

Mikä nimi on ollut kaikkein suosituin?

Suosiolukumenetelmä etunimien suosionvaihtelun tutkimuksessa

LASSE HÄMÄLÄINEN

1 Johdanto

Tämän katsauksen aiheena on suomalaisten etunimien suosionvaihtelu. Käsittelen ai-
hetta varsin teoreettisesta ja menetelmäkeskeisestä näkökulmasta. Tärkeimpiä kysy-
myksiäni ovat, mitä etunimen suosioksi kutsutulla ilmiöllä tarkoitetaan ja millaisella
tutkimusmenetelmällä se tulisi määrittää.

Kipinän tälle katsaukselle on antanut Maija Sartjärven pro gradu -tutkielma (2011),
joka käsittelee etunimien yleistymistä ja suosionvaihtelua kaksitavuisten *ija-* ja *nja-*
päätteisten naistenimien kautta. Sartjärvi (2013) esittelee ja syventää tutkielmansa tu-
loksia tässä Virittäjässä. Etunimien suosionvaihtelun tutkimusteoriaa käsitellessään
Sartjärvi tekee kiinnostavan havainnon, jonka hän ilmaisee tutkielmansa tiivistelmässä
seuraavasti:

Etunimien suosionvaihtelututkimuksissa on tähän asti tukeuduttu prosentti-
lukuihin: mitä suuremman osuuden vuosittain syntyneistä lapsista kunkin nimen
haltijat muodostavat, sitä suosituimpi nimi on. Tutkielmassa kyseenalaistetaan näin
yksiselitteinen nimen suosion määritelmä. Vuosittain annettujen erilaisten nimien
määrä on kasvanut merkittävästi 1900-luvun puolivälistä lähtien. Samalla vali-
koima, josta nimet valitaan, on kasvanut. Tämä suuri muutos nimenannossa pi-
täisi ottaa huomioon etunimien suosionvaihtelututkimuksissa. Vuoden yleisimmin
valittu nimi 2000-luvulla ei voi saada yhtä suurta prosenttiosuutta kuin vuoden
kärkinimi 1950-luvulla, mutta se ei välttämättä tee nimestä vähemmän suosittua.
(Sartjärvi 2011: tiivistelmä.)

Tästä havainnosta huolimatta Sartjärvi käyttää *Aija-* ja *Anja-*tyyppien suosionvaihtelua tutkiessaan pääosin samaa, prosenttilukuihin perustuvaa menetelmää, johon suomalainen etunimitutkimus perinteisestikin nojaa. Sartjärven (2011: 30–31) esittämän kyseenalaistuksen ja muutaman havainnollisen laskuesimerkin pohjalta voi kuitenkin kysyä, voidaanko etunimen suosiota sen ajallisessa kontekstissa kuvata laske-
malla pelkästään jonkin tietyn nimen saaneiden lasten suhteellinen osuus kaikista tietynä vuonna syntyneistä lapsista.

Esitän tässä katsauksessa etunimien suosionvaihtelun tutkimushistorian ja Sartjärven tutkielman pohjalta kehittämäni mallin, jonka avulla etunimien suosiota voidaan arvioida nimeämiskulttuurissa tapahtuneet muutokset huomioon ottaen. Luvussa 2 esittelen taustaksi etunimien suosionvaihtelun tutkimushistoriaa, ja luvussa 3 esitän kehittämäni mallin sekä perustelut sen käytölle. Lopuksi pohdin lyhyesti, millä tavoin ja millaisiin tarkoituksiin sitä on mahdollista soveltaa käytännön tutkimuksessa.

2 Etunimien suosionvaihtelututkimuksen historiaa

Etunimien tutkimus on ollut pitkään keskeinen osa suomalaista nimistöntutkimusta. Henkilönnimet on perinteisesti ollut nimistöntutkimuksen tutkituin pääkategoria heti paikannimien jälkeen ja etunimet taas henkilönnimien tutkituin alakategoria. 1970- ja 1980-lukujen vaihteessa tutkimusresurssien tarve väheni huomattavasti, kun tutkijat saivat Väestötietokeskuksen nimitietokannat käyttöönsä (Kiviniemi 1982: 7). Suurten ja tarkkojen aineistojen käyttöön saaminen ei kuitenkaan ole vähentänyt etunimitutkimuksen määrää vaan mahdollistanut sen siirtymisen toisenlaisia, teoreettisesti haasteellisempia näkökulmia kohti.

Suomalaisen etunimitutkimuksen merkittävin tutkija Eero Kiviniemi on käsitellyt etunimien suosionvaihtelua monipuolisesti 1980-luvulta 2000-luvulle. Hänen tutkimustensa keskeisin anti on koottu teokseen *Suomalaisten etunimet* (Kiviniemi 2006). Etunimiä on tutkittu myös innovaatioiden diffuusiotutkimuksen näkökulmasta, muihin keksintöihin rinnastettavana kulttuurisena ilmiönä (esim. Mustakallio 1996). Lisäksi oman kirjallisuusperinteensä muodostavat hakuteosmaiset etunimikirjat, joita on ilmestynyt useita (esim. Lempiäinen 2004; Saarikalle & Suomalainen 2007; Teppo & Vilkkuna 1947; Vilkkuna 2005 [1976]).

Tutkimusten teoreettinen puoli on painottunut vahvasti nimien suosionvaihtelun tutkimiseen. Eniten tähän lienee vaikuttanut Kiviniemi, joka keskittyy aihepiiriin paitsi edellä mainitussa teoksessa myös teoksessa *Rakkaan lapsen monet nimet* (1982) sekä etunimien ajallista vaihtelua tarkastelevassa artikkelissaan (2001). Tutkimuskohteina ovat niin yksittäisten nimien suosionvaihtelu yli vuosisadan mittaisella aikavälillä kuin kunkin aikakauden suosikinimet.

Kuten Sartjärvi edellä esitetyssä lainauksessa toteaa, on etunimien suosionvaihtelun tutkimus tähän saakka perustunut tietyn nimen määrän suhteelliseen osuuteen kaikkien annettujen nimien määrästä. Tämän menetelmän on ensimmäisenä ottanut käyttöön Kiviniemi, joka sanoo tarkoittavansa etunimien suosionvaihtelulla ”ajallista ja

alueellista vaihtelua eri etunimien valinnassa” (Kiviniemi 2006: 168). Tämä lyhyt määritelmä ei suoranaisesti ota huomioon nimien antokertoja, vaikka ne käytännössä ovatkin suosionmäärittelyn pohjana.

Tällainen suosion ja suosionvaihtelun määrittely saattaa johtua siitä, että suosion ja antokertojen yhteyttä on helppo pitää itsestään selvänä. Määritelmä perustuu siihen, mitä arkijärki sanelee: nimi, joka valitaan lapselle nimeksi usein, on suosittu. Tätä seikkaa on vaikea kiistää. Toisaalta suosio on ihmisen luoma abstraktio, joka tieteellisessä käytössä kaipaa tuekseen jonkin reaalityodellisuuden nojaavan määritelmän. Ei ole syytä olettaa, että jokainen tekstin vastaanottaja ymmärtää suosion käsitteen samalla tavoin kuin kirjoittaja.

Sartjärven (2011) mukaan suhteellisiin osuuksiin perustuvan määritelmän ongelma on se, että se ei ota huomioon nimeämiskulttuurissa tapahtunutta muutosta. Erilaisten vuosittain annettujen etunimien lukumäärä on ajan saatossa pääosin kasvanut¹, minkä vuoksi nimien on nykyisin aiempaa vaikeampi ”yltää” suuriin suhteellisiin osuuksiin. Näin ollen ei voida yksiselitteisesti todeta, että vuonna 1950 2,5 % kaikista sukupuolelle S annetuista etunimistä kerännyt nimi olisi ollut aikanaan suosittu kuin vuonna 2000 2,0 % kaikista sukupuolelle S annetuista etunimistä kerännyt nimi. (Mts. 30–31.)

Sartjärvi ehdottaa ratkaisuksi toisenlaista tapaa määrittellä nimen suosio: Klassisen todennäköisyysslaskennan periaatteen mukaisesti tietyn nimityypin suosio voidaan määrittellä sen perusteella, mikä on suotuisten alkeistapausten (vuonna *a* annettujen, tyyppiä edustavien erilaisten nimien lukumäärä) suhde kaikkiin alkeistapauksiin (kaikkien vuonna *a* annettujen erilaisten nimien lukumäärä). Esimerkiksi vuonna 1950 on annettu 1 226 erilaista naisennimeä, joista 21 edustaa tyyppiä *-ija*. Todennäköisyys saada *ija*-tyypin nimi vuonna 1950 siis on tämän mallin mukaan $\frac{21}{1226}$ eli noin 1,7 %. Vastaavasti vuonna 2000 on annettu 2 374 naisennimeä, joista yhdeksän edustaa *ija*-tyyppiä. Näin todennäköisyys saada *ija*-tyypin nimi vuonna 2000 on $\frac{9}{2374}$ eli noin 0,4 %. (Sartjärvi 2011: 30–31.)

Tämä laskentamalli sisältää kuitenkin kaksi ongelmaa. Ensinnäkään se ei sovellu yksittäisten nimien suosion tarkasteluun. Jos sen avulla halutaan selvittää esimerkiksi *Aija*-nimen suosiota vuosina 1950 ja 2000, ovat luvut $\frac{1}{1226}$ ja $\frac{1}{2374}$. Tämän perusteella *Aija*-nimen saaminen on ollut todennäköisempää vuonna 1950 kuin vuonna 2000. Tästä voitaneen päätellä, että *Aija* on ollut suosittu nimi vuonna 1950 kuin vuonna 2000, mikä on toki totta. Jos tilalle vaihdetaan esimerkiksi 2000-luvun suosikinimi *Veera*, joka oli edellisen vuosisadan puolivälissä varsin harvinainen, todennäköisyydet ovat kuitenkin samat kuin nimen *Aija* kohdalla: vuonna 1950 $\frac{1}{1226}$ ja vuonna 2000 $\frac{1}{2374}$. Silti ei voitane päätellä, että *Veera* on ollut suosittu nimi vuonna 1950 kuin vuonna 2000.

Sartjärvi ei välttämättä ole tarkoittanut klassisen todennäköisyysslaskennan mallia sovellettavaksi yksittäisten nimien suosionvaihtelun tutkimiseen. Edellä esitetty ongelma ei suoranaisesti estäkään sen käyttöä nimiryhmien suosion tarkastelussa. Mallin toinen ongelma sen sijaan koskettaa niin yksittäisten nimien kuin nimimuottienkin

1. Vuosien 1930–1940 paikkeilla erilaisten vuosittain annettujen nimien määrä tilapäisesti laski, kuten taulukoista 1 ja 2 jäljempänä voidaan havaita.

suosion tarkastelua. Siinä jokainen erilainen nimi huomioidaan samanlaisena, todennäköisyyteen ja näin nimen suosioon yhtä lailla vaikuttavana tapauksena:

Klassisen todennäköisyyslaskennan periaatteen mukaan **oletetaan**, että perusjoukko on äärellinen ja **että kaikki alkeistapahtumat esiintyvät yhtä usein** eli ovat symmetrisiä [Holopainen – Pulkkinen 2006: 103]. – – Alkeistapahtumat ovat symmetrisiä, kun **oletetaan, että kaikki nimivalinnat ovat yhtä todennäköisiä**. (Sartjärvi 2011: 30; lihavointi tämän katsauksen kirjoittajan.)

On ongelmallista olettaa, että kaikki nimivalinnat ovat yhtä todennäköisiä. Nimivalinnan lähtötilanteessa toki on näin: teoriassa kaikilla keksityillä (tai toistaiseksi keksimättömillä) nimillä on yhtä suuri todennäköisyys tulla valituksi. Käytännössä tätä oletusta kuitenkin häiritsevät erilaiset tekijät, kuten erikoisuutta tavoittelevien nimenantajien taipumus samojen muotinimien valitsemiseen (Kiviniemi 1982: 169) tai yksinkertaisesti se, että kaikki nimenantajat eivät tunne kaikkia mahdollisia nimiä, jolloin vähemmän tunnettujen nimien todennäköisyys tulla valituksi pienenee ja tunnetumpien vastaavasti kasvaa. Nimervalinta ei ole prosessi, joka suoritetaan matemaattisessa tyhjiössä, vaan siihen vaikuttaa suuri määrä erilaisia tekijöitä.

Käytännössä ongelma ilmenee seuraavalla tavalla: Poimitaan satunnaisesti tiettyä vuodelta erittäin suosittu ja erittäin harvinainen nimi ja lasketaan niiden todennäköisyydet tulla valituksi. Esimerkiksi vuoden 2000 tyttöjen kärkinimen *Jenna* ja satunnaisen uniikinimen *A* todennäköisyys on sama, $\frac{1}{2374}$. On kuitenkin ilmeistä, että nimiä ei voida tämän perusteella pitää yhtä suosittuina. Tämä ongelma ei koske pelkästään yksittäisiä nimiä. Voidaan kuvitella nimityypit *B* ja *C*. Nimityyppi *B* sisältää viisi erittäin usein annettua suosikkiniemeä. Tyyppi *C* sisältää 20 uniikinimeä. Klassisen todennäköisyyslaskennan kaavan mukaan tyyppi *C* on näin neljä kertaa suosittumpi kuin tyyppi *B*, vaikka tyyppiä *B* edustavien nimien esiintymien määrä reaali maailmassa onkin selvästi suurempi.

Edellä esitetyistä syistä Sartjärven mallia on vaikea soveltaa sellaisenaan etunimien suosion tutkimukseen. Mallin perusidea on kuitenkin lähtökohdiltaan oikeansuuntainen ja tarjoaa erinomaisen pohjan lisäpohdintoille.

3 Suosiolukumenetelmä

Sekä perinteisestä etunimien suosion tutkimuksen mallista että Sartjärven mallista on löydettävissä ongelmakohtia. Toisaalta molemmissa on myös hyvät puolensa, eikä niitä ole syytä hylätä kevyin perustein. Kuvaan seuraavaksi, millä tavoin näiden menetelmien hyvät puolet voisi yhdistää.

Aiemman tarkastelun perusteella perinteinen, kaikkien yksittäisten nimiesiintymien osuukien tarkasteluun pohjautuva menetelmä on vähemmän ongelmallinen. Se kuitenkin tarvitsee korjauskertoimen, joka ottaa huomioon annettujen nimien määrässä tapahtuneen ajallisen vaihtelun ja näin korjaa vaihtelusta aiheutuvan prosenttilukujen vääristymän.

Perinteisessä mallissa nimen suosiota kuvaava suhteellinen osuus x lasketaan seuraavasti:

$$x = \frac{A_\beta}{N_\beta}$$

A_β = Tiettyinä vuonna a annettujen nimien A lukumäärä (kpl)

N_β = Kaikkien tiettyinä vuonna a annettujen nimien lukumäärä (kpl)

A = Suosion tarkastelun kohteena oleva nimi tai nimiryhmä

Sartjärven mallissa todennäköisyys y tietyn nimen tai nimiryhmän antamiselle lasketaan seuraavasti:

$$y = \frac{A_\gamma}{N_\gamma}$$

A_γ = Tiettyinä vuonna a annettujen **erilaisten** nimien A lukumäärä (kpl)

N_γ = Kaikkien tiettyinä vuonna a annettujen **erilaisten** nimien lukumäärä (kpl)

A = Suosion tarkastelun kohteena oleva nimi tai nimiryhmä

Näiden tarkastelutapojen ero on se, että siinä missä perinteisessä mallissa lasketaan nimien A yleisyys reaaliaimailmassa, Sartjärven mallissa lasketaan nimien A yleisyys selaisessa hypoteettisessa tilanteessa, jossa kaikkia nimiä annettaisiin yhtä monta kertaa.

Nimien todelliseen suosioon on mahdollista päästä käsiksi vertailemalla lukuja x ja y keskenään. Kuinka suuri on nimenannon reaalistuma, käytännön nimiesiintymien määrä suhteessa siihen, kuinka paljon nimiä tulisi olla klassisen todennäköisyyden mukaan? Jos nimiesiintymien määrä on käytännössä suurempi kuin sen teoriassa tulisi olla, eikö nimeä voida tällöin pitää ainakin jossain määrin suosittuna?

Lukuja voidaan vertailla keskenään jakolaskulla. Näin ollen esitän, että nimen suosio vuonna a voidaan kuvata seuraavasti:

$$\alpha = \frac{x}{y} = \frac{A_\beta}{N_\beta} \div \frac{A_\gamma}{N_\gamma}$$

α = Nimen suosio vuonna a

x = Nimien A osuus kaikista annetuista nimistä tiettyinä vuonna a

y = Todennäköisyys nimien A antamiselle tiettyinä vuonna a

A_β = Tiettyinä vuonna a annettujen nimien A lukumäärä (kpl)

N_β = Kaikkien tiettyinä vuonna a annettujen nimien lukumäärä (kpl)

A_γ = Tiettyinä vuonna a annettujen **erilaisten** nimien A lukumäärä (kpl)

N_γ = Kaikkien tiettyinä vuonna a annettujen **erilaisten** nimien lukumäärä (kpl)

A = Suosion tarkastelun kohteena oleva nimi tai nimiryhmä

Mitä kaava tuottaa lopputulokseksi? Kaikkien α :n tekijöiden yksikkö on sama (kpl), joten yksiköt supistuvat jakolaskuissa pois. Jäljelle jää pelkkä positiivinen rationaali-

luku², joka itsessään ei kerro paljoakaan. Jos luku on suurempi kuin 1 eli osoittaja on suurempi kuin nimittäjä, nimi on käytännössä yleisempi kuin sen tulisi olla, mikäli kaikki nimivalinnat olisivat yhtä todennäköisiä. Vastaavasti jos luku on pienempi kuin 1, nimi on käytännössä harvinaisempi kuin sen klassisen todennäköisyyden mukaan tulisi olla. Yksinkertaistaen voidaan todeta, että mitä suurempi suosioluku on, sitä suositumpi nimi on.³ Ilman tarkempaa tuntumaa on kuitenkin vaikea sanoa, onko esimerkiksi suosioluvun 10 saava nimi melko suosittu vai erittäin suosittu.

Tämä ei kuitenkaan aiheuta ongelmaa, sillä kaava on muodostettu nimien suosion vertailuun, ei sen absoluuttiseen arvioimiseen. Suosioluku tarvitsee analyysin tueksi muita suosiolukuja, joihin sitä verrataan. Voidaan esimerkiksi laskea suosioluvut yhden nimen osalta sadan vuoden ajalta ja lukuja vertailemalla tutkia, milloin nimi on ollut suosituimmillaan. Esittämäni malli toimii lähinnä omassa, sisäisessä järjestelmässä, ja sen ulkopuolella sillä ei ole juurikaan informaatioarvoa. Tätä voi pitää mallin heikkoutena, mutta toisaalta vuosittain annettujen nimien määrän huomioon ottavaa, jo itsessään ymmärrettävän lopputuloksen antavaa mallia ei liene mahdollista luoda. Saman kritiikin voisi esittää myös perinteistä mallia kohtaan: voiko etunimitutkimusta tuntematon ihminen intuitiivisesti ymmärtää, miten yleinen on nimi, joka annetaan yhdelle prosentille sukupuolta S olevista syntyneistä lapsista?

Edellä esitetyt huomiot käsittelevät tapauksia, joissa käsittelyn kohteena oleva kaavan A on yksi nimi. A voi kuitenkin sisältää myös useamman nimen. Tutkittavana ryhmänä voivat olla yhtä lailla *ija-* ja *nja-*tyyppien nimet, kaikki T -kirjaimella alkavat nimet tai vaikkapa kaikki kyseisenä vuonna sukupuolta S edustaville lapsille annetut nimet (jolloin lopputulos tosin on 1). Saatava luku on tutkittavan ryhmän nimien keskiarvo, joten päätelmien tekeminen sen perusteella saattaa olla vaikeampaa. On toki hyvä huomioida, että ryhmään kuuluvien nimien määrän kasvun myötä myös laskutoimitus vaikeutuu. Teknisten apuvälineiden ansiosta tämä tosin ei ole enää nykyisin ylitsepääsemätön ongelma.

Uuden tutkimusmenetelmän esittelyn yhteydessä on monesti tarpeen ottaa käyttöön uutta terminologiaa. Jo aiemmin tässä luvussa olen käyttänyt termejä *suosioluku* ja *suosiolukumenetelmä*. *Suosioluvulla* tarkoitan lukua, jonka esittämäni kaava tuottaa lopputulokseksi ja jonka tarkoitus on siis kuvata numeerisesti tietyn nimen suosiota tietyssä vuonna. Käytän suosioluvun lyhenteenä kirjainta α . Menetelmää kokonaisuudessaan kutsun yksinkertaisesti *suosiolukumenetelmäksi*.

2. Kaikkien tekijöiden on oltava positiivisia kokonaislukuja, joten niiden osamäärien osamäärän on näin oltava positiivinen rationaaliluku.

3. Vertailu olisi mahdollista suorittaa myös toisinpäin, vaihtamalla x nimittäjäksi ja y osoittajaksi. Tällöin tilanne kuitenkin kääntyisi pääläelleen myös lukujen tulkinnan suhteen: nimi olisi sitä suositumpi, mitä pienempi sen suosioluku olisi. Nähdäkseni on selkeämpää, että suuri suosioluku kuvaa suurta suosiota ja pieni suosioluku pientä suosiota.

4 Käytännön sovelluksia

Uudenlaisen tutkimusmenetelmän kehittämisestä ei ole juurikaan hyötyä, mikäli menetelmä ei toimi. Niinpä olen testannut menetelmää pienellä tutkimuksella.

Minua on jo pitkään kiinnostanut kysymys siitä, mikä nimi on ollut oman suosionsa huipulla Suomen historian suosituin, jos huomioon otetaan kaikki nimet kaikkina aikoina. Tarkkaan tilastoidun historian osalta, 1800-luvun lopulta alkaen, kysymykseen on voitu perinteisesti antaa vastaukseksi 1960-luvun loppupuolelta nimet *Sari* ja *Mika*. *Sarin* osuus syntyneistä tyttölapsista oli vuonna 1966 7,4 prosenttia, ja *Mikan* osuus vuoden 1969 poikalapsista 7 prosenttia (Kiviniemi 2006: 24–25).

Näiden nimien huippuvuodet olivat kuitenkin hyvin erilaisen nimeämiskulttuurin aikaa. Nimen yksilöllisyys ei ollut tuolloin lainkaan yhtä muodikasta kuin nykyään, joten harvinaisia ja uniikkeja nimiä oli vähemmän, samoin annettuja nimiä ylipäänsä. Kaikkia nykyisin annettuja nimiä ei ollut tuolloin keksittykään, joten voidaan perustellusti sanoa, että nimillä oli suurempi todennäköisyys tulla valituksi kuin nykyään.

Tämä epätasapainoisuus perinteisessä laskentatavassa on korjattavissa suosiolukumenetelmän avulla. Taulukoissa 1–2 ja kaavioissa 1–3 olen tarkastellut tyttöjen ja poikien kärkinimien suosiota aikavälillä 1920–2000 sekä prosentuaalisten osuuksien että suosiolukujen valossa.

Taulukko 1.

Tyttöjen vuosittaiset kärkinimet 1920–2000.

| Vuosi | Kärkinimi | Annettu | Syntyneitä | Eri nimiä | Suosio (%) | Suosioluku |
|-------|-------------|---------|------------|-----------|------------|-------------|
| 1920 | Aino | 1 165 | 30 340 | 922 | 3,8 | 35,4 |
| 1925 | Aino | 1 057 | 29 842 | 926 | 3,5 | 32,8 |
| 1930 | Eila | 1 268 | 30 503 | 868 | 4,2 | 36,1 |
| 1935 | Anja | 1 161 | 28 712 | 886 | 4,0 | 35,8 |
| 1940 | Ritva | 1 433 | 27 377 | 824 | 5,2 | 43,1 |
| 1945 | Tuula | 1 760 | 40 972 | 1 164 | 4,3 | 50,0 |
| 1950 | Pirjo | 1 891 | 43 884 | 1 146 | 4,3 | 49,4 |
| 1955 | Pirjo | 1 676 | 40 357 | 1 153 | 4,2 | 47,9 |
| 1960 | Päivi | 1 773 | 38 073 | 1 276 | 4,7 | 59,4 |
| 1965 | Sari | 2 349 | 36 815 | 1 434 | 6,4 | 91,5 |
| 1970 | Sari | 1 732 | 31 239 | 1 731 | 5,5 | 96,0 |
| 1975 | Katja | 1 157 | 31 513 | 1 759 | 3,7 | 64,6 |
| 1980 | Heidi | 1 027 | 30 824 | 1 814 | 3,3 | 60,8 |
| 1985 | Laura | 902 | 30 367 | 2 098 | 3,0 | 62,3 |
| 1990 | Laura | 844 | 31 244 | 2 545 | 2,7 | 68,7 |
| 1995 | Laura | 687 | 29 327 | 2 540 | 2,3 | 59,5 |
| 2000 | Jenna | 438 | 25 963 | 2 394 | 1,7 | 40,4 |

Taulukko 2.

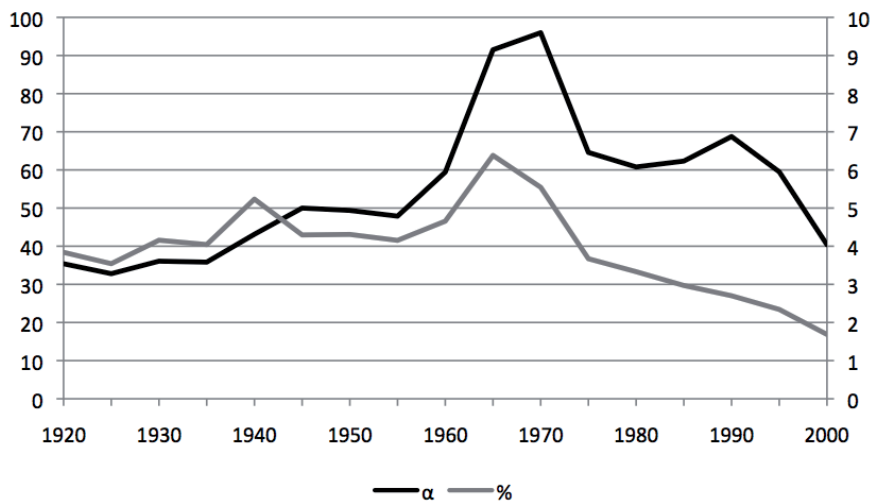
Poikien vuosittaiset kärkinimet 1920–2000.

| Vuosi | Kärkinimi | Annettu | Syntyneitä | Eri nimiä | Suosio (%) | Suosioluku |
|-------|-------------|---------|------------|-----------|------------|-------------|
| 1920 | Veikko | 1 129 | 25 299 | 794 | 4,5 | 35,4 |
| 1925 | Veikko | 1 226 | 28 539 | 815 | 4,3 | 35,0 |
| 1930 | Pentti | 1 480 | 31 074 | 858 | 4,8 | 40,9 |
| 1935 | Pentti | 1 298 | 29 863 | 739 | 4,3 | 32,1 |
| 1940 | Matti | 1 182 | 28 024 | 689 | 4,2 | 29,1 |
| 1945 | Matti | 2 004 | 43 119 | 776 | 4,6 | 36,1 |
| 1950 | Markku | 1 991 | 45 639 | 758 | 4,4 | 33,1 |
| 1955 | Kari | 2 096 | 42 366 | 717 | 4,9 | 35,5 |
| 1960 | Jari | 2 203 | 39 572 | 794 | 5,6 | 44,2 |
| 1965 | Jari | 2 582 | 38 351 | 923 | 6,7 | 62,1 |
| 1970 | Mika | 2 271 | 32 608 | 1 154 | 7,0 | 80,4 |
| 1975 | Jani | 1 750 | 33 296 | 1 298 | 5,3 | 68,2 |
| 1980 | Mikko | 1 563 | 32 327 | 1 304 | 4,8 | 63,0 |
| 1985 | Antti | 1 105 | 31 565 | 1 465 | 3,5 | 51,3 |
| 1990 | Ville | 1 013 | 32 624 | 1 784 | 3,1 | 55,4 |
| 1995 | Ville | 803 | 30 621 | 1 837 | 2,6 | 48,2 |
| 2000 | Aleksi | 650 | 27 682 | 1 799 | 2,3 | 42,2 |

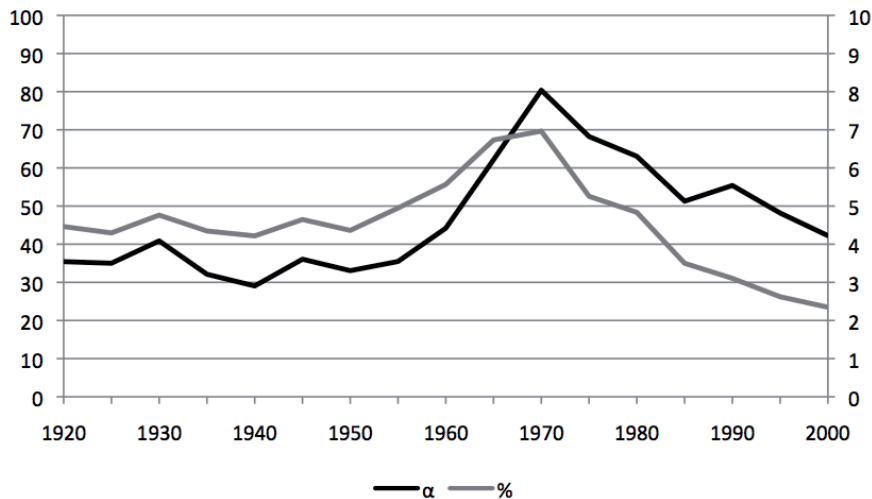
Tutkimusaineisto on kerätty Kotimaisten kielten keskuksen hallinnoimista Väestörekisterikeskuksen aineistoista (VRK). Tutkimusaineisto on rajattu siten, että mukana ovat kaikki suomenkielisille lapsille syntymässä annetut ensimmäiset etunimet. Kiviniemi (2006: 21) on määritellyt erilaisiksi etunimiksi kaikki eri tavoin kirjoitetut nimiasut, ja on käytännöllistä pitäytyä tässä määritelmässä. Olen laske-
nut tiedot vain joka viidenneltä vuodelta, sillä nimeämishistoriassa tapahtuvat muutokset sekä eroavaisuudet suosionvaihtelun määrittämisen menetelmissä käyvät ilmi näillä tiedoin. Tutkittava aikaväli 1920–2000 on valittu sen perusteella, että VRK:n aineistossa ovat vain kaikki vuonna 1965 elossa olleet ja sen jälkeen syntyneet suomalaiset, ja aineisto loppuu vuoden 2002 alkuun. Vuotta 1920 edeltävät tiedot olen katsonut jokseenkin puutteellisiksi, sillä osa 1900-luvun alkuvuosina syntyneistä henkilöistä on jo ehtinyt kuolla ennen vuotta 1965. Tällä ei välttämättä olisi merkitystä näiden vuosien kärkinimien suosiolukuihin, mutta toisaalta tarkastelemani aikaväli lienee riittävän laaja.

Mitkä nimet siis ovat olleet omana aikanaan kaikkein suosituimmat? Kuten taulukoista 1 ja 2 nähdään, vastaukset tähän kysymykseen eivät muutu siirtymällä suhteellisten osuuksien tarkastelusta suosiolukumenetelmään. *Sarin* (1965 ja 1970) ja *Mikan* (1970) huippuvuodet ovat keränneet selvästi suurimmat suosioluvut. *Sari* on annettu nimeksi vuonna 1970 syntyneille tyttölapsille peräti 96 kertaa useammin kuin sitä annettaisiin kaikkien nimien ollessa yhtä todennäköisiä.

Kaavioihin 1 ja 2 olen sijoittanut kulkemaan vuosittaisten kärkinimien suosioikäyrät sekä prosenttiosuuksien että suosiolukujen perusteella. Niillä ei toki ole samaa yksikköä tai samansuuruisia lukuarvoja, mutta sovitetuissa kahden pystyakselin kaavioissa voidaan kuvaajien kulkusuuntia vertailemalla tarkastella sitä, millaisia eroja eri menetelmin saaduissa tuloksissa on.



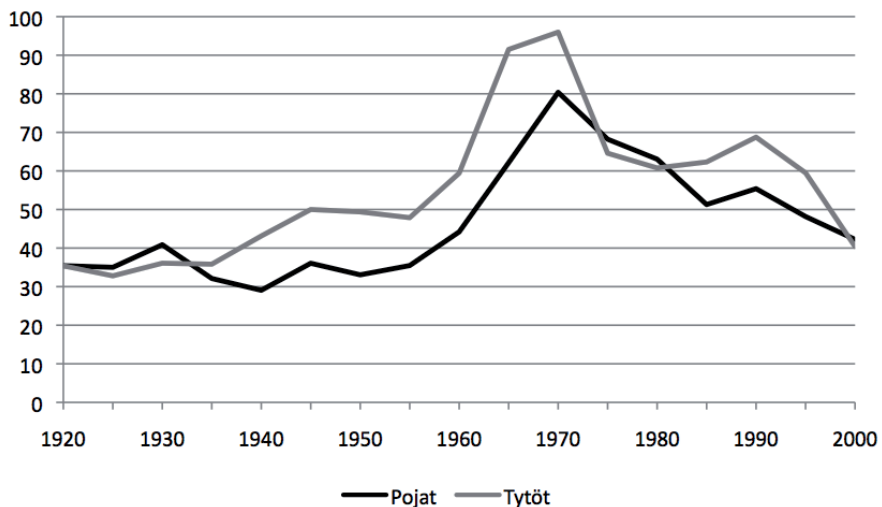
Kaavio 1.
 Tyttöjen kärkinimien suosio 1920–2000 prosentuaalisesti ja suosiolukujen perusteella.



Kaavio 2.
 Poikien kärkinimien suosio 1920–2000 prosentuaalisesti ja suosiolukujen perusteella.

Sekä tyttöjen että poikien kärkinimissä kuvaajien kulku on pääosin samansuuntainen. Vuosisadan alkupuolen prosenttiosuuksia kuvaavat käyrät kulkevat jokseenkin samansuuntaisesti suosiolulukäyrien kanssa, mutta 1960- ja 1970-lukujen taitteessa ne taipuvat selvästi suosiolulukäyrien alle. Tämä tarkoittaa sitä, että siinä missä prosenttein mitattuna vuosittaisten kärkinimien suosio on romahtanut sitten 1970-luvun, ei muutos ole suosioluvuin mitattuna lainkaan yhtä suuri. Menetelmien erot ovat siis selvät. Poimittakoon esimerkiksi tyttöjen kärkinimet vuosilta 1940 ja 1990: Nimien *Ritva* ja *Laura* suhteelliset osuudet ovat olleet 5,2 % ja 2,7 % (kaavio 1). Näiden lukujen perusteella vaikuttaa siltä, että *Ritva* on ollut vuonna 1940 lähes kaksi kertaa suosituampi kuin *Laura* vuonna 1990. Suosiolukujen perusteella (ks. taulukko 1) tilanne kuitenkin kääntyy päinvastaiseksi: *Lauran* luku on 68,7 ja *Ritvan* 43,1. Tyttöille vuosittain annettujen eri nimien määrä tarkastellulla aikavälillä on lähes kolminkertaistunut, ja kasvaneeseen nimivalikoimaan nähden *Lauran* suosio vuonna 1990 onkin merkittävämpi kuin *Ritvan* vuonna 1940.

Keskeisimmät huomiot kaavioista 1 ja 2 voidaan siis tehdä niissä kohdin, joissa kuvaajat eivät kulje samansuuntaisesti. Sekä tyttöjen että poikien suosiolulukäyrissä on vuonna 1990 selvästi havaittava nousu, mutta prosenttikäyrät jatkavat tasaista laskuaan. Ilmiön taustalla on se, että vuosien 1985 ja 1990 välillä erilaisten annettujen nimien määrä on lisääntynyt huomattavasti. Samankaltainen poikkeama on tyttöjen nimissä vuosien 1940 ja 1945 välillä: kärkinimien prosentuaalinen osuus laskee, mutta suosioikäyrä jatkaa nousuaan annettujen nimien määrässä tapahtuneen huomattavan kasvun vuoksi.



Kaavio 3.

Tyttöjen ja poikien kärkinimien suosio 1920–2000 suosiolukujen perusteella.

Kaavioon 3 olen koonnut sekä tyttöjen että poikien etunimien suosiolulukäyrät vertaillakseni niitä keskenään. Jo aiemmin on tiedetty, että tytöille annetaan enemmän erilaisia nimiä (Kiviniemi 2006: 32–35), kun taas erot tyttöjen ja poikien kärkinimien suh-

teellisissä osuuksissa ovat varsin pienet (taulukot 1 ja 2). Näin ollen on ilmeistä, että tyttöjen suosioluvut ovat pääosin jonkin verran poikien lukuja suurempia. Poikkeuksena on kuitenkin 1970–1980-luku, jolloin poikien kärkinimien suhteelliset osuudet pienenevät selvästi tyttöjen kärkinimiä hitaammin, sekä 1920-luku, jolloin tyttöjenkin nimivalikoima oli vielä varsin suppea ja toisaalta kärkinimien suhteelliset osuudet melko pieniä.

5 Lopuksi

Tässä katsauksessa olen pyrkinyt osoittamaan, että etunimien suosionvaihtelun tutkimuksen kannalta ei ole yhdentekevää, millä tavoin suosio määritellään. Kuten Sartjärvin (2013: 66) toteaa, käsitteitä ”suosittu” ja ”yleinen” käytetään lähes synonyymisesti ja niiden välille tulisi tehdä selvä ero. Määrittelyn tekeminen ei kuitenkaan ole helppoa, eikä sitä tulekaan tehdä liian yksinkertaisesti. Määrittelyssä törmätään nimistöntutkimuksen ulkopuolellekin ulottuviin, käsitteellisen ajattelun aiheuttamiin ongelmiin: Mitä suosiolla oikeastaan tarkoitetaan? Mitkä tekijät vaikuttavat siihen? Tulisiko meidän puhua suosion ja suosionvaihtelun sijaan neutraalimmin esimerkiksi frekvenssistä ja sen vaihtelusta?

Tiedossani ei ole akateemisia lähteitä, joissa *suosio* määriteltäisiin yksiselitteisesti. Kenties sellaisia ei olekaan. *Kielitoimiston sanakirja* osoittaa käsitteen monivivahteisuuden:

suosio³ 1. erityinen suopeus, myötämielisyys, hyvántahtoisuus jkta t. jtak kohtaan; suosiossa olo, suosittuus. – 2. *Suosiolla* sovinnolla, hyvällä, vastustelematta. – (KS s. v. *suosio*.)

Asiaa on ehkä helpompi lähestyä käytännönläheisen esimerkin avulla: Jos sanomme, että karkkipussi *X* on suosittu, tarkoitamme oletettavasti sitä, että monet ihmiset pitävät *X*:stä ja siksi kaupan karkkihyllyllä valitsevat sen. On kuitenkin ilmeistä, että *X*:n suosittuuteen vaikuttaa sen ostavien ihmisten lukumäärän lisäksi se, kuinka paljon kilpailevia tuotteita karkkihyllyllä on. Jos *X* olisi ainoa markkinoilla oleva karkkipussi ja sitä ostettaisiin paljon, ei välttämättä olisi erityisen mielekästä kutsua *X*:ää suosituksi. Enemminkin voitaisiin sanoa, että karkit ovat suosittuja. Niinpä mikäli *Matti* olisi ainoa olemassa oleva miehennimi ja se annettaisiin kaikille poikalapsille, sanoisimme pikemminkin nimeämisen olevan suosittua kuin nimen *Matti* olevan suosittu.

Karamelliesimerkin ja nimeämisen analogia osoittaa, että suosio on varsin hyvä termi kuvaamaan etunimien esiintymistiheyttä. On kuitenkin syytä pohtia, onko järkevintä pitäytyä yksinkertaisessa määrittelyssä, kun Väestörekisterikeskuksen tietojen pohjalta on mahdollista muodostaa useampiakin näkökulmia huomioon ottavia määritelmiä. Suosiolukumenetelmä ei varmastikaan ole ainoa mahdollinen tapa uudistaa käsitystämme etunimien suosiosta: pidän sitä ennemminkin keskustelunavauksena mahdolliselle tulevalle metodologiselle pohdinnalle.

On mielenkiintoista pohtia, voiko tämänkaltaista teoreettisen ja käytännöllisen esiintymistiheyden yhdistävää menetelmää käyttää myös nimistöntutkimuksen ulkopuolella. Suosion tutkimus lienee keskeisemmässä asemassa esimerkiksi kauppa- kuin kielitieteiden alalla. En ole tässä katsauksessa voinut tarkastella muilla tieteenaloilla

vastaavanlaisissa tutkimuksissa käsitellyjä menetelmiä, joten en tiedä, onko esittämäni kaltainen malli jo käytössä muualla. Uskoisin sen kuitenkin olevan käyttökelpoinen muissakin suosiontutkimuksissa, joissa tutkittavan ilmiön esiintymiseen vaikuttaa erilaisten vaihtoehtojen määrä ja niiden vaihtelu ajan saatossa.

Toisaalta on hyvä muistaa käsitteen monikäyttöisyys: jos tutkimuskohteena on esimerkiksi ihmisyksilöiden suosio (esim. Hakkarainen 1980; Soutamo 2007), olisi matemaattisen mallin soveltaminen arviointia tukevien lukuarvojen asettamisen kannalta hankalaa ja eettisesti arveluttavaa. Suosiolukumenetelmä on tarkoitettu kuvaamaan tilanteita, joissa entiteetin suosiota määrittää ihmisten kvantitatiivinen suhtautuminen siihen, toisin sanoen entiteetin yleisyys.

Lähteet

- HAKKARAINEN, PENTTI 1980: *Suosio ja torjunta koululuokassa. The most and least popular pupils in the class*. Jyväskylä: Kasvatustieteiden tutkimuslaitos.
- HOLOPAINEN, MARTTI – PULKKINEN, PEKKA 2006: *Tilastolliset menetelmät*. Helsinki: WSOY.
- KIVINIEMI, EERO 1982: *Rakkaan lapsen monet nimet. Suomalaisen etunimet ja nimenvalinta*. Espoo: Weilin & Göös.
- 2001: Etunimien ajallisen vaihtelun selvittelyä. – Kaija Mallat, Terhi Ainiala & Eero Kiviniemi (toim.), *Nimien maailmasta* s. 214–235. Kieli 14. Helsinki: Helsingin yliopiston suomen kielen laitos.
- 2006: *Suomalaisen etunimet*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- KS = *Kielitoimiston sanakirja. Osa 3*. Eija-Riitta Grönros (päätoim.), Minna Haapanen, Tarja Riitta Heinonen, Leena Joki, Liisa Nuutinen & Marjatta Vilkkamaa-Viitala. Helsinki: Kotimaisten kielten tutkimuskeskus 2006.
- LEMPIÄINEN, PENTTI 2004: *Suuri etunimikirja*. 3., tarkistettu painos. Helsinki: WSOY.
- MUSTAKALLIO, SARI 1996: *Sisko ja sen Veli*. – Eero Kiviniemi & Sari Mustakallio (toim.), *Nimet, aatteet, mielikuvat. Kolme näkökulmaa nimiin* s. 9–98. Kieli 11. Helsinki: Helsingin yliopiston suomen kielen laitos.
- SAARIKALLE, ANNE – SUOMALAINEN, JOHANNA 2007: *Suomalaiset etunimet Aadasta Yrjöön*. Helsinki: Gummerus.
- SARTJÄRVI, MAIJA 2011: *Aija ja Meija, Anja ja Minja. Kaksitavuisten ija- ja nja-päätteisten naistenimien yleistymisen ja suosionvaihtelu*. Pro gradu -tutkielma. Helsingin yliopiston suomen kielen, suomalais-ugrilaisen ja pohjoismaisten kielten ja kirjallisuuksien laitos.
- 2013: Kahden etunimityypin yleistymisen ja suosionvaihtelu. – *Virittäjä* 117 s. 44–69.
- SOUTAMO, SAARA 2007: *Henkisyttä ja karismaa. Dalai laman suosio suomalaisten kuulijoiden näkökulmasta*. Pro gradu -tutkielma. Helsinki: Helsingin yliopiston uskontotieteen laitos.
- TEPPO, HANNES – VILKUNA, KUSTAA 1947: *Etunimikirja*. Tietolipas 5. Helsinki: Suomalaisen Kirjallisuuden Seura.
- VILKUNA, KUSTAA 2005 [1976]: *Etunimet*. 4., uudistettu laitos. Toimittanut Pirjo Mikkonen. Helsinki: Otava.

Kirjoittajan yhteystiedot (address):
etunimi.j.sukunimi@helsinki.fi