–48; Mann & Stewart 2000, 187–188; Herring 2001a). Kieli on jatkuvassa muutoksessa muutenkin, mutta internetin, matkapuhelinten ja muun uuden viestintäteknologian tuleminen on saattanut vauhdittaa muutosta ja luonut kenties uudenlaisia tapoja käyttää kieltä.

Valtaosa verkkokeskustelujen kielenkäyttöä käsittelevästä tutkimuksesta kohdistuu englannin kieleen (Herring 2010a), joka verkkoviestinnän kokonaisuutta ajatellen hallitsee verkkokeskusteluja. Viestintää tapahtuu kuitenkin hyvin monilla kansallisilla

kielillä. Onkin esitetty, että tutkimuksia tarvitaan jatkossa enemmän muista kulttuureista ja kielistä kuin vain englannista (ma.). Esimerkiksi kiina on lähes yhtä käytetty kieli kuin englanti (ks. esim. Gerrard 2007), vaikka länsimaissa emme siihen juuri törmää. Myös suomi on ahkerassa käytössä suomalaisten keskinäisessä verkkoviestinnässä.

Suomen kielen käytöstä verkossa ei tiedetä yhtä paljon kuin englannista, koska verkkoviestinnän kielen tutkimus on ollut pääasiassa ulkomaista. Suomen kieltä on silti tutkittu muuten ahkerasti, esimerkiksi puheen osalta. Yksi suosittu alue on ollut sociolinguvistinen puhekielen ja murteiden tutkimus. Tässä tutkimuksessa yhdistetäänkin perinteisen sociolinguvistisen tutkimuksen näkökulmia verkkokeskustelun tutkimukseen, sillä suomen puhekielen piirteiden tutkimuksesta saadut tulokset muodostavat kiintoisan lähtökohdan myös verkkokeskustelun piirteiden tarkasteluun. Perinteisen sociolinguistiikan tutkimuskohteena on ollut erityisesti puhuttu kieli, tarkemmin ottaen puhutun kielen tilanteinen tai sosiodemografinen vaihtelu (ks. esim. Downes 1988). Verkkoviestinnässä esiintyvä kielenkäyttö kuitenkin luontevasti rikkoo perinteisen puheen ja kirjoituksen välisen rajan, sillä verkossa puhe ja kirjoitus limittyvät monin tavoin toisiinsa (Crystal 2002, 24–48). Monet ihmiset kokevat *kirjoittavansa* sähköposteja mutta *juttelevansa* chateissa ja virtuaalimaailmoissa. Useissa verkon keskustelumuodoissa on esimerkiksi samanlaisia aikarajoitteita kuin puhuessa kasvotusten, sillä toisen reaktiota saatetaan odottaa välittömästi. Toisaalta viestintä käytännössä tapahtuu näppäimistöllä kirjoittamalla, jolloin joudutaan huomioimaan oikeinkirjoituksen kaltaisia asioita.

Tieteellisen tiedon lisäämisen lisäksi verkkokeskustelun kielentutkimuksella voidaan nähdä käytännön sovelluksia. Markkinointiviestinnän kannalta on arvokasta oppia tuntemaan ihmisten viestintätapoja ja kielenkäyttöä verkossa paremmin. Sitä voidaan tällöin kohdistaa tehokkaammin puhuttelemalla vastaanottajia heidän ”omalla kielellään”. Nykymaailmassa mainontaa on pystyttävä kohdentamaan entistä tarkemmin (ks. esim. Malmelin 2003, 125–126, 163), ja kielenkäyttö voidaan nähdä yhtenä keinona siihen. Samalla mainostaminen on siirtymässä yhä enemmän internetiin, mikä luo uudenlaisia mahdollisuuksia. Esimerkiksi Facebookissa käyttäjät

antavat palvelulle itsestään valtavasti tietoa. Markkinoijilla on varsin rajattu pääsy tällaiseen tietoon, mutta muuten sen avulla mainontaa voisi kohdentaa ennennäkemättömän tarkasti (ks. Salmenkivi & Nyman 2008, 122). Kun tähän yhdistetään osallistaminen, asiakkaat voidaan saada jopa mukaan markkinointiponnisteluihin (mts. 221–225). Keskeistä on sujuva vuoropuhelu asiakkaiden ja yrityksen välillä, missä kielenkäyttö nousee tärkeään asemaan.

Toinen mahdollinen sovellusalue on vaalikampanjointi, jossa yhtä lailla on herätty sosiaalisen median hyödyntämiseen henkilökohtaisemman lähestymistavan avulla (ks. esim. Utz 2009). Tätä kirjoitettaessa käytävät eduskuntavaalit osoittavat myös, että ehdokkaat käyttävät valtavasti Facebookia vaalikampanjoissaan, sekä erillisten mainosten että omien profiilisivujensa kautta. Voidaan esimerkiksi pohtia, antaako ehdokas itsestään vaikeasti lähestyttävän kuvan, jos hän käyttää vain kirjakieltä omassa profiilissaan, tai otetaanko häntä vakavasti poliitikkona, jos hän käyttää siellä puhekieltä.

1.1 Tutkimuksen tavoite

Tieteellinen tutkimus suomen kielen käytöstä verkkokeskusteluissa on vielä lapsenkengissään, joten siihen liittyvän tiedon lisääminen on tavoite itsessään. Tässä tutkielmassa työnnytään tuolle alueelle soveltamalla perinteistä puhekielen tutkimusta verkkokeskusteluihin hyödyntäen samalla englanninkielisestä verkkoviestinnän tutkimuksesta saatua tietoa. Sosiolingvistinen kysymyksenasettelu voi nimittäin olla verkossa hyvin samanlainen, vaikka ympäristö kasvatusten ja internetissä keskustelemisen välillä on täysin erilainen (Nuolijärvi 2000). Näiden erojen selvittäminen onkin yksi tuoreista tutkimushaasteista.

Tutkimuksen kohdeilmionä on tapa käyttää puhekieltä jäljittelevää kieltä kirjoitetussa suomenkielisessä arkisessa verkkoviestinnässä. Tutkimusmenetelmä perustuu kvantitatiiviseen suomen puhekielen tutkimusperinteeseen. Tavoitteena on ensisijaisesti selvittää

- **kuinka paljon ja millä tavoin arkinen verkkokeskustelu on puhekielistä.**

Tämä käsittää ensinnäkin sen, mitkä puhekielen ilmiöt siirtyvät verkkokeskusteluun ja missä määrin. Lisäksi pyritään selvittämään, onko verkkokeskustelussa jotain säännönmukaisia kielenkäyttötapoja, joiden mukaan keskustelijoita voisi luokitella. Pää tavoitteeseen liittyvinä alatavoitteina

- selvitetään, kuinka läheisesti verkkokeskustelu pyrkii jäljittelemään puhe- tai yleiskieltä, vai onko se aivan erillinen tapa käyttää kieltä,
- ryhmitellään keskustelijoita heidän kielenkäyttötapansa mukaan,
- vertaillaan erityyppisiä verkkokeskusteluja sen selvittämiseksi, millaisia keskinäisiä eroja niillä on, ja
- arvioidaan metodologista soveltuvuutta, eli empiirisessä puhutun kielen tutkimuksessa käytetyn variaationanalyysimenetelmän sopivuutta verkkokeskustelun tutkimukseen.

Kielen käyttötapaan vaikuttavia tekijöitä voivat olla muun muassa puhujan tausta, viestintäkanava ja keskustelun tyyppi sekä oletettu yleisö tai vastaanottava yhteisö. Esimerkiksi Nuolijärvi (2006, 326) on osoittanut, että murteellisuus vaihtelee samalla puhujalla sen mukaan, missä päin Suomea hän puhuu, eli keskustelukumppani tai yleisö vaikuttaa siihen (ks. myös Hurta 1999, 59). Niinpä helsinkiläisen voi arkisesti internetissä keskustellessaan otaksua käyttävän omaa helsinkiläistä keskustelutyyliään, kun vastaanottajatkin ovat pääkaupunkiseutulaisia. Vastaavasti hän saattaa käyttää yleiskieltä tai mahdollisimman yleisluontoista puhekieltä silloin, kun hän olettaa yleisökseen keitä tahansa suomalaisia.

Tutkimus tapahtuu vertaamalla verkosta kerättyä aineistoa aiempaan puhekielen tutkimustietoon. Vertailukohtana on pääkaupunkiseudun puhekieli, jota voidaan ajatella Suomessa hallitsevana puhekielen tyyppinä. Lähtökohtana on osoittaa, että kansainvälisessä kirjallisuudessa esitetty väite (Crystal 2002, 25–47) siitä, että verkkokeskustelun puhekielisyys jää yleiskielen ja puhekielen väliin, pätee myös

suomen kieleen.

Vertailukohtana käytetään kotimaisia Helsingin puhekielen tutkimuksia. Koska aineistot eivät ole suoraan vertailukelpoisia, valitaan aikaisemmista tutkimuksista ”minimiarvo”, joka kuvaa helsinkiläisten puhekielisyyden vähimmäistasoa, ja vastaavasti tässä tutkimuksessa verkkokeskusteluista valitaan sellaisia, joissa puhekielisyyys ilmeni maksimaalisesti. Jos tämä maksimi jää selvästi annetun minimin alle, voidaan päätellä, että puhekielisyyden asteissa on todennäköisesti merkittävä ero.

H1. Verkkokeskustelu sisältää keskustelun tyyppistä riippumatta puhekielisyyttä mutta toisaalta myös yleiskielen variantteja silloin, kun puhekielinen muoto olisi sekä mahdollinen että yleensä puheessa käytetty, jääden keskimäärin selvästi perinteisen kirjoitetun ja puhutun kielen väliin.

Yleisen puhekielisyyden asteen selvittämistä ei voida pitää riittävänä, sillä puhekieli koostuu suuresta joukosta erilaisia piirteitä, joiden yleisyys puhutussakin kielessä vaihtelee runsaasti (ks. esim. Paunonen 2006, 71–90; Nuolijärvi 1986). Tässä tutkimuksessa lähtökohtana on, että verkkokeskustelu pyrkii pitkälti jäljittelemään puhekieltä, eikä ole perusteita olettaa, että yksittäisten piirteiden yleisyydet vaihtelisivat näiden välillä, sillä aikaisempaa tutkimustietoa juuri tästä ei ole. Jos eroja kuitenkin löytyisi, se tukisi väitettä verkkokeskustelun ainutlaatuisesta kielenkäytön luonteesta.

H2. Puhekielen piirteistä samat ovat yleisiä verkkokeskustelussa kuin puheessakin.

Tiedetään, että jo nuoret hallitsevat yleensä yleiskielen normit, vaikka he eivät niitä puheessa käyttäisikään (ks. esim. Hurttu 1999, 89–90). Niinpä he tekevät tiedostettuja tai tiedostamattomia valintoja kirjoittaessaan suomen kieltä. Nämä valinnat poikkeavat yksilöiden välillä, ja ihmisillä on hyvin vaihtelevia tapoja käyttää kieltä jopa pienten alueiden sisällä. Esimerkiksi Helsingin puhetapojen tiedetään vaihdelleen paljon kaupunginosien välillä läpi 1900-luvun (Paunonen 2006, 71–73). Oletuksena on siis suuri yksilöiden välinen vaihtelu, joten ei voitane löytää yhtä

yksittäistä tapaa käyttää puhekieltä verkossa siten, että se yhdistäisi kaikkia keskustelijoita edes yhdenlaisen keskustelumuodon sisällä.

Puhekielen tutkimuksesta tiedetään kuitenkin, että on olemassa muutamia erilaisia kielenkäyttötapoja, joihin puhujat voidaan luokitella. Eräs tunnettu tapa on jakaa pääkaupunkiseudun kielenkäyttäjät neljään ryhmään: (1) yleiskieliset, (2) jonkin verran puhekielisiä variantteja käyttävät, (3) monipuolisesti pääkaupunkiseudun puhekielisyttä ilmentävät sekä (4) osittain alueellisia puhekielisyksiä käyttävät (Nuolijärvi & Tiittula 2000, 237–247, 277). Tässä tutkimuksessa ei pystytä yhtä monipuoliseen jakoon, erityisesti koska alueellisia puhekieliä tai murteita ei tarkastella. Puhujia voidaan kuitenkin pyrkiä ryhmittelemään akselilla yleiskielinen–puhekielinen kolmen ensimmäisen kohdan mukaisesti.

H3. Keskustelijat voidaan jakaa keskustelutyypistä riippumatta luokkiin heidän puhekielisyytensä asteen mukaan.

Verkossa on lukuisia erilaisia kanavia ja tapoja keskustella. Esimerkkejä näistä ovat reaaliaikaiset chatit, Facebookin seinäkirjoitukset, Twitterin viestisyötteen, keskustelupalstat ja blogit. Näiden sisälläkin skaala voi olla laaja, ja muun muassa verkkoportaleissa sekä uutisartikkelien ja blogien yhteydessä on merkittävästi toisistaan poikkeavia keskustelupalstoja. Keskustelutapojen kielenkäyttö poikkeaa toisistaan useiden eri tekijöiden mukaan (Crystal 2002, 6–10). Se, mitkä tekijät saavat aikaan variaatiota verkkokeskustelun kielessä, selviää vain empiirisen tutkimuksen kautta (Herring 2007).

Hakulinen (1989) jakaa suullisen ja kirjoitetun kielen skaalalle, jossa toisessa ääripäässä ovat hyvin vakiintuneet tekstit, kuten uskonnolliset kirjoitukset, ja toisessa nopeasti katoavat, spontaanit arkikeskustelut. Toisena ulottuvuutena hän pitää institutionaalisuutta, jonka vastakohtana ovat jälleen arkiset keskustelut. Jos verkkoviestinnässä halutaan tutkia puhekielisyttä, sitä varmasti ilmenee parhaiten epäinstitutionaalisissa, spontaaneissa, arkipäiväisissä verkkokeskusteluissa.

Erilaisista keskustelutyypeistä valittiin tähän tutkimukseen kolme: reaaliaikaiset (eli synkroniset) chatit, ei-reaaliaikaiset keskustelupalstat sekä jollakin tapaa näiden

väliin jäävät Facebookin seinäkirjoitukset. Perusteena on niiden edustavuus sekä erityisesti oletettu läheisyys puhekieleen, minkä voidaan olettaa syntyvän toisaalta teknisten tekijöiden kuten reaaliaikaisuuden kautta ja toisaalta keskusteluyhteisön luonteen ja teeman kautta (ks. Herring 2007). Sosiaaliset ja temaattiset tekijät on pyritty valitsemaan mahdollisimman samankaltaisiksi, jotta teknisten tekijöiden vaikutus tulisi esiin, erityisesti tärkeänä pidetyn synkronisuuden (ks. esim. Herring 2001a). Siksi kaikista keskustelutyypeistä on pyritty löytämään samanlaista arkipäiväisiä aiheita käsittelevää epämuodollista keskustelua.

H4. Verkkokeskustelun tyyppi vaikuttaa puhekielisyyteen siten, että lähimpänä reaaliaikaista olevat viestit ovat myös lähempänä puhekieltä: reaaliaikaiset keskustelut ovat lähempänä kuin Facebook-tilapäivitykset, jotka puolestaan ovat lähempänä kuin keskustelupalstojen viestit.

1.2 Tutkielman rakenne

Tutkielma rakentuu seuraavasti. Teoriaosan ensimmäinen ja suurempi kokonaisuus käsittelee suomalaista puhekieltä (luku 2). Liikkeelle lähdetään määritelmistä ja puhekielen elementeistä. Tässä tutkielmassa keskitytään pääkaupunkiseudulla puhuttavaan puhekieleen, jonka erityispiirteitä ja historiaa kuvataan seuraavaksi. Luvun lopussa on yleisempi katsaus olennaiseen kotimaiseen sociolinguvistiseen puhekielentutkimukseen. Luku 3 kuvailee teorian toista osuutta eli verkkoviestintää. Luku keskittyy verkkoviestinnän ominaisuuksien kuvaamiseen, sillä niillä on merkittävä vaikutus kielenkäyttöön verkossa (ks. esim. Crystal 2002). Yllä esitettyihin hypoteeseihin vastataan analyysin yhteydessä luvussa 3.5. Viimeisenä on yhteenveto (luku 5), jossa analyysissä saatuja tuloksia pohditaan peilaten niitä tutkimuksen tavoitteisiin. Lisäksi arvioidaan tutkimusmenetelmää sekä esitetään jatkotutkimusehdotuksia.

2 Puhekieli

Tämän tutkielman keskeisimpiä käsitteitä on puhekieli. Sananmukaisesti kyse on puhutusta kielestä, mutta asia ei ole niin yksinkertainen, kuten tämän tutkielman aihekin antaa ymmärtää. Tässä luvussa esitelläänkin ensin puhekieleen liittyviä oleellisia termejä. Toisena asiana kuvataan, millaisista asioista puhekieli muodostuu ja miten se eroaa yleiskielestä. Tässä tutkimuksessa tarkastelussa on pääkaupunkiseudun puhekieli, joka on Suomessa myös hallitseva puhekielen tyyppi. Siksi siihen paneudutaan seuraavaksi tarkemmin ensin sen syntyhistorian kautta ja kuvaamalla sitten tärkeimpiä sille ominaisia piirteitä. Luvun lopuksi luodaan katsaus Suomessa tehtyyn aikaisempaan puhekielen ja sosiolingvistiikan tutkimukseen.

2.1 Käsitteitä

Puhekielellä tarkoitetaan yleensä puhuttua kieltä. Jaottelun voi tehdä joko tuottotavan tai ilmiäsuun mukaan. Tuottotavan mukaan jaoteltuna vain puhuttu kieli on puhekieltä, ja kaikki muu on kirjoitettua tai kirjakieltä, jota kutsutaan myös yleiskieleksi. Ilmiäsuun mukaan sen sijaan puhekieltä voi myös kirjoittaa, ja vastaavasti kirjakieltä voi puhua. Puhekieli ja puhuttu kieli eivät siis tarkoita samaa. Muutenkin jaottelu on siinä mielessä ongelmallinen, että yleiskieli perustuu normeihin. Puhekieliä taas on erilaisia, minkä takia juuri tuottotapa on ollut luontevampi lähtökohta. Toisaalta puhekieli voidaan määrittellä standardista eli yleiskielestä poikkeamiseksi. (Hurta 1999, 56–58.)

Tämän tutkimuksen lähtökohta on, että puhekieltä voi myös kirjoittaa, joten ilmiäsuun mukainen jakotapa on ainoa mahdollinen. Termiä *puhekieli* siis käytetään ilmaisemaan sitä, että kielen ilmiäsuu muistuttaa puhuttua kieltä.

Hurta erottelee yleiskielen ja sitä vastaavan puhekielen ääripään eli murteen väliin myös kaksi muuta apusermiä, joiden kanssa saadaan jatkumo yleiskieli–yleispuhekieli–aluepuhekieli–murre. Tässä *yleispuhekieli* tarkoittaa puhekieltä, joka sisältää koko Suomessa puheessa yleisiä poikkeamia yleiskielestä, mutta perustuu

silti laajalti yleiskielen normeihin. Siihen eivät siis kuulu selvästi paikallistettavat murteelliset ilmaisut ollenkaan. *Aluepuhekieli* puolestaan on tasoittunut murre, josta tunnistaa alueen, jossa sitä puhutaan, mutta josta tarkan lähtöalueen paljastavat ilmaukset ovat hioutuneet pois. Esimerkiksi Tampereen puhetta voi kutsua aluepuhekieleksi. Tällaista jakoa Hurtta kutsuu rakenteellis-alueelliseksi kontrastina edellä esiteltyyn ilmiösuun perustuvaan jaotteluun. (Hurtta 1999, 59–60; ks. myös Lehikoinen 2005, 90–91.)

Monissa tutkimuksissa on käytetty hieman poikkeavia termejä likimain samasta asiasta. Sen mitä Hurtta edellä kutsuu yleispuhekieleksi tai aluepuhekieleksi, on Nuolijärvi väitöskirjassaan (1986) nimennyt helsinkiläispuhekieleksi. Termi ei toki ole täysin synonyyminen, sillä helsinkiläispuhekielellä viitataan helsinkiläislähtöiseen tapaan puhua, vaikka se onkin jossain määrin levinnyt kaikkialle Suomeen. Myöhemmin Nuolijärvi ja Tiittula (2000) kutsuvat samaa kieltä eteläsuomalaiseksi kaupunkipuhekieleksi ja Lappalainen (2004) pääkaupunkiseudun (johon hän lukee Helsingin lisäksi Espoon, Vantaan ja Kauniaisen) puhekieleksi, jotka viittaavat jo sanoinakin yleisempään kieleen kuin vain Helsingissä puhuttuun. Jokseenkin eriävä näkemys on Mantilalla (1997), jonka mukaan yleispuhekieltä ei ole olemassa. Tämä näyttäisi kuitenkin olevan määrittelykysymys, sillä Mantila ei ilmeisesti hyväksy termiä siksi, että koko Suomelle yhteisistä piirteistä huolimatta kaikkialla on ainakin jonkin verran alueellista vaihtelua. Hän esittää (ma.), että samaan aikaan, kun puhekielen alueelliset erot tasoittuvat, syntyy tilalle uusia variantteja, jotka poikkeavat yhtä paljon kuin aikaisemmat ennen (ks. kielen muutosmekanismeista tarkemmin Paunonen 2003). Hurttakin (1999, 67–68, 96) huomauttaa, että yleispuhekielen käsite on mahdollisesti vain tutkijoiden haave ideaalitulanteesta. Lisäksi tiedetään, että kielellistä variaatiota esiintyy myös Helsingin sisällä (Paunonen 2006; Lappalainen 2004, 59). Kaiken kaikkiaan luontevinta on tulkita pääkaupunkiseudun puhekieli aluepuhekieleksi, joka on kuitenkin levinnyt niin paljon, että se on puhutuista kielimuodoista lähimpänä yleispuhekieltä. Katso terminologiasta myös Mielikäinen 1980.

Toinen tärkeä huomio liittyy siihen, että tutkimuksissa kielimuodot oletetaan usein selvärajaisiksi: jokin piirre joko kuuluu kielimuotoon tai ei, ja tietty kielimuoto sisältää

juuri tietyt piirteet (Hurtta 1999, 96). Jotta asioita voidaan käsitellä, joitakin tämäntapaisia yleistyksiä joudutaan luonnollisesti tekemään. Käytännössä ihmisten puheessa piirteet sekoittuvat ainutlaatuisiksi kokonaisuuksiksi, joten on pidettävä mielessä, että tutkimustulokset ovat vain suuntaa antavia yleistyksiä.

Tutkimuksellisesta näkökulmasta puhekielet voidaan jakaa murteiden ja *kaupunkikielen tutkimukseen*, joista tämä tutkielma keskittyy jälkimmäiseen. Murteentutkimuksessa tavoitellaan kokonaiskuvaa murteesta, mutta kaupunkikielen tutkimuksessa pyritään selvittämään vain muutamia sosiaaliselta kannalta erottelevia kielenpiirteitä. Tämä perustuu siihen, että maalaisyhteisöissä tilanteissa ja keskustelukumppaneissa ei ole paljon vaihtelua, jolloin erilaisia kielenkäyttötapoja ei välttämättä tarvita. Kaupungeissa sen sijaan erilaisia kielellisiä vaatimuksia ja tilanteita on paljon enemmän, jolloin vaihtelua on enemmän yksilöidenkin puheessa. (Paunonen 2006, 14–17.)

Tässä tutkielmassa käytetään termiä *pääkaupunkiseudun puhekieli* tarkoittamaan koko Suomessa laajalle levinnyttä helsinkiläislähtöistä puhekielen muotoa, jonka voidaan katsoa olevan hallitseva puhekielen muoto. Koska puhekielissä on Suomessa paljon vaihtelua, on perusteltua valita sellainen, joka on mahdollisimman laajalle levinnyt, koska tutkimuskohteena on puhekielisyys yleisesti. Vaikka pääkaupunkiseudun puhekieltä haluaisi tulkita kapea-alaisesti aluepuhekielenä, sitä voidaan silti pitää yleisimpänä puhekielen muotona puhujien määrässä mitattuna. Niinpä se on melko varmasti internetissä käytetyistä puhekielen muodoista käytetyin.

2.2 Puhekielen elementit

Yleensä kirjakieltä on pidetty oikeampana ja arvostetumpana tapana käyttää kieltä. Kirjoitettu kieli (yleiskieli) on kuitenkin muodostettu alun perin puheesta laatimalla säännöt (Hakulinen 1989; Karlsson 1995). Puhuminen on siis paljon vanhempi kielenkäyttötapa. Tutkimuksen lähtökohdat ovat perinteisesti silti olleet kirjoitetussa kielessä, vaikka edellä mainitulla perusteella puhe ei ole ”köyhempi” tapa käyttää kieltä. Kieli muuttuu koko ajan, ja muutos näkyy nimenomaan puheessa.

Kirjoitukseen eli normeihin muutokset tulevat suurella viiveellä, vasta kun ne näkyvät jo kaikkien puheessa (Kuiri 1989; Paunonen 1995b).

Lähtökohtana puheen ja kirjoituksen yhteneväisyyksissä on se, että ne voivat olla samankaltaisia vain kielen muodon osalta (Tiittula 2001). Sen sijaan esimerkiksi puheen ääni ja kirjoituksen visuaalinen ulkoasu eivät ole vertailukelpoisia. Seuraavassa on eritelty kielen muotoja, joissa kirjakieli ja puhekieli ovat vertailtavissa ja eroavat toisistaan.

Vaikka puhe noudattaisi täysin yleiskielen normeja, siinä on seuraavat neljä eroavaisuutta kirjoitukseen nähden (Lehikoinen 2005, 92–94). Äng-äänteelle ei ole omaa kirjoitusasua (*kenkä – kenkä*). N assimiloituu usein toisen konsonantin edellä (*tulenpas – tulempas*). Sanojen lopussa voi esiintyä ääntämyksessä kahdennusta, jota ei kirjoiteta (*tulepa – tuleppa*). Kahden vokaalin väliseen tavurajaan saattaa puheessa tulla ääntämistä helpottava konsonantti (*kauan – kauvan*). Nämä erot tulevat esiin lähinnä puheen litteroinnissa. Joskus niitä näkee myös kirjoituksessa, mutta tällöin ne luultavammin kielivät väärin opituista oikeinkirjoitussäännöistä kuin puhekielisyydestä.

Lappalainen (2004) pitää olennaisina puhekielestä analysoitavina tekijöinä ensinnäkin fonologisia, morfologisia ja morfosyntaktisia piirteitä. Lähinnä tämänkaltaisiin piirteisiin (ks. luku 2.3.2) perustuu tämänkin tutkielman empiirinen analyysi, ja niitä tutkitaan tyypillisessä variaationanalyysissä.

Toisena kokonaisuutena on sanasto, jota pidetään kuitenkin tutkimuksen kannalta vaikeampana kohteena (Lappalainen 2004, 15–16). Helpoimmin sanastollisesti erottuu slangi. Sanaluokista jossain määrin substantiivit, adjektiivit, sti-päätteiset adverbit ja konjunktiot, mutta erityisesti pronominit ovat puheessa yleisempiä kuin kirjoituksessa. Diskurssipartikkeleilla, kuten *no*, *joo* ja *niin*, voidaan muun muassa osoittaa puheenvuoron alkua tai antaa palautetta keskustelukumppanille. Tämä osoittaa, että keskustelussa on käyttöä funktioille, joita kirjoituksessa ei tarvita, mikä osaltaan selittää eroja. Enemmän tarvetta on myös epäröintiä ilmaiseville, lieventäville ja vahvistaville ilmauksille, kuten *ehkä*, *melko*, *kamalan* ja *hirveän*. Muita sanoihin liittyviä eroja ovat vielä lyhempien sanojen ja suppeamman sanaston

suosiminen puheessa. (Lappalainen 2004, 92; Lehikoinen 2005, 95–104; Tiittula 1992, 54–68.)

Muut eroavaisuudet liittyvät etupäässä lauserakenteeseen. Puheessa käytetään lyhempiä lauseita (tai *vuoroja*) kuin tekstin virkkeet. Lauseiden rajatkaan eivät erotu puheessa niin selvästi, vaan niitä voidaan esimerkiksi yhdistää toisiinsa pitkiksi ketjuiksi konjunktoiden ja partikkelien avulla. Puhuttujen lauseiden lyhydestä huolimatta kirjoitettu ilmaisu on tiiviimpää, mikä tekee siitä monimutkaisempien lauserakenteiden ohella raskaampaa. Puheessa on toisaalta paljon toistoa mutta toisaalta myös ellipsejä. Lopuksi puheen luonteesta johtuen siinä esiintyy keskeytyksiä, korjauksia, uudelleenaloituksia ja taukoja. (Lehikoinen 2005, 95–104; Tiittula 1992, 82–98.)

Eri elementit eivät välttämättä anna yhteneväistä kuvaa puhujan puhekielisyydestä. Nuolijärvi ja Tiittula (2000, 230–231) esittävät esimerkkinä omasta tutkimuksestaan Esko Ahon puhumisen, jonka morfologia tekee puhekielisemmäksi mutta samaan aikaan syntaksi yleiskielisemmäksi.

Tiittula (2001) pohtii vielä puheen esittämisestä kirjoitettuna seuraavaa. Puhekielistä vaikutelmaa ei voi luoda kirjoittamalla äännteitä täysin tarkasti. Litteraatiot, joihin niihinkin on vaikea tallentaa puheen kaikkia ominaisuuksia (kuten äänensävyjä), ovat selvästi vaikealukuisempia kuin puhekieliseltä vaikuttava ”tavallisempi” teksti. Puhekielisyyttä luodaan mielellään varsinkin kirjallisuudessa käyttämällä leimautumattomia (siis yleispuhekielisiä) piirteitä. Eri kielissä puheenomaisuuden tuottaminen toimii eri lailla, mikä tekee muun muassa kääntämisestä hankalaa. On selvää, että samasta syystä puhekielisyyttä on tutkittava eri kielissä erikseen.

2.3 Pääkaupunkiseudun puhekieli

Helsingissä ja sen ympäristössä käytetty ja sieltä muuallekin Suomeen levinnyt puhekieli on erityinen suomalaisessa kielihistoriassa. Useimmilla alueilla on ollut paljon pidempään omat murteensa, mutta suomen kieli on Helsingissä yleistynyt vasta huomattavan myöhäisessä vaiheessa. Niinpä se on ensin ottanut vaikutteita eri

murteista kaupunkiin muuttaneiden mukana. Myöhemmin asetelma on kääntynyt päinvastaiseksi siinä mielessä, että Helsingin puhekieli on alkanut vaikuttaa muualla Suomessa puhuttuun kieleen. Nuolijärven ja Tiittulan (2000, 222–223) mukaan tämä vaikutus toimii siten, että erityisesti pääkaupunkiseudun puhekielen leimautumattomat variantit, jotka siis eivät vahvasti luo vaikutelmaa puhujan helsinkiläisyydestä tai murteellisuudesta, ovat yleistyneet laajasti koko Suomessa (vrt. luvun 2.1 *yleispuhekieli*).

Teoreettisesti pääkaupunkiseudun puhekielen yleistymistä voisi luonnehtia sosiolingvistiikassa tunnetulla aaltoteorialla (teoriasta ks. esim. Hudson 1996, 39–41). Siinä kielenpiirteiden leviäminen rinnastetaan veden pinnalla etenevään aaltoon, joka syntyy heittäessä kiven veteen. Joutuessaan kauemmas lähtöpisteestä aalto heikkenee, ja variantin leviäminen loppuu jossain vaiheessa, mikä riippuu sen vaikutuksen vahvuudesta. Aaltoteoria on historiallisesti ymmärretty maantieteellisenä ilmiönä, mutta nykypäivänä olisi luontevampaa ajatella sosiaalisia verkostoja. Erityisesti internet yhdistää ihmisiä sijainnista riippumatta ja paljon vahvemmin sosiaalisten tekijöiden mukaan, jolloin aallot etenevät verkostoja eivätkä maan pintaa pitkin. Tämän voi ajatella edistäneen eri puhekielten leviämistä koko Suomen alueelle.

Levinneisyytensä ansiosta pääkaupunkiseudun puhekieltä voidaan siis pitää nykyään hallitsevimpana puhekielenä, ja siksi sitä voi lähtökohtaisesti pitää myös internetissä yleisenä keskusteluissa. Niinpä se on tärkeässä asemassa myös tämän tutkielman kannalta, joten on syytä esitellä sen historiaa syntyvaiheista lähtien. Lisäksi kuvaillaan itse kieltä siinä esiintyvien piirteiden kautta, keskittyen erityisesti yleisiin ja laajalle levinneisiin variantteihin.

2.3.1 Kehitysvaiheet

Helsinki oli ollut perustamisestaan, vuodesta 1550, lähtien alun perin pitkälti ruotsinkielistä aluetta. 1700-luvulle tultaessa suomenkielisiä tosin oli Helsingissä huomattavasti, mutta seuraavan sadan vuoden aikana ruotsi muuttui selvästi vallitsevaksi, mikä johtui muun muassa ruotsin arvostuksesta. Niinpä noilla

aikaisilla suomenkielisillä helsinkiläisillä ei ollut juurikaan vaikutusta siihen, millaiseksi suomen kieli kaupungissa myöhemmin muotoutui. Suomi alkoi lopullisesti yleistyä vasta 1850-luvulta alkaen. (Paunonen 1995a, 2–3; 2006, 19–23.)

Ensimmäiset suomen puhujat Helsingissä olivat toisaalta muualta muuttaneita murteiden puhujia ja toisaalta sivistyneistöä, joka puhui kirjakieltä, joka ei suoraan pohjautunut mihinkään murteeseen. Sivistyneistö tosin suosi ruotsin käyttöä hyvin pitkään. Helsingin sivistyneistö alkoi myös ajamaan kirjakieltä ainoaksi oikeaksi tavaksi puhua suomea. Kirjakieli valikoitui käyttöön siksi, että Helsingissä ei ollut vanhastaan selkeää puhekielen tapaa, ja toisaalta suomalaisuusaate halusi tehdä selvän eron vieraisiin kieliin. Murteellisuuksien lisäksi puheesta haluttiin karsia vieraiden kielten vaikutus, erityisesti ruotsin. Tämä levisi suomen kielen opetuksessa kaikkialle Suomeen. Nämä fennomaanit ajoivat intohimoisesti suomen asemaa ja saivat vastaanansa yhtä kiihkomielisiä ruotsin kielen kannattajia, mikä johti suuriin kieliriitoihin. Samalla kuitenkin alkoi kehittyä työläisten keskuudessa stadin slangia, joka otti sekä sanastoa että ääntämistä erityisesti ruotsista mutta myös venäjältä. Myöhemmin slangista ovat harventuneet vierasperäiset ääntämykset, esimerkiksi *gimma* → *kimma* 'nainen'. Näin eri yhteiskuntaluokilla oli varsin erilainen suhtautuminen ja tapa käyttää kieltä. Työväki näet pyrki myös enemmän ymmärtämään toisiaan ja oli valmis käyttämään taitonsa mukaan molempia kieliä, eikä itsepäisesti pitänyt toista kieltä parempana. (Paunonen 1995a, 12–13, 16–23; 2006, 33–36, 44; ks. myös Paunonen 1989.)

Työväestön eli murteita puhuvien helsinkiläisten kieli on aluksi, 1800-luvun loppupuolella, sisältänyt lähinnä läntisten murteiden piirteitä, erityisesti länsiusmaalaisia ja etelähämäläisiä. Näistä osa on sittemmin kadonnut helsinkiläisten puheesta. Itämurteilla ei siis alussa ollut sijaa, vaan niitä alkoi kuulua vasta 1920-luvulta lähtien. Vasta sotien jälkeen 40-luvulla idästä muutti paljon väkeä Helsinkiin, erityisesti siirtokarjalaisia. (Paunonen 1995a, 13–16; 2006, 41.)

Helsingin puhekieli on kaiken kaikkiaan saanut Paunosen (2006, 36–48) mukaan vaikutteita viidestä eri murreryhmästä. Ensimmäinen ovat etelähämäläiset ja länsiusmaalaiset piirteet, joiden yleisyyttä selittää ennen kaikkea maantieteellinen

läheisyys. Tähän kuuluvat muun muassa Nurmijärvi ja Lohja. Esimerkkejä näistä piirteistä ovat joidenkin vokaaliyhtymien muuttuminen pitkiksi vokaaleiksi (*istuu* pro *istua*), A:n loppuheitto partitiivissa ja paikallissijoissa (*isoi* pro *isoja*, *siel* pro *siellä*) ja kysyvänä pronomininä *ketä* pro *kuka*. Toinen ryhmä ovat yleishämäläiset piirteet, joihin lukeutuvat ts:n korvaaminen tt:llä (*itte* pro *itse*) sekä pronominiin *tuo* ja *nuo* vastineet *toi* ja *noi*. Kolmanneksi yleislänsimurteiset piirteet ovat vaikuttaneet Helsingin puhekieleen, esimerkiksi *hiano* pro *hieno*. Itämurteista puolestaan on tullut sellaisia piirteitä kuin d:n kato (*yhen* pro *yhden*) ja ts:n korvaaminen t:llä (*kato* pro *katso*). Näistä jälkimmäinen on esiintynyt myös joissakin länsimurteissa. Tästä päästään viidenteen ryhmään, joka sisältää yleissuomalaisia piirteitä, joita on siis jo vanhastaan esiintynyt sekä läntisissä että itäisissä murteissa. Näihin kuuluu tiettyjen vokaaliyhtymien korvautuminen pitkällä vokaalilla (eri vokaaliyhtymät kuin edellä, esimerkiksi *korkee* pro *korkea*), i: loppuheitto (*yks* pro *yksi*), persoonapronominiin pikamuodot *mä* pro *minä* ja *sä* pro *sinä*, vastaavasti *tää* pro *tämä* ja *nää* pro *nämä*, mon. 1. persoonan inkongruenssi (*me mennään*) sekä vastaavasti mon. 3. persoonassa (*lapset leikki*). Kuri (1989) mainitsee 3. persoonan *se* pro *hän* esiintyvän myös useimmissa murteissa. Lisäksi hän väittää edellä mainitun mon. 1. persoonan passiivin olevan peräisin Savosta, joten joidenkin piirteiden kohdalla niiden yleisyydestä ja alkuperästä saattaa olla eriäviäkin näkemyksiä.

On hyvä huomata, että on piirteitä, jotka esiintyvät suunnilleen kaikkialla Suomessa vanhastaan, joten ei ole oikein väittää, että pääkaupunkiseudun puhekieli olisi levittänyt ne kaikki muualle. Ne vain sattuvat kuulumaan myös siihen ja ovat siis oikeammin sanottuna yleispuhekieltä eli eroavat yleiskielestä mutta eivät alueellisesti. Tällaisia ovat muun muassa monikon 1. ja 3. persoonan subjektin ja predikaatin inkongruenssi (esim. *me menemme* – *me mennään*) sekä pronominiin *se* ja *ne* käyttö ihmisistä kolmannessa persoonassa (Mantila 1997). Toisaalta on paljon piirteitä, jotka ovat ensin tulleet helsinkiläisten puheeseen ja sieltä sitten levinneet muualle Suomeen. Tässä välissä piirteet ovat saattaneet myös muuttua, eli niillä on murteista poikkeava helsinkiläinen leima (ks. Paunonen 1995a, 14–15).

Murreryhmien lisäksi on kuudeskin ryhmä, ja se sisältää puheessa esiintyneet kirjakieliset piirteet (Paunonen 2006, 47–48). Nämä ovat luonnollisesti

murrepiirteiden ”vastakohtia” eli yleiskielisinä säilyneitä variantteja. Esimerkkejä ovat A:n ja i:n säilyminen sanan lopussa, persoonapronomini minä ja mon. 1. ja 3. persoonan kongruentit muodot. Paunonen toteaa näiden enimmäkseen vastaavan yleissuomalaisen piirteiden joukkoa. Siksi ne ovat saaneet vertailussa kirjakielisen leiman, mikä on luultavasti vaikuttanut niiden harvinaistumiseen puheessa.

Myös ruotsilla on ollut vaikutuksensa Helsingin puhekieleen, mutta vaikutus rajoittuu lähinnä slangiin. Paunosen (2006, 48) mukaan ruotsi ei ole vaikuttanut nykyiseen helsinkiläispuheeseen näkyvästi eli esimerkiksi puhevarianttien kautta kuten murteet, vaan ”kyse on eräänlaisesta piilovaikutuksesta”.

Voidaan siis todeta, että Helsingissä ovat sulautuneet yhteen yleiskieli, eri murteet (joista vahvimmin länsi- ja hämäläiset murteet) sekä ruotsin kielen vaikutus muun muassa stadin slangin kautta. Vaikka Helsingin sisäistä vaihtelua puhujissa korostetaan tutkimuksissa, voitaneen kieltä pitää sulautumisten kautta nykyään yhtenäisempänä kuin vaikkapa sata vuotta sitten. Tällä perusteella voidaan siis puhua pääkaupunkiseudun puhekielestä, jolla tarkoitetaan joukkoa kielellisiä piirteitä, jotka ilmenevät erityisesti helsinkiläisten ja ympäryskuntalaisten puheessa.

1900-luvulla yleiskielestä ja murteista lähteneen kielen suunta on ollut yhtenäisempään puhekieleen päin (esim. Paunonen 2006, 91–92). Suunta on pysynyt samana tai ei ainakaan ole kääntynyt takaisinpäin. Tämä on peruste sille, miksi puhekielisyyttä on hyödyllistä tutkia nimenomaan helsinkiläisten puheesta ja miksi satunnaisestikin heistä valitun otoksen voi olettaa antavan yhteneviä tuloksia. Samalla on muistettava, että vaikka puhekielisyyys yleistyy yleiskieleen verrattuna, puhekieli itsessään muuttuu jatkuvasti eikä siis ole kiinteä maali, jota kohti oltaisi matkalla.

Koska pääkaupunkiseudun puhekieli sisältää piirteitä hyvin monesta Suomen murteesta, siitä voinee missä tahansa päin maata asuva löytää oman mahdollisen murteensa kanssa jotain yhteistä, joka kuitenkin poikkeaa kirjakielestä. Tämä saattaisi edesauttaa myös muiden helsinkiläisten kielenpiirteiden leviämistä näiden tuttujen siivellä. Voi edelleen pohtia, että tällainen vaikutus näkyisi erityisen hyvin internetissä, jossa kielivariantit pääsevät tänä päivänä kaikkein tehokkaimmin

kosketuksiin toistensa kanssa ja sekoittumaan. Tämä saattaisi selittää pääkaupunkiseudun puhekielen oletettua yleisyyttä verkossa. Toisaalta syy voi olla niinkin yksinkertainen, että kyseisen variantin puhujia on eniten.

2.3.2 Kielellisiä piirteitä

Kuten aikaisemmin todettiin, kaupunkikielen tutkimuksessa pyritään selvittämään lähinnä tärkeimpiä sosiaalisesti erottelevia kielenpiirteitä. Perinteisesti näitä piirteitä on selvitetty ryhmistä, joita yhdistää muun muassa ikä, sukupuoli, koulutus ja asuinseutu. Tässä tutkimuksessa näistä lähinnä vain sukupuoli on tiedossa, mutta ajatuksena on, että internet mediana saattaisi olla uusi yhdistävä taustatekijä.

Pääkaupunkiseudun puhekieltä käsitteleviä tunnettuja tutkimuksia ovat tehneet muun muassa Nuolijärvi, Paunonen ja Lappalainen. Tässä tutkimuksessa käytettävät piirteet perustuvat mainittujen tutkijoiden kolmeen tutkimukseen (Nuolijärvi 1986, 99–205; Paunonen 2006, 67–69, ks. myös 1995a, 29–232; Lappalainen 2004, 60–63). Piirteet ovat luonteeltaan fonologisia, morfologisia ja morfosyntaktisia, ja ne esitellään seuraavaksi sellaisessa järjestyksessä, että yleisimmin puhekielistä varianttia ilmentävät piirteet tulevat ensin. Samaa listaa ja järjestystä käytetään myöhemmin analyysissa (luku 3.5). Nuolijärven, Paunosen ja Lappalaisen käyttämät piirreluettelot ovat pitkälti yhteneviä, mutta Paunonen on jakanut piirteet pienempiin osiin, yhteensä 25 kategoriaan. Lappalaisella niitä on 15 ja Nuolijärvellä vain yhdeksän. Tässä esitellään piirteet Paunosen jaon mukaan, jotta tuloksia voitaisiin verrata hänen tutkimukseensa. Myös siinä on joitakin eroja, mitkä kaikki tapaukset lasketaan kuhunkin piirteeseen. Vaikka tässä tutkielmassa pääasiassa sovelletaan Paunosen jaottelua, myös joitakin eroja muihin jaotteluihin esitellään lyhyesti.

1) i:n loppuheitto supistumaverbien imperfektissä

Sanan lopussa oleva i voi puheessa kadota monissa erilaisissa tapauksissa. Paunonen tarkasteli kuutta tällaista tapausta, ja ensimmäisessä i katoaa supistumaverbeissä yks. 3. persoonan imperfektimuodossa (*hyppäsi – hyppäs*) (Paunonen 2006, 68). Paunosen aineistossa tämä piirre esiintyi ainoana 100 %

puhekielisenä. Myös muut tapaukset ovat piirteiden puhekielisimmästä päästä.

2) i:n loppuheitto konditionaalissa

Toinen i:n katotapaus koskee konditionaalia, jossa voi kadota joko vain lopussa oleva i tai molemmat s:n ympäriltä (*antaisi – antais, antas*) (Paunonen 2006, 68).

3) NUT-partisiipit

NUT-partisiipilla tarkoitetaan aktiivin 2. partisiipin muotoa verbistä. Puhekielisessä muodossa sanan lopusta katoaa t (*oli tullut – oli tullu, on tehnyt – on tehny*) (Paunonen 2006, 69).

4) i:n loppuheitto diftongiin päättyvässä imperfektissä

Kolmantena i:n katotapauksena ovat yks. 3. persoonan imperfektimuodot, jotka päättyvät diftongiin (*antoi – anto*) (Paunonen 2006, 68).

5) *minä* ja *sinä* puhemuodot

Paunonen (2006, 69) on valinnut tarkasteltavaksi persoonapronominiksi yks. 1. persoonan taivutusmuotoineen (*minä – mä, minulla – mulla*). Lappalainenkin (2006, 295) toteaa, että mä- ja mu-muodot ilmenevät helsinkiläisten puheessa lähes aina. Yks. 2. persoonan esiintymät Paunonen on ilmeisesti yhdistänyt 1. persoonaan (ks. Paunonen 1995a, 152–153), koska niitä esiintyy niin vähän hänen aineistossaan. 2. persoona ei siis ole syytä jättää pois tarkastelusta.

6) i:n loppuheitto translatiivissa

Translatiivimuotoisista eli ksi-päätteisistä nomineista i voi jäädä puheessa pois (*aluksi – aluks*) (Paunonen 2006, 68).

7) inen-adjektiivit

Inen-loppuisista adjektiiveista i katoaa luonnollisesti päätteen alusta eikä sanan lopusta (*punainen – punanen*) (Paunonen 2006, 68). Kyse on kuitenkin vastaavalla tavalla i:n kadosta. Tällaisista tapauksista pronomini *jokainen* on jätetty pois, koska

Paunosen mukaan se esiintyy vain i:n kanssa, sekä pronominaalit, kuten *tällainen* tai *sellainen*, jotka puolestaan esiintyvät lähes ainoastaan ilman i:tä.

8) i:n loppuheitto nominatiivissa

Viimeinen i:n loppuheittotapauksista on usein esiintyvän sanan perusmuoto, kuten yksi (*yks*), uusi (*uus*), ensi (*ens*), vuosi (*vuos*) tai täysi (*täys*) (Paunonen 2006, 68). Nuolijärvi (1986, 163) huomauttaa, että juuri s:n perässä i katoaa usein helsinkiläisten puheessa (kuten piirteissä 1, 2 ja 6 edellä).

Kuuden mainitun i:n katotapauksen lisäksi Lappalainen (2004, 61) mainitsee vielä yks. 2. persoonan omistusliitteen (*autosi – autos*), oin-adverbit (*milloin – millon*), isin-adverbit (*takaisin – takasin*) sekä seuraavat i-loppuiset diftongin jälkikomponentin kadot: -isA (*valoisa – valosa*), -oittaa (*sekoittaa – sekottaa*) ja -Aise (*läpäise – läpäse*).

Tämän ja muiden piirteiden kohdalla on tärkeää huomata, että kaikkia (esimerkiksi) i-loppuisia sanoja ei voida analysoida, koska niistä ei puheessa esiinny erilaista muotoa. Esimerkiksi sanasta *oli* ei esiinny Helsingissä lyhyttä muotoa *ol*. Tulee siis tietää, minkä sanojen kohdalla variantit ovat mahdollisia.

9) *tämä* ja *nämä* puhemuodot

Demonstratiivipronomineilla *tämä* ja *nämä* on puhemuodot *tää* ja *nää* (Paunonen 2006, 69) sekä sanoilla *tuo* ja *nuo* muodot *toi* ja *noi* (Nuolijärvi 1986, 188, 193). Sanojen *tämä* ja *nämä* taivutusmuodot eivät puhekielessä eroa yleiskielestä, joten niistä vain perusmuotojen analysointi tulee kyseeseen.

10) mon. 3. persoonan kongruenssi

Ensimmäinen kahdesta tilanteesta, joissa subjekti ja predikaatti eivät ole samassa sijamuodossa, esiintyy mon. 3. persoonassa (Paunonen 2006, 69). Tällöin predikaatti vaihtuu yksikkömuotoon (*lapset leikkivät – lapset leikkii*). Perusteena 3. persoonan inkongruenssille Nuolijärvi (1986, 199–200) arvelee, että subjekti ilmaisee jo monikollisuuden, joten predikaatissa sitä ei enää tarvita. 3. persoonan

kongruenssista katso lisää myös Paunonen 1995b.

11) *he* puhemuoto

3. persoonan monikossa henkilöistä on usein tapana käyttää sanaa *ne* sanan *he* sijaan (Paunonen 2006, 69).

12) *olla*-verbin puhemuodot

Olla-verbistä käytetään puheessa lyhyempiä (pikapuhe)muotoja (*olen – oon, en ole – en oo* tai *en o*). Paunonen (2006, 69) ei jostain syystä kuitenkaan huomioi konditionaalimuotoja (*olisin – oisin*). Lappalainen (2004, 62) tutki myös verbejä *mennä, panna* ja *tulla*, joista tavataan vastaavasti pikapuhemuotoja.

13) mon. 1. persoonan kongruenssi

Inkongruenssia voi esiintyä toisena tapauksena mon. 1. persoonassa (*me menemme – me mennään*) (Paunonen 2006, 69). Tässä verbi siis vaihtuu passiivimuotoon. Paunonen lisää myös, että ellipsitapaukset (eli subjekti jää pois) voidaan lukea mukaan, jos käyttöyhteydestä ilmenee selvästi, että on tarkoitettu kyseisiä persoonamuotoja (*menemme – mennään*).

14) yks. 1. persoonan omistusmuodot

Paunonen (2006, 67, 69) on huomionut tutkimuksessaan omistusmuodoista vain yks. 1. persoonan, jossa esimerkiksi *minun kirjani* tai *kirjani* muuttuu muotoon *mun kirjani, mun kirja* tai harvoin *minun kirja*. Näistä omistusmuodoista katso lisää myös Paunonen 1995b.

15) eA ja oa-nominit ja niiden johdokset

Seuraavat neljä tapausta liittyvät vokaaliyhtymien assimiloitumiseen (eli monoftongiutumiseen). Niissä kyse on kahden peräkkäisen eri vokaalin muuttumisesta pitkäksi ensimmäiseksi vokaaliksi. Ensimmäisenä tapauksena ovat eA- (eli ea- tai eä-) ja oa-loppuiset nominit ja niiden johdokset, kuten *oikeestaan – oikeestaan* (Paunonen 2006, 68). Lappalainen (2004, 61) lukee tähän ryhmään

nominien lisäksi myös vastaavanlaiset verbit, esimerkiksi *lukea – lukee*. Yhtä lailla hänen mukaansa muihinkin jäljempänä esiteltäviin assimilaatiotapauksiin voidaan liittää soveltuvat verbit.

16) *hän* puhemuoto

Vastaavasti kuin monikossa (piirre 11) 3. persoonan yksikössä *hän* muuttuu muotoon *se* (Paunonen 2006, 69).

17) eA ja OA-partitiivit

Samoin kuin perusmuotoisten sanojen kohdalla edellä, myös eA ja OA-loppuiset partitiivimuodot voivat assimiloitua, esimerkiksi *ovea – ovee* (Paunonen 2006, 68).

18) yleiskielen ts:n vastineet

Sekä Paunonen (2006, 68), Nuolijärvi (1986, 115, 119) että Lappalainen (2004, 62) käsittävät ts:n vastineet yksittäiseksi piirteeksi. Siinä ts voi korvautua t:llä tai tt:llä (*itse – ite* tai *itte*), mutta lähinnä vain usein esiintyvissä sanoissa, joiksi mainitaan *itse, seitsemän, katsoa, katsella, viitsiä, etsiä, metsä* ja *Ruotsi*. Lappalaisen mukaan muista poiketen myös *tarvita* voidaan lukea listalle.

19) yleiskielen hd:n vastineet

Yleiskielen d:n vastineet jakautuvat kahdeksi piirteeksi, joista ensimmäisestä voidaan puhua tarkemmin hd:n vastineina (Paunonen 2006, 68). Näitä ilmenee lähinnä d:n katona usein esiintyvissä sanoissa, kuten *yhden – yhen* tai *lähdetään – lähetään*.

20) MA-infinitiivin illatiivi

MA-infinitiivillä tarkoitetaan 3. infinitiiviä. Tämän illatiivimuoto (eli MAAN-muoto) ilmenee puheessa usein lyhennettynä, esimerkiksi *tekemään – tekee* tai *tekeen* (Paunonen 2006, 69).

21) A:n loppuheitto adessiivissa

A:n loppuheitolla tarkoitetaan sanan lopusta pois jäävää a:ta tai ä:tä. A:n loppuheitto-

eli katotapauksista Paunonen (2006, 68) käsittelee vain nomineja ja niistäkin vain kahta sijamuotoa. Ensimmäinen näistä on adessiivi (*paikalla – paikal*).

22) A:n loppuheitto inessiivissä

Toinen sijamuodoista on inessiivi (*talossa – talos*). Nuolijärvi (1986, 154–155) ja Lappalainen (2004, 61) ottavat mukaan laajemman valikoiman tapauksia lisäten edellisiin essiivin (*päivänä – päivän*), elatiivin (*joukosta – joukost*), ablatiivin (*kaupungilta – kaupungilt*), tA-muotoisen partitiivin (*kahta – kaht*), paikallissijaiset adverbit (*edessä – edes*) sekä partikkelit *että, jotta, mutta* ja *kanssa*.

23) yleiskielen d:n vastineet

Toisena d:n vastinetapauksena ovat me, te ja he -pronominien genetiivit ja akkusatiivit sekä tietää-verbi, eli niin sanotut VVdV-tapaukset, joissa d:tä edeltää kaksi vokaalia ja sen perässä on yksi: *meidän – meiän* tai *meijän, en tiedä – en tiä* tai *en tiiä* (Paunonen 2006, 67). Selvästi vierasperäisissä sanoissa (esimerkiksi *video*) vastineet eivät ole mahdollisia. Myös yhden vokaalin perässä d voi kadota, mutta Helsingissä vain sanojen *edes, edessä* ja *odottaa* kohdalla (Nuolijärvi 1986, 107; Lappalainen 2004, 61–62). Kadon lisäksi d voi joskus korvautua r:llä, kuten *meirän*, mutta se on Helsingissä hyvin harvinaista. Nuolijärvi (1986, 99–100) lisää vielä, että varsinkin tämän piirteen kohdalla puhekielessä d:n säilyminen on myös hyvin yleistä. On siis muistettava, että puhekielisyyys ei aina tarkoita poikkeamista yleiskielestä, vaikka se olisikin mahdollista.

24) UA-partitiivit

Viimeiset kaksi piirrettä liittyvät partitiivimuotoisten sanojen lopussa olevien vokaaliyhtymien assimiloitumiseen. Ensimmäiset ovat UA-päätteisiä, esimerkiksi *lyhtyä – lyhtyy* (Paunonen 2006, 68).

25) iA-partitiivit

Viimeiseen assimilaatioryhmään kuuluvat iA-loppuiset partitiivimuodot, kuten *kahvia – kahvii* (Paunonen 2006, 68). Näiden lisäksi Lappalainen (2004, 61) lukee mukaan sanat *takia* ja *äkkiä*. Neljä edellä esiteltyä assimilaatiopiirrettä (15, 17, 24 ja 25)

voidaan kaiken kaikkiaan tiivistää seuraaviin vokaalimuunnoksiin: iA → ii, UA → UU, eA → ee ja OA → OO (Nuolijärvi 1986, 128, 137).

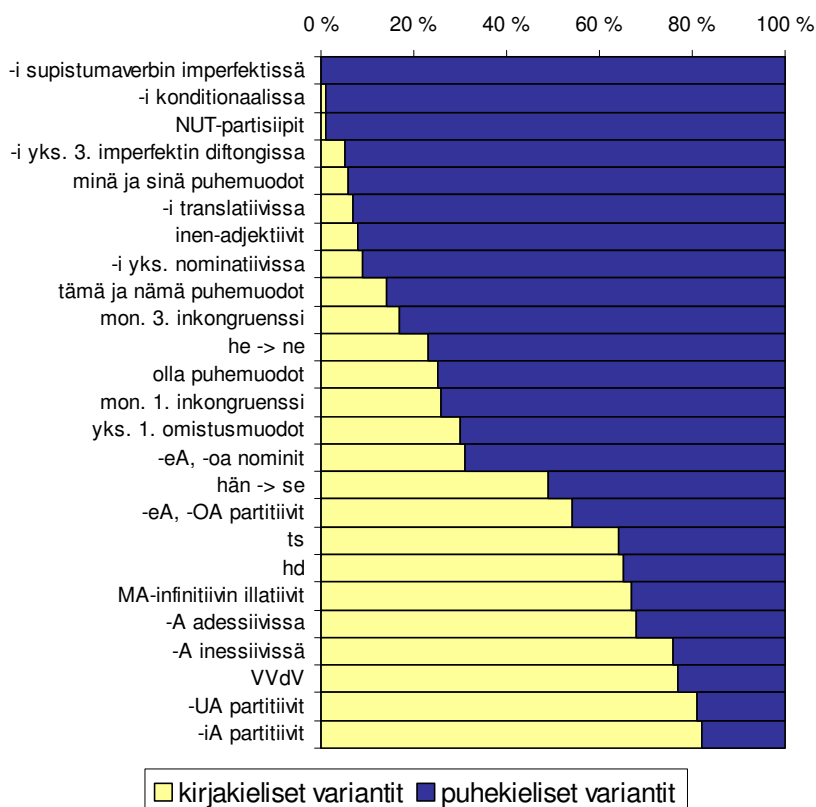
Jostakin syystä Paunonen (2006) on jättänyt diftongien avartumisen pois, vaikka se hänellä aikaisemmassa tutkimuksessa (1995a) oli mukana, samoin kuin Nuolijärvellä (1986). Piirre poikkeaa muista jonkin verran siinä mielessä, että avartumisella on paljon mahdollisia asteita, eli kysymyksessä ei ole joko täysin suppea tai täysin avartunut muoto (syödä – syädä). Kirjoituksessa kuitenkin vain ääripäät ovat mahdollisia. Voidaan kenties olettaa, ettei henkilö käytä kirjoittaessaan avartunutta muotoa, ellei hänen puheensakin ole juuri siitä ääripäästä, ja siksi kyseinen piirre ei tässä tutkimuksessa olisi muutenkaan riittävän informatiivinen. Toinen listalta puuttuva piirre on pronominisubjektin ilmipano 1. ja 2. persoonassa (*istun – mä istun*) (ks. Lappalainen 2004, 63).

Kaikki edellä esitellyt piirteet eivät ole samanluonteisia keskenään, vaan ne ovat erilaajuisesti yleisiä Suomessa. Erityisesti osa esiintyy kaikkialla maassa ja osa lähinnä pääkaupunkiseudulla tai ovat vasta leviämässä sieltä muualle. Mantila (1997) jakaa kaikki suomen kielessä esiintyvät puhekieliset piirteet seuraaviin viiteen luokkaan. Ensimmäisessä ovat yleiset ja neutraalit piirteet, jotka esiintyvät niin laajalti ja ovat niin vanhoja puhekielessä, että ne olisivat voineet tulla valituiksi kirjakielen normeiksi. Näitä ovat mon. 1. ja 3. persoonan kongruenssit, *se* ja *ne* käyttäminen henkilöistä 3. persoonassa sekä i:n kato -inen-päätteisissä adjektiiveissa, -si-päätteisissä sanoissa sekä i-loppuiseen diftongiin päättyvissä sanoissa. Kaikki nämä lukeutuvat myös 25 edellä esiteltyyn piirteeseen. Toisena ovat myös hyvin laaja-alaiset puhekielisyydet, joissa ei ole itä- ja länsimurteiden välistä eroa. Niiden leviäminen on kuitenkin pysähtynyt, ja tavallisin käyttöympäristö on maaseutu, sillä niitä pidetään leimaavina kaupungeissa. Tähän ryhmään kuuluu yleisgeminaatio (*tulee – tullee*) sekä svaa-vokaali (*kolme – kolome*). Myöskään kolmannen ryhmän piirteissä ei ole eroa idän ja lännen välillä, mutta ne ovat vauhdilla leviäviä. Niissä on selvä sosiaalinen variaatio, ne ovat tunnettuja erityisesti pääkaupunkiseudulla. Esimerkkinä tästä ryhmästä mainittakoon -ea ja -OA-päätteisten sanojen assimilaatio. Tämä on toinen merkittävä ryhmä, johon tämän tutkimuksen piirteet kuuluvat. Neljänteen ryhmään lukeutuvat piirteet ovat selvästi leimautuneita ja

kartettavia aluepuhekielisyyksiä. Ne osoittavat tiettyyn maakuntaan ja maaseudulle, eikä niitä paljon kuule edes kyseisten alueiden kaupungeissa. Viimeisenä Mantilan ryhmistä ovat niin sanotut elävät paikallisuudet. Nekin ovat vahvasti alueellisia, eikä niitä välttämättä tunneta kaikkialla, mutta ne voivat silti olla leviämässä. Näistä esimerkkinä ovat *mie* ja *sie*-pronomiinien käyttö.

Piirteiden järjestys perustuu kuvan 1 piirrehierarkiaan. Kuvassa vaalea väri edustaa yleiskielisiä variantteja ja tumma puhekielisiä. Näistä keskiarvon laskien Paunonen on saanut puhekielisyyssprosentiksi 65 %. Kuva koskee töölöläisiä, joita pidetään yleiskielisempinä kuin monissa muissa Helsingin kaupunginosissa asuvia, erityisesti Sörnäsissä, joka on ollut tutkimuksissa vastinparina asuinalueista. Paunosen aineisto on 1990-luvulta, ja hänen vastaavassa aineistossaan 1970-luvulta puhekielisyyden asteeksi tuli vain 38 % (Paunonen 2006, 88–90). Tämä osoittaa selvästi trendin kohti puhekielen suosimista. Kuvasta nähdään, että Nuolijärven ja Lappalaisen suppeampien ryhmittelyjen mukaiset piirteet (esimerkiksi kaikki *i:n* tai *A:n* katotapaukset) ovat yleensä lähellä toisiaan hierarkiassa eivätkä koskaan vastakkaisissa ääripäissä, mikä puolustaa mahdollisuutta käsitellä niitä myös suurempina ryhminä.

Kuva 1. Helsinkiläisten kielenpiirteiden puhekielisyshierarkia Paunosen mukaan töölöläisten puheessa 1990-luvun alussa.



Lähde: Paunonen 2006, 90

Yksi tärkeä syy, miksi suurin osa piirteistä esiintyy jopa yksilöiden kohdalla vaihtelevasti sekä yleis- että puhekielisinä, on kielenkäyttötilanne. Kulloinkin valikoituva variantti määräytyy muun muassa keskustelukumppanin, puheenaiheen ja puheenvuoron luonteen mukaan (Nuolijärvi & Tiittula 2000, 228). Tässä yhteydessä ei kuitenkaan mennä syvemmälle näihin syihin, sillä tutkimuksen lähtökohta on, että ainakin puheenaihe onnistutaan vakioimaan.

Nuolijärvi ja Tiittula (2000, 233) huomauttavat tarkastellessaan esimerkkinä Esko Ahon televisiopuhetta, että laaja heterogeeninen yleisö voi saada puhujan tasapainottelemaan varianttien käytön kanssa kaikkien miellyttämiseksi. Myös internetissä on yleensä laaja heterogeeninen yleisö, mikä voi saada ihmiset valitsemaan vähemmän leimallisia variantteja kuin omassa puheessa. Toisaalta ei ole tavatonta, että huomiohakuisuus, aggressiivisuus tai oman identiteetin

korostaminen saivat tahallaan käyttämään murteellisia variantteja.

2.4 Sosiolingvistinen puhekielentutkimus Suomessa

Sosiolingvistiikka tutkii kielenulkoisten tekijöiden vaikutusta kielenkäyttöön, eli esimerkiksi sosiaalisia ja yhteiskunnallisia syitä siihen, miksi kieltä käytetään eri tavoin. Kotimainen sosiolingvistinen variaationtutkimus on erityisesti menetelmällisesti pohjana tälle tutkimukselle, vaikka tutkielman luonne on muuten hieman erilainen. Seuraavassa käsitellään lyhyesti sosiolingvistiikan kotimaista tutkimusta ja taustamuuttujien roolia siinä sekä pohditaan variaationtutkimuksen saamaa kritiikkiä.

Variaationtutkimus on lähtenyt Suomessa liikkeelle 1970-luvulla pian sen jälkeen, kun se Yhdysvalloissa kehitettiin (Lappalainen 2004, 13–14). Valtaosa tutkimuksesta on Lappalaisen mukaan hajallaan sadoissa aihepiiristä tehdyissä pro gradu -tutkielmissa (ks. myös Mantila 1997). Muutamia laajempia ja pääkaupunkiseudun puheeseen liittyviä tutkimuksia ovat tehneet erityisesti aiemmin mainitut Paunonen, Nuolijärvi ja Lappalainen. Paunonen (1995a; 2006) on tutkinut helsinkiläisen kielen vaiheita ja vertaillut erityisesti eri-ikäisiä töölöläisiä ja sörnäisläisiä helsinkiläisiä sekä heidän kielensä muuttumista ajan kuluessa. Nuolijärvi on selvittänyt muualta Suomesta Helsinkiin muuttaneiden kielellistä sopeutumista. Lappalaisen työ käsittelee variaationtutkimuksen ja taustojen perusteella kieltä selittävän perinteen laajentamista tilanteisella näkökulmalla. Hän yhdistää aiheeseen muun muassa keskustelunanalyysin menetelmiä. Pro gradu -työt puolestaan ovat keskittyneet pääasiassa paikallismurteisiin eri puolilla Suomea, eivätkä ne siksi ole tässä yhteydessä niin kiinnostavia.

Jonkin verran tätä vastaavan pro gradu -tutkielman on tehnyt Määttänen (2007). Hän selvitti kirjoitetun tekstin muuttumista puhekielimmäisemmäksi yleisemmin, verraten toisiinsa kaunokirjallisia tekstejä, elokuvien ja tv-ohjelmien tekstityksiä sekä internetin keskustelupalstojen viestejä, joista viimeinen on tutkimuksen kohteena tässäkin tutkielmassa. Hän tutki kielenpiirteiden lisäksi muun muassa syntaksia ja sanastoa.

Analyysi oli kuitenkin melko kevyt, ja Määttänen totesi verkkokeskustelusta lähinnä, että puhekielisyysaste jäi selvästi 0–100 %:n välille ja että muissa hänen tutkimissaan teksteissä prosentti ylsi monissa piirteissä sataan, koska ne oli kirjoitettu harkitusti täysin puhekielisiksi. Hän piti sanastoa (esimerkiksi slangi) tärkeimpänä elementtinä puheenomaisen vaikutelman luomisessa.

Taustamuuttujista ikä on kiistattomin ja merkittävin selittävä tekijä sosiolingvistisille eroavuuksille kielenkäytössä. Yhtäältä nuoremmat sukupolvet ovat puhekielisempiä kuin vanhat (Paunonen 2006, 77–81), ja toisaalta vanhat sukupolvet ovat ajan kuluessa muuttaneet puhettaan tuoreempaan suuntaan (mts. 81–88). Myös Nuolijärvi ja Tiittula (2000, 224) korostavat iän tärkeyttä, sillä heidän mukaansa nuorten puhe ennustaa tulevaisuuden kielenkäyttöä. Kuiri (1989) on selityksenä tälle esittänyt, että kielen muuttuessa nuoret oppivat uudet tuulet ensimmäisenä tapanaan käyttää kieltä, ja vanhat normit tuntuvat vierailta, mutta vanhemmat pitäytyvät itse nuorena oppimisessaan jo vanhentuneissa kielenkäyttötavoissa. Jos siis haluaa selvittää kielen muuttumista, kannattaa tutkia nuoria. Tässäkin tutkielmassa kiinnostus kohdistuu nykyiseen ja tulevaan kielenkäyttöön, ja aineisto on kerätty enimmäkseen nuorilta aikuisilta.

Tarkastellaan seuraavaksi sukupuolta, sillä se on tämän tutkimuksen empiirisen aineiston ainoa tunnettu taustamuuttuja. Sukupuolen merkitys voidaan tiivistää siten, että naiset suosivat standardin mukaista kieltä ja prestiisivariantteja, miehet alueellista ja sosiaalisesti leimaavampaa kieltä (Lappalainen 2004, 37–39). Nuolijärvi (1986, 218–219) päättelee, että naiset sopeutuvat uuteen kieliympäristöön miehiä paremmin. Lappalainen (2004, 37–39) jatkaa kuitenkin, että sukupuoli ei välttämättä ole todellinen taustamuuttuja vaan sosiaaliset verkostot, jotka yleensä eroavat toisistaan miehillä ja naisilla. Lisäksi sukupuolen merkitys on usein todettu mitättömäksi muihin tekijöihin verrattuna. Nuolijärvin (2000) kritisoi sukupuolta taustamuuttujana, sillä vaikka joitakin eroja sukupuolten välillä on löydetty, ne auttavat hyvin vähän sosiaalisen todellisuuden hahmottamisessa.

Sukupuolen lisäksi muitakin taustamuuttujia pidetään ikää alttiimpina kritiikille (Nuolijärvi & Tiittula 2000, 43; Lappalainen 2004, 58). Muita Suomessa tärkeitä

muuttujia ovat alueellinen tausta, koulutus ja ammatti. On helppoa ajatella, että maantieteellinen sijainti merkitsisi eniten, mutta Nuolijärvi ja Tiittula (2000, 222–223) huomauttavat, että esimerkiksi ammatti voi määrätä kielenkäyttöä enemmän kuin asuinpaikka. Lisäksi he ovat havainneet, että taustamuuttajat voivat korreloida keskenään. Esimerkkinä annetaan ammatti ja sukupuoli, mikä selittyy sillä, että naiset ja miehet ovat töissään toisaalta saman sukupuolen kanssa tekemisissä, ja toisaalta mies- ja naisvaltaisilla aloilla työtehtävät voivat asettaa erilaisia odotuksia ja rajoituksia kielenkäytölle. Lappalainen (2004, 34–41) nostaa perinteisten muuttujien joukkoon tilanteisen vaihtelun, jota on vaikea tutkia kvantitatiivisesti, ja käyttää itse apuna laadullisia menetelmiä, erityisesti keskusteluanalyysiä.

Sosiolingvistisen tutkimuksen ”peruskiveä”, variaationtutkimusta, johon tämäkin tutkielma perustuu, on varsinkin 2000-luvun puolella alettu myös kritisoida vahvasti. Nuolijärven (2000) mukaan se on erityisesti liian yksipuolista, eli se tarvitsisi tuekseen muita menetelmiä voidakseen antaa syvällisemmän kuvan tarkasteltavista ilmiöistä. Tällaisen kritiikin voi tulkita siten, että kypsyvän tieteenalan on pystyttävä vastaamaan muuttuvan maailman vaihtuviin ja kasvaviin vaatimuksiin. Se, mikä riitti 1970-luvulla, ei enää tyydytä selitysvoimallaan 2000-luvulla.

Lappalainen (2004, 15–16) puolestaan jakaa kritiikkinsä viiteen eri kohtaan. (1) Taustamuuttajat korreloivat vain silloin, kun tarkastellaan ryhmien keskiarvoja, eivät yksilöiden kohdalla. (2) Aineiston keruutapa on lähes poikkeuksetta ollut haastattelu, jolloin tilanteinen vaihtelu ja arkikeskustelunomainen vuorottelu jäivät toteutumatta. Tähän Lappalainen itse otti ratkaisuksi keskusteluanalyysin liittämisen tutkimukseen (ks. tästä myös Nuolijärvi 2000). (3) Tarkastelun kohteena ovat lähinnä äänne- ja muoto-opilliset piirteet, eivät sanasto ja syntaksi. Ongelma on siis kapea-alaisuus. Tosin Lappalainen myöntää, että muiden piirteiden tutkiminen on vaikeampaa. (4) Painotus on kvantitatiivisuudessa johtuen muun muassa suurista aineistoista, jolloin taustamuuttujien vaikutus selviää, mutta yksilölliset tekijät eivät. (5) Suuri osa varsinkin pro gradu -töistä keskittyy maaseutupaikkakuntien kieleen, vaikka suurin osa ihmisistä asuu eteläsuomalaisissa kaupungeissa.

Tämän tutkimuksen osalta Lappalaisen kohtiin voidaan vastata seuraavasti. (1)

Kiinnostus kohdistuu etupäässä ryhmiin, mutta toisaalta taustamuuttajat eivät ole edes tiedossa. (2) Tilanteista vaihtelua ei voida kontrolloida, mutta havainnoivasta datankeruutavasta johtuen arkikeskustelunomaisuus toteutuu, mikä on merkittävä etu perinteiseen tutkimukseen verrattuna. (3) Soveltaessa vanhaa menetelmää uudelle alueelle on varmintä turvautua olemassa olevaan tutkimusperinteeseen, muun muassa metodologisen vertailtavuuden vuoksi. (4) Kuten kohdassa yksi, kiinnostus on ryhmiä ja suuria kokonaisuuksia kohtaan. (5) Tässä keskitytään nimenomaan mahdollisimman yleiseen ja laajalle levinneeseen kaupunkikieleen.

3 Verkkokeskustelu

Kuten edellä todettiin, sociolingvistiikka tutkii tyypillisesti yhteiskunnallisia ja sosiaalisia syitä variaatioon kielenkäytössä. Yhtä lailla yhteiskunnallisten tekijöiden ohella voidaan olla kiinnostuneita muista kielenulkoisista tekijöistä. Median kehittymisen myötä tärkeäksi kategoriaksi muodostuvat tekniset tekijät. Matkaviestimet, televisio ja internet kaikki sekä asettavat erilaisia rajoituksia että avaavat mahdollisuuksia käyttää kieltä uusilla tavoilla.

Verkkokeskustelun tutkimus on toistaiseksi keskittynyt pitkälti englannin kieleen. Oleellisin ero juuri tämän tutkimuksen kannalta onkin juuri kieli. Siksi aikaisemmassa tutkimuksessa esitettyjä englannin kieltä koskevia väitteitä pyritään osin vahvistamaan sellaisenaan tässä tutkielmassa suomen kielen osalta. Kielellisten erojen vuoksi ulkomaiseen kirjallisuuteen voi kuitenkin tukeutua rajallisesti, ja jonkin verran kriittisesti on syytä suhtautua myös kulttuurierojen takia. Siksi luku 2 pohjautuikin kotimaiseen tutkimukseen. Tässä luvussa puolestaan keskitytään yleisiin verkkokeskustelun piirteisiin, joita ulkomainen tutkimus käsittelee selvästi laajemmin.

Verkkokeskustelussa on keskeistä, millaisia rajoituksia ja mahdollisuuksia se aiheuttaa keskustelulle erityisesti teknisten ominaisuuksiensa kautta. Näitä teknisiä ominaisuuksia eritellään tässä luvussa ensimmäiseksi. Tämän jälkeen pohditaan, mitä se merkitsee kielenkäytön kannalta. Verkkokeskustelulla on nimittäin useita sille ominaisia kielellisiä tekijöitä, joita ei perinteisessä kirjoitetussa tai puhutussa kielessä ole esiintynyt, ainakaan merkittävässä määrin. Tekniikka ei kuitenkaan peitä alleen sosiaalisia seikkoja, jotka ovat aina merkittäviä ihmisten välisessä viestinnässä, samoin kuin vaikkapa perinteisissä kirjeissä tai sanomalehdissä. Jaottelu teknisiin, kielellisiin ja sosiaalisiin ominaisuuksiin on viime kädessä veteen piirretty viiva, sillä kaikki vaikuttavat toisiinsa, mutta asiaan pyritään silti tuomaan selkeyttä jonkinlaisen ryhmittelyn kautta. Kaiken edellä esitellyn pohjalta kuvaillaan tämän jälkeen toisistaan poikkeavia tapoja keskustella verkossa. Näistä keskustelutyypeistä sopivimmat valitaan empiirisen analyysin pohjaksi. Lopuksi tehdään vielä lyhyt katsaus kansainvälisestä verkkokeskustelun variaationtutkimuksesta tehtyihin havaintoihin.

3.1 Teknisiä ominaisuuksia

Herring (2007) on luonut verkkokeskustelun tutkimusta varten luokittelujärjestelmän, joka karkeasti ottaen jakautuu lukuisiin eri teknisiin, sosiaalisiin ja tilanteisiin tekijöihin, joista kaikilla on havaittu jonkinlaista vaikutusta verkkokeskusteluun. Teknisiä tekijöitä ovat synkronisuus, yksi- tai kaksisuuntaisuus, pysyvyys, viestin koko, kanavat (eli mediat), anonymiteetti, yksityisyys, suodatus, lainaaminen sekä viestin muoto. Muihin kuin teknisiin tekijöihin palataan myöhemmin. Luokittelujärjestelmä on tarkoitettu mahdollisimman yleisluontoiseksi, joten sitä ei pidä noudattaa sokeasti vaan tilanteen ja tarkoituksen mukaan (ma.). Mainituista tekijöistä tarkoituksenmukaisimpia tässä tutkimuksessa ovat kanavat, synkronisuus sekä jossain määrin anonymiteetti.

Teksti on aina ollut internetissä pääasiallinen viestintäkanava. Aikaisemmin syynä olivat tekniset rajoitukset, joiden vähettyä muut kanavat ovat yleistyneet. Teksti hallitsee silti useimpia keskustelutyyppejä, kuten chattejä, keskustelupalstoja, blogeja ja sähköpostia (Herring 2010a). Sen tuottaminen ja siirtäminen on helppoa ja nopeaa. Äänen sekä nonverbaalisen viestinnän mahdollistavan videon käyttäminen ovat siis edelleen vasta yleistymisvaiheessa. Esimerkiksi videopuhelut mahdollistavat hyvin samankaltaisen keskustelun kuin kasvotusten, jolloin kielellisiä eroja saattaa olla vähemmän. Sen sijaan tekstimuotoinen keskustelu verkon välityksellä on pakottanut ihmiset sopeuttamaan kielenkäyttöään muun muassa muiden yllä mainittujen tekijöiden suhteen. On kuitenkin syytä mainita erityisenä viestintäkanavana virtuaalimaailmat, joissa on mahdollista viestiä teoilla ja muulla puhtaasti sanattomalla viestinnällä, kuitenkin siten, että henkilö itse ei fyysisesti tuota näitä toimintoja vaan erilaisten valmiina annettujen kommentojen kautta.

Ratkaisevin ja yleisesti hyväksytyin ominaisuus on keskustelun synkronisuus (Herring 2001a). Se tarkoittaa reaaliaikaisuutta, eli ovatko keskustelijat verkossa samaan aikaan vuorovaikutuksessa keskenään. Synkronisia viestintäkanavia ovat muun muassa web-pohjaiset chatit, IRC (Internet Relay Chat), ICQ ja Windows Messenger. Tekstimuotoiseen keskusteluun lisäksi myös puhelut ja videoneuvottelut kuuluvat tähän ryhmään. Asynkronisia muotoja, jotka ovat tavallisempia, ovat muun muassa

sähköposti, blogit ja keskustelupalstat. Niissä lähettäjällä on halutessaan aikaa valmistella viestinsä huolella, ja vastaanottaja ei ole tavallisesti paikalla, kun viesti saapuu. Keskustelun kannalta synkronisuus tarkoittaa palautteen välittömyyttä, kirjoitetuksi kieleksi nopeaa tempoa ja lyhyempää aikaa viestien suunnitteluun (ma.). Se on siis ominaisuuksiltaan lähempänä keskustelua kasvotusten.

Anonymiteetti on muodostunut internetissä usein käytetyksi mahdollisuudeksi, jolloin oma identiteetti voidaan aktiivisesti joko piilottaa tai näyttää. Täydellinen anonymiteetti nimimerkin suojissa on kuitenkin osittain harhaa. Ajatukset siitä, että verkossa henkilöstä näkyy aina vähemmän ja vuorovaikutus on siksi aina laimeampaa, on osoitettu vääräksi, ja se riippuukin pitkälti ympäristöstä ja muista tekijöistä (ks. esim. Mann & Stewart 2000, 180). Herring (2001a) mainitsee seuraavanlaisia esimerkkejä. Sosiolingvistiikan periaatteiden mukaan kieli kertoo käyttäjästänsä tahtomattakin. Huoliteltu kieli muiden keskustelijoiden käyttäessä puheenomaisempaa kieltä herättää huomiota ja synnyttää oletuksen korkeasti koulutetusta henkilöstä. Ikä paljastuu sellaisten viestien sisällöstä, jotka kertovat elämäntilanteesta tai -kokemuksista. Jos sukupuoli tuodaan esiin (ja monet nimimerkit viittaavat usein kirjoittajan sukupuoleen), se voi korostua paljon tavallista enemmän, koska muuta tarttumapintaa on vähemmän. Esimerkiksi virtuaalimaailmoissa sukupuoli on yleensä selvemmin esillä ja vaikuttaa siten käyttäytymiseen ja kieleenkin (Mann & Stewart 2000, 179). Myös keskustelun sisältö voi antaa tahattomasti viitteitä viestijän sukupuolesta (Herring 2001b). Ihmisiä jopa selvästi häiritsee, jos he eivät tiedä keskustelukumppaninsa sukupuolta (Crystal 2002, 51).

Esitellään vielä lyhyesti muut mainitut tekijät (ks. näistä tarkemmin Herring 2001a; 2007). Yksisuuntainen viesti on sellainen, joka lähetetään kokonaisuudessaan vastaanottajalle vasta sitten, kun se on valmis. Kaksisuuntainen viesti siirtyy koko ajan sitä mukaa, kun sitä tuotetaan, näppäimen painallus kerrallaan. Sitä voisi siis ajatella synkronisuuden äärimuotonakin. Valtaosa keskustelutavoista on yksisuuntaisia, sillä esimerkiksi chateissa viestit yleensä lähetetään vasta valmiina enteriä painamalla. Viestit voivat säilyä luettavissa eripituisia aikoja, jolloin puhutaan niiden pysyvyydestä. Sähköposti säilyy, kunnes vastaanottaja poistaa sen, mutta

chatissä viesti saattaa kadota, kun se vierii pois kuvaruudulta. Pysyvyys mahdollistaa asynkronisen viestinnän ja helpottaa samanaikaisten keskustelujen seuraamista. Viestien kokoa on rajoitettu muun muassa tekstiviesteissä ja mikroblogipalvelu Twitterissä tiettyyn merkkimäärään. Tämä kannustaa ihmisiä tiiviiseen ilmaisuun ja lyhenteiden käyttöön. On kuitenkin todettu, että kirjoittamisen hitaus on merkittävämpi syy tiiviin ilmaisuun käyttämiseen (ks. esim. Herring 2001a). Viestien yksityisyys määrittää, kuka voi vastaanottaa niitä. Sähköpostit ovat teknisiltä lähtökohdiltaan yksityisiä, mutta Twitterissä oletusasetuksena on julkisuus. Vastaanoton suodatus estää viestinnän, sillä suodatetut viestit eivät päädy vastaanottajalle ollenkaan. Tämä on tärkeä ominaisuus varsinkin roskapostin ja muiden ei-toivottujen viestien poissulkemisessa. Aikaisempien viestien lainaamisella voidaan esimerkiksi luoda vaikutelma viestien peräkkäisyydestä ja liittymisestä toisiinsa, vaikka niiden välissä olisi paljon aikaa tai muita viestejä. Viimeisenä tekijänä viestin muoto määrittää sen visuaalisen ulkoasun. Esimerkiksi sähköposteissa tämä tarkoittaa viestien ilmaantumisjärjestystä ruudulle sekä lähettäjän, lähetysajan ja otsikon erillistä näyttämistä viestin yhteydessä.

Kuten tämänkin tutkimuksen osalta myöhemmin nähdään, usein muutama ominaisuus riittää erottelemaan keskustelutyypit toisistaan yksittäisen tutkimuksen tarpeisiin. Esimerkiksi Baron (2010) on tutkiessaan verkkokeskustelun vuorotteluilmiöitä käyttänyt vain kahta muuttujaa, synkronisuutta ja yksityisyyttä. Herringin luokittelun yleisyydestä kertoo Baronin jaottelun sopiminen siihen, vaikkei sitä ole tietoisesti käytetty.

3.2 Kielenkäyttö

Internetin kielenkäytön osalta esitellään ensin uskomuksia sen luonteesta. Tämän jälkeen Gricen keskustelun periaatteita peilataan verkkomaailmaan, ja lopuksi pohditaan järjestelmällisemmin, miten kieli verkossa eroaa sekä yleis- että puhekielestä.

Verkossa tapahtuvaan kielenkäyttöön liittyy paljon uskomuksia ja ennakkoluuloja,

joista kaikki eivät pidä paikkaansa. Julkisuudessa esitetyt kuvat verkon kielenkäytöstä eivät usein perustu tutkimukseen, vaan niitä muodostetaan kärjistetyksi erilaisten päämäärien ajamiseen (Thurlow 2006). Voidaan esimerkiksi kauhistella ja lietsoa pelkoa kielen rappioitumisesta tavoitteena säilyttää vanhat normit. Thurlow (ma.) havaitsi seuraavat neljä pääasiallista tapaa suhtautua mediassa kielenkäyttöön verkossa.

Uusi erilainen tapa käyttää kieltä. Verkko on saanut aikaan kielellisen vallankumouksen. Sen erot perinteiseen tapaan käyttää kieltä korostuvat, ja siitä käytetään mielellään uusia termejä kuten *netspeak*.

Ihmettely nopeasti kasvavaa kielenkäyttötapaa kohtaan. Esitetään paljon superlatiiveja ja miljardiluokan lukuja internetissä ja muissa viestimissä kulkevista viestimääristä. Taloudellisissa yhteyksissä näihin liitetään arvioita suurista summista, joita kyseisessä viestinnässä (potentiaalisesti) liikkuu.

Pelko kielen rappiosta. Kieli korruptoituu internetissä vellovassa holtittomassa kielenkäytössä, jossa kielen oikeellisuuteen ei kiinnitetä huomiota. Erityisesti nuoriso on vaarassa olla oppimatta kunnollista kieltä ollenkaan.

Ulkonäöllisten erityispiirteiden korostaminen. Erityisesti lyhenteet ja hymiöt määrittävät kielenkäyttöä verkossa. Tämäkin nähdään yleensä kielteisessä valossa.

Kaikki neljä tapaa ovat selvästi kärjistettyjä sekä yksipuolisia, vaikka jokaisessa on myös todellisuuspohjaa. Ne ovat näkökulmaltaan ulkopuolisia, eikä yksikään ole välttämättä verkossa viestivien itsensä tapa nähdä asiaa. Voisi olettaa, että heille sekä itse verkkomedialla että kielellä on lähinnä välinearvo sosiaalisten tavoitteiden saavuttamiseen. Mitä tulee puhekielimäisyyden tutkimiseen verkkokeskustelussa, ensimmäinen kohta on kaikkein kiinnostavin. On selvää, että verkkokeskustelu poikkeaa puhekielestä sekä yleiskielestä ainakin neljännen kohdan kuvaamalla tavalla, mutta puhekielimäisyyden osalta tietoa ei juuri ole, minkä selvittäminen on yksi tämän tutkielman lähtökohdista. Tämän tärkeyttä puolestaan voi perustella juuri nopealla kasvulla eli Thurlow'n toisella kohdalla.

Kolmannen kohdan eli kielen rappion suhteen voidaan todeta, että samaa pelkoa on

historiassa lietsottu puhekielen suhteesta yleiskieleen. Jos tilanteen näkee analogisena ja ottaa esimerkiksi huomioon, että kirjakielen normit muuttuvat hitaasti puhekielen suuntaan, voisi olettaa verkon kielimuotojen vastaavalla tavalla ohjaavan tulevaisuudessa normitetun kielen kehitystä yhdessä puhutun kielen kanssa, ja kenties ohjaavan jopa puhekielen kehitystä.

Toinen tapa lähestyä kieltä on Gricen tunnettujen keskustelun periaatteiden kautta. Ne ovat laatu (todenmukaisuus), määrä (informatiivisuus), yhtenäisyys (asiassa pysyminen) ja tapa (selkeys). Yleisesti hyväksytään, että kuulija olettaa puhujan toimivan näiden periaatteiden mukaan.

Periaatteet eivät kuitenkaan täysin päde verkossa. Todenmukaisuutta on usein syytä epäillä, jos puhuja on tuntematon, ja jotkut syöttävät tarkoituksellisesti väärää tietoa saadakseen aikaan reaktioita muissa. Jotkut ihmiset haluavat vain tarkkailla toisia eivätkä itse osallistua keskusteluun, jolloin heidän osaltaan määrä on häiritsevän vähäistä. Toisessa ääripäässä muille lähetetään heidän tahtomattaan valtavasti viestejä, esimerkiksi roskapostia. Ne voivat myös sisällöltään olla häiritseviä, kuten seksuaalisia. Asiassa pysymisen periaate rikkoutuu siinä, että toisinaan ihmiset hakevat internetistä vain huomiota eivätkä välitä ollenkaan siitä, onko heidän viesteillään sisällöllistä merkitystä (ks. tästä esim. Arpo 2005, 166–171). Lopuksi monet verkkosivut ovat vaikeaselkoisia eivätkä siis selkeitä. (Crystal 2002, 48–58.)

Viimeinen huomio ei liity keskusteluihin vaan verkkoviestintään yleisemmin, joten se voidaan tässä sivuuttaa. Usein ihmiset vakiintuvat käyttämään tiettyjä verkkopalveluja säännöllisemmin kuin toisia. Näissä heille kehittyy jonkinlaisia sosiaalisia suhteita keskustelukumppaneihin, jotka eivät siis ole enää täysin tuntemattomia. Voitaneen olettaa, että tällaisten verkkotuttavien välillä syntyy vähemmän tarvetta syöttää toiselle väärää tietoa. Tässä tutkimuksessa valittiin henkilöitä, jotka ovat viestimääriensä perusteella vakiintuneita käyttäjiä valituissa verkkokeskustelupalveluissa. Niinpä heidän voi olettaa sekä tuntevan muita samankaltaisia käyttäjiä että olevan itse tunnettuja noissa palveluissa. Poikkeuksen saattavat muodostaa muut satunnaiset kävijät, joiden kanssa nämä henkilöt mahdollisesti myös keskustelevat. Todenmukaisuuden periaatteen voi siis katsoa

olevan tilannesidonnainen.

Määrän suhteen on olemassa erilaisia sääntelymahdollisuuksia. Osallistumattomuus keskusteluun ei useimmiten käytännössä ole ongelma, sillä siihen on totuttu, ja hiljaiset tarkkailijat voidaan vain jättää huomiotta. Liikaa viestivien varalle on kehitetty suodattimia, tai viestien pääsyä muiden nähtäväksi voidaan moderoida.

Yhtenäisyys on ehkäpä kiinnostavin periaate. Mikä katsotaan asiassa pysymiseksi, jos samaan aikaan käydään useita rinnakkaisia keskusteluja vaihtelevista aiheista vaihtelevien henkilöiden kesken, ja jatkuvasti tapahtuu eksymisiä aiheesta mitä kummallisimmille sivuraiteille? Crystal (2002, 58) toteaa, että asiassa pysymisen periaatetta tulisikin tulkita verkossa uudella tavalla. Siellä korostuvat erilaiset sosiaaliset merkitykset ja tavoitteet sisältöä enemmän. Niinpä huomiohakuisuus ja epäloogiselta vaikuttava keskustelunaiheiden vaihtelu on syytä nähdä osana internetin keskustelukulttuuria eikä totuttujen sääntöjen rikkomisena.

Tarkastellaan seuraavaksi konkreettisemmin, mitä yhtäläisyyksiä verkkokeskustelulla on sekä kirjoitettuun että puhuttuun kieleen. Crystal (2002, 25–47) vertailee näitä kolmea toisiinsa seitsemän ominaisuuden mukaan: aikariippuvuus, spontaanisuus, läsnäolo, rakenteen jäykkyys, sosiaalinen vuorovaikutus, korjattavuus ja ilmaisun rikkaus. Puhe on aikariippuvaa, häipyvää ja sillä on tietty yleisö, ja kirjoitus vastaavasti tilariippuvaa, pysyvää ja etäällä yleisöstä. Puheessa ei ole viivettä tuottamisen ja vastaanottamisen välillä, mutta kirjoituksessa väistämättä on. Puhetta tuotetaan spontaanimmmin ja väljemmin rakentuneesti, mutta kirjoituksessa harvemmin muun muassa sen pysyvyyden vuoksi, eli teksti voidaan vastaanottaa moneen kertaan. Puheen kasvokkainen vuorovaikutus mahdollistaa nonverbaalisten ja deiktisten ilmaisujen (kuten demonstratiivipronominit) käytön, joista jälkimmäisiä on kirjoittaessa varottava epäselvyyksien välttämiseksi. Puhe sisältää paljon ilmauksia ja ilmaisutapoja, joita ei voida tai ole tapana kirjoittaa, kuten slangia tai täytesanoja, ja kirjoitus puolestaan on rakenteeltaan hienostuneempaa ja jäykempää. Puhe soveltuu sosiaalisten suhteiden ylläpitämiseen ja kirjoitus faktojen esittämiseen ja tallentamiseen. Puheetta ei voi muuttaa tuottamisen jälkeen. Vaikka tämä on kirjoitetussa tekstissä mahdollista, jäljelle jäävät virheet ovat toisaalta pysyviä.

Puheessa ilmaisun rikkaus liittyy ääntämismahdollisuuksien laajaan kirjoon, kuten voimakkuuteen, taukoihin ja sävyyn. Tekstissä vastaava vahvuus on graafinen ilmiasu.

Verkkokeskustelussa näiden ominaisuuksien asettuminen puhetta tai kirjoitusta lähemmäs riippuu usein keskustelun tyypistä. Aikariippuvuus, spontaanius, rakenteen löyhyys ja sosiaalinen vuorovaikutus muistuttavat varsinkin synkronisissa keskustelumodoissa enemmän puhetta. Samaan aikaan verkkokeskustelut ovat väistämättä tilariippuvia, joten kyse ei ole yksinkertaisesta joko-tai-jaosta. Lähellä kirjoitusta ollaan tilariippuvuuden lisäksi läsnäolossa ja faktuaalisuudessa. Sen sijaan ilmaisun rikkaus puuttuu usein sekä äänneiden että graafisen ilmiasun suhteen. (Crystal 2002, 28–48.)

Kokonaisuutena kieli internetissä on lähempänä kirjoitettua kuin puhuttua (Crystal 2002, 47), mutta tällöin huomioidaan kaikki sen sisältö eikä vain keskusteluja. Tavallisten verkkosivujen kaltainen yksisuuntainen viestintä on nimittäin lähes kaikin tavoin lähellä kirjoitettua päätä. Varsinaisissa keskustelutavoissa on niin paljon hajontaa, ettei vastaavaa yleistystä voida tehdä, kuten ominaisuuksien hajonnasta edellä todettiin.

Myös Mann ja Stewart ovat vertailleet puhetta, kirjoitusta ja verkkokeskusteluja toisiinsa. Heidän mukaansa verkkokeskustelussa suhteessa toisiin muotoihin voidaan nähdä seuraavat suorat analogiat (Mann & Stewart 2000, 181–182). Se on keskustelunomaista, ja varsinkin synkronisessa viestinnässä ihmiset kokevat 'puhuvansa' eivätkä 'kirjoittavansa'. Siinä voidaan kirjoittamisen tavoin harkita omia viestejä ja vastauksia toisten viesteihin, varsinkin sähköpostin tapaisissa asynkronisissa muodoissa. Vuorovaikutukseltaan se on verrattavissa puhelimeen, jossa ihmiset nopeasti vaihtavat kuulumisia. Lisäksi kirjoitus verkossa sisältää paljon lyhenteitä ja muita ominaisuuksia, jotka ovat tuttuja muistiinpanojen tekemisestä paperille.

Näiden yhtäläisyyksien rinnalla Mann ja Stewart (2000, 184–187) näkevät joitakin huonoja ja hyviä puolia verrattuna puheeseen ja kirjoitukseen. Kirjoitukseen verrattuna heikkoutena on se, että verkkokeskustelu johtaa helposti

harkitsemattomaan ja pinnalliseen ajatustenvaihtoon, jolloin keskustelulle ei synny sisällöllistä merkitystä. Hyvä puoli on se, että kirjoituksen tavoin (varsinkin asynkroninen) verkkokeskustelu mahdollistaa harkitun viestien tuottamisen ja vastaanottamisen itselle sopivaan, rauhalliseen tahtiin. Lisäksi viestit jäävät talteen kuten kirjoituksessa, jolloin niihin voi palata myöhemmin, voidaan käsitellä suurempia viestimääriä ja ottaa osaa useampiin keskusteluihin samanaikaisesti kuin kasvotusten olisi mahdollista. Puheeseen verrattuna puolestaan vuorovaikutuksen luonnollisuudessa menetetään paljon, ja syntyy helpommin väärinymmärryksiä. Myönteisenä seikkana viestintä voi olla luonteeltaan keskustelunomaista samoin kuin puhuessa.

Verkkokeskustelulla on myös paljon ominaisuuksia, jotka eivät juonnu suoraan kirjoituksesta eivätkä puheesta. Tavoitteena on viestinnän taloudellisuus, eli halutaan saada sanomat lähetettyä mahdollisimman nopeasti ja pienellä vaivalla. Se johtaa kieliopin ja oikeinkirjoitussääntöjen rikkomiseen, suureen virhemäärään sekä runsaaseen lyhenteiden käyttöön. Kielioppisäännöistä poiketaan kuitenkin siten, että viestien ymmärrettävyys säilyy. Tavallisia rikkeitä ovat isojen kirjaimien kirjoittaminen pienellä, välimerkkien jättäminen pois ja kirjoitusvirheiden korjaamatta jättäminen, kuten esimerkiksi 1 alla nähdään. Tämä kaikki on erityisen yleistä synkronisessa keskustelussa. Asynkronisessa sen sijaan näillä tekijöillä voidaan varioida kieltä tilanteen mukaan. Sähköpostista pystyy tekemään epämuodollisemman juuri tällaisilla keinoilla, tai vastaavasti muodollisemman pitäytymällä yleiskielessä, kirjeen kirjoittamisen normeissa ja oikeinkirjoitussäännöissä. Merkittävä huomio onkin, että kyseessä ei ole pelkästään teknisten seikkojen aiheuttama kielellinen muutos vaan itse asiassa sosiaalisesti varioiva ilmiö. (Mann & Stewart 2000, 187–188; Crystal 2002, 81–91; Herring 2001a.)

Esimerkki 1, oikeinkirjoitussäännöistä poikkeamista (informantti #12, muiden keskustelijoiden viestit poistettu):

ei me mennä treffeille

no ei vitussa menn

no en mä sit lähe

#I3: en oo, siin on iso ero

Verkkokeskustelu vaatii puheesta ja kirjoituksesta poiketen myös teknistä tietokoneen käyttötaitoa. Kieli näkyy keskustelijan kokemuksessa sekä internetistä että sen kulttuureista, ja uudet käyttäjät leimautuvat tässä suhteessa helposti aloittelijoiksi. Lisäksi nonverbaalisia vihjeitä lisätään tekstiin ja keskustelun vuorottelusäännöt rikkoutuvat, ja näitä käsitellään seuraavaksi tarkemmin osana verkkokeskustelun sosiaalisia ominaisuuksia. (Mann & Stewart 2000, 187–188; Herring 2001a.)

3.3 Sosiaalisia ominaisuuksia

Verkkokeskustelulla on samoja sosiaalisia funktioita kuin puhutulla keskustelulla (Herring 2010a), ja se heijastelee muutenkin käyttäjiensä sosiaalista todellisuutta (Herring 2001a). Se ei siis vain täytä lukuisia keskustelun ulkoisia tunnusmerkkejä, kuten äänenpainojen imitointia tekstillä, vuorottelua ja keskustelunomaista koodinvaihtoa.

Teknisesti samantyyppisiin keskusteluihin on syntynyt varsin erilaisia keskustelukulttuureita (Herring 2001a). Esimerkiksi web-pohjaiset chatit ja IRC ovat teknisesti hyvin samankaltaisia keskustelumuoja, mutta niitä käyttävät yleensä erilaiset ihmiset, ja niihin on muodostunut erilaisia kirjoittamattomia normeja esimerkiksi kielenkäytön suhteen. Vaikka – kuten aiemmin on todettu – tekniset, kielelliset ja sosiaaliset tekijät sitoutuvat helposti toisiinsa, ne vaikuttavat joissakin tapauksissa siis selvästi irrallaan toisistaan. Tässä luvussa käsitellään viimeistä ryhmää näistä eli sosiaalisen vuorovaikutuksen tekijöitä.

Luultavasti tärkein erottava tekijä puheen ja verkkokeskustelun välillä on nonverbaalisten elementtien lähes täydellinen puuttuminen, jolloin sosiaalinen vuorovaikutustilanne on ratkaisevasti erilainen. Toinen poikkeava ilmiö on keskustelijoiden välinen vuorottelu, jonka ominaisuuksia tekniset rajoitteet muuttavat.

Keskustelun sääntöjä ajatellen institutionaalisuus on vielä internetissä kiinnostava kysymys, sillä sen määrittely on hankalampaa kuin tavallisesti.

Palataan kuitenkin ensiksi Herringin (2007) luvussa 3.1 mainittuun luokittelujärjestelmään. Tekniset tekijät olivat siitä vasta ensimmäinen puoli. Toiseen lukeutuvat osallistumistekijät, taustatekijät (esimerkiksi sukupuoli, ikä, kielitaito, tietotekninen taito ja arvoasema), viestinnän tarkoitus, aihepiiri, viestien sävy, viestinnän toiminnallinen tarkoitus (kuten väittely tai vitsailu), normit sekä kieli. Osallistumistekijöitä ovat esimerkiksi yksityisyys, anonymiteetti ja monenvälisyys, eli tapahtuuko keskustelu kahden kesken vai suuremman ryhmän välillä. Kiinnostavaa on havaita yksityisyys ja anonymiteetti sekä teknisissä että muissa tekijöissä. Niitä voidaan siis katsoa kahdesta näkökulmasta: miten keskustelu ympäristön teknisesti mahdollistaman anonymiteetin käyttö näkyy vuorovaikutuksessa, ja millä tavoin tuntemattomuus vaikuttaa sosiaalisesti ja kielellisesti keskustelukäyttäytymiseen. Muutenkaan tekijät eivät ole irrallisia vaan yleensä korreloivat useiden toisten tekijöiden kanssa.

Asiaa voidaan tarkastella myös esimerkiksi vastaanottajien suhteen. Julkisesti viestiminen verkossa muistuttaa muita eräällä tapaa yksisuuntaisia massamedioita kuten televisiota, joissa yleisönä on koko Suomi (televisiokeskusteluista ks. Nuolijärvi & Tiittula 2000, 225). Yleisö vaikuttaa viestimiseen ja kielivalintoihin muun muassa siten, että yksityinen viestintä on usein epämuodollista ja julkinen vahvasti institutionaalista ja normitettua. Kiinnostavinta verkossa on kuitenkin yksityisen ja julkisen rajakohta. Siellä useat keskustelutyypit ovat sellaisia, että viestit on kohdistettu rajatulle vastaanottajajoukolle, mutta kuka tahansa voi silti lukea niitä. Tämän tutkielman aineisto on kerätty lähtökohtaisesti juuri tällaisista palveluista. Yleisön rajoittaminen saattaa tehdä viestinnästä epämuodollisempaa, vaikka foorumi teknisesti ottaen on julkinen, ja viestijä voi jopa olettaa tietynasteisen yksityisyyden viesteilleen (ks. Arpo 2005, 65).

On selvää, että sanaton viestintä teknisesti ottaen puuttuu tekstimuotoisesta verkkokeskustelusta. Käytännössä sitä voidaan osittain korvata verbaalisesti, jolloin nonverbaaliset vihjeet muodostuvat eri tavoin, kuin on totuttu (Walther, Loh & Granka

2005). Tunteita ja ilmeitä ilmaistaan muun muassa sanoin tai hymiöiden avulla. Seuraavan esimerkin kolmessa viestissä on käytetty tavanomaista silmää iskevää hymiötä, tykkäämistä ilmaisevaa sydänsymbolia sekä 3. persoonassa ilmaistua viestin lähettäjän tekemää toimintaa, jonka tunnusmerkki viestin alussa oleva '*' tässä tapauksessa on. Sanattoman viestinnän korvaamiseen on siis melko monipuolisiakin mahdollisuuksia, vaikka viestit täytyy tuoda eksplisiittisesti esiin.

Esimerkki 2, nonverbaalista viestintää hymiöiden ja sanojen avulla (informantti #I3):

seuraan vähän jälkijunassa touhuja ;)

nyt on sentään melkein jouluaatto <3

* #I3 irroittelee tarroja

Kasvokkaisen keskustelun päällekkäisyyden välttämisen säännöt eivät päde sellaisenaan verkkokeskustelussa. Kaikki voivat jutella samaan aikaan ja tulla silti kuulluiksi, ja samassa paikassa voi samanaikaisesti olla useita keskusteluja käynnissä päällekkäin. Onkin esitetty, että puhutun keskustelun vuorotteluperiaatteista kannattaa poiketa, sillä tekniset ominaisuudet mahdollistavat tehokkaampia keskustelustrategioita. Esimerkkinä tällaisesta on viestien kohdentamisen nimellä (ks. esimerkki 1 aiemmin), mikä on muissa viestintämuodoissa yleensä vähemmän tarpeellista ja käytettyä, mutta verkkokeskustelussa auttaa erottamaan oikean keskustelun useiden samanaikaisten joukosta. (Mann & Stewart 2000, 187–188; Herring 2001a; Anderson, Beard & Walther 2010.)

Verkkokeskustelussa pyritään kuitenkin pitkälti puheesta tuttujen keskustelunormien ja -käytäntöjen noudattamiseen (Herring 2010a; Anderson ym. 2010). Yksi näistä on puheenvuoron vaihtuminen keskustelukumppanien välillä. Esimerkiksi pitkiä hiljaisia taukoja halutaan välttää, ja kysymyksen esittäminen toiselle synnyttää odotuksen tältä saatavasta vastauksesta. Verkkokeskustelussa nämä muodostuvat ongelmallisiksi teknisten rajoitteiden takia (Herring 2001a). Verkossa saattaa esimerkiksi olla viivettä, jolloin viestien perille menemiseen kuluu odotettua pidempi aika, jonka vastaanottaja saattaa tulkita tarkoituksellisena pitkänä taukona. Vaikka

vuorojen välissä on väistämättä enemmän taukoa verkossa, niiden on kuitenkin havaittu olevan suhteessa yhtä pitkiä kasvokkaisessa ja (asynkronisessa) verkkokeskustelussa (Kalman, Ravid, Raban & Rafaeli 2006).

Pohditaan seuraavaksi keskustelun institutionaalisuutta, jonka Nuolijärvi ja Tiittula (2000, 13–16, 215, 217–218) määrittelevät seuraavasti. Julkisella foorumilla käyty keskustelu on aina institutionaalista, ja siinä on aina mukana vähintään yksi osallistuja, jolla on määriteltävissä oleva instituution antama tehtävä tai tavoite. Keskustelijoilla on epäsymmetrinen suhde toisiinsa. Institutionaalinen keskustelu voi näyttää spontaanilta ja olla silti suunniteltu etukäteen.

Tässä määritelmässä ei selvästi ole huomioitu verkkokeskustelua. Esimerkiksi useimmat chatit ovat julkisia foorumeita, mutta niiden luonne on silti hyvin arkikeskustelunomainen. Asia onkin lopulta melko tulkinnanvarainen. Jos keskustelukanavalla on moderaattori, hänellä voidaan katsoa olevan institutionaalinen rooli pitää yllä järjestystä. Hänellä on myös valtaa muihin nähden, eli roolit ovat epäsymmetriset. Muut keskustelijat voivat silti pääosin jättää moderaattorin (eli ainoan institutionaalisen osapuolen) huomiotta. Tällöin keskustelun institutionaalisuus vaihtelisi sen mukaan, osallistuuko moderaattori keskusteluun. Institutionaalisuuden tiedetään vaikuttavan kielenkäyttöön ja nimenomaan puhekielisyteen (ks. Nuolijärvi & Tiittula 2000), joten kysymys on kiinnostava. Myöskään muissa keskustelutyypeissä asia ei ole yksiselitteinen.

Cho (2010) vertasi sähköpostin käyttöä muistiona perinteisiin paperimuistioihin työyhteisössä. Hän totesi, että sähköpostit ovat vähemmän muodollisia ja sisältävät enemmän puhemaisia ominaisuuksia kuin paperimuistiot. Lisäksi Chon (ma.) tutkimus osoitti, että sähköposteilla on paperimuistioita vahvempi sosiaalinen funktio. Tutkimustulosten voikin katsoa kertovan siitä, että institutionaalisessakin yhteydessä voi verkossa esiintyä puhekielisyyttä.

Tässä tutkimuksessa tehdään yleistäviä luonnehdintoja keskustelutyypin institutionaalisuudesta yksinkertaisuuden vuoksi. Institutionaalisenä tyyppinä voidaan pitää erityisesti blogia, tapauksesta riippuen jossain määrin sähköpostia sekä keskustelupalstaa, ja vastaavasti hyvin arkikeskustelunomaisina ainakin chattejä.

Blogin ylläpitäjällä on valta valita keskustelunaiheet, sillä lähtökohtaisesti vain hänen kirjoituksensa toimivat keskustelunavauksina. Muissa institutionaalisuus tulee esiin heikommin esimerkiksi lähetettyjen viestien moderoinnin kautta, ja keskustelujen lukeminenkin saattaa vaatia ylläpitäjän hyväksynnän. Sähköpostilla hoidetaan kenties muita enemmän institutionaalista viestintää, kuten kutsuja, työhakemuksia ja tiedottamista.

3.4 Keskustelutyypit

Edellä on esitelty verkkokeskustelun ominaisuuksia erilaisista näkökulmista. Näiden perusteella keskustelut voidaan jakaa muutamiin selvästi toisistaan poikkeaviin tyyppeihin. Tällaisia tyyppiluokituksia käydään seuraavaksi joitakin läpi, ja lopuksi kaiken esitetyn pohjalta luokitellaan verkkokeskustelut tämän tutkimuksen kannalta sopiviin tyyppeihin. Jako on pitkälti tulkinnanvarainen, sillä on olemassa esimerkiksi erilaisia hybridimuotoja. Lisäksi on useita marginaalisia erityistapauksia, joihin ei tässä paneuduta.

Crystal (2002) on jakanut verkkoviestinnän viiteen kielenkäytöltään eroavaan kategoriaan: sähköposti, chatit, keskustelupalstat, virtuaalimaailmat ja web. Hieman tarkempaan jakoon ovat päätyneet Mann ja Stewart (2000, 8–14). Edellisten ohella he erottavat joukosta omiksi tyypeikseen sähköpostilistat, uutisryhmät, multimediamyymälät ja jopa online-fokusryhmät. Näistä sähköpostilistat ja uutisryhmät voitaisiin katsoa lähinnä webissä toimivien keskustelupalstojen kaltaisiksi, vaikka niillä on yhtäläisyyttä myös tiedottavaan yksisuuntaiseen viestintään ja sähköpostilistoilla nimensä mukaisesti sähköpostiin. Syynä eroihin luokittelussa saattaa olla Mannin ja Stewartin tutkimuksen tekemisen näkökulma, kun Crystal keskittyi kielen variaatioon.

Näistä jaoista voidaan tehdä kaksi tärkeää huomiota. Ensiksikin ne käsittävät kaiken verkkoviestinnän. Varsinkin web poikkeaa muista viestinnän yksisuuntaisuudellaan, eikä sitä voida pitää keskustelumuotona. Toiseksi luokittelut ovat osin vanhentuneita. Osa keskustelutavoista on menettänyt suosiotaan, ja erityisesti sosiaalinen media on

tuonut uusia keskustelupalveluita ihmisten ulottuville. Näistä on syytä mainita ainakin Facebook, Twitter ja blogit.

Facebook on kiinnostavan monimuotoinen tapaus. Se sisältää chatin, asynkroniset pikaviestit (tilapäivitykset) sekä yksityisviestit (sähköposti). Viimeisimpänä kehityksenä chat ja yksityisviestit on vielä osittain sulautettu toisiinsa. Periaatteessa siis pelkästään Facebookista saisi monipuolisen tutkimuskohteen. Lisäksi tapahtumasivut, fanisivut, pelit ja muut sosiaalista vuorovaikutusta sisältävät ominaisuudet tekevät Facebookista vaikeasti luokiteltavan ja hahmotettavankin kokonaisuuden.

Herring, Scheidt, Wright ja Bonus (2005) näkevät blogit kotisivuja ja verkkokeskustelua yhdistävänä ja niiden väliin jäävänä genrenä. He vertaavat näitä kolmea kolmen tekijän mukaan: päivitysten tiheys, viestinnän symmetrisyys ja multimodaalisuus. Verkkokeskustelu on yleisesti jatkuvasti päivittyvää, viestinnältään kaksisuuntaista ja pelkkään tekstiin perustuvaa. Web-sivut puolestaan päivittyvät harvoin, ovat lähinnä tekijänsä yksisuuntainen viestintäkanava ja sisältävät multimediaa. Yksisuuntaisuuden takia kotisivuja ei voida nähdä verkkokeskusteluna, mutta blogit lähestyvät keskustelun luonnetta erityisesti siksi, että lukijoilla on niissä yleensä mahdollisuus kommentoida kirjoituksia ja siten käydä keskustelua kirjoittajan kanssa.

Keskustelutyyppejä voidaan myös käyttää eri tavoin niiden ominaisuuksista riippumatta (Baron 2010). Esimerkiksi sähköposti voi lähestyä synkronista viestintää, jos osapuolet ovat samaan aikaan paikalla ja vastaavat viesteihin välittömästi. Vastaavasti reaaliaikaisia pikaviestipalveluita (instant messaging) voidaan käyttää asynkronisesti, eli viestit menevät perille, vaikka vastaanottaja ei samalla hetkellä olisikaan paikalla. Kaiken kaikkiaan jako on siis hyvinkin subjektiivinen, kuten edellä todettiin. Tämän tutkimuksen kannalta järkeväksi osoittautui taulukon 1 mukainen luokittelu.

Taulukko 1. Verkkokeskustelutyypit ja niiden ominaisuuksia.

Keskustelutyyppi	Kanavat	Synkroninen	Institutionaalinen
reaaliaikainen keskustelu, chat	teksti	kyllä	ei
ei-reaaliaikaiset pikaviestit	teksti	ei, mutta voi lähestyä sitä	ei
keskustelupalsta	teksti	ei	lievästi
sähköposti	teksti	ei	vaihtelee
blogi	teksti	ei	kyllä
multimedia	monimediainen (video, ääni)	kyllä	ei yleensä
virtuaalimaailma	monimediainen (virtuaalimaailma, teot)	kyllä	lievästi

Reaaliaikaisiin keskusteluihin luetaan web-chatit, eri Messengerit, IRC, ICQ ja muut vastaavat palvelut. Ne ovat verkkokeskustelun toisessa ääripäässä erityisesti synkronisuutensa vuoksi. Niitä voi luonnehtia viestintäkanaviltaan (vain tekstimuotoisina) kapeiksi, epäinstitutionaaliksi, usein anonyymeiksi, epämuodollisiksi ja viestinnältään taloudellisuuteen pyrkiviksi. Viestintä voi olla yksityistä kahdenvälistä tai julkisempaa ja jopa kohdistamatonta omien ajatusten purkamista keskustelunavauksina.

Toisen ryhmän muodostavat *asynkroniset pikaviestit*, joihin lukeutuvat lähinnä mikroblogipalvelu Twitter ja Facebookin tilapäivitykset. Molemmissa valitun lähettäjäjoukon viestejä voidaan lukea kronologisesti yhdestä viestisyöttestä ja vastata niihin omilla kommenteilla. Viestien tuoreus on tärkeä tekijä, eikä useita päiviä vanhoihin viesteihin yleensä kiinnitetä enää huomiota. Tässä suhteessa palvelut lähestyvät reaaliaikaisuutta, vaikka ne eivät teknisesti sitä ole. Palvelut eivät ole anonyymeja tai juurikaan institutionaalisia. Vastaanottajajoukko saattaa olla tiedossa, mutta viestejä ei yleensä kohdisteta yksittäisille henkilöille. Siksi Twitter ja Facebookin tilapäivitykset ovatkin enemmän oman persoonan ja mielipiteiden esiintuomista kuin tasavertaista dialogia.

Keskustelupalstoissa korostuu niiden sisällöllinen monimuotoisuus, vaikka teknisesti ne ovat hyvin samankaltaisia keskenään. Sisältö ja tarkoitus sekä niiden kautta osallistujien luonne ratkaisevatkin kielenkäyttöä oletettavasti eniten. Keskustelupalstoihin voidaan lukea myös joitakin teknisesti poikkeavia muotoja, kuten uutisryhmät. Keskustelupalstoja voi luonnehtia asynkronisiksi, usein julkisiksi ja anonyymeiksi, edellisistä muodoista poiketen jossain määrin institutionaalisiksi sekä viestinnältään edellisiä vaihtelevammin muodolliseksi.

Sähköposti on asynkronista, usein muodollista, vaihtelevasti institutionaalista ja lähtökohtaisesti yksityistä viestintää. Keskustelijat siis tuntevat toisensa, eikä muista muodoista poiketen ulkopuolisilla ole koskaan mahdollisuutta edes lukea keskusteluja.

Kuten edellä on todettu, *blogit* alkavat jo hieman etäännyä varsinaisista verkkokeskusteluista. Blogaus on tavallista omalla nimellä, joten se ei usein ole anonyymiä. Sen luonne on valta-asetelmaltaan vahvasti institutionaalista, ja se on kaikista keskustelutyypeistä keskimäärin julkisinta. Edellä esiteltyjä pikaviestejä selvemmin esiin nousee omien mielipiteiden yksipuolinen esiintuonti. Tekstin lisäksi multimedia voi olla blogeissa rajoitetusti mahdollista.

Multimediakeskusteluksi on luokiteltu tässä kaikki puhumista sisältävä keskustelu. Kaksi selvää päämuotoa ovat äänikeskustelut esimerkiksi Skypeen kaltaisten internetpuhelupalveluiden kautta sekä videokeskustelut kuten yhteydenpito web-kameran avulla, joista jälkimmäisessä äänen lisäksi osapuolet siis näkevät toisensa. Keskustelumuoto on synkroninen ja yksityinen, mutta ei anonyymi eikä yleensä institutionaalinen. Poikkeuksena ovat esimerkiksi töissä käytävät videoneuvottelut.

Viimeisenä ja erikoisimpana keskustelutyypinä ovat *virtuaalimaailmat*. Virtuaalimaailmat voivat olla tekstipohjaisia tai multimodaalisia, ja ne ovat erityisen tavallisia tietokonepeleissä. Niissä käyttäjät pystyvät liikkumaan kuvitteellisessa tilassa ja suorittamaan erilaisia toimintoja. Näitä voi käyttää apuna viestinnässä, joka verbaaliselta osaltaan usein tapahtuu kuitenkin tekstinä. Virtuaalimaailmat ovat synkronisia mutta anonyymeja, ja institutionaalisia lähinnä pääsyn rajoittamisen ja moderoinnin osalta. Tosin niiden virtuaalisesta luonteesta johtuen on esimerkiksi

mahdollista luoda institutionaalisia keskustelutilanteita.

Kanavat eli käytetyt keskustelumediat ovat tärkeä rajaava ominaisuus, koska tässä tutkimuksessa fokuksena ovat vain tekstimuotoiset keskustelut. Niinpä multimediatekstit ja virtuaalimaailmat jätetään empiirisen analyysin ulkopuolelle, vaikka virtuaalimaailmoissa sanallinen keskustelu usein tapahtuukin kirjoittaen. Oletettavasti tärkein ominaisuus keskustelutyypeissä on tämän tutkimuksen kannalta synkronisuus, ja muiden avulla voidaan ensisijaisesti korostaa muita keskinäisiä eroja. Institutionaalisuuden todettiin olevan verkossa tavallista monimutkaisempi käsite, mutta yleisesti ottaen vähemmän institutionaaliset keskustelut ovat puhekielisempiä.

Synkroniset ja epäinstitutionaaliset keskustelut on todettu arkikeskustelunomaisiksi. Tällä perusteella reaaliaikaisista keskustelutyypeistä on luonnollista valita jokin analysoitavaksi. Asynkroniset pikaviestit ovat kiehtovalla tavalla lähellä niitä, vaikka ne erottuvatkin selvästi omaksi tyypikseen. Niistä on myös toistaiseksi tehty suhteellisen vähän kielitieteellistä tutkimusta (Androutsopoulos & Beißwenger 2008), mikä puoltaa niiden valintaa. Jäljelle jäävistä kolmesta blogit ovat selvimminkin etäällä arkikeskusteluista ja muutenkin luonteeltaan hyvin rajallisesti keskustelunomaisia. Sähköposti on aineistona muita hankalammin saatavissa, erityisesti mitä tulee arkiseen viestittelyyn. Niinpä kolmanneksi vertailukohdaksi empiiriseen analyysiin valikoituvat keskustelupalstat, jotka voi ominaisuuksiensa pohjalta olettaa kuitenkin vähemmän puhekielisiksi kuin kaksi muuta valittua keskustelutyyppeä.

3.5 Verkkokeskustelun variaationtutkimus

Sosiolingvistiikkaan pohjautuvaa variaationtutkimusta on jonkin verran tehty lähinnä internetissä käytetystä englannin kielestä, joten on syytä luoda lyhyt katsaus myös siihen. Siinä perinteinen äänneiden tutkiminen on vaihtunut puheen kirjoitetun ilmiön tutkimukseen. Yleisesti ottaen variaationtutkimus verkossa on keskittynyt teknologian vaikutukseen kieleen. Tutkimuskohteista ovat jonkin verran muita suosittumia olleet synkroniset kirjoitetut keskustelumuodot. (Androutsopoulos

tulossa.)

Variaationanalyysin soveltamisessa verkossa on useita rajoitteita, joista osa pohjautuu sen ongelmiin jo puhutun kielen tutkimuksessa (Androutsopoulos tulossa):

1) Variaationanalyysi on luontevinta lähempänä kasvokkaista keskustelua oleville keskustelumuuodoille. Uusissa web 2.0 -ympäristöissä kieli integroituu kuitenkin hyvin hajanaisiin, visuaalisiin ja multimodaalisiin yhteyksiin, eikä sitä voida välttämättä analysoida pelkkänä kielenä vaan monimutkaisina semioottisina rakenteina. Siksi onkin luontevaa nähdä nämä uudet verkkomedian muodot juuri laadullisempaa tutkimusotetta vaativina.

2) Variaationanalyysi soveltuu parhaiten yksittäisen kielen analysointiin, mutta verkossa monikielinen koodinvaihto sekoittaa kieliä toisiinsa.

3) Variaationanalyysi jättää huomiotta ilmiöitä, kuten hymiöt ja kielen valinta, joita on vaikea käsitellä kielellisten muuttujien kautta.

4) Kvantitatiivinen lähestymistapa ei sovi harvoin esiintyville ilmiöille, vaikka ne voivat olla merkityksellisiä.

5) Riippuvaisuus kielenulkoisista muuttujista selittävinä tekijöinä voi ohjata tutkimaan dataa, joista ne ovat helposti saatavilla, ja vaikeammin sosiaalisesti tai demografisesti luokiteltavat kohteet jäävät vähemmälle. Iän ja sukupuolen kaltaiset taustatekijät eivät välttämättä ole aina merkityksellisiä.

Näitä ongelmia on Androutsopoulosen (tulossa) mukaan voitu osittain kiertää valitsemalla tutkimuskohteiksi sopivimpia keskustelutyypppejä, mikä on johtanut jossain määrin yksipuolisuuteen. Myös tässä tutkimuksessa voidaan nähdä joidenkin kohtien olevan ongelmallisia. Edellä esitetyn listan kohdat 2, 3 ja 4 eivät kuitenkaan ole tällaisia. Puhekieli on ilmiönä nimenomaan usein esiintyvä sekä sellainen, jota kielellisten muuttujien avulla on luontevinta tarkastella. Tarkastelussa on vain yksi kieli, sillä puhekielisyyttä on vaikea tutkia samaan aikaan useasta kielestä. Aineistossa tosin ilmeni koodinvaihtoa muihin kieliin (esimerkki 3 alla), minkä myös Kotilainen (2002) on havainnut olevan vakiintunutta suomalaisissa chat-

keskusteluissa, varsinkin erilaisten fraasien käytössä. Ensimmäisen kohdan suhteen on todettava, että varsinkin Facebook edustaa uusia web-ympäristöjä mainitulla tavalla. Tutkimuksen fokus on kuitenkin niin kapea, ettei esimerkiksi multimodaalisuutta ole tarkoitukseen huomioida, sillä oletuksena tutkittava ilmiö ei riipu siitä. Muutoin keskustelumuodot onkin valittu synkronisemmasta eli oikeaa puhetta lähempänä olevasta päästä. Listan viimeinen kohta voidaan nähdä ongelmana. Taustatekijöitä käytetään hyvin vähän, mutta tutkimuskohteet on valittu osittain juuri helppouden mukaan käyttäen selittävänä tekijänä niiden synkronisuutta, vaikka se ei välttämättä ole ainoa ratkaiseva tekijä.

Esimerkki 3, koodinvaihtoa suomesta englantiin (informantti #C3):

Ei taida tarvita kysyä että "was it as good for you as for me?" :)

Yleensä otsikon lausahduksen jälkeen on saanut todeta että "I'll get mi coat"

Verkossa kielestä puuttuu muun muassa institutionaalista normien säätelyä, jota esimerkiksi painetussa tekstissä on. Verkkokeskustelun tutkimukset ovatkin osoittaneet, että kieli verkossa on hyvin hajanaista ja vaihtelevaa, mikä tekee tutkimuksestakin vaikeampaa. Se tulisi nähdä erilaisina sosiaalisina toimintoina eikä vain teknologiaan purkitettuna puhemuotona. (Androutsopoulos tulossa.)

Variaationtutkimusta yksinään pidetään joiltakin osin yksinkertaistavana ja ongelmallisena metodina jo puheen osalta (Nuolijärvi 2000). Esimerkiksi joidenkin taustamuuttujien merkitys voi olla kiistanalaista, tai niiden perusteella rakennettavia ryhmiä ei voida pitää aitoina sosiaalisina ryhminä. Toisaalta vastakkaisiakin näkemyksiä on: Androutsopoulosin (2007) mukaan variaationtutkimusta menetelmänä ei pidä väheksyä, sillä verkossakin kielivalinnoilla (kuten standardinmukaisuus, puheenomaisuus tai populaarikulttuurien suosima sanasto) luodaan sosiaalista identiteettiä ja tullaan osaksi haluttua kulttuuria.

Nykyään puheen variaationtutkimusta pyritään täydentämään muilla menetelmillä (esim. Lappalainen 2004), ja samaan tulisi pyrkiä verkkokeskustelussakin, kuten

muun muassa Androutsopoulos (2007) ja Siebenhaar (2008) ovat tehneet. Sosiolingvistinen kysymyksenasettelu voi nimittäin verkossa olla hyvin samanlainen, vaikka ympäristö on täysin erilainen (Nuolijärvi 2000). Esimerkiksi kvantitatiivisella pohjatutkimuksella voidaan selvittää suuresta datamassasta, millaista dataa kvalitatiivista tutkimusta varten kannattaa hankkia, sillä kvalitatiivinen lähestymistapa pakottaa keskittymään pieniin ryhmiin tai kokonaisuuksiin, jolloin riski niiden epäedustavuudesta kasvaa (Siebenhaar 2008). Viimeisimpien joukossa verkkokeskustelun tutkimukseen on ehdotettu heteroglossista lähestymistapaa, joka pyrkii huomioimaan juuri sen monet tasot, kuten multimodaalisuuden ja pinnan alla kamppailevat erilaiset sosiaaliset ja kulttuurilliset tekijät (Androutsopoulos tulossa).

Kaikki samat menetelmät eivät silti välttämättä sovi sekä puheeseen että verkkokeskusteluun niiden erilaisesta luonteesta johtuen. Menetelmistä onkin todettu, että perinteisiä kielentutkimuksen menetelmiä on yritetty soveltaa sellaisenaan uuteen verkkokeskustelu-ympäristöön, vaikka siihen olisi järkevämpää kehittää aivan uusia korvaavia menetelmiä (Androutsopoulos tulossa). Menetelmien soveltuvuuserojen selvittäminen onkin yksi uusista tutkimushaasteista.

Eräs huomionarvoinen asia vielä on internetin luonne suhteessa muuhun maailmaan. Internetiä voidaan pitää joko sidoksissa muuhun olevana tai täysin erillisenä maailmana ja tutkimuskohteena (Mann & Stewart 2000, 203–207). Tässä tutkielmassa painottuu ensimmäinen näkökulma, mikä on luontevaa juuri kielentutkimuksessa ja sosiolingvistiikassa, varsinkin jos halutaan tehdä vertailua verkon ulkopuolisiin ilmiöihin tai selittää niillä verkossa näkyvää toimintaa.

4 Analyysi

Tämän empiirisen osion aluksi esitellään käytetty aineisto lähtien sen keräämisperusteista ja päätyen sen kuvailuun. Menetelmäosassa kuvaillaan analysointimenetelmää ja kunkin tutkimushypoteesin ratkaisutapaa. Viimeisessä eli tulososiossa hypoteesit ratkaistaan yksitellen ja pohditaan muita aineistosta ilmeneviä kiinnostavia seikkoja.

4.1 Aineisto

Herring (2004) näkee empiirisessä verkkoviestinnän tutkimuksessa kuusi pääasiallista aineistonvalintaperustetta. Näistä yleisimmin käytetyt ovat ajan ja aineiston teeman mukaan valitseminen. Aikaperuste antaa mahdollisuuden pitkittäistutkimuksiin ja säilyttää sisällöllisen rikkauden. Myös temaattinen valintaperuste mahdollistaa pitkittäistutkimukset jossain määrin, ja lisäksi sillä saadaan analyysiä helpottavaa sisällöllistä koherenssia. Muut neljä perustetta ovat täysi satunnaisuus, tietty ilmiö, tietty yksilö tai ryhmä sekä helppo saatavuus. Tässä tutkimuksessa ensisijaiseksi valintaperusteeksi muodostui keskustelujen teema, joka rajaamalla pyrittiin parantamaan aineiston eri osien vertailukelpoisuutta yhtenäistämällä niitä aihepiirin osalta.

Aineiston keräämisessä Lappalainen (2004, 16–17) pohtii haastattelun ja havainnoinnin eroja. Haastattellessa ryhmä muodostuu tutkijan kutsumana, mutta havainnoinnissa se on luonnollinen sosiaalinen ryhmä. Jälkimmäisessä säilytetään ryhmän aitouden vuoksi sen sisäistä sosiaalista vuorovaikutusta, mikä mahdollistaa varsinkin laadullisen tutkimuksen paremmin. Tosin sosiaalisen ryhmän ollessa pieni tuloksia on vaikeampi yleistää, ja kvalitatiivinen tarkka analyysi estää kovin suurten ryhmien tutkimisen. Lappalainen puhuu osallistuvasta havainnoinnista, jonka rinnalla vaihtoehtona on havainnointi ilman vuorovaikutusta informanttien kanssa, mikä on kenties internetin myötä tullut vasta varteen otettavaksi vaihtoehdoksi. Tällöin toki informaatiota saadaan vähemmän, mutta aineiston kerääminen on huomattavasti kevyempää. Myös Paunonen (1989) toteaa haastattelujen olevan ongelmallinen

aineistonkeruutapa, sillä haastattelutilanne vaikuttaa kielenkäyttöön, jonka nimenomaan haluttaisiin pysyvän luonnollisena. Yhdistelmä osallistuvaa ja ei-osallistuvaa havainnointia oli Arpolla (2005) väitöskirjassaan, jossa hän tutki internetin keskusteluryhmiä. Hän keräsi aineiston aikaisemmin kirjoitetuista keskusteluviesteistä, jolloin hän ei itse vaikuttanut niiden sisältöön, mutta oli tämän jälkeen yhteydessä joihinkin kirjoittajiin epäselväksi jääneiden asioiden selvittämiseksi. Myös ei-osallistuvaa havainnointia verkossa käytetään (esim. van Compernelle 2008).

Tässäkin tutkimuksessa aineisto kerättiin julkisista keskusteluista havainnoimalla ilman vuorovaikutusta keskustelijoiden kanssa. Tämä oli vaivattomin keräystapa sekä takasi objektiivisuuden siinä mielessä, ettei tutkija voi vaikuttaa aineiston tuottamiseen. Selvimpänä haittapuolena monet informanttien taustatekijät jäivät hämäräksi. Perinteisemmässä sosiolingvistiksessä tämä ei olisi ollut hyväksyttävää, mutta tässä tapauksessa taustat olivat vähemmän merkityksellisiä.

Arpo (2005, 102–103) ottaa verkkokeskustelujen analysoinnin apuvälineeksi puheen rekisterit. Jokainen ihminen luonnostaan hallitsee lukemattomia rekistereitä, joista hän käyttää tilanteeseen sopivinta. Arpo käyttää itse kolmitasoista tapaa analysoida keskustelijoiden rekistereitä. Ensimmäisellä ja korkeimmalla tasolla on puheen kenttä eli jokin sosiaalisen elämän alue, kuten politiikka tai perhe. Toisena on puheen moodi alamoodeineen, missä oleellisin ero on puhutun ja kirjoitetun kielen välillä. Alamoodeja esimerkiksi kirjoitetussa kielessä voivat olla sanomalehtikieli ja mainoskieli. Kolmantena puheen tyyli kertoo suhteesta keskustelijoiden välillä. Näitä suhtautumisia ovat muun muassa myötäilevä, huoleton tai hyökkäävä.

Tähän tutkielmaan peilaten kaikki aineisto on pyritty saamaan samalta elämänalueelta, jota voisi luonnehtia yleisesti arkielämäksi. Myös puheen moodi on pyritty saamaan samaksi eli tuttavalliseksi ja rennoksi arkikieleksi. Eniten eroja aineistossa oletettavasti syntyy kolmannella tasolla. Chatissä julkisetkin keskustelut ovat joko kahdenkeskisiä tai rajatun ryhmän sisäisiä, jolloin keskustelijoiden suhteista voidaan puhua, vaikka ne jäisivätkin ulkopuoliselle epäselviksi. Facebookissa oletettu yleisö on viestijän kaverit, jolloin heidän välillään on vähintään jonkinlainen

tuttavuussuhde, mutta lukijakunta on laajempi ja viestejä ei ole usein kohdistettu kenellekään tietyille henkilöille. Keskustelupalstalla tilanne on samankaltainen, mutta nyt lukijana voi olla kuka tahansa, vaikka viestejä viestiketjuissa kohdistetaan Facebookia enemmän tietyille henkilöille. Tämänkaltaiset erot saattavat vaikuttaa puheen rekisteriin (eli puhekielisyteen). Toisin sanoen keskustelun tyyppi vaikuttanee ainakin jollakin tavoin puhekielisyden asteeseen, kuten hypoteesissa H4 oletetaan. Tässä Arpon esittämän jaottelun perusteella joka tapauksessa oletetaan, että rekisterin kaksi ensimmäistä tasoa ovat keskustelutyyppeiden välillä karkeasti ottaen samat.

Tavoitteena oli saada aineisto kerättyä vaivattomasti kuitenkin siten, että se olisi mahdollisimman kattava tutkimuskysymysten ratkaisemiseksi. Keskustelupalstoista ihmissuhteisiin ja nuorisotrendeihin painottuva City.fi tarjosi mahdollisuuden selata viestejä käyttäjäkohtaisesti, mikäli käyttäjä sen sallii, ja siksi se valittiin. Siellä myös käyttäjät saavat halutessaan ilmoittaa asuinpaikkansa. Koska helsinkiläisiä ei helposti löytynyt riittävästi, mukaan otettiin myös espoolaiset ja vantaalaiset.

Facebookissa luontevin aineisto ovat niin sanotut seinäkirjoitukset, eli pääasiassa tilapäivitykset, joissa henkilö voi jakaa tuoreita tuntojaan, ja niihin kirjoitetut kommentit. Tyypillistä on, että yhteen tilapäivitykseen tulee kommentteja käyttäjän kavereilta, joihin käyttäjä itse vastaa vielä omalla kommentilla. Facebookissa näkyvien tietojen yksityisyys on ollut laajan keskustelun kohteena, mutta osa käyttäjistä antaa edelleen kenen tahansa selata omaa profiiliaan vapaasti. Aineistoksi otettiin tällaisia henkilöitä, jotka olivat ilmoittaneet kotipaikaksi Helsingin, mitä käytettiin hakuehtona.

Chateistä hyvin suosittuja ovat pitkään olleet web-selaimessa toimivat keskustelukanavat, joita on useimmissa suurissa verkkoportaaleissa, esimerkiksi suomi24.fi ja City.fi-palveluissa. Niissä keskustelujen tallentaminen on kuitenkin työlästä, joten aineisto kerättiin IRC-keskusteluista. IRC on vastaavanlainen reaaliaikainen, mutta paljon vanhempi keskustelujärjestelmä. Se vaatii yleensä erillisen ohjelman, mutta kyseisissä ohjelmissa on tavallisesti sisäänrakennettu mahdollisuus keskustelujen tallentamisen.

IRC:ssä on keskustelukanavia asuinpaikan mukaan eri kaupungeille. #helsinki-kanavalla ei kanavan ohjeiden mukaan sallittu keskustelujen tallentamista muuhun kuin yksityiseen käyttöön, joten kanavista valittiin #espoo. Tällä kanavalla keskustelua tallennettiin, kunnes sopivalta määrältä keskustelijoita oli saatu kerättyä riittävä määrä sanoja. Vain julkisia viestejä kerättiin.

Aineistoksi valikoitui IRC:ssä keskustelijoita, jotka sattuivat aineiston keräämisen aikana keskustelemaan kanavilla aktiivisimmin. Keskustelupalstan ja Facebookin kohdalla sen sijaan vanhoja viestejä pystyi selaamaan vuosia taaksepäin, ja niistä kertyi viestejä vuosilta 2007–2011. Ei ole syytä olettaa, että yksilön tai yhteisön kielenkäyttö muuttuisi näin lyhyessä ajassa merkittävästi, joten ajanjaksojen vaihtelusta ei ole haittaa.

Aineistoon valittiin kriteerit täyttävien henkilöiden keskusteluista satunnaisesti, kymmenen henkilön viestit kustakin keskustelutyypistä. Keskustelupalstan ja Facebookin kymmenestä henkilöstä viisi oli miehiä ja viisi naisia. IRC:ssä kaikkien henkilöiden sukupuolesta ei ole varmuutta, ja naisia informanteista oli kahdesta neljään. Vaikka informanttien valinnassa pyrittiin satunnaisuuteen, mikä on tärkeää erityisesti tilastollisten menetelmien kannalta, ei tästä ole varmuutta, sillä valinnanvaraa henkilöissä oli hyvin vähän. On mahdollista, että julkisesti viestejään nähtäväksi asettavia henkilöitä yhdistää jokin tekijä, mikä ei koske keskimääräistä internetin käyttäjää.

Sukupuolten välillä tiedetään olevan eroa kielenkäytössä puheen lisäksi myös verkkokeskustelussa (ks. esim. Herring 2001b; Huffaker & Calvert 2005; Mann & Stewart 2000, 167–168), mutta puhekielisyyden osalta tästä ei ole tietoa. Suomalaisen puheessa puolestaan tiedetään olevan käytettyjen puhekielisten piirteiden yleisyydessä eroa miesten ja naisten välillä (ks. esim. Paunonen 2006, 70–90; Lappalainen 2004, 37–39). Niinpä lähtökohdaksi tässä tutkimuksessa otetaan suomen puhekielentutkimuksen erot, eli sukupuoli voi olla merkittävä taustatekijä myös verkkokeskustelun puhekielisyydessä. Siksi molempia sukupuolia pyrittiin saamaan aineistoon tasapuolisesti. Sukupuoli onkin ainoa taustatekijä, joka edes melkein kaikista informanteista tiedetään.

Kaupunkikielen tutkimuksissa yleensä käytetty noin sataa informanttia (Paunonen 2006, 64–65), ja Paunonen itse käytti tutkimuksissaan 96:ta ja 45:tä, Nuolijärvi (1986) 48:aa ja Lappalainen (2004) 16:ta, vaikka viimeksi mainittu käyttikin tukena kvalitatiivista analyysiä. Niinpä tämän tutkimuksen 30 kuulostaa alimitoitetulta, mutta on huomioitava, että muissa tutkimuksissa henkilöitä on tarkasteltu pienempinä ryhminä useampien taustamuuttujien mukaan, jolloin yksittäisen tällaisen ryhmän kokona kymmenen henkilöä vastaa jo karkeasti aikaisempia tutkimuksia.

Aineiston määrän voi valita joko ajanjakson tai sanamäärän mukaan. Puhetta tutkiessa on päädytty usein ajanjaksoon (Lappalainen 2004, 55–56), jolloin henkilöiden analysoitavat sanamäärät poikkeavat toisistaan. Tässä tutkimuksessa oli luontevinta pyrkiä samoihin sanamääriin, koska varsinkin ei-reaaliaikaisessa keskustelussa aika menettää merkitystään. Vaihtelua tuli siitä syystä, että joidenkin henkilöiden kirjoittamat viestit kokonaisuudessaan eivät käsittäneet enempää sanoja ja toisaalta ei ollut mielekasta rajoittaa määrää pienimmän mukaan. Lappalainen käytti omassa tutkimuksessaan (mts. 55) informantikohtaisina sanamäärinä pääosin 1100–1700 sanaa arkikeskustelua analysoidessaan, mutta Nuolijärvi (1986) selvästi suurempia määriä, noin 2000–7000. Tässä hieman pienemmät määrät ovat useimpien tarkasteltavien piirteiden kannalta riittäviä. Suuremmassa aineistossa muutamat harvinaiset piirteet olisivat saattaneet esiintyä useammin, jolloin niistä saatu kuva olisi ollut luotettavampi. Taulukkoon 2 on koottu tietoja kerätystä aineistosta.

Taulukko 2. Tietoja tutkimusaineistosta.

Keskustelutyyppe	Aineistolähde	keskustelijoita yhteensä	sukupuoli-jakauma	aineiston koko keskustelijaa kohti (sanoina)
IRC (chat)	#espoo-IRC-kanava	10	6–8 miestä, 2–4 naista	900–1300
Facebookin tilapäivitykset	henkilön omat tilapäivitykset ja niiden kommentit	10	5 miestä, 5 naista	800–1600
Keskustelupalsta	kaikkiin keskusteluaiheisiin kirjoitetut viestit ja vastausviestit	10	5 miestä, 5 naista	1000–2000

Vaikka voidaan pitää mahdollisena, että verkossa käytetty kieli on jotakuinkin yhdenmukaista koko Suomessa, yhdenmukaisuutta pyrittiin varmistamaan valitsemalla henkilöt Helsingin seudulta. Osa on viestien sisällöstä päätellen muuttanut sinne muualta Suomesta, mutta kaikki ovat varmasti ainakin ehtineet olla kosketuksissa pääkaupunkiseudun puhekielen kanssa ja omaksua siitä mahdollisesti joitakin aineksia omaan kielenkäyttöön. Osa on puolestaan hyvin todennäköisesti syntyperäisiä pääkaupunkiseutulaisia jo valmiiksi. Asuinpaikan lisäksi kieltä voidaan omaksua nykypäivänä myös mediasta, erityisesti juuri internetistä, mikä lisää kosketuspintaa pääkaupunkiseudun puhekieleenkin. Aineistossa esiintyi muutaman kerran murteellisia *mie*-muotoja, mutta näilläkin henkilöillä helsinkiläiset piirteet olivat hallitsevia, joten kyse lienee satunnaisesta kotiseutuun viittaavasta koodinvaihdosta. Muuten mikään ei viitannut murteisiin.

Koska ikä on ollut tärkein taustamuuttuja sosiolingvistisessä tutkimuksessa, olisi ollut hyvä tietää henkilöiden iät. Niissä esiintyi kuitenkin runsaasti vaihtelua siltä osin, kun henkilöt olivat sen julkisesti ilmoittaneet. Vaihtelua oli noin 20:stä 50:een ikävuoteen, painotuksen ollessa selvästi nuorena päässä. Iän merkitys taustatekijänä korostuisi, jos vastaavia tutkimuksia olisi tehty ennen, ja tuloksia haluttaisiin verrata niihin (Hurtta 1999, 69), mutta tässä tapauksessa se ei ole oleellinen. IRC:n osalta kenenkään ikä ei ollut tiedossa, joten siellä on saattanut olla myös teini-ikäisiä. Keskustelujen sisällöistä päätellen iät liikkuvat noin 20:n ja 30:n välillä.

Aihepiireissä oli keskustelutyypin välillä selvästi yhteistä, mutta ne myös vaihtelivat odotettua enemmän. Kaikki olivat luonteeltaan sellaisia, että niihin kirjoitettiin enimmäkseen hetken mielihoiteesta tai oman vapaa-ajan viihteeksi, päätellen viestien sisällöstä. Näin ne täyttävät ainakin pääosin vaatimuksen keskustelujen arkisesta luonteesta. Viestien keskimääräiset pituudet vaihtelivat odotetusti siten, että chatissä viestit olivat hyvin lyhyitä ja keskustelupalstalla toisinaan useita tekstikappaleita käsittäviä. Taulukossa 3 kuvataan viestien pituutta ja aiheita.

Taulukko 3. Luonnehdintaa keskustelutyypin viesteistä.

Keskustelutyyppi	Viestien tyypillinen pituus	Viestien tyypillisiä aiheita
IRC (chat)	lyhyt, paljon yhden tai kahden sanan viestejä	arkiset tekemiset, ajankohtaiset asiat
Facebookin tilapäivitykset	melko lyhyt, yleensä yksi lause	arkiset tekemiset, työ, harrastukset, omat lapset
Keskustelupalsta	suhteellisen pitkä, usein pari kappaletta tekstiä	arkiset tekemiset, ihmissuhteet, seksi

Koska henkilöiden anonymiteetti on keskeinen eettinen kysymys tutkimuksessa, kaikkien henkilöiden nimet muutettiin koodeiksi siten, että kirjain kertoo keskustelutyyppin ja numero henkilön järjestysnumeron kyseisessä aineistonosassa, esimerkiksi #F4 on Facebook-aineiston neljäs informantti. Myöskään nimimerkkejä ei säilytetty, ja keskusteluissa esiintyvät nimet on muutettu, joten henkilöitä on käytännössä mahdoton tunnistaa aineistosta.

4.2 Tutkimusmenetelmä

Herring (2004) on pohtinut verkkokeskustelun yleistä tutkimusmetodiikkaa ja yrittänyt siten tuoda selkeyttä muuten hajanaiseen tutkimuskenttään. Hänen mukaansa tutkimuksen kannalta samat lähtökohdat pätevät internetin keskusteluissa ja diskursseissa kuin sen ulkopuolellakin, mutta sen lisäksi teknisillä ominaisuuksilla on yleensä, mutta ei aivan välttämättä, vaikutusta. Empiirisen tutkimuksen selvitettäväksi jää, missä määrin tällaista vaikutusta esiintyy. Tässä tutkimuksessa pyritään selittämään teknisistä tekijöistä ensisijaisesti synkronisuudella verkkokeskustelujen keskinäistä kielellistä vaihtelua, mutta paremman luokittelun aikaansaamiseksi myös muilla ominaisuuksilla.

Tutkimusmenetelmä on lähtökohtaisesti kvantitatiivinen, kuten variaationtutkimuksessa yleisestikin. Päävertailukohta on Paunosen 2006 julkaistu tutkimus, ja siksi tutkittavat piirteet valittiin samoiksi. Muihin tutkimuksiin vertaileminen samanaikaisesti on hankalaa, koska niissä on käytetty erilaisia ja eri

tavoin ryhmiteltyjä piirreluetteloita. Paunosenkaan tutkimuksen kanssa ei voida tehdä tilastollista vertailua, koska populaatiot eroavat melko paljon toisistaan, eikä Paunosen käyttämä data ollut käytössä, vaan vain hänen julkaisemansa lopputulokset. Paunosen tuloksia voidaan verrata kuitenkin suuntaa-antavasti. Erityisesti hänen lähes kaksikymmentä vuotta vanhaan aineistoon perustuvien ja kirjakielisempään väestönosaan kohdistuneiden tulostensa voidaan olettaa edustavan nyky-Helsingin puhekielisyyden minimitasoa, sillä puhekielisyyden tiedetään sekä olevan yleisempää nuorempien keskuudessa että kasvavan ajan myötä. Koska analysoitavat verkkokeskustelut on puolestaan valittu mahdollisimman puhekieltä suosivasta päästä, riittää tämä ensimmäisen hypoteesin H1 selvittämiseksi, eli jääkö verkkokeskustelun puhekielisyyys alle varsinaisen puheen puhekielisyydystason. Tosin myös verkkokeskustelun epämuodollisuuden on todettu kasvavan ajan myötä (Herring 1998), mikä voi tarkoittaa juuri puhekielisyyden lisääntymistä.

Hypoteesi H2 ratkaistaan vertaamalla Paunosen (2006, 90) saamaa piirteiden puhekielisyyshierarkiaa tässä tutkimuksessa saatuun kaikkien keskustelutyypin yhdistettyyn hierarkiaan. Tilastollinen vertailu ei tässäkään ole mahdollinen. Tuloksissa on huomioitava, että Paunosen tutkimus alkaa olla lähes kaksikymmentä vuotta vanhan aineiston vuoksi vanhentunut. Yksi kysymys nousee siitä, vaihtuuko piirteiden hierarkia ajan kuluessa. Verrattuna Paunosen aikaisempaan 70-luvun tutkimukseen (mts. 88–90) huomataan, että hierarkia on pysynyt karkeasti ottaen samankaltaisena, ja muutkin tutkimukset (esim. Lappalainen 2004) tukevat sitä, ettei muutosta kovin nopeasti tapahdu.

Hypoteesiin H3 liittyen on tavoitteena jakaa henkilöt ryhmiin heidän kielenkäyttötapansa mukaan. Tämä tapahtuu yksinkertaisimmillaan katsomalla, muodostuuko heistä selkeitä ryhmiä heille lasketun puhekielisyydsasteensa mukaan. Tämän lisäksi voidaan tarpeen mukaan käyttää esimerkiksi matemaattisia menetelmiä tarkempien erojen selvittämiseksi.

Viimeinen tutkimushypoteesi H4 voidaan analysoida yksinkertaisin tilastollisin menetelmin (ANOVA ja parittaiset t-testit) selvittämällä, poikkeavatko

keskustelutyypin puhekielisyyskeskiarvot merkitsevästi toisistaan ja oletettuun suuntaan.

Yksi huomionarvoinen yksityiskohta menetelmässä on prosentuaalisten keskiarvojen laskeminen. Yksittäisen henkilön yksittäisen piirteen prosenttiosuus voidaan laskea puhekielisten varianttien osuutena kaikista kyseisen piirteen ilmenemistapauksista. Kun halutaan tietää kyseisen henkilön kaikkien piirteiden vaikutus eli yleinen puhekielisyysaste, laskemistapoja on kaksi: joko lasketaan kaikki puhekieliset ilmenemistapaukset yhteen ja lasketaan prosentti jälleen aivan kaikista tapauksista (painotettu keskiarvo), tai sitten lasketaan (painottamaton) keskiarvo aiemmin lasketuista prosenteista. Nuolijärvi (1986) on laskenut keskiarvot yksittäisten frekvenssien summasta (siis painotetut keskiarvot), myös kaikkien yhdistetyn keskiarvon. Paunonen (2006) näyttää laskeneen yhdistetyn puhekielisyysasteen yksittäisten piirteiden (painottamattomana) keskiarvona. Aikaisemmassa tutkimuksessaan (1995) hän on vältellyt keskiarvojen laskemista keskiarvoista kokonaan.

Tässä tutkimuksessa kaikkien informanttien data katsotaan samanarvoiseksi. Ei ole mielekästä, että informantti, jolta on dataa 2000 sanaa saisi kaksinkertaisen painoarvon 1000 sanan informanttiin nähden. Vastaavasti kaikki tutkittavat piirteet katsotaan samanarvoisiksi. Koska niiden ilmenemismäärissä on valtavia eroja, nolosta useisiin kymmeneen henkilöä kohden, tulisi toisille piirteille muuten moninkertainen painoarvo. Tällaista valintaa voisi puolustaa sillä, että yksikin hyvin usein käytetty variantti voinee leimata puhujaa vahvasti yleis- tai puhekieliseksi. Tällä kertaa lähtökohtana kuitenkin on, että kokonaiskuva tulee kaikista piirteistä yhdessä ja että piirteet on valittu edustavasti siten, että kaikki ovat tärkeitä. Niinpä tässä keskiarvojen keskiarvot lasketaan aina ilman painotuksia.

Kokonaispuhekielisyysaste voi laskea keskiarvona joko piirteiden tai henkilöiden mukaan. Kysymys lienee siitä, pitääkö henkilöitä vai piirteitä edustavampina. Tässä tutkimuksessa käytettiin tutkimushypoteesista riippuen molempia laskutapoja. Laskutavan vaikutus tuloksiin vaihtelisi tässä tutkimuksessa 0–3 %-yksikön välillä, mutta ei vaikuttaisi hypoteesien hyväksymiseen.

4.3 Tulokset

Taulukosta 4 nähdään tarkasteltujen keskustelutyyppeiden puhekielisyyden aste. Liitteessä 1 on esitetty piirrekohtaiset puhekielisyydsasteet, joista nämä keskiarvot on laskettu. Pyrkimyksenä oli selvittää, jääkö puhekielisyyden aste kokonaisuutena puheen puhekielisyyden asteen alapuolelle (H1). Keskustelutyyppeiden keskiarvoksi saatiin 43 %. Kun tätä verrataan Paunosen (2006, 90) saamaan 65 %:iin, huomataan eron olevan varsin suuri. Myös verkkokeskustelun ääripäätä edustavaan IRC:hen on eroa useampi prosenttiyksikkö. Vaikka aineistot eivät ole suoraan vertailukelpoisia, edustaa Paunosen arvo puhekielen minimitasoa ja tutkimukseen valitut keskustelutyyppit verkkokeskustelun puhekielisyyden oletettua maksimipäätä. Puhutun kielen puhekielisyydsaste lienee siis todellisuudessa selvästi yli 65 %. Tällä perusteella hypoteesi H1 voidaan hyväksyä.

Taulukko 4. Verkkokeskustelutyyppeiden puhekielisyydet.

Keskustelutyyppi	IRC	Facebook	Keskustelupalsta	Yhteensä	Puhe (Paunonen 2006, 90)
Puhekielisyyds	61 %	45 %	23 %	43 %	65 %

Seuraavassa on lähellä keskiarvoa oleva näyte kustakin kolmesta keskustelutyypistä erojen havainnollistamiseksi. Tarkastelun alaiset puhekieliset ilmaisut on lihavoitu ja yleiskieliset kursivoitu.

Esimerkki 4, näyte IRC-viesteistä, puhekielisyyds 66 % (informantti #14, muiden keskustelijoiden viestit poistettu):

<#I4> Oli *tuossa Hansatiellä Expressbus ojassa*, melkeen kyljellään.

<#I4> ei mitää *hajua*

<#I4> se oli sillo viiden aikaan siinä

<#I4> siinä koulun *kohdalla*, tai sen mäen alla

<#I4> siis radan ja tien *välissä*

<#I4> jaa se oliki 18

<#I4> se dösä, joka **kaatu**

<#I4> vähä mietinki et mitä expressbus **ois kökkelis tehny**
<#I4> 165:ssa oli ainakin **yks** aamu/iltapäivä
<#I4> aamudösät varsinki on iha jäätä
<#I4> tosin 165 on aika **lämpönen ollu** aina
<#I4> mut 18.t on sellaisia koppeja *joissa* on pyörät ja penkit
<#I4> ni ei ihme et on *jäätäviä*, 213:lla en **oo** kovin **useesti**
ajanut, mutta **eikö** nekin **oo** ihan "oikeita" busseja silti

IRC-otteesta huomataan, että puhekieliset ja yleiskieliset variantit ovat molemmat yleisiä riippuen piirteestä. Katkelmasta saa melko puhekielisen kuvan muutenkin kuin tutkittujen piirteiden osalta. Facebook-aineisto oli melko kaksijakoinen, kuten myöhemmin huomataan. Seuraava ote on yleiskielisemmän ryhmän puhekielisimmän informantin tilapäivityksistä ja kommentaista muiden vastauksiin. Selvästi edellistä yleiskielisempää vaikutelmaa luovat varianttien lisäksi myös isojen alkukirjainten ja välimerkkien käyttö, vaikka hymiöt pistävätkin silmiin.

Esimerkki 5, näyte Facebook-viesteistä, puhekielisyys 33 % (informantti #F9):

Nyt äkkiä **hakee** ranskalaista *herkkujuustoa* narinkkatorilta! ;) *itselle tarttui* 4 erilaista *juustoa*.. =P

Onko **susta** tullut itse Master Chef? :]

Tässä on ainakin **yks** sielu joka haluaa painaa dislike-*nappia* pakkasen ja viiman kohdalle.

Jenkit **kusee** omaansa ja Lähi-idän heebot **kusee** toisten kuppeihin. *Näyttäisi* vain *Aasialaisilla* olevan jonkin asteista *pidätyskykyä*..

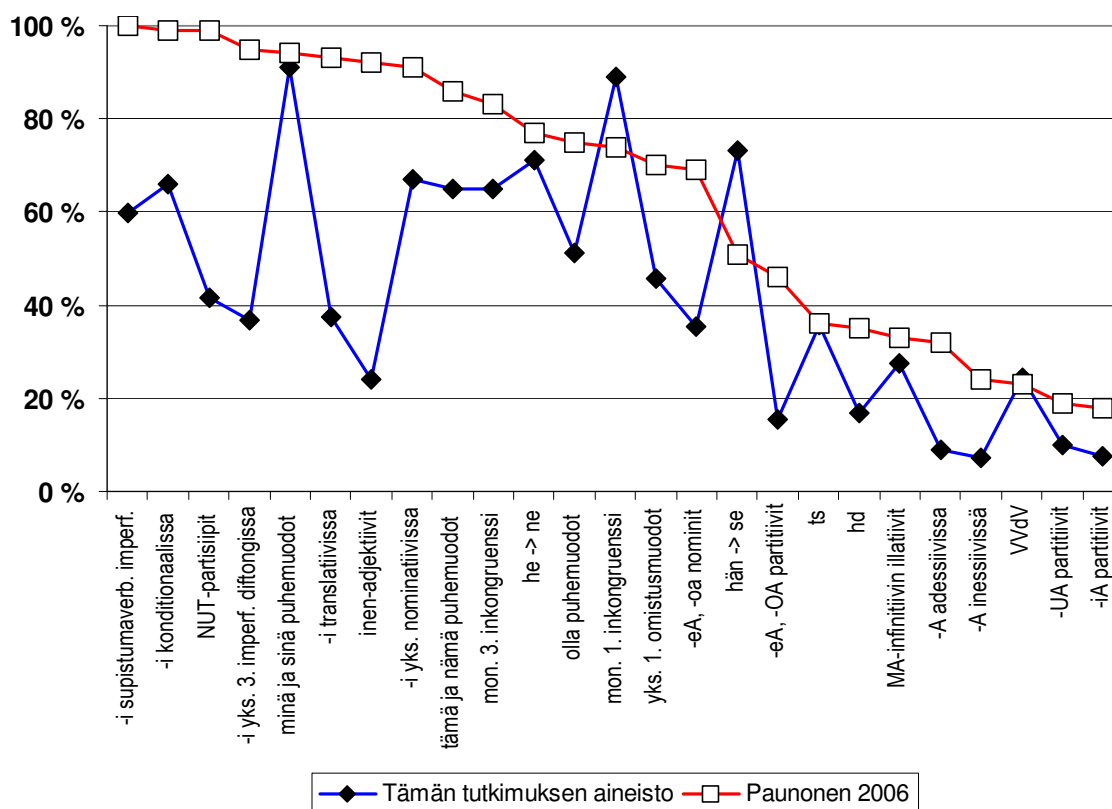
Kolmantena on City.fi-aineistosta poimittu yksi viesti. Viestin pituusero edellisiin tulee selvästi näkyviin. Alussa olevasta naurahduksesta, muutamista puhekielisistä ilmaisuista, huutomerkeistä ja yksittäisestä hymiöistä huolimatta ero Facebook-esimerkkiin on nähtävissä. Pituuden lisäksi muun muassa omistusliitteiden ja kongruenssin säilyttäminen saavat sen vaikuttamaan rennon lisäksi asialliselta.

Esimerkki 6, näyte City.fi-keskustelupalstan viesteistä #C6, puhekielisyys 24 % (informantti #C6):

Hahaha, **mä** just näin yhden tapauksen sattumalta *baarissa 2 viikkoa* sitten ja **katoin** sitä hetken epäuskoisena, ensimmäiset ajatukset *vilistivät päässäni*: idiootin ilmeet, **kauheet** hiukset ja *junttimeininki*. Miten **mä oon** ikinä **voinu** olla *tuohon ihastunut!* Enkä *ole* edes kovin turhamainen. :D Hyvät naurut sain kyllä! Aika monesti tulee ihmeteltyä jälkikäteen *ihastuksia, pitkäaikainen exäni* saa kyllä *kredittiä* ansioistaan, *tiedän* tasan tarkkaan miksi *olimme yhdessä vuosia* ja myös tasan tarkkaan miksi *erosimme*. Hyvä niin.

Toisena hypoteesina (H2) oli selvittää, ovatko samojen piirteiden puhekieliset variantit yleisiä sekä puheessa että verkkokeskustelussa. Tätä varten piirteet on koottu puheen mukaiseen yleisyysjärjestykseen kuvassa 2, jossa toiset pisteet edustavat Paunosen (2006, 90) tutkimustuloksia (kuten kuvassa 1 aiemmin) ja toiset kaikkien tutkittujen verkkokeskustelutyypin keskiarvoa. Vaikka samanlainen laskeva suunta on nähtävissä molemmissa, pomppivat verkkokeskustelun piirteet huomattavasti verrattuna puheeseen.

Kuva 2. Puheen ja verkkokeskustelun piirteiden puhekielisten varianttien yleisyys.



Erojen vertaamisen helpottamiseksi voidaan piirteet ensin ryhmitellä niiden puhekielisten varianttien yleisyyden mukaan joko vahvasti puhekielisiin, pääasiassa yleiskielisiin tai näiden väliin jääviin varioiviin piirteisiin. Tällaisessa ryhmittelyssä Lappalainen (2004, 67) veti rajan noin 85 %:n puhekielisyyssasteeseen puhekielisten ja varioivien piirteiden välille, ja selvästi yleiskielisiä hänellä ei ollut ollenkaan. Koska verkkokeskustelussa puhekielisyyys todettiin kokonaisuutena vähäisemmäksi, voidaan rajatkin asettaa alemmaksi. Katsomalla yksittäisten informanttien viestejä voidaan arvioida, että noin 70 %:n puhekielisyyssaste antaa vielä vahvasti puhekielisen vaikutelman. Toisaalta verkkokeskustelussa ilmeni myös useimmiten yleiskielisinä pysyviä piirteitä, joiden raja voidaan vastaavasti asettaa 15 %:iin. Näin piirteet jakautuvat kolmeen luokkaan taulukon 5 mukaisesti.

Taulukko 5. Tutkitut kielenpiirteet niiden puhekielisyyden mukaan ryhmiteltynä.

Puhekieliset piirteet	Varioivat piirteet	Yleiskieliset piirteet
minä, sinä → mä, sä he → ne mon. 1. persoonan kongruenssi hän → se	-i supistumaverbien imperfektissä -i konditionaalissa -NUT partisiipit -i yks. 3. persoonan imperfektin diftongissa -i translatiivissa -inen adjektiivit -i yks. nominatiivissa tämä, nämä → tää, nää mon. 3. persoonan kongruenssi olla pikamuodot yks. 1. persoonan omistusmuodot -eA, -oa nominit -eA, -OA partitiivit ts:n vastineet hd:n vastineet MA-infinitiivin illatiivit VVdV-muotoiset d:n vastineet	-A adessiivissa -A inessiivissä -UA partitiivit -iA partitiivit

Muutama piirre erottuu selvästi puhekielisinä, eivätkä nämä vastaa aikaisempien tutkimusten puhekielisimpiä piirteitä. Koska suomenkielinen internet kattaa koko maan, on luontevaa etsiä selitystä piirteiden mahdollisesta yleissuomalaisesta luonteesta. Kuten luvussa 2.3.2 esiteltiin, yleissuomalaisiksi piirteiksi voidaan nimetä mon. 1. ja 3. persoonan inkongruenssit, *se* ja *ne* käyttäminen 3. persoonassa *hän* ja *he* sijaan sekä i:n kato inen-adjektiiveissa, supistumaverbien imperfekteissä ja yks. 3. persoonan imperfektin loppudiftongeissa. Tässä aineistossa puhekielisimmät olivat

yllä olevan taulukon mukaan *he* → *ne*, mon. 1. persoonan inkongruenssi sekä *hän* → *se*, jotka kaikki ovat yleissuomalaisen piirteiden listalla. Tosin kuten myöhemmin nähdään, keskustelupalstalla *hän* ja *he* eivät korvautuneet pronomineilla *se* ja *ne* kovin usein. Neljäntenä puhekielisenä piirteenä oli vielä yks. 1. ja 2. persoonan pronomien pikamuodot.

Toisaalta jos katsotaan tuloksista muita yleissuomalaisia piirteitä, niistä osa on huomattavan harvoin puhekielisiä, nimittäin kaikki mainitut i:n katotapaukset. Näistäkin on se poikkeus todettava, että IRC:ssä i:n katotapausten puhekieliset variantit olivat toisista poiketen yleisiä. Ei kuitenkaan voida todeta, että piirteiden yleissuomalaisuus yksinään määräisi niiden puhekielisyyttä verkossa.

Sen sijaan yleissuomalaisuuden voidaan jollain tavalla tulkita vaikuttavan, koska kolme neljästä vahvasti puhekielisestä piirteestä kuului tähän joukkoon. Neljäs piirre, *minä*, *sinä* → *mä*, *sä* on luonteeltaan sellainen, että kyseiset pronominit korvautuvat murteissakin yleensä jollain variantilla, eli *minä* antaa helposti yleiskielisen leiman puhujalle. Oleellisempaa saattaa siis olla korvata se millä tahansa puhekieliselä variantilla kuin jättää yleiskieliseksi.

Taulukon neljä yleiskielistä piirrettä olivat myös Paunosen tutkimuksessa viiden yleiskielisimmän joukossa. Tämä viittaa siihen, että edellä tehdystä huomiosta huolimatta piirteiden puhekielisyyssuuntaus on jossain määrin samanlainen verkossa kuin puheessakin.

Kokonaisuutena voidaan tehdä tulkinta, että verkossa suositaan jonkinlaista ”kevyttä puhekielisyyttä”, joka on osin luonteeltaan erilainen kuin puhe. Tärkein ominaisuus näyttäisi olevan, että leimallisesti yleiskielistä vaikutelmaa ei haluta antaa, mikä syntyisi juuri muutaman ratkaisevan kielenpiirteen kautta. Loput yleissuomalaiset piirteet saattavat olla luonteeltaan vähemmän leimallisia, vaikka ne puheessa säännönmukaisesti esiintyisivätkin. Tärkeä päätelmä tästä kuitenkin on, että verkkokeskustelu ei puhekielisyydeltään ole välttämättä niin selvästi pääkaupunkiseutulaista, kuin olisi voinut odottaa.

Yleissuomalaisen piirteiden yleisyydessä voidaan nähdä kiinnostava yhtäläisyys

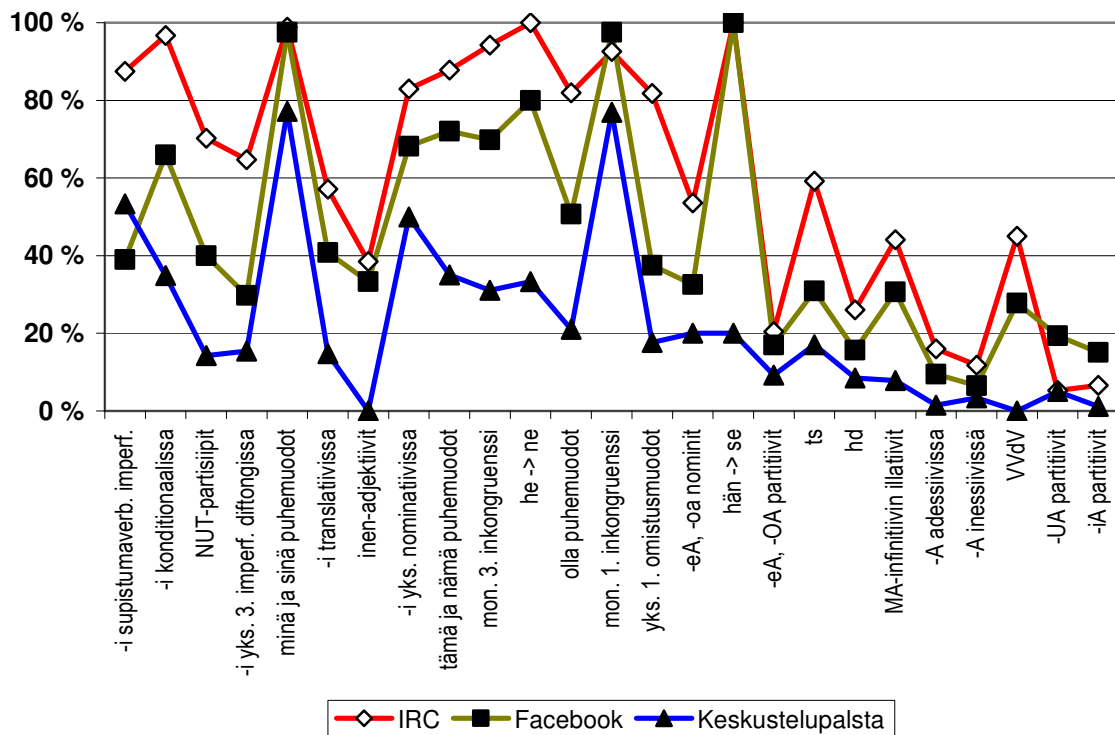
Nuolijärven tutkimukseen (1986). Siinä missä Paunonen on tutkinut syntyperäisiä helsinkiläisiä, Nuolijärvi selvitti väitöskirjassaan muualta Helsinkiin muuttaneiden kielen omaksumista. Myös hän totesi, että mahdollisimman laajalevikkiset puhekielen piirteet omaksutaan helpoimmin (sekä tutkittujen pohjalais- että savolaisperäisten informanttien keskuudessa) (mts. 208–210). Tämän tutkimuksen ulkopuolelle jää kuitenkin pohtia, voidaanko pääkaupunkiseudun puhekielen omaksumisesta verkossa ja Helsinkiin muuttaneiden keskuudessa tehdä näin yhtäläisiä päätelmiä.

Kuvasta 2 edellä huomataan lisäksi, että kaksi piirrettä (mon. 1. persoonan kongruenssi ja pronominin 'se' käyttäminen henkilöstä yks. 3. persoonassa) on selvästi korkeammalla kuin Paunosen vertailuarvo. Jos verkkokeskustelu on vähemmän puhekielistä kuin puhe, ei yhdenkään piirteen kuuluisi selvästi ylittyä tässä aineistossa verrattuna Paunosen tutkimukseen. Selitys voisi löytyä aineistojen erilaisuudesta, mutta luontevampaa on tulkita tämän osoittavan, että kieli on yleisesti puhekielistynyt, kuten trendin tiedetään olevan. Niinpä tänä päivänä tehtävä tutkimus puheesta luultavasti jättäisi nämäkin piirteet alleen kuvaajassa. Koska Paunosen luvut edustavat puhekielisyyden minimimäärää puheessa, niiden voi olettaakin ylittyvän tässä tutkimuksessa.

Kaikkien keskustelutyyppeiden puhekielisyydsasteiden keskiarvo ei kuitenkaan kerro koko kuvaa, sillä joissakin piirteissä niiden välillä on huomattavia eroja, kuten kuvasta 3 nähdään. Pääosin puhekielisyyds on kaikissa samansuuntaista, mutta eroa on erityisesti i:n kadossa konditionaalissa, mon. 3. persoonan kongruenssissa, *se* ja *ne* käyttämisessä sanojen *hän* ja *he* sijaan, olla-verbin pikamuodoissa sekä yks. 1. persoonan omistusmuodoissa. Näissä ero puhekielisimmän (IRC:n) ja yleiskielisimmän (keskustelupalstan) välillä on yli 60 %-yksikköä eli huomattavasti enemmän kuin näiden keskimääräinen ero 38 %-yksikköä. Lisäksi havaitaan, että keskustelupalstalla tietyt piirteet esiintyvät lähes yksinomaan yleiskielisten, mitä kahdessa muussa ei ilmene oikeastaan ollenkaan. Nämä ovat -inen-päätteiset adjektiivit, -A:n säilyminen adessiivissa ja inessiivissä, d:n säilyminen vokaalien ympärillä sekä -UA ja -iA-päätteiset partitiivit. Nämä kertovat siitä, että keskustelutyypeissä saattaa olla myös keskinäisiä eroja puhekielisten piirteiden

suhteen.

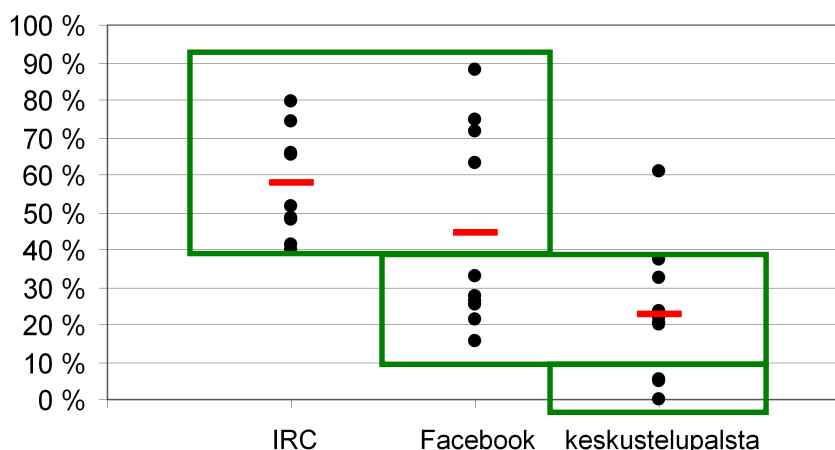
Kuva 3. Eri verkkokeskustelutyypin piirteiden puhekielisten varianttien yleisyys.



Kaiken kaikkiaan verkkokeskustelussa esiintyvien yksittäisten piirteiden puhekielisyydestä on vedettävä johtopäätös, että puheen ja verkkokeskustelun välillä on joitakin selviä eroja, jotka näyttävät ilmenevän verkkokeskustelun tyypistä riippumatta. Hypoteesia H2 niiden yhdenmukaisuudesta ei siis voida vahvistaa.

Kolmantena hypoteesina (H3) oli, että keskustelijoita voidaan jakaa selvärajaisiin ryhmiin heidän kielenkäyttönsä mukaan. Kuvaan 4 on merkitty kunkin informantin puhekielisyyssaste pisteenä keskustelutyypin mukaan eriteltynä. Liitteenä 2 on keskustelutyypeittäin kaikkien informanttien piirrekohtaiset puhekielisyyssasteet, joista nämä on laskettu keskiarvona. Kuvassa poikkeviivat edustavat koko keskustelutyypin keskiarvoa, aikaisemmasta poiketen nyt henkilöiden eikä piirteiden keskiarvona. Huomataan, että kaikki IRC-keskustelijat ovat melko puhekielisiä ja vastaavasti yhtä lukuun ottamatta kaikki keskustelupalstan keskustelijat vain lievästi puhekielisiä, jos ollenkaan. Facebookin kohdalla sen sijaan henkilöt jakautuvat selvästi kahtia jompaankumpaan ryhmään edellisten kanssa.

Kuva 4. Yksittäisten henkilöiden puhekielisydet keskustelutyypin mukaan.



Kuvan perusteella henkilöt voidaan jakaa melko selvästi kolmeen johdannossa ehdotettuun ryhmään, eli (1) yleiskielisiin, (2) jonkin verran puhekielisiä variantteja käyttäviin ja (3) monipuolisesti pääkaupunkiseudun puhekielisyttä ilmentäviin. Tällöin rajat voidaan vetää suoraan puhekielisyysasteen mukaan 40 %:iin ja 10 %:iin (laatikot kuvassa). Nyt huomataan, että puhekielisyys kasvaa selkeissä portaissa keskustelutyypin mukaan. On kuitenkin syytä huomata, että kolmenkymmenen informantin joukko on melko pieni tällaista ryhmittelyä varten. Suurempi otos tarvittaisiin sen selvittämiseksi, jakautuuko Facebookissa käytetty kieli näin selvästi kahtia ja asettuvatko rajat mainittuihin arvoihin.

Ryhmittelyn apumenetelmänä kokeiltiin klusterointia (k-means klusterimäärillä 2, 3, ja 4), joka jakaa samankaltaisia piirreprosentteja saaneet henkilöt ryhmiin matemaattisesti. Tällä olisi voinut löytää piileviä samankaltaisuuksia muissakin suhteissa kuin vain yksinkertaisella yleiskielinen–puhekielinen-akselilla. Klustereille ei kuitenkaan löytynyt mielekäästä tulkintaa, joten ryhmittelyssä turvauduttiin vain edellä kuvattuun silmämääräiseen arviointiin.

IRC:ssä lienee kehittynyt kulttuuri, joka ohjaa kaikkia käyttämään puhekielistä ilmaisua. Facebook on paljon moninaisempi kokoelma erilaisia palveluita, joten se antaa käyttäjälleen enemmän vapauksia kielen suhteen hänen oman käyttötarkoituksensa ja yleisönsä mukaan. Keskustelupalstalla ilmennyt yksittäinen selvästi muita puhekielisempi informantti saattaa olla poikkeustapaus. Keskimäärin

siellä käytetään kuitenkin selvästi muita hillitympää linjaa, joskin täysin yleiskielisetkin tapaukset ovat harvinaisia. Silti jotkut saattavat nähdä sen asiallisena ja muodollisena viestintätapana, joka edellyttää yleiskielistä kieliasua.

Palataan kuvaan 4 yllä myös neljännen hypoteesin (H4) selvittämiseksi, ja tarkastellaan kuvaan merkittyjä keskiarvoja. Hypoteesin mukaan IRC:n pitäisi olla selvästi puhekielisempi kuin Facebook, joka puolestaan jäisi selvästi edellä mainitun ja keskustelupalstan väliin. Huolimatta Facebook-aineiston kaksijakoisuudesta nyt tarkastellaan siis keskiarvoa eli koko keskustelutyypistä saatavaa yleiskuvaa.

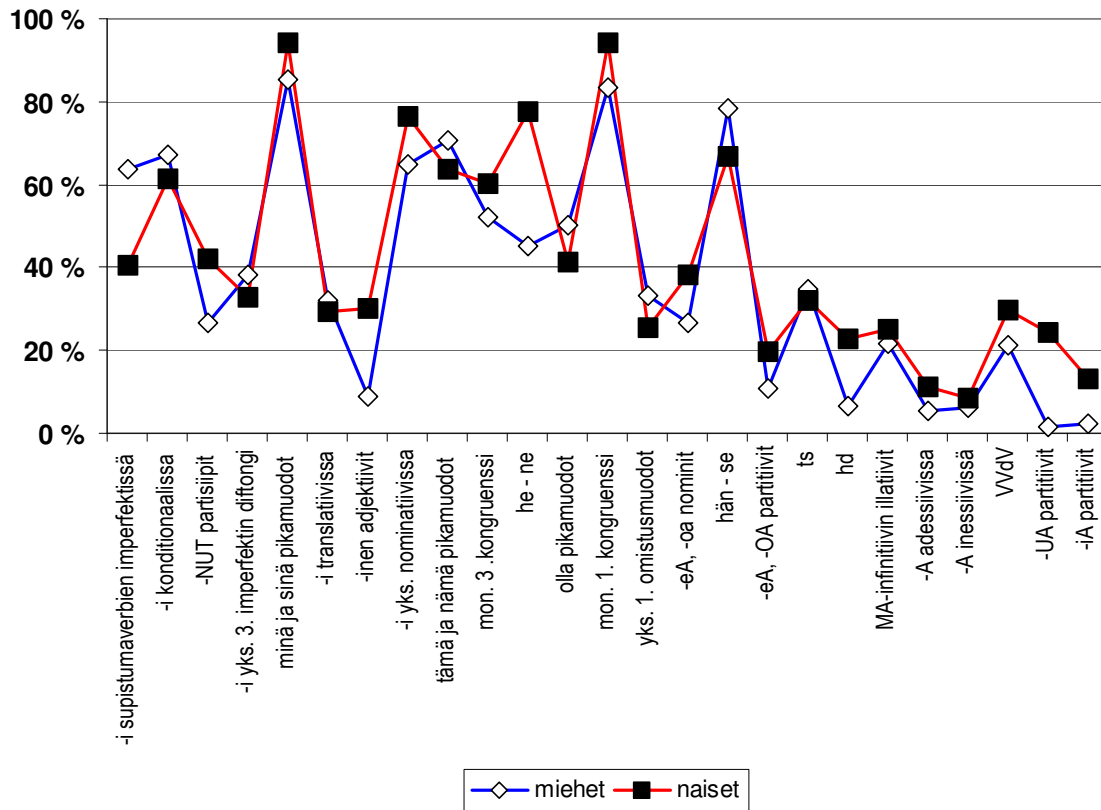
Kuvan perusteella oletukset saattaisivat hyvin toteutuakin. Myös ANOVA:n mukaan ryhmien puhekielisyyssosuudet eroavat toisistaan tilastollisesti merkitsevästi, $F(2, 30) = 9,52$, $p < 0,001$. Parittaiset t-testit osoittavat, että IRC-ryhmän puhekielisyyssosuus ($\bar{x} = 0,58$) ei kuitenkaan ole merkitsevästi suurempi kuin Facebook-ryhmän ($\bar{x} = 0,45$), $t(14) = 1,42$, $p = 0,089$, mutta Facebook-ryhmän (ja siten myös IRC-ryhmän) arvo on merkitsevästi suurempi kuin keskustelupalstaryhmän ($\bar{x} = 0,23$), $t(16) = 2,15$, $p = 0,023$. Teknisesti tämä selittynee Facebook-keskustelijoiden huomattavalla hajonnalla, sillä varianssi 0,063 oli yli kaksi kertaa suurempi kuin toisissa ryhmissä. Molemmat p-arvot ovat joka tapauksessa sellaista suuruusluokkaa, että suurempi otanta olisi ollut hyödyllinen. Tällä aineistolla H4 voidaan hyväksyä vain osittain.

Vaikka aineistossa informanttien taustat eivät olleet kaikkien kohdalta tiedossa, sukupuoli oli yksi tekijä, joka kahta IRC-aineiston henkilöä lukuun ottamatta tunnettiin. Sukupuolten välillä on kielenkäytössä tiettyjä tunnettuja eroja sekä puheessa että verkkokeskustelussa (Huffaker & Calvert 2005). Erot ovat kuitenkin melko pieniä, mitä kuvaa sekin, ettei kahden tuntemattomaksi jääneen informantin sukupuolia pystytty kohtalaisen suuresta viestimäärästä tunnistamaan. Kuitenkin internetissä miesten kieli on yleensä aggressiivisempaa ja kiroilevampaa, kun taas naiset ovat kohteliaampia ja myötäilevämpiä (Herring 2001b). Puhekielisyydestä ei tiedetä mitään, ja siksi sitä on kiintoisaa katsoa, vaikkei se ole tutkimuksen pääkysymyksiä.

Piirrekohtaisesti erot olivat pieniä (kuva 5). Vaikka suurimmillaan ero oli yksittäisen piirteen kohdalla 30 %-yksikköä, piirteiden puhekielisyyssasteet kulkivat paljon

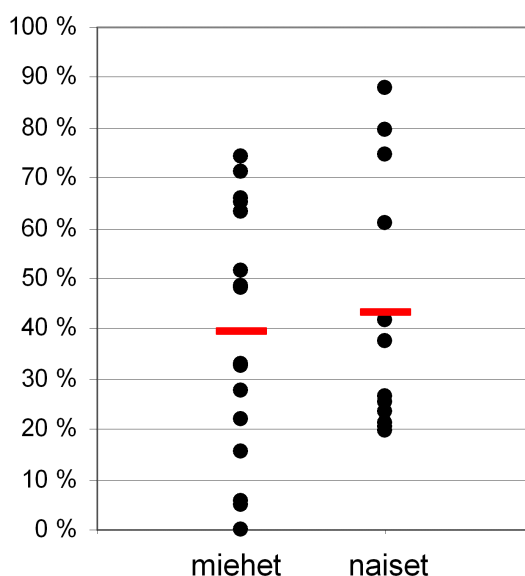
selvemmin käsi kädessä kuin esimerkiksi keskustelutyyppeiden välillä.

Kuva 5. Miesten ja naisten väliset piirrekohtaiset puhekielisyyserot.



Kuvasta 6 nähdään informanttien henkilökohtaiset puhekielisyyssasteet sukupuolen mukaan. Keskiarvoissa (poikkiviivat kuvassa) on pieni ero, mutta jos sitä verrataan hajontaan, nähdään selvästi, ettei ero ole merkitsevä (parittainen t-testi: $t(26) = 0,405$, $p = 0,344$). Sukupuolten välillä ei siis näyttäisi olevan selvää eroa puhekielisyyden suhteen verkossa, vaikka joissakin muissa suhteissa sitä on havaittu (ks. esim. Baron 2010; Herring 2010b).

Kuva 6. Yksittäisten henkilöiden puhekielisyydet sukupuolen mukaan.



Aineistossa ilmeni myös useita sellaisia helsinkiläiseen kieleen kuuluvia puhekielen piirteitä, joita tässä ei tutkittu, koska piirrevalikoiman oli tarkoitus olla vain edustava valikoima. Siinä sen voidaan katsoa onnistuneen.

Eräs huomiota herättävä ongelma menetelmän luotettavuudessa havaittiin. Informantti #F10 oli analyysin mukaan huomattavan yleiskielinen (16 %), mutta käytännössä näin ei ollut. Se ilmeni runsaana hymiöiden käyttönä, englanninkielisinä lausahduksina ja puhekielisenä sanastona, joista yhtäkään käytetty menetelmä ei havaitse. Muiden kohdalla menetelmä antoi todennäköisesti huomattavasti oikeamman kuvan henkilön kielenkäytöstä. Vaikka tuo yksittäinen informantti antaa heikon kuvan menetelmän soveltumisesta verkkokielenkäytön analyysiin, sitä voidaan keskimäärin pitää näiden tulosten perusteella sopivana. Ongelmana menetelmän luotettavuudessa on siis se, ettei se havaitse kaikkia puhekielisyyden piirteitä. Tästä voidaan päätellä, että menetelmä ei sovellu hyvin yksilön puhekielisyyden arviointiin, mutta joukon kyllä. Tällaisenaan sitä olisi hyvä tukea muiden tekijöiden, erityisesti sanaston, analyysillä.

Esimerkki 7, poikkeavaa puhekielisyyttä Facebook-aineistossa, mitattu puhekielisyyss 16 % (informantti #F10):

Parille bisselle La Lunaan ja dinnerille Routaan :) Nice!

Erittain makeita fotoja tyhjyydestä, suosittelen tsiigaamaan!
:)

Flunssa, day three. Vikat jouluostokset piti jättää apujoukkojen hoidettavaksi. Kohta luovuttamaan parit taulut tilaajalleen ja sitten lähtee sukulointi tour de Finland käyntiin. :)

Taulukkoon 6 on koottu vielä yhteen varsinaiset tutkimushypoteesit sekä niiden hyväksyttävyyt. H1 ja H3 voitiin hyväksyä täysin, mutta H2 hylättiin. H4 ei pätenyt IRC:n ja Facebookin välillä, mutta keskustelupalsta osoittautui merkittävästi vähemmän puhekieliseksi kuin kaksi muuta. Tästä voisi vetää johtopäätöksen, että Facebook on jollakin tavalla huomattavan lähellä synkronisia keskustelumuotoja, tai sitten jokin tuntematon tekijä vaikuttaa siihen.

Taulukko 6. Yhteenveto tutkimushypoteeseista.

#	Hypoteesi	Hyväksyminen
H1	Verkkokeskustelun puhekielisyys jää selvästi yleiskielen ja puheen väliin.	hyväksytään
H2	Samojen kielenpiirteiden puhekieliset variantit ovat yleisiä puheessa ja verkkokeskustelussa.	hylätään
H3	Keskustelijat voidaan jakaa selvärajaisiin ryhmiin heidän puhekielen käyttönsä mukaan.	hyväksytään
H4	IRC on puhekielisempi kuin Facebook, joka puolestaan on puhekielisempi kuin keskustelupalsta.	hyväksytään osittain

5 Päätelmät

Tässä tutkielmassa on selvitetty arkisen suomalaisen verkkokeskustelun puhekielisyyttä pääasiassa perinteisen sosiolingvistisen variaation tutkimuksen keinoin. Päättökysymyksenä oli selvittää, kuinka paljon ja millä tavoin arkinen verkkokeskustelu on puhekielistä.

Verkkokeskustelu on joissakin tapauksissa huomattavan puhekielistä, mutta se jää kaikissa tilanteissa keskimäärin alle puheen puhekielisyysasteen. Kokonaisuutena yhtäläisyyttä puheen ja verkkokeskustelun välillä on enemmän kuin eroa, joten puhekielisuuden suhteen ei voida puhua kovin erilaisesta tavasta käyttää kieltä, vaikka tällaista näkemystä tunnetusti puoltavat monet muut verkkokeskustelun tekijät, kuten hymiöiden ja lyhenteiden käyttö. Aivan samankaltaista verkkokeskustelun puhekielisyys ei kuitenkaan näyttäisi olevan pääkaupunkiseudun puheeseen verrattuna, vaan kaikkialla Suomessa esiintyvät piirteet, kuten subjektin ja predikaatin inkongruenssi mon. 1. ja 3. persoonassa, korostuvat siellä, ja jotkut lähinnä Helsingin seudulla esiintyvät piirteet ovat harvinaisempia. Kiehtovana yksityiskohtana samansuuntaista puhekielen omaksumista on havaittu muualta Suomesta Helsinkiin muuttaneiden keskuudessa (Nuolijärvi 1986), mutta tämän tutkielman rajoissa ei voida vetää johtopäätöksiä mahdollisista yhtäläisyyksistä.

Keskustelijoita voidaan pyrkiä luokittelemaan heidän puhekielen käyttönsä mukaan, vaikka kovin luotettavia rajoja luokkien välille ei voida muodostaa tämän tutkielman aineiston suppeuden vuoksi. IRC:n tapaisissa reaaliaikaisissa keskusteluissa yleiskielisyyttä ei näyttäisi esiintyvän, vaan niissä on vahvasti puhekielinen keskustelukulttuuri. Facebookin tilapäivityksissä kielellistä vaihtelua on enemmän, mikä voi johtua sen monipuolisuudesta, jolloin erilaisia käyttötarkoituksia ja käyttäjiä on enemmän. Facebook näyttikin jakautuvan kahtia toisaalta vahvasti ja toisaalta vain kevyesti puhekielisiin käyttäjiin. Facebookia voisikin kuvailla vaihtelevaksi mediaksi, sillä tavallaan se lähestyy reaaliaikaisuutta, vaikka monet haluavat pitää siellä yllä asiallista ja siten melko yleiskielistä profiilia. Tästä vaihtelevuudesta huolimatta verrattuna IRC:hen näiden kahden välillä ei ilmennyt tilastollisesti merkitsevää eroa puhekielisydessä. Keskustelupalstat sen sijaan poikkesivat näistä

kahdesta ollen keskimäärin vähemmän puhekielisiä. Toisin kuin kahdessa muussa ryhmässä, osa henkilöistä käyttää jopa täysin yleiskielistä ilmaisua, mikä johtunee samankaltaisista keskustelunaiheista huolimatta keskustelupalstan asynkronisuudesta.

Tutkimuksessa haettiin tarkoituksellisesti verkkokeskustelun puhekielistä ääripäätä, joten tuloksia on yleistettävä varoen. Samoin aineiston pieni koko antaa syytä varovaisuuteen, erityisesti tilastollisten tunnuslukujen osalta. Internetin eri keskustelutyypit voidaan luokitella lukuisilla eri tavoilla, ja niiden moninaisuudesta johtuen jako on aina joko jonkinlainen kompromissi todellisuudesta tai liian yksityiskohtainen ollakseen käytännöllinen. Käytännöllisyyden vuoksi keskustelutyypeistä ja erityisesti niiden keskinäisestä samankaltaisuudesta tehtiinkin yleistäviä oletuksia. Ei kuitenkaan tiedetä, että eri tavat keskustella tietyn keskustelutyypin sisällä olisivat samanlaisia (tässä tapauksessa puhekielisyyden suhteen) tai että valittu tapa olisi keskustelutyypinsä sisällä edustavin. Ei ole siis sanottua, että chateissä puhekielisyys olisi samanlaista kuin IRC:ssä, tai Twitterissä samanlaista kuin Facebookissa, vaikka monia muita yhtäläisyyksiä erityisesti teknisestä näkökulmasta onkin. Tässä tutkielmassa otettiin pitkälti lähtökohdaksi tekniset ominaisuudet, koska niihin on helpompi päästä käsiksi kuin vaikkapa sosiaalisiin tai kulttuurillisiin tekijöihin, joiden tiedetään kuitenkin yhtä lailla vaikuttavan kieleen.

Tulosten saavuttamisessa tutkimus onnistui melko hyvin, ja vain kaksi rajoitusta kohosi ilmeiseksi. Aineisto oli turhan pieni joidenkin tutkimushypoteesien selvittämiseen, ja pelkkä puhekielisyyden selvittäminen antaa melko kapean kuvan verkkokeskustelun kielestä, sen muusta luonteesta puhumattakaan.

Tutkimus onnistui myös menetelmän soveltamisen osalta. Variaationtutkimus näyttäisi menetelmänä soveltuvan verkkokeskustelun tutkimiseen, sillä verrattuna sen soveltamisessa puheen tutkimukseen mitään selviä esteitä ei tullut esiin. Päinvastoin se voi olla työmäärältään kevyempää. On kuitenkin tärkeää huomata, että menetelmää voidaan käyttää vain melko kapea-alaisten tutkimuskysymysten selvittämiseksi. Toisin kuin tässä tutkimuksessa, olisi suotavaa huomioida vielä

informanttien taustat, mikä onnistuisi melko helposti esimerkiksi haastattelemalla jälkikäteen. Tosin joissakin tapauksissa internetin anonyymi luonne saattaa vaikeuttaa sitä (Herring 2001a).

Pohdittaessa yleisesti verkkokeskustelun tutkimusta internetissä on syytä huomata, että internetiin liittyvissä tutkimuksissa on yleensä keskenään erilainen kysymyksenasettelu, aineisto, teoreettinen lähtökohta ja tutkimusperinne (Arpo 2005, 297). Siksi on vaikea arvioida tutkimusta suhteessa aiempaan verkkokeskustelun tutkimukseen, ja tämän kaltaiselle soveltavalle ja osin vertailevalle tutkimukselle saattaa olla tiedemaailmassa kysyntää.

Tämä tutkielma on ollut vain pieni katsaus yhteen kiinnostavaan aiheeseen verkkokeskustelun tutkimuksen hajanaisessa kentässä. Se nostaakin esiin monensuuntaista tarvetta jatkotutkimukselle.

Kaksi asiaa on joka tapauksessa syytä huomioida verkkokeskustelun tutkimuksessa tämän tutkimuksen kokemusten perusteella. Ensiksikin käytössä tulisi olla laajempi tutkimusmenetelmien valikoima, sillä variaationtutkimusta on kritisoitu juuri yksipuolisuudesta ja riittämättömyydestä yksinään. Muut menetelmät valikoituvat tutkittavien ilmiöiden mukaan, eikä siksi mitään yksittäistä voida suositella rinnalle. Kuitenkin yleisesti ottaen laadulliset menetelmät ovat olleet suosittuja. Toiseksi aineiston koon on oltava riittävä. Koska aineiston hankinta ja käsittely verkossa voivat olla helpompia kuin puheen tutkimuksessa, olisi perusteltua käyttää jopa suurempia aineistomääriä kuin perinteisesti on tehty. Erityisesti informanttien määrään kannattaa kiinnittää huomiota. Jos tutkitaan puhekielen eri piirteitä, suurempi informanttikohtainen sanamäärä varmistaisi, että harvinaisemmatkin piirteet esiintyvät riittävän useasti.

Jos pitäydytään tämän tutkielman tutkimuskysymyksissä, eri keskustelutyyppejä voisi tutkia vielä tarkemmin. Mukaan voisi ottaa enemmän keskustelutyyppejä, esimerkiksi tässä yhteydessä samaan tyyppiin luokiteltuja, kuten keskenään erilaisia keskustelupalstoja. Voidaan nimittäin olettaa, että keskustelun aihe vaikuttaa puhekielisyyteen (esim. politiikka, uskonto, harrastukset), samoin kuin keskusteluyhteisö. Tällöin laadullisiakin tutkimusmenetelmiä voitaisiin hyödyntää

tehokkaasti esimerkiksi erilaisten keskustelukulttuurien analysointiin.

Puhuttu kieli muuttuu jatkuvasti, joten luotettavinta olisi vertailla samana aikana ja samanlaisesta populaatiosta kerättyjä aineistoja puheen ja verkkokeskustelun osalta toisiinsa. Tässä tutkimuksessa jouduttiin tuoreemman tiedon puutteessa tekemään (perusteltuja) oletuksia kielen muutoksesta, ja tällöin päästään vain suuntaa-antaviin arvioihin yhtäläisyyksistä, eikä tarkka vertaaminen ole mahdollista. Samoin kuin puheen kohdalla, pitäisi verkossakin selvittää kielenkäytön muuttumista ajan myötä (Herring 2010a). Suomen kielen kohdalla tämä vaatisi ensimmäisten tällaisten tutkimusten tekemistä mahdollisimman pian, jotta pitkittäistutkimus olisi myöhemmin mahdollista.

Siitä ei ole juurikaan tutkimuksia, miten sama henkilö käyttää kieltä eri paikoissa ja tilanteissa verkossa, ja edelleen erityisesti internetissä verrattuna puhuttuun kieleensä. Hyvää esimerkkiä tämän mahdollisuuksista näyttävät kuitenkin Condon ja Čech (2010), jotka ovat verranneet toisiinsa puhetta sekä synkronista ja asynkronista verkkokeskustelua vuorotteluilmiöiden suhteen.

On esitetty, että verkossa pitäisi tutkia myös puhetta eikä vain kirjoitettua kieltä (Herring 2004; 2010a). Tästä olisi mahdollista mennä vielä askel pidemmälle, ja liittää mukaan virtuaalimaailmojen sanattoman viestinnän ulottuvuus. Tällöin päästäisiin lopulta yhtä moninaisiin viestintäkeinojen kokonaisuuksiin, kuin verkon ulkopuolisessa maailmassa on totuttu.

Kvantitatiivista tutkimusotetta on mahdollista soveltaa sosiolingvistiikassa myös muilla tavoin kuin on totuttu. Tilastotieteen menetelmissä on käyttämättömiä mahdollisuuksia. Toistaiseksi niitä hyödyntävät tutkimukset näyttäisivät olevan tilastotieteellisesti melko kevyitä (esim. van Compernelle 2008; Cho 2010; Baron 2010). Hypoteettisena tulevaisuuden esimerkkinä voisi mainita henkilöiden taustojen perusteella kielenkäyttöä ennustavan regressiomallin kehittämisen. Tällaista voitaisiin käyttää vaikkapa automaattisissa asiakaspalvelujärjestelmissä, joissa rennompia kieliä on hyväksyttävää (esim. nuorten populaarikulttuuri).

Monille yritysten työntekijöille sosiaalisen median ja varsinkin sähköpostin kaltaisten

verkkoviestintätapojen käyttö on arkipäivää. Institutionaalista lähtöasetelmasta käsin on luontevaa käyttää asiakasrajapinnassa ja muiden sidosryhmien kanssa aina huoliteltua ja asiallista yleiskieltä. Myös rennompia kieliä saattaa kuitenkin olla läheisempien suhteiden luomiseksi joissakin erityistilanteissa tarpeellista. Tällöin työntekijöitä voi olla tarpeen ohjeistaa kielenkäytöstä tarkemmin, jolloin tarvitaan jonkinlaista etukäteistietoa siitä, millaista kieltä vastaanottajien voi olettaa käyttävän. Käytännön esimerkkinä voidaan nähdä keskustelupalstojen ja chattien moderaattorit, joiden tehtävä on asiakaspalveluluonteinen. He joutuvat päättämään, sopeutuvatko muiden kieleen vai pitävätkö yllä asiallista yleiskielistä linjaa. Useimpiin chatteihin voisi sopia puhekielinen tuttavallinen helposti lähestyttävä moderaattori, mutta keskustelupalstoille ei välttämättä, muun muassa vahvemman institutionaalisuuden vuoksi.

Tätä kirjoitettaessa eduskuntavaalit ovat käsillä. Vaalityössä sosiaalinen media on otettu innolla käyttöön, ja muun muassa Facebook ja Twitter ovat olleet pullollaan ehdokkaiden viestejä ja profiileja. He joutuvat tekemään viestinnässään valintoja kielensä suhteen (niin verkossa, toreilla kuin joukkoviestimissäkin), ja tämä voi olla yksi heidän menestykseensä vaikuttava tekijä. Niinpä tällaisten henkilöiden kielenkäytön mahdollinen vaikutus suosioon ja menestykseen voisi olla laajemminkin yhteiskunnallisesti kiinnostava kysymys.

Verkkokeskustelussa käytetyn kielen tutkimukselle on siis nähtävissä erilaisia yhteiskunnallisia ja liikkeenjohdollisia hyötyjä. Vaikka ne toistaiseksi näyttävät vähäisiltä, internetin merkityksen jatkuvasti korostuessa niitä voi ilmaantua nopeastikin lisää. Tärkeintä on kuitenkin tieteelliseltä kannalta soveltaa tuttuja menetelmiä uusille alueille ja avata uusia uria tuleville tutkimuksille.

Lähteet

Aineisto

10 mahdollisuuksien mukaan satunnaista informanttia kustakin seuraavasta lähteestä siten, että jokaiselta on kerätty 800–2000 sanaa aikavälillä syyskuu 2010 – tammikuu 2011:

City.fi-portaalin (www.city.fi) keskustelupalsta: henkilön kirjoittamat viestit julkisesti selattavissa henkilön profiiliin kautta, kotipaikaksi ilmoitettu Helsinki, Espoo tai Vantaa

Facebook-tilapäivitykset: henkilön kirjoittamat tilapäivitysviestit julkisesti selattavissa henkilön profiiliin kautta, kotipaikaksi ilmoitettu Helsinki

IRC #espoo-kanava

Kirjallisuus

Anderson Jeffrey F., Beard Fred K., & Walther Joseph B. 2010. Turn-Taking and the Local Management of Conversation in a Highly Simultaneous Computer-Mediated Communication System. *Language@Internet*. Vol 7, article 7. http://www.languageatinternet.de/articles/2010/2804/index_html/, 13.4.2011.

Androutsopoulos Jannis 2007. Style online: Doing hip-hop on the German-speaking Web. Teoksessa *Style and Social Identities: Alternative Approaches to Linguistic Heterogeneity*, toim. Auer Peter. De Gruyter, Berliini. 279–317.

Androutsopoulos Jannis tulossa. From Variation to Heteroglossia in the Study of Computer-Mediated Discourse. Teoksessa *Digital Discourse: Language in the New Media*, toim. Thurlow Crispin & Mroczek Kristine. OUP, Lontoo. <http://jannisandroutsopoulos.files.wordpress.com/2010/09/variation-to-heteroglossia-in-cmc-2nd-draft.pdf>, 12.4.2011.

Androutsopoulos Jannis, Beißwenger Michael 2008. Introduction: Data and Methods in Computer-Mediated Discourse Analysis. *Language@Internet*. Vol 5, article 2. <http://www.languageatinternet.de/articles/2008/1609/introduction.pdf/>,

15.4.2011.

Arpo Robert 2005. *Internetin keskustelukulttuurit*. Joensuun yliopiston humanistisia julkaisuja 39. Joensuun yliopistopaino, Joensuu.

Baron Naomi S. 2010. Discourse Structures in Instant Messaging: The Case of Utterance Breaks. *Language@Internet*. Vol 7, article 4.
http://www.languageatinternet.de/articles/2010/2651/index_html/, 13.4.2011.

Cho Thomas 2010. Linguistic Features of Electronic Mail in the Workplace: A Comparison with Memoranda. *Language@Internet*. Vol 7, article 3.
<http://www.languageatinternet.de/articles/2010/2728/cho.pdf/>, 13.4.2011.

Condon Sherri L., Čech Claude G. (2010). Discourse Management in Three Modalities. *Language@Internet*. Vol 7, article 6.
http://www.languageatinternet.de/articles/2010/2770/index_html/, 13.4.2011.

Crystal David 2002. *Language and the Internet*. Cambridge University Press, Cambridge.

Downes William 1988. *Language and Society*. Fontana Paperbacks, Lontoo.

Gerrard Peter 2007. Estimating Linguistic Diversity on the Internet: A Taxonomy to Avoid Pitfalls and Paradoxes. *Journal of Computer-Mediated Communication*. Vol 12, issue 4. 1298–1321.

Hakulinen Auli 1989. Puheen ja kirjoituksen pragmatiikkaa. Teoksessa *Kielen käyttö ja käyttäjä. Äidinkielen opettajain liiton vuosikirja XXXVI*, toim. Kauppinen Anneli ja Keravuori Kyllikki. Äidinkielen opettajain liitto, Helsinki. 11–22.

Herring Susan C. 1998. Le style du courrier électronique: variabilité et changement. *Revue d'aménagement linguistique 84–85*. 9–16.
<http://ella.slis.indiana.edu/~herring/term.html>, 26.3.2011.

Herring Susan C. 2001a. Computer-mediated discourse. Teoksessa *The Handbook of Discourse Analysis*, toim. Schiffrin Deborah, Tannen Deborah ja Hamilton

Heidi E. Blackwell, Oxford. 612–634. <http://odur.let.rug.nl/~redeker/herring.pdf>, 26.3.2011.

Herring Susan C. 2001b. Gender and power in online communication. *Center for Social Informatics Working Papers*. <http://www.slis.indiana.edu/csi/WP/WP01-05B.html>, 8.1.2011.

Herring Susan C. 2004. Computer-Mediated Discourse Analysis: An Approach to Researching Online Behavior. Teoksessa *Designing for Virtual Communities in the Service of Learning*, toim. Barab S. A., Kling R. ja Gray J. H. Cambridge University Press, New York. 338–376.
<http://ella.slis.indiana.edu/~herring/cmda.pdf>, 14.4.2011.

Herring Susan C. 2007. A Faceted Classification Scheme for Computer-Mediated Discourse. *Language@Internet*. Vol 4, article 1.
http://www.languageatinternet.de/articles/2007/761/Faceted_Classification_Scheme_for_CMD.pdf/, 14.4.2011.

Herring Susan C. 2010a. Computer-Mediated Conversation: Introduction and Overview. *Language@Internet*. Vol 7, article 2.
<http://www.languageatinternet.de/articles/2010/2801/herring.intro.pdf/>, 13.4.2011.

Herring Susan C. 2010b. Who's Got the Floor in Computer-Mediated Conversation? Edelsky's Gender Patterns Revisited. *Language@Internet*. Vol 7, article 8.
http://www.languageatinternet.de/articles/2010/2857/index_html/, 13.4.2011.

Herring Susan C., Scheidt Lois Ann, Wright Elijah, Bonus Sabrina 2005. Weblogs as a bridging genre. *Information Technology & People*. Vol 18, issue 2. 142–171.

Hudson Richard A. 1996. *Sociolinguistics*. Cambridge University Press, Cambridge.

Huffaker David, Calvert Sandra 2005. Gender, Identity, and Language Use in Teenage Blogs. *Journal of Computer-Mediated Communication*. Vol 10, issue 2.
<http://jcmc.indiana.edu/vol10/issue2/huffaker.html>, 8.1.2011.

- Hurta Heikki 1999. Variaation tutkimuksen myytit ja stereotypiat. Teoksessa *Kirjoituksia sosiolingvistiikasta*, toim. Määttä Urho, Pälli Pekka ja Suojanen Matti K. Tampereen yliopistopaino, Tampere. 53–101.
- Kalman Yoram, Ravid Gilad, Raban Daphne, Rafaeli Sheizaf 2006. Pauses and Response Latencies: A Chronemic Analysis of Asynchronous CMC. *Journal of Computer-Mediated Communication*. Vol 12, issue 1. 1–23.
- Karlsson Fred 1995. Normit, kielenkäyttö ja kieliopit. Teoksessa *Tutkimuksen etulinjassa*, toim. Rydman Jan. WSOY, Helsinki. 161–172.
- Kotilainen Lari 2002. Moi taas, ai äm päk: Lauseet, tilanteet ja englanti suomenkielisessä chat-keskustelussa. Teoksessa *Äidinkielen merkitykset*, toim. Herlin Ilona, Kalliokoski Jyrki, Kotilainen Lari, Onikki-Rantajääskö Tiina. SKS, Helsinki. 191–209.
- Kuiri Kaija 1989. Puheen lauseita kirjoituksessa? Teoksessa *Kielen käyttö ja käyttäjä. Äidinkielen opettajain liiton vuosikirja XXXVI*, toim. Kauppinen Anneli ja Keravuori Kyllikki. Äidinkielen opettajain liitto, Helsinki. 187–195.
- Lappalainen Hanna 2004. *Variaatio ja sen funktiot: Erään sosiaalisen verkoston jäsenten kielellisen variaation ja vuorovaikutuksen tarkastelua*. SKS, Helsinki.
- Lappalainen Hanna 2006. Kelan asiointikeskustelut Helsingin puhekielen variaation kuvastajina – Esimerkkinä yksikön 1. ja 2. persoonan persoonapronominien käyttö. Teoksessa *Helsinki kieliyhteisönä*, toim. Juusela Kaisu ja Nisula Katariina. Helsingin yliopiston suomen kielen ja kotimaisen kirjallisuuden laitos, Helsinki. 275–298.
- Lehikoinen Laila 2005. *Suomea ennen ja nyt. Suomen kielen kehitys ja vaihtelu*. Finn Lectura, Helsinki.
- Malmelin Nando 2003. *Mainonnan lukutaito*. Gaudeamus, Helsinki.
- Mann Chris, Stewart Fiona 2000. *Internet Communication and Qualitative Research*.

SAGE Publications Ltd., Lontoo.

- Mantila Harri 1997. Johdanto. Teoksessa *Pohjoissuomalaisen puhekielen sosiolingvistinen variaatio*, toim. Makkonen Seija ja Mantila Harri. Suomen ja saamen kielen ja logopedian laitoksen julkaisuja 8. Oulun yliopistopaino, Oulu. 1–23.
- Mielikäinen Aila 1980. Nykysuomalaisen puhekielen tutkimuksen taustaa. Teoksessa *Nykysuomalaisen puhekielen murros. Jyväskylän osatutkimus. Raportti 1*, toim. Mielikäinen Aila. Jyväskylän yliopiston suomen kielen ja viestinnän laitoksen julkaisuja 20, Jyväskylä. 1–40.
- Määttänen Leena 2007. Puheenomaisten piirteiden ilmeneminen erityyppisissä suomalaissa kirjoitetuissa teksteissä. Pro gradu -tutkielma. Groningenin yliopisto.
<http://scripties.let.eldoc.ub.rug.nl/FILES/root/Master/DoorstroomMasters/FinoegrischeTalenenC/2008/L.M.Mttnen/Ma-1096974-L.M.Mttnen.pdf>, 9.11.2010.
- Nuolijärvi Pirkko 1986. *Kolmannen sukupolven kieli: Helsinkiin muuttaneiden suurten ikäluokkien eteläpohjalaisten ja pohjoissavolaisten kielellinen sopeutuminen*. SKS, Helsinki.
- Nuolijärvi Pirkko 2000. Sosiolingvistiikka kielentutkimuksen kentässä. Teoksessa *Kieli, diskurssi & yhteisö*, toim. Sajavaara Kari ja Piirainen-Marsh Arja. Jyväskylän yliopistopaino, Jyväskylä. 13–37.
- Nuolijärvi Pirkko 2006. Helsingin kieliolot 2000-luvun alussa. Teoksessa *Helsinki kieliyhteisönä*, toim. Juusela Kaisu ja Nisula Katariina. Helsingin yliopiston suomen kielen ja kotimaisen kirjallisuuden laitos, Helsinki. 318–335.
- Nuolijärvi Pirkko, Tiittula Liisa 2000. *Televisiokeskustelun näyttämöllä*. SKS, Helsinki.
- Paunonen Heikki 1989. Muuttuvat puhesuomen muodot. Teoksessa *Kielestä kiinni. Tietolipas 113*, toim. Aalto Seija, Hakulinen Auli, Laalo Klaus, Leino Pentti ja Lieko Anneli. SKS, Helsinki.

- Paunonen Heikki 1995a. *Suomen kieli Helsingissä. Huomioita Helsingin puhekielen historiallisesta taustasta ja nykyvariaatiosta*. Helsingin yliopiston suomen kielen laitos, Helsinki.
- Paunonen Heikki 1995b. Puhekielen kielioppia etsimässä. Teoksessa *Tutkimuksen etulinjassa*, toim. Rydman Jan. WSOY, Helsinki. 183–201.
- Paunonen Heikki 2003. Suomen kielen morfologisista muutosmekanismeista. Teoksessa *Muotojen mieli: kirjoituksia morfologiasta ja variaatiosta. Kieli 15*, toim. Laitinen Lea, Lappalainen Hanna, Markkola Päivi ja Vaattovaara Johanna. Helsingin yliopiston suomen kielen laitos, Helsinki. 187–248.
- Paunonen Heikki 2006. Vähemmistökielestä varioivaksi valtakieleksi. Teoksessa *Helsinki kieliyhteisönä*, toim. Juusela Kaisu ja Nisula Katariina. Helsingin yliopiston suomen kielen ja kotimaisen kirjallisuuden laitos, Helsinki. 13–99.
- Salmenkivi Sami, Nyman Niko 2008. *Yhteisöllinen media ja muuttuva markkinointi 2.0*. Talentum, Helsinki.
- Siebenhaar Beat 2008. Quantitative Approaches to Linguistic Variation in IRC: Implications for Qualitative Research. *Language@Internet*. Vol 5, article 4. <http://www.languageatinternet.de/articles/2008/1615/siebenhaar.pdf/>, 15.4.2011.
- Thurlow Crispin 2006. From Statistical Panic to Moral Panic: The Metadiscursive Construction and Popular Exaggeration of New Media Language in the Print Media. *Journal of Computer-Mediated Communication*. Vol 11, issue 3. 667–701.
- Tiittula Liisa 1992. *Puhuva kieli. Suullisen viestinnän erityispiirteitä*. Finn Lectura, Helsinki.
- Tiittula Liisa 2001. Puhuttu kieli ja sen illuusio kirjallisuudessa. *Kielikuvia* 1/01. 5–16.
- Utz Sonja 2009. The (Potential) Benefits of Campaigning via Social Network Sites. *Journal of Computer-Mediated Communication*. Vol 14, issue 2. 221–243.

van Compernelle Rémi A. 2008. Language Variation in Online Personal Ads from Quebec: The case of *ne*. *Language@Internet*. Vol 5, article 1.
<http://www.languageatinternet.de/articles/2008/1666/vc5.pdf/>, 13.4.2011.

Walther Joseph, Loh Tracy, Granka Laura 2005. Let Me Count the Ways: The Interchange of Verbal and Nonverbal Cues in Computer-Mediated and Face-to-Face Affinity. *Journal of Language and Social Psychology*. Vol 24, issue 1. 36–65.

Liite 1: Keskustelutyypin puhekielisyysfrekvenssit ja -asteet piirteittäin

(Kunkin kohdan 1. sarakkeessa ovat puhekielisten varianttien määrät, 2. sarakkeessa puhe- ja yleiskielisten varianttien yhteismäärät ja 3. sarakkeessa näistä laskettu puhekielisyysaste.)

Piirre	IRC			Facebook			Keskustelupalsta			Kaikki yht.	Paunonen 2006, 90	
	# Selitys	puh.	yht.	%	puh.	yht.	%	puh.	yht.			%
1	-i supistumaverb. imperf.	14	16	88 %	16	41	39 %	8	15	53 %	60 %	100 %
2	-i konditionaalissa	145	150	97 %	93	141	66 %	52	149	35 %	66 %	99 %
3	-NUT partisiipit	33	47	70 %	12	30	40 %	10	70	14 %	41 %	99 %
4	-i yks. 3. imperf. diftongi	22	34	65 %	14	47	30 %	4	26	15 %	37 %	95 %
5	minä ja sinä pikamuodot	255	258	99 %	119	122	98 %	179	232	77 %	91 %	94 %
6	-i translatiivissa	28	49	57 %	40	98	41 %	17	115	15 %	38 %	93 %
7	-inen adjektiivit	5	13	38 %	5	15	33 %	0	27	0 %	24 %	92 %
8	-i yks. nominatiivissa	34	41	83 %	30	44	68 %	8	16	50 %	67 %	91 %
9	tämä ja nämä pikamuodot	173	197	88 %	62	86	72 %	48	137	35 %	65 %	86 %
10	mon. 3. kongruenssi	82	87	94 %	51	73	70 %	51	164	31 %	65 %	83 %
11	he - ne	9	9	100 %	4	5	80 %	6	18	33 %	71 %	77 %
12	olla pikamuodot	150	183	82 %	36	71	51 %	45	214	21 %	51 %	75 %
13	mon. 1. kongruenssi	25	27	93 %	39	40	98 %	20	26	77 %	89 %	74 %
14	yks. 1. omistusmuodot	9	11	82 %	6	16	38 %	12	68	18 %	46 %	70 %
15	-eA, -oa nominit	30	56	54 %	14	43	33 %	17	85	20 %	35 %	69 %
16	hän - se	55	55	100 %	9	9	100 %	5	25	20 %	73 %	51 %
17	-eA, -OA partitiivit	9	44	20 %	10	59	17 %	7	76	9 %	16 %	46 %
18	ts	29	49	59 %	13	42	31 %	14	82	17 %	36 %	36 %
19	hd	6	23	26 %	5	32	16 %	5	59	8 %	17 %	35 %
20	MA-infinitiivin illatiivit	26	59	44 %	34	111	31 %	8	102	8 %	28 %	33 %
21	-A adessiivissa	36	225	16 %	28	296	9 %	5	331	2 %	9 %	32 %
22	-A inessiivissä	27	229	12 %	18	275	7 %	11	328	3 %	7 %	24 %
23	VVdV	9	20	45 %	5	18	28 %	0	19	0 %	24 %	23 %
24	-UA partitiivit	2	38	5 %	12	62	19 %	2	40	5 %	10 %	19 %
25	-iA partitiivit	9	136	7 %	28	186	15 %	3	251	1 %	8 %	18 %
	keskiarvo			61 %			45 %			23 %	43 %	65 %

Liite 2: Informanttien puhekielisyyshäiriöt ja -asteet piirteittäin

IRC-aineisto (Ylimmällä rivillä on kunkin informantin tunniste, hänen sukupuolensa m tai n sekä hänen viestiensä yhteenlaskettu sanamäärä.)

Piirre	#1	?	899	#12	m	988	#13	n	1040	#14	m	995	#15	m	854	#16	n	1007	#17	?	1135	#18	m	1290	#19	m	1062	#10	m	1204
# Selitys	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%
1 -i supistumaverb. imperf.	1	1	100 %	0	0	-	1	1	100 %	4	4	100 %	1	1	100 %	2	2	100 %	1	1	100 %	0	0	-	4	4	100 %	0	2	0 %
2 -i konditionaalisissa	10	11	91 %	17	17	100 %	6	7	86 %	17	17	100 %	12	13	92 %	22	22	100 %	9	10	90 %	28	28	100 %	14	14	100 %	10	11	91 %
3 -NUT partisiipt	0	1	0 %	5	5	100 %	1	3	33 %	4	5	80 %	1	4	25 %	10	11	91 %	8	9	89 %	4	4	100 %	0	1	0 %	0	4	0 %
4 -i yks. 3. imperf. diftongi	0	0	-	1	1	100 %	2	4	50 %	5	5	100 %	1	5	20 %	4	4	100 %	3	3	100 %	2	2	100 %	2	6	33 %	2	4	50 %
5 minä ja sinä pikamuodot	9	9	100 %	33	33	100 %	59	60	98 %	14	14	100 %	10	10	100 %	28	29	97 %	28	28	100 %	18	18	100 %	19	20	95 %	37	37	100 %
6 -i translatiivissa	0	5	0 %	1	1	100 %	0	6	0 %	1	1	100 %	2	7	29 %	2	2	100 %	8	8	100 %	7	7	100 %	0	5	0 %	7	7	100 %
7 -inen adjektiivit	0	0	-	0	0	-	0	2	0 %	1	1	100 %	0	0	-	2	2	100 %	1	1	100 %	0	1	0 %	0	0	-	1	6	17 %
8 -i yks. nominatiivissa	4	4	100 %	4	4	100 %	1	4	25 %	3	5	60 %	6	7	86 %	6	6	100 %	2	2	100 %	1	1	100 %	0	1	0 %	7	7	100 %
9 tämä ja nämä pikamuodot	3	19	16 %	15	16	94 %	9	11	82 %	18	18	100 %	17	18	94 %	7	7	100 %	18	19	95 %	23	24	96 %	16	18	89 %	47	47	100 %
10 mon. 3. kongruenssi	3	6	50 %	13	13	100 %	2	3	67 %	6	7	86 %	4	4	100 %	7	7	100 %	12	12	100 %	18	18	100 %	10	10	100 %	7	7	100 %
11 he - ne	0	0	-	2	2	100 %	0	0	-	0	0	-	1	1	100 %	1	1	100 %	3	3	100 %	2	2	100 %	0	0	-	0	0	-
12 olla pikamuodot	7	11	64 %	33	33	100 %	18	26	69 %	10	11	91 %	3	9	33 %	25	26	96 %	26	26	100 %	4	10	40 %	10	10	100 %	14	21	67 %
13 mon. 1. kongruenssi	2	2	100 %	1	1	100 %	6	7	86 %	1	2	50 %	1	1	100 %	12	12	100 %	2	2	100 %	0	0	-	0	0	-	0	0	-
14 yks. 1. omistusmuodot	0	0	-	0	0	-	3	4	75 %	0	0	-	0	0	-	2	2	100 %	0	1	0 %	0	0	-	3	3	100 %	1	1	100 %
15 -eA, -oa nominit	0	4	0 %	3	6	50 %	5	9	56 %	6	6	100 %	3	6	50 %	3	3	100 %	5	8	63 %	1	3	33 %	2	4	50 %	2	7	29 %
16 hän - se	2	2	100 %	3	3	100 %	4	4	100 %	5	5	100 %	12	12	100 %	2	2	100 %	7	7	100 %	11	11	100 %	4	4	100 %	5	5	100 %
17 -eA, -OA partitiivit	0	2	0 %	0	2	0 %	2	8	25 %	3	6	50 %	0	3	0 %	0	4	0 %	0	3	0 %	3	6	50 %	0	4	0 %	1	6	17 %
18 ts	0	7	0 %	1	1	100 %	0	1	0 %	4	7	57 %	0	3	0 %	6	8	75 %	1	2	50 %	3	4	75 %	4	6	67 %	10	10	100 %
19 hd	0	1	0 %	1	3	33 %	1	5	20 %	0	1	0 %	0	1	0 %	2	3	67 %	1	3	33 %	0	0	-	1	3	33 %	0	3	0 %
20 MA-infinitiivin illatiivit	3	4	75 %	2	2	100 %	1	11	9 %	2	2	100 %	1	4	25 %	5	8	63 %	4	5	80 %	7	10	70 %	1	2	50 %	0	11	0 %
21 -A adessiivissa	1	25	4 %	11	18	61 %	3	20	15 %	2	27	7 %	2	14	14 %	10	24	42 %	5	21	24 %	1	26	4 %	1	29	3 %	0	21	0 %
22 -A inessiivissä	1	32	3 %	9	26	35 %	0	18	0 %	7	25	28 %	1	20	5 %	6	21	29 %	0	26	0 %	0	18	0 %	3	25	12 %	0	18	0 %
23 VVdV	0	0	-	1	2	50 %	0	3	0 %	0	2	0 %	1	1	100 %	6	7	86 %	0	2	0 %	1	1	100 %	0	2	0 %	0	0	-
24 -UA partitiivit	0	2	0 %	0	3	0 %	0	3	0 %	0	5	0 %	0	2	0 %	1	3	33 %	0	3	0 %	0	6	0 %	1	8	13 %	0	3	0 %
25 -iA partitiivit	0	3	0 %	2	14	14 %	0	20	0 %	1	17	6 %	1	11	9 %	1	10	10 %	2	15	13 %	0	15	0 %	2	19	11 %	0	12	0 %
keskiarvo			40 %			74 %			41 %			66 %			51 %			79 %			65 %			65 %			48 %			49 %

Facebook-aineisto

Piirre	#F1	n	857	#F2	n	1344	#F3	m	1095	#F4	n	1366	#F5	m	1511	#F6	n	961	#F7	m	808	#F8	n	867	#F9	m	1380	#F10	m	1432	
#	Seelys	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%	p.	yht.	%
1	-i supistumaverb. imperf.	2	4	50 %	0	6	0 %	1	1	100 %	6	6	100 %	1	2	50 %	1	3	33 %	1	2	50 %	0	10	0 %	3	4	75 %	1	3	33 %
2	-i konditionaalissa	9	9	100 %	0	12	0 %	17	17	100 %	24	25	96 %	8	14	57 %	5	16	31 %	14	14	100 %	3	10	30 %	10	15	67 %	3	9	33 %
3	-NUT partisiipit	1	1	100 %	0	4	0 %	1	1	100 %	7	7	100 %	1	3	33 %	0	5	0 %	2	3	67 %	0	2	0 %	0	3	0 %	0	1	0 %
4	-i yks. 3. imperf. diftongi	2	7	29 %	0	6	0 %	2	2	100 %	5	5	100 %	0	10	0 %	1	4	25 %	4	4	100 %	0	6	0 %	0	2	0 %	0	1	0 %
5	minä ja sinä pikamuodot	12	12	100 %	30	30	100 %	23	24	96 %	13	13	100 %	8	8	100 %	11	13	85 %	7	7	100 %	7	7	100 %	5	5	100 %	3	3	100 %
6	-i translatiivissa	9	9	100 %	0	15	0 %	12	15	80 %	11	11	100 %	1	11	9 %	0	10	0 %	4	8	50 %	0	9	0 %	3	7	43 %	0	3	0 %
7	-inen adjektiivit	3	3	100 %	0	1	0 %	0	0	-	1	2	50 %	0	3	0 %	0	0	-	0	0	-	0	2	0 %	1	2	50 %	0	2	0 %
8	-i yks. nominatiivissa	4	4	100 %	0	2	0 %	8	8	100 %	10	10	100 %	2	4	50 %	0	0	-	4	4	100 %	0	2	0 %	2	2	100 %	0	8	0 %
9	tämä ja nämä pikamuodot	4	4	100 %	5	10	50 %	20	20	100 %	14	15	93 %	6	11	55 %	5	6	83 %	2	4	50 %	1	3	33 %	3	6	50 %	2	7	29 %
10	mon. 3 .kongruenssi	0	0	-	4	6	67 %	7	7	100 %	10	10	100 %	5	9	56 %	3	11	27 %	6	6	100 %	4	4	100 %	10	17	59 %	2	3	67 %
11	he - ne	0	0	-	0	0	-	0	0	-	1	1	100 %	0	0	-	0	0	-	2	2	100 %	1	2	50 %	0	0	-	0	0	-
12	olla pikamuodot	2	2	100 %	3	12	25 %	6	7	86 %	11	12	92 %	4	4	100 %	0	7	0 %	8	10	80 %	1	4	25 %	0	7	0 %	1	6	17 %
13	mon. 1. kongruenssi	2	2	100 %	3	3	100 %	2	2	100 %	14	14	100 %	6	6	100 %	1	1	100 %	8	8	100 %	2	2	100 %	1	1	100 %	0	1	0 %
14	yks. 1. omistusmuodot	0	0	-	3	5	60 %	1	1	100 %	0	0	-	0	3	0 %	0	0	-	0	0	-	0	2	0 %	1	3	33 %	1	2	50 %
15	-eA, -oa nominit	4	4	100 %	0	7	0 %	1	4	25 %	4	4	100 %	0	2	0 %	1	5	20 %	1	1	100 %	1	9	11 %	2	7	29 %	0	0	-
16	hän - se	0	0	-	3	3	100 %	1	1	100 %	2	2	100 %	0	0	-	2	2	100 %	0	0	-	1	1	100 %	0	0	-	0	0	-
17	-eA, -OA partitiivit	3	5	60 %	0	11	0 %	1	4	25 %	5	6	83 %	1	9	11 %	0	3	0 %	0	5	0 %	0	6	0 %	0	8	0 %	0	2	0 %
18	ts	1	4	25 %	0	1	0 %	2	2	100 %	4	4	100 %	1	8	13 %	1	4	25 %	3	4	75 %	1	1	100 %	0	8	0 %	0	6	0 %
19	hd	1	1	100 %	0	3	0 %	1	2	50 %	2	3	67 %	0	2	0 %	0	5	0 %	1	1	100 %	0	1	0 %	0	10	0 %	0	4	0 %
20	MA-infinitiivin illatiivit	8	11	73 %	1	11	9 %	10	10	100 %	7	7	100 %	0	8	0 %	0	12	0 %	4	9	44 %	1	9	11 %	3	6	50 %	0	28	0 %
21	-A adessiivissa	6	20	30 %	0	29	0 %	1	23	4 %	14	33	42 %	0	49	0 %	0	18	0 %	7	19	37 %	0	20	0 %	0	27	0 %	0	58	0 %
22	-A inessiivissä	1	14	7 %	0	33	0 %	0	25	0 %	8	18	44 %	0	46	0 %	0	20	0 %	9	30	30 %	0	20	0 %	0	31	0 %	0	38	0 %
23	VVdV	0	0	-	0	2	0 %	3	4	75 %	2	3	67 %	0	4	0 %	0	2	0 %	0	1	0 %	0	1	0 %	0	1	0 %	0	0	-
24	-UA partitiivit	4	9	44 %	0	1	0 %	0	2	0 %	8	8	100 %	0	11	0 %	0	9	0 %	0	2	0 %	0	6	0 %	0	8	0 %	0	6	0 %
25	-iA partitiivit	13	18	72 %	0	24	0 %	0	12	0 %	14	19	74 %	0	28	0 %	0	12	0 %	1	15	7 %	0	11	0 %	0	22	0 %	0	25	0 %
	keskiarvo			75 %			21 %			71 %			88 %			28 %			25 %			63 %			26 %			33 %			16 %

City.fi-keskustelupalsta-aineisto

Piirre	#C1	m	1472	#C2	m	1938	#C3	m	1210	#C4	n	1937	#C5	m	1566	#C6	n	1945	#C7	n	990	#C8	n	1455	#C9	n	1276	#C10	m	1870	
#	Selitys	puh.	yht.	%	puh.	yht.	%	puh.	yht.	%	puh.	yht.	%	puh.	yht.	%	puh.	yht.	%	puh.	yht.	%	puh.	yht.	%	puh.	yht.	%	puh.	yht.	%
1	-i supistumaverb. imperf.	0	1	0 %	0	0	-	1	1	100 %	0	1	0 %	0	2	0 %	1	1	100 %	1	1	100 %	1	1	100 %	0	1	0 %	4	6	67 %
2	-i konditionaalissa	0	24	0 %	0	10	0 %	4	9	44 %	5	18	28 %	0	12	0 %	9	22	41 %	10	11	91 %	14	15	93 %	4	14	29 %	6	14	43 %
3	-NUT partisiipit	0	3	0 %	0	4	0 %	0	6	0 %	0	5	0 %	0	4	0 %	1	14	7 %	5	6	83 %	4	7	57 %	0	4	0 %	0	17	0 %
4	-i yks. 3. imperf. diftongi	0	1	0 %	0	1	0 %	1	3	33 %	0	2	0 %	0	4	0 %	0	4	0 %	1	1	100 %	0	0	-	1	6	17 %	1	4	25 %
5	minä ja sinä pikamuodot	0	16	0 %	7	32	22 %	6	6	100 %	19	24	79 %	1	1	100 %	25	26	96 %	11	11	100 %	35	35	100 %	16	22	73 %	59	59	100 %
6	-i translatiivissa	0	9	0 %	0	18	0 %	6	6	100 %	2	11	18 %	0	18	0 %	0	5	0 %	6	6	100 %	2	12	17 %	0	13	0 %	1	17	6 %
7	-inen adjektiivit	0	0	-	0	6	0 %	0	2	0 %	0	1	0 %	0	9	0 %	0	5	0 %	0	0	-	0	0	-	0	2	0 %	0	2	0 %
8	-i yks. nominatiivissa	0	1	0 %	0	1	0 %	0	2	0 %	0	1	0 %	0	1	0 %	2	2	100 %	3	3	100 %	1	2	50 %	2	2	100 %	0	1	0 %
9	tämä ja nämä pikamuodot	0	6	0 %	1	18	6 %	7	21	33 %	7	12	58 %	0	17	0 %	3	19	16 %	7	7	100 %	5	10	50 %	3	6	50 %	15	21	71 %
10	mon. 3 .kongruenssi	0	28	0 %	2	29	7 %	5	12	42 %	12	23	52 %	2	9	22 %	9	16	56 %	5	8	63 %	4	6	67 %	7	17	41 %	5	16	31 %
11	he - ne	0	4	0 %	1	3	33 %	1	2	50 %	3	4	75 %	0	0	-	0	0	-	0	0	-	1	1	100 %	0	0	-	0	4	0 %
12	olla pikamuodot	0	21	0 %	0	24	0 %	3	11	27 %	2	37	5 %	0	8	0 %	4	29	14 %	12	15	80 %	5	24	21 %	1	10	10 %	18	35	51 %
13	mon. 1. kongruenssi	0	1	0 %	0	1	0 %	4	4	100 %	0	0	-	0	0	-	1	3	33 %	2	2	100 %	7	7	100 %	0	0	-	6	8	75 %
14	yks. 1. omistusmuodot	0	0	-	1	7	14 %	0	0	-	2	4	50 %	0	2	0 %	2	17	12 %	0	8	0 %	0	5	0 %	2	8	25 %	5	17	29 %
15	-eA, -oa nominit	0	8	0 %	0	16	0 %	2	4	50 %	2	13	15 %	0	8	0 %	2	13	15 %	5	5	100 %	5	7	71 %	1	7	14 %	0	4	0 %
16	hän - se	0	4	0 %	2	7	29 %	0	0	-	2	4	50 %	0	0	-	0	0	-	0	0	-	0	5	0 %	0	1	0 %	1	4	25 %
17	-eA, -OA partitiivit	0	5	0 %	0	3	0 %	1	8	13 %	2	13	15 %	0	15	0 %	3	11	27 %	0	6	0 %	0	2	0 %	1	6	17 %	0	7	0 %
18	ts	0	7	0 %	0	7	0 %	2	4	50 %	3	13	23 %	0	7	0 %	3	22	14 %	4	5	80 %	0	4	0 %	0	5	0 %	2	8	25 %
19	hd	0	8	0 %	0	13	0 %	0	1	0 %	1	6	17 %	0	8	0 %	0	7	0 %	3	9	33 %	0	3	0 %	1	2	50 %	0	2	0 %
20	MA-infinitiivin illatiivit	0	9	0 %	0	12	0 %	0	4	0 %	0	9	0 %	0	10	0 %	2	21	10 %	1	5	20 %	4	13	31 %	1	7	14 %	0	12	0 %
21	-A adessiivissa	0	27	0 %	0	37	0 %	0	28	0 %	1	72	1 %	0	33	0 %	0	28	0 %	4	26	15 %	0	27	0 %	0	22	0 %	0	31	0 %
22	-A inessiivissä	0	43	0 %	0	47	0 %	1	29	3 %	1	35	3 %	0	34	0 %	0	43	0 %	8	29	28 %	1	15	7 %	0	23	0 %	0	30	0 %
23	VVdV	0	1	0 %	0	0	-	0	3	0 %	0	1	0 %	0	1	0 %	0	2	0 %	0	1	0 %	0	4	0 %	0	1	0 %	0	5	0 %
24	-UA partitiivit	0	3	0 %	0	2	0 %	0	3	0 %	0	6	0 %	0	3	0 %	0	3	0 %	1	3	33 %	0	5	0 %	1	6	17 %	0	6	0 %
25	-iA partitiivit	0	33	0 %	0	19	0 %	0	23	0 %	0	33	0 %	0	31	0 %	0	29	0 %	3	18	17 %	0	24	0 %	0	15	0 %	0	26	0 %
	keskiarvo			0 %			5 %			32 %			20 %			6 %			24 %			61 %			38 %			20 %			22 %