



FRED KARLSSON ja MATTI WIBERG

# Puolueohjelmien kieliopillinen kompleksisuus

## *Johdanto*

Millaisia *tekstejä* eduskuntapuolueiden periaateohjelmat<sup>1</sup> ovat? Artikkelissamme Karlsson ja Wiberg (2010) olemme analysoineet Suomen kahdeksan eduskuntapuolueen ohjelmia leksikaalisesti: mitä sanoja eri puolueet käyttävät (tai eivät käytä) puolueohjelmissaan ja mitä johtopäätöksiä tästä voidaan tehdä ohjelmien samanlaisuudesta tai erilaisuudesta. Nyt käsillä olevan artikkelin tavoitteena on analysoida ohjelmien kieliopillista, erityisesti morfologista ja syntaktista kompleksisuutta.

Puoluelain (10/1969, §§1–2) mukaan puolueella tarkoitetaan oikeusministeriössä pidettävään puolueräkisteriin merkittyä rekisteröityä yhdistystä, jolla on säännöistä ilmenevän tarkoituksen toteuttamiseksi laadittu, yhdistyksen valtiollisessa toiminnassa noudatettavat periaatteet ja tavoitteet ilmaiseva yleisohjelma. Ohjelma on oleellinen puolueen luonteen elementti.

Puolueohjelmia on monenlaisia. Olavi Borg (1965, 17–19) jakaa ne arvoperiaatteita ilmentäviin periaateohjelmiin, lyhyehköä ajanjaksoa koskeviin, konkreettisiin ja monesti yksityiskohtaisia toimintatavoitteita esitteleviin ja propagoiviin toimintaohjelmiin sekä periaate- ja toimintaohjelman välimuotoon sijoittuviin yleisohjelmiin. Ohjelmat ovat puolueiden ideologisten sitoumusten ja periaatteitten kirjauksia, jotka ovat läpikäyneet laajan järjestökäsittelyn ja yleensä varsin laajan jäsenkeskustelun.

Ohjelmien nimet, tavoitteet, hyväksymisajankohdat sekä jatkossa käytettävät lyhenteet ovat aakkosjärjestyksessä seuraavat:

KANSALLISEN KOKOOMUKSEN PERIAATEOHJELMA. Hyväksytty Joensuun puoluekokouksessa 9.–11.6.2006. (Kok)

KD:N PERIAATEOHJELMA. Suomen Kristillisdemokraatit r.p. 10.11.2006.

<sup>1</sup> Käytämme Yhteiskuntatieteellisen tietoarkiston Pohtiva-puolueohjelmatietokantaan arkistoituja tuoreimpia periaateohjelmia.

Kristillisdemokraattien tavoiteohjelma vaalikaudelle 2007–2011. (KD)

KESKUSTAN PERIAATEOHJELMA. Kohti tasa-arvoista, luonnonmukaista sivistisyhteiskuntaa. Hyväksytty Oulun puoluekokouksessa 10.6.2006. (Kesk)

PERUSSUOMALAISTEN PERIAATEOHJELMA. Perussuomalaisten lähiajan tavoiteohjelma. Kokkola 2005. (PS)

RKP:N PUOLUEOHJELMA. Askeleen edellä. Hyväksytty Vaasan puoluekokouksessa 11.6.2006. (RKP)

SOSIALIDEMOKRATIAN PERIAATTEET OIKEUDENMUKAISEEN MAAILMAAN. Hyväksytty XXXVIII puoluekokouksessa Turussa 29.5.1999. (SDP)

VASEMMISTOLIITON PERIAATEOHJELMA. Vasemmiston tie oikeudenmukaiseen maailmaan. Hyväksytty 5. puoluekokouksessa 16.6.2007. (Vas)

VIHREIDEN PERIAATEOHJELMA. Vihreän liiton periaateohjelma Tehtävänä reilu muutos. Puoluekokouksen 28.5.2006 hyväksymä periaateohjelma. (Vihr)

Ohjelmaotos on sama kuin Jukka Paastelan ja Heikki Paloheimon antologiassa (2006).

### *Tutkimustehtävä ja -menetelmä*

Tarkoituksenamme on erityisesti morfologisen ja syntaktisen analyysin keinoin analysoida, poikkeavatko puolueiden ohjelmat toisistaan kieliopilliselta kompleksisuudeltaan. Neutraalina verrokkitekstinä käytämme Helsingin Sanomien viittä poliittista uutista 7.2.2009 (yhteensä 757 sanetta, lyhennettynä HS).

Tutkimusmenetelmänä sovellamme puolueohjelmien sanojen frekvenssianalyysia sekä automaattista morfologista analyysia, jolla palautamme taivutetut sanat perusmuotoihinsa. Yhdistääksemme yhteen kuuluvat taivutusmuodot ja päästäksemme käsiksi lekseemeihin (perusmuotoihin) ja niiden frekvensseihin sekä saneiden muoto-opillisiin piirteisiin olemme käyttäneet Lingsoft Oy/Ab:n automaattista suomen kielen lauseenjäsennintä FINCG (”Finnish Constraint Grammar”; Karlsson (1990) selvittää sen yleisiä toimintaperiaatteita). Ohjelma ratkoo pienen määrän morfologisia homografioita väärin ja eräissä hankalissa tapauksissa vaihtoehtoja jää useita, esim. omistusliitteellisten objektien numerus (yksikkö vai monikko: Näin hänen *kenkänsä*). Virheet ja monitulkintaisiksi jääneet tapaukset on korjattu käsin.

Esimerkiksi perussuomalaisten puolueohjelman ensimmäinen virke *Vaihtoehtojen puute on suoranainen uhka todelliselle kansanvallalle* saa seuraavan analyysin:

”<Vaihtoehtojen>”  
 ”vaihto#ehto” N GEN PL  
 ”<puute>”  
 ”puute” N NOM SG  
 ”<on>”  
 ”olla” COP V PRES ACT SG3

”<suoranainen>”  
 ”suoranainen” A POS NOM SG  
 ”<uhka>”  
 ”uhka” N NOM SG  
 ”<todelliselle>”  
 ”todellinen” A POS ALL SG  
 ”<kansanvallalle>”  
 ”kansan#valta” N ALL SG  
 ”<>”

Tämä formaatti tarkoittaa, että sana (tarkemmin sane, vrt. alh.) *vaihtoehtojen* kuuluu lekseemiin *vaihto#ehto*, joka on kaksiosainen yhdyssana (yhdyssanan rajan osoittimena on merkki '#') ja substantiivi, N(oun). Lisäksi tämä sane on monikon (PL) genetiivimuoto (GEN). Ohjelma antaa siis paitsi lekseemin eli perusmuodon myös saneen morfologisen rakenteen. Sane on on saanut analyysin ”olla” COP V PRES ACT SG3, mikä tarkoittaa, että se kuuluu lekseemiin olla, joka on verbi (V), mikä tässä tapauksessa on toteutunut preesensin aktiivin yksikön kolmannessa persoonassa. Ohjelma ratkoo morfologiset monitulkintaisuudet (kuten <tulen> ”tuli” N GEN SG vs. <tulen> ”tulla” V PRES ACT SG3), mutta ei suorita täydellistä konstituenttirakenne-, dependenssi- tai lauseenjäsänenalyysia.

Tällä tavalla aikaansaatuja morfologisesti analysoituja puoluekohtaisia perusaineistoja vertaamme kieliopillisten yhtäläisyyksien ja erojen toteamiseksi. Suomen runsaan ilmimorfologian ansiosta muoto-opillisesta analysoidusta korpuksesta voidaan päätellä paljon myös virkkeiden lauseopista, vaikkei ohjelma täydellistä syntaktista analyysia annakaan.

Kielioppiterminologian osalta noudatamme Auli Hakulisen et al. *Ison suomen kieliopin* (2004) käytänteitä.

### *Saneiden ja lekseemien määrät*

**Sane** on juoksevan tekstin sanaesiintymä (*token*), joita tässä virkkeessä on 11. Sane on alfanumeeristen merkkien (eli kirjainten ja/tai numeroiden) jono, jonka kummallakin puolella on välilyönti, välimerkki tai kappaleen alku- tai loppumerkki. Välimerkki ei kuulu saneeseen. Saneet ovat **sanetyypinsä** esiintymiä. Tämän kappaleen ensimmäisessä virkkeessä sanetyppi *on* edustuu kaksi kertaa. **Lekseemi** eli leksiikaalinen sana on sana epäsuorasti mukaan luettuina sen kaikki taivutus- sekä omistusliitteelliset ja liitepartikkelilliset muodot, jotka yhdessä konstituivat lekseemin sanetyypit. Lekseemiin TALO kuuluvat siten sanetyypit *talon*, *taloissa*, *talommeko*, *talojemmepa* jne. – yli 2000 sanetyyppiä.

Taulukko 1. Saneiden, sanetyyppien ja lekseemien määrät sekä *type/token*-suhteet.<sup>2</sup>

	saneet	sanetyypit	tyyp/saneet	lekseemit	leks/saneet
KD	13666	6463	0,47	3579	0,26
Vihr	6218	3156	0,51	1808	0,29
Vas	2857	1591	0,56	1032	0,36
RKP	2707	1587	0,59	986	0,36
Kok	2607	1417	0,54	819	0,31
Kesk	1792	1124	0,63	716	0,40
SDP	1187	768	0,65	512	0,43
PS	624	459	0,74	376	0,60
summa	31658	16565	9828		
keskiarvo	3957	2071	1229		
mediaani	2657	1502	903		

KD:n sanemäärä on 22 kertaa suurempi kuin PS:n. Sanemäärien keskiarvo on vajaa 4000, mediaani 2657. KD:n ohjelma on pitkä ja vihreidenkin pitkänpuoleinen, PS:n lyhyt ja SDP:nkin lyhyenpuoleinen. Vas, RKP ja Kok ovat lähellä mediaaniarvoa.

Sanetyyppien määrän suhde sane-esiintymiin kuvaa tekstin saneiden toisteisuutta. Mitä pidempi teksti on, sitä enemmän siinä on toistoa. Sanetyyppejä esiintyy yhteensä 16 565 (= kaikkien puolueiden käyttämien sanetyyppien summa eli kaikki eri puolueilla esiintyneet sanetyypit yhteen laskettuina), joista 12 112 on erilaisia. Korrelaatio ”suurempi sanemäärä – alhaisempi *type/token*-suhde” pätee kaikkiin paitsi Kok:een, jonka teksti on toisteisin.

PS tarvitsee vain 376 eri leksikaalista sanaa ja 624 sane-esiintymää perustavoitteidensa kuvaukseen, KD 10-kertaisen määrän lekseemejä ja 20-kertaisen määrän saneita. Lekseemien määrän mediaani on 903, jota lähinnä ovat Vas, Kok ja RKP. Oikeanpuoleinen sarake vahvistaa, että Kok:lla on muita suurempi taipumus puhua toistuvasti samoista asioista.

Saneiden pituuksissa (kirjaimina) ei ole merkittäviä eroja. Sanepituuksien keskiarvojen vaihteluväli on hyvin pieni (8,47–8,86) ja mediaani on kaikilla 8 paitsi PS:llä 9. Käyttämässämme Helsingin Sanomien verrokkitekstissä (757 sanetta) sanepituuden keskiarvo on 8,3 ja mediaani niin ikään 8. Puolueohjelmien kieliasu saneiden tasolla on siis kompleksisuudeltaan samankaltaista ja kohtalaisen lähellä yleistajuttavuuteen pyrkivän sanomalehden verrokkiaineistoa.

Taulukon 1 lekseemien summaluku 9828 tarkoittaa eri puolueiden käyttämiä kaikkia (samoja tai eri) lekseemejä. Käytettyjen eri lekseemien määrä on 5421. Kolmasosa lekseemeistä (n = 1816) esiintyy enemmän kuin yhden puolueen ohjelmassa,

<sup>2</sup> Taulukon 1 olemme lainanneet artikkelistamme Karlsson ja Wiberg (2010), koska taulukon sisältämät tiedot ovat oleellisia myös kieliopilliselle kompleksisuusanalyysille. Jatkossa esitettävät taulukot 2–5 ovat ennen julkaisemattomia.

vain 59 eli yksi prosentti kaikkien puolueiden teksteissä. Peräti 3605 eli 2/3 lekseemeistä esiintyy vain yhden puolueen ohjelmissa, mikä vihjaa että ohjelmien väliset sanastolliset erot olisivat erittäin suuret. Tätä vaikutelmaa vääristävät kuitenkin tekstien pituuserot, erityisesti KD:n ohjelma, jonka osuus kaikista saneista on 43,2 % ja lekseemeistä 36,4 %. Puoluekohtaisista leksikaalisista eroista ks. tarkemmin Karlsson ja Wiberg (2010).

### *Kieliopillinen kompleksisuus*

Selvitämme nyt puolueohjelmien muoto- ja eräitä lauseopillisiäkin kompleksisuuseroja. Perussääntö on tietenkin: mitä yksinkertaisempi muoto- ja lauserakenne, sitä luettavampi ja ymmärrettävämpi teksti.

Keskeisenä lauseopillisena tarkasteluyksikkönä on (kirjoitetun kielen) **virke**, jonka määrittelemme näin: virke on saneiden jono, jonka kummallakin puolella on merkki joukosta { . ? ! ¶ }, missä piste ’.’ viittaa vain sellaisiin pisteisiin, jotka aidosti lopettavat virkkeen (ei esim. järjestyslukusanojen jälkeiseen pisteeseen kuten 2. *yritys* – oikeiden virkkeenlopetuspisteiden löytymisestä huolehtii jäsenysohjelmamme), ja merkki ’¶’ on kappaleiden aloitus- tai lopetusmerkki (tietyissä tekstinkäsittelyohjelmissa). Virkkeen sisällä voi siis olla pilkkuja, kaksoispisteitä, puolipisteitä, luettelaviivoja, ajatusviivoja ym. ja niitä seuraavaa ainesta, mutta ei merkkejä joukosta { . ? ! ¶ }.

Muoto-opillisiä eli morfologisia eroja vertaamme kahden yksinkertaisen muutujan avulla:

- (i) yhdyssanojen osuus kaikista saneista,
- (ii) produktiivisesti käytettyjen johdinten lukumäärä saneissa.

Mitä suurempi yhdyssanojen osuus kaikista saneista on ja mitä enemmän saneissa keskimäärin on johtimia, sitä mutkikkaammasta sanojen rakenteesta on kyse.

Lauseopillisen eli syntaktisen kompleksisuuden<sup>3</sup> operationaalistamme kuudella virkkeiden rakenteeseen liittyvällä muuttujalla:

- (i) virkkeiden sanemäärä (pitkät virkkeet ovat mutkikkaampia kuin lyhyet),
- (ii) verbien määrä per virke (verbit ovat funktioita, jotka saavat argumentteja eli subjekteja, objekteja jne; suurempi verbimäärä tarkoittaa kompleksisempaa virkerakennetta); verbeiksi laskemme verbisaneet lukuun ottamatta VA-partisiippeja eli preesensin partisiippeja kuten *uudistava*, jotka on jäsennetty adjektiiveiksi – näin siksi, että rajanveto produktiivisten partisiippien ja aitojen adjektiivien välillä on sumea,

<sup>3</sup> Tässä artikkelissa emme voi käsitellä syntaktisen kompleksisuuden koko problematiikkaa esim. upotettujen lauseiden, partisiippiattribuuttien tai (muiden) nominaalistusten osalta.



Morfologisen kompleksisuuden osalta yllättää, että PS:llä on eniten yhdyssanoja – enemmän kuin joka kuudes sana on yhdyssana, ja moni niistä liittyy työhön: *työvoima, työvoimavaltainen, työvoimatarve, työvoimakoulutus, työnteko, työnantajakapula, työnantajamaksu*. Muuten PS:n ohjelma on muita ytimekkäämpi.

Suhteellisesti eniten johdoksia<sup>4</sup> löytyy Kok:lta, joka innokkaasti viljelee verbien substantivointeja *minen*-johtimen avulla ja niihin liittyviä yksinkertaisten peruslauseiden uudelleen järjestelyjä: *olosuhteiden kehittäminen, heikommista huolehtiminen, kielten ja kulttuurien vaaliminen, itsenäisyyden turvaaminen, osaamisen vaaliminen* (kaksi substantivointia!), *uuden osaamisen muodostuminen* (samoin), *paremman yhteiskunnan kehittäminen, huolenpidon toteutuminen, tulevaisuuden uskon vahvistaminen, Euroopan integraatioon osallistuminen, politiikan haasteiden loppuminen, elämän sosiaalisten perusedellytysten täytyminen* jne. Sananjohdoksen suosio kokoomuksella, erityisesti substantivoiminen ja siihen liittyvä kieliasun komplisoituminen kompensoituu osin sillä, että yhdyssanojen osuus on selvästi pienempi kuin muilla puolueilla.

Kahden morfologisen muuttujan osalta vihreät erottuvat joukosta muita yksinkertaisemmalla kielenkäytöllään: heillä on vähänpuoleisesti sekä yhdyssanoja että johdettuja sanoja. Vertailu Helsingin Sanomien uutisiin kertoo kuitenkin selvästi, että kaikissa puolueohjelmissä käytetään huomattavasti (kaksi kertaa) enemmän johdoksia kuin uutisteksteissä. Yhdyssanojen osalta yhtä kirkasta eroa ei ole. Taulukko 3 selvittää tarkemmin, mitkä johtimet ovat erityisen ahkerassa käytössä (taajaan esiintyvä adjektiivien *sti*-johdin on melkein sijan luonteinen eikä ole mukana laskelmissa).

Taulukko 3. Viisi yleisintä johdinta.

<b>KD</b>	<b>n</b>	<b>%</b>	<b>Kesk</b>	<b>n</b>	<b>%</b>	<b>Kok</b>	<b>n</b>	<b>%</b>
v-minen	490	19,6	a-uus	67	19,7	a-uus	134	22,2
v-us	460	18,4	n-llinen	63	18,5	v-minen	123	20,4
a-uus	283	11,3	v-minen	54	15,9	n-llinen	70	11,6
v-u	254	10,2	v-us	41	12,1	v-us	64	10,6
n-llinen	215	8,6	v-u	27	7,9	v-u	57	9,4

4 Sananjohdon problematiikkaa ei olemassa olevissa suomen kielen automaattisissa jäsentimissä kuten FINCG:ssä ole dokumentoitu kovin tarkkaan. Tyydymme analysoimaan ohjelman tuottamia, taulukossa 3 käsiteltyjä yleisimpiä produktiivisia johtimia ja niiden määriä. Emme ota kantaa monien muiden johdinten hankaliin leksikaalistumisen asteiden arviointiin (ovatko esim. sanat *kahvila* ja *pätevä* synkronisessa mielessä johdettuja). Monijohtimiset sanat kuten *rauha+llis+uus* lisäävät kaikkien osajohdintensa määrää. Hypoteesimme on se, että produktiivisten johdinten erityisen runsas esiintyminen lisää kompleksisuutta, koska niiden hetkellinen kokoaminen ja purkaminen edellyttävät sekä sanoman tuottajalta että vastaanottajalta todellista prosessointia. Leksikaalistuneiden johdosten kohdalla tällaista prosessointia ei ole.

<b>Perus</b>	<b>n</b>	<b>%</b>	<b>RKP</b>	<b>n</b>	<b>%</b>	<b>SDP</b>	<b>n</b>	<b>%</b>
v-minen	19	16,5	v-us	98	21,1	a-uus	62	23,4
v-us	15	13,0	a-uus	73	15,7	n-llinen	43	16,2
a-uus	13	11,3	n-llinen	50	10,8	v-minen	31	11,7
v-tta	10	8,7	v-minen	46	9,9	a-us	31	11,7
v-u	9	7,8	a-us	45	9,7	v-u	28	10,6
<b>Vas</b>	<b>n</b>	<b>%</b>	<b>Vihr</b>	<b>n</b>	<b>%</b>	<b>HS</b>	<b>n</b>	<b>%</b>
v-minen	79	16,3	a-uus	199	18,2	v-u	18	23,1
a-uus	66	13,6	v-minen	187	17,1	v-ja	12	15,4
v-us	63	13,0	v-us	129	11,8	v-tta	7	9,0
a-us	57	11,7	v-u	110	10,0	a-uus	7	9,0
v-us	56	11,5	n-llinen	101	9,2	v-us	6	7,7

Johdinten nimityksissä ensimmäinen kirjain tarkoittaa kannan sanaluokkaa (v = verbi, a = adjektiivi, n = substantiivi), jota seuraa itse johdin (esimerkkejä: v-minen *ostaminen*, v-tta *vauhdittaa*, v-u *kysely*, v-us *puolustus*, v-ja *johtaja*, a-uus *suomalaisuus*, a-us *vapaus*, n-llinen *suhteellinen*). Luettelot ovat hyvin samankaltaisia koostumukseltaan, vaikka järjestyksessä on eroja: v-minen ja a-uus esiintyvät kaikilla yleisinä, v-u ja v-us seitsemällä, n-llinen kuudella. HS-vertailuluvut kertovat puolueohjelmien verrattain raskaasta tyylistä, eritoten nominaalistuksen suosiosta. Substantivoiva v-minen on kaikissa ohjelmissa yleinen, mutta uutistekstissä paljon harvinaisempi (järjestysluvultaan seitsemäs). Toisaalta aktiivinen tekijännimen johdin v-ja (*johtaja*, *tulija*, *haastaja*) on uutisteksteissä toiseksi yleisin johdin, mutta kaikissa ohjelmissa paljon harvinaisempi (järjestysnumeroltaan eri puolueilla 7, 7, 9, 9, 10, 10, 11, 12), mikä tietenkin lisää ohjelmien abstraktiutta.

Palaamme taulukkoon 2, eritoten sen lauseopillisiin muuttujiin. PS:n virkkeet ovat selkeästi lyhyempiä kuin muilla, kahdeksan sanaa, kun ne muilla sisältävät 10–11 sanaa ja HS:ssakin 11. Kristillisillä ja vihreillä on huomattavasti enemmän verbejä per virke kuin muilla eli monipolvisempi tapa ilmaista tekemiset ja niiden keskinäiset suhteet (enemmän eri aikamuotoja, enemmän infinitiivejä tyyppiä (*haluamme*) *vaikuttaa*, ym.). Pilkkuja per virke on selvästi eniten SDP:llä. Ne tekevät tekstistä helposti luettelomaisen:

Ihmisarvosta, ihmisoikeuksista ja yksilön vastuuntunnosta lähtevä joukkoliike kohotti ihmiset kansalaisiksi, joille kuuluvat valtiolliset, sivistykselliset, sosiaaliset ja työelämän oikeudet. (SDP)

Relatiivilauseita useissa ohjelmissa viljellään 2–3 kertaa enemmän kuin uutisteksteissä, erityisesti RKP:llä. Ne lisäävät kuvailevuutta ja vähentävät pääasian näkyvyyttä. Keskusta käyttää silmiinpistävän vähän relatiivilauseita. Alistuskonjunktioita (esim. *että*, *jos*, *jotta*, *koska*, *kun*, *vaikka*) on eniten pisimmissä ohjelmissa eli

kristillisillä ja vihreillä. Alistuskonjunktiot, erityisesti adverbiaalina toimivan lauseen aloittavat alistuskonjunktiot, ovat tyypillisiä argumentoiville teksteille, missä pääargumentteja ei niinkään ilmoiteta totuuksina, vaan niille esitetään myös lauseen muotoisia syitä, ehtoja, varauksia, ajallisia rajoituksia, päämääriä, tapaa jne. (vrt. Hakulinen et al. 2004: 1061). Erityisen tyypillisiä ovat runsaat tapaa ilmaisevat tai muissa tehtävissä olevat *että*-lauseet:

Hallituksen on kolmikannassa vaikutettava siihen, *että* jatketaan maltillisia tuloratkaisuja. Kolmikantaa on laajennettava *niin, että* yrittäjien valtaenemmistö pääsee mukaan. (KD)

Energiantuotannossa on painotettava hajasijoitusta, energiansäästöä ja uusiutuvia energianlähteitä *siten, että* vuoteen 2050 mennessä 2/3 energiasta syntyy uusiutuvista energialähteistä ja energiantuotannon kasvihuonekaasupäästöt vähentyvät 80 % vuoteen 1990 verrattuna. (Vihr)

Yllättävästi HS-tekstissä on kirkaasti eniten sivulauseita, jotka alkavat alistuskonjunktiolla. Tämä johtuu siitä, että uutistekstissä esiintyy paljon sanomis- ja ajattelemisverbejä (*ilmoitti että...*; *arveli että...*), joita puolestaan on vähän persoonattomissa puolueohjelmissa ja joita ei voi pitää suoraviivaisesti kompleksisuutta lisäävinä. Pelkkää sivulauseiden määrää funktion katsomatta ei siis voi pitää tekstin kompleksisuuden mittarina.

Rinnastuskonjunktiot lisäävät pilkkujen tapaan luettelomaisuutta. Myös rinnastuskonjunktioita on eniten SDP:llä, enimmillään kymmenkunta yhdessä kappaleessa:

Yhteiskunnan mitta on ihminen; sen rakenteet *ja* menettelytavat ovat ihmistä varten. Ne suojaavat piittaamattomuudelta, joka hajottaa yhteisön. Ne ehkäisevät rajoittamatonta markkinakilpailua, jonka vuoksi heikommat kärsivät *ja* elinympäristö tuhoutuu. Oikeusvaltio määrittää oikeudet *ja* velvollisuudet suojaten yksilöä *ja* yhteisöjä kapitalismin itsekkyydeltä. Ihmisten on oltava tasa-arvoisia lain edessä. Yhteiskunnan on oltava vakaa *ja* turvallinen, jotta ihmiset *ja* yhteisöt, miehet *ja* naiset tasavertaisina voivat suunnitella tulevaisuuttaan *ja* kehittyä. Sosialidemokratialle lait *ja* sopimukset ovat edistyksen välineitä yhteiskunnallisten ristiriitojen ratkaisemiseksi *ja* eriarvoisuuden purkamiseksi. (SDP)

PS käyttää rinnastuksia poikkeuksellisen vähän.

Yksinkertainen tapa vertailla puolueohjelmien kokonaiskompleksisuutta käytettyjen muoto- ja lauseopillisten muuttujien ja niiden arvojen avulla (mukaan luettuna kokonaissanemäärä) on panna puolueet muuttujakohtaiseen järjestykseen. Muuttujan suurimman arvon omaava puolue saa yhden pisteen, toiseksi suurimman arvon omaava kaksi pistettä jne. Pisteet ynnätään ja eniten pisteitä saanut on yksinkertaisin, vähiten pisteitä saanut kompleksisin. Tämä menetelmä antaa taulukosta 4 ilmenevän tuloksen.

Taulukko 4. Ohjelmien morfologinen ja syntaktinen kompleksisuusjärjestys (pohjana taulukon 2 datat).<sup>5</sup>

”pisteet”	1	2	3	4	5	6	7	8	9
sanemäärä	kd	vih	vas	rkp	kok	kes	sdp	hs	per
yhdyssanojen osuus saneista	per	kd	vas	rkp	hs	kes	sdp	vih	kok
johtimet: määrä per sane	kok	sdp	kes	kd	per	vih	rkp	vas	hs
virkkeet: pituus saneina	kd	hs	rkp	vih	sdp	kes	kok	vas	per
verbit: määrä per virke	kd	vih	rkp	hs	per	kok	vas	sdp	kes
pilkut: määrä per virke	sdp	rkp	hs	kd	vas	kes	vih	kok	per
relatiivipron. määrä per virke	rkp	sdp	vas	kd	per	vih	kok	kes	hs
alistuskonj. määrä per virke	hs	kd	vih	rkp	kok	per	vas	sdp	kes
rinnastuskonj. määrä per virke	kok	sdp	rkp	kes	kd	vih	vas	per	hs
<b>yksinkertainen – kompleksinen</b>	<b>kes</b>	<b>per</b>	<b>vas</b>	<b>hs</b>	<b>kok</b>	<b>vih</b>	<b>sdp</b>	<b>rkp</b>	<b>kd</b>
	57	57	51	50	49	44	42	31	24

Tämän karkean mittarin mukaan puolueohjelmat jakaantuvat neljään ryhmään: rakenteellisesti yksinkertaisimmat eli Kesk ja PS, sitten vähän mutkikkaampina Vas ja Kok, edelleen mutkikkaampina Vihr ja SDP ja vihdoin mutkikkaimpien ryhmä RKP ja KD, joista jälkimmäinen eroaa huomattavasti kaikista muista. Poliittiseen uutiskieleen verrattuna (HS) viimeksi mainitut neljä puolueohjelmaa ovat rakenteeltaan selkeästi mutkikkaampia.

PS:n rakenteellinen yksinkertaisuus koskee ennen kaikkea virkerakennetta. Seuraavassa katkelmassa ei ole yhtään sivulausetta, pilkkua eikä rinnastuskonjunktioita:

Työttömyyttä ylläpitävä työnantajapula on poistettava. Ilman omaa syytään luottokelvottomaksi joutuneiden yrittäjien toimintaedellytykset on palautettava. Kannatamme yritystoimintaa ja elinkeinovapautta. Yritysten kasvuedellytysten turvaaminen on työllisyyden hoitamisen ehto. Työttömyyden poistamiseen tarvitaan tehokkaita valtion ohjaustoimia. (PS)

Toisaalta katkelman substantiivilausekkeet (substantiivit määritteineen) ovat melko raskaita: on paljon yhdyssanoja ja substantivoiteja kuten *yritysten kasvuedellytysten turvaaminen*, *työllisyyden hoitaminen* ja *työttömyyden poistaminen*. Seuraavassa näyte keskustan puolueohjelman suhteellisen yksinkertaisesta virkerakenteesta, vaikka siinäkin useampi substantiivilauseke on raskaahko:

Päätökset tulee tehdä lähellä ihmistä. Kunnallinen ja maakunnallinen itsehallinto kannustaa ihmisiä oma-aloitteisuuteen ja yhteistyöhön kotiseudun tulevaisuuden rakentamiseksi. Kansalaisille tulee luoda uusia mahdollisuuksia osallistua ja vaikuttaa päätöksentekoon. Rohkaisemme kansalaisia vapaaehtoiseen kansalaistoimintaan. (Kesk)

5 Pilkkujen osalta emme ole käsin poistaneet kielenhuoltonormien vastaisia pilkkuvirheitä.

*Eräiden kielioppikategorioiden esiintyminen*

Tässä käsittelemme eräiden keskeisten, tekstien kokonaisilmeeseen ja merkitykseen vaikuttavien kielioppikategorioiden esiintymistä, erityisesti sellaisia, jotka edustuvat merkittävästi eri tavoin ainakin muutamilla puolueilla. Taulukossa 5 on kahdenkymmenen tällaisen muuttujan puoluekohtaiset absoluuttiset arvot sekä arvon osuus suhteessa puolueohjelman virkkeiden määrään. Joka muuttujan kohdalla suurin suhteellinen arvo on merkitty puolilihavalla ja kursiiivilla (**0,21**, siis noin joka 5. virkkeessä), pienin arvo alleviivauksella (0,02) eli joka 50. virkkeessä. Jos useammalla puolueella on sama suurin tai sama pienin arvo, kaikki merkitään. Suurimmasta ja pienimmästä käyttäjästä ja käyttöarvosta puhumme jatkossa termeillä SUURKÄYTTÄJÄ (-KÄYTTÖ) ja PIENKÄYTTÄJÄ.

Taulukko 5 vahvistaa, että PS:n ohjelma poikkeaa kielellisesti eniten muista yksinkertaisuudeltaan ja iskevyydeltään. Perussuomalaiset ovat peräti kahdentoista kielellisen kategorian pienkäyttäjiä: verbit yleensä, A-infinitiivit, imperfektimuodot, substantiivit, adjektiivit, lukusanat, intensifikaattorit kuten *erittäin, hyvin, niin (pitkä)* sekä kaikki genetiivissä, elatiivissa ja partitiivissa taipuvat sanat.

Verbien yleensä ja samoin A-infinitiivien vähäisyys tarkoittaa erityisesti virkkeiden yksiverbisyyttä ja samalla lyhyyttä – jo edellähän todettiin, että PS:llä virkkeet ovat selkeästi lyhyempiä kuin muilla (taulukko 2). Substantiivien eli yleisimmän sanaluokan vähäisyys implikoi yleistä vähäsanaisuutta ja myös sitä, että virkkeillä on selkeähkö fokus: samassa virkkeessä ei puhuta monesta eri asiasta. Myös pronomien ja lukusanojen vähäisyys tarkoittaa, että päähuomio kohdistuu substantiivipääsanoihin. Samassa hengessä genetiivisijojen vähäisyys tarkoittaa ennen kaikkea, että pääsanana toimivissa substantiivilausekkeissa ei ole niin paljon muita substantiivilausekkeita etumäärteinä, kuten on neljä seuraavassa virkkeessä: *Suome/n lähitulevaisuude/n keskeinen haaste on väestö/n huoltosuhtee/n muutos* (KD). Partitiivisijan keskeisin merkitystehtävä on rajaamattoman kvantiteetin ilmaiseminen (*kahvi/a*), ja siksi myös partitiivisijan vähäisyys myötävaikuttaa tekstin täsmällisyyteen: ei puhuta paljon epämääräisesti kvantifioiduista tarkoitteista. Imperfektien vähäisyys tarkoittaa tekstin liikkumista enemmän tässä-ja-nyt-dimensiossa. Adjektiivien, komparatiivimuotojen (*halve/mpi*) ja intensifikaattorien vähäisyys tarkoittaa vähempää subjektiivista arviointia.

Toisaalta PS on kahden kieliopillisen kategorian suurkäyttäjä: kieltoverbin ja passiivin. Nämä ovat tehokkaan poliittisen retoriikan palveluksessa. Passiivit ovat yleisiä ennen kaikkea toimintasitoumuksissa eli direktiiveissä tyyppiä *Kuntien on lopetettava aiheeton pätkätöiden teettäminen*. Kieltoverbi on PS:illa viisi kertaa yleisempi kuin esim. keskustalla, erityisesti seuraavan tyyppisissä suorasukaisissa virkeissä:

*Emme hyväksy ihmisarvoa loukkaavaa toimintaa.* – – Puolustusvoimien määrärahoja *ei* saa harkitsemattomasti karsia. – – Puolueemme *ei* hyväksy perhe-

käsitteen luonnotonta laajentamista. – – Perussuomalaiset *eivät* hyväksy kultaisia kädenpuristuksia. – – Suomalaisen päätösvaltaa *ei* saa siirtää ylikansallisille elimille. – – Perussuomalaiset *eivät* hyväksy puoluetukien korottamista. (PS)

Suur- ja pienkäyttöarvoilla mitattuna toiseksi suoraviivaisin ohjelma on keskustalla, jolta löytyy viisi pienkäyttöarvoa, mutta ei yhtään suurkäyttöarvoa. Pienlukuja ovat kieltoverbi, passiivi, *olla*-verbin muodot, intensifikaattorit ja monikkomuodot. Keskusta käyttää siis runsasmerkityksisiä verbejä (siis muita kuin *olla*) monipuolisemmin kuin muut, ja verbinmuodot ovat leimallisesti aktiivisia (passiivimuotojen osuus 2–5 kertaa pienempi kuin muilla). Vertaa verbimerkitysten ja aktiivisuuden osalta Kesk:n ohjelmaa *olla*-maksimin omaaviin vihreisiin ja passiivipitoiseen PS:ään (kaikki kappaleet yhtäjaksoisia):

Yrittäjyys, yhteistyö ja sopiminen *muodostavat* perustan taloudelle ja työllisyydelle. *Kehitämme* eettisesti kestäväää markkinataloutta. *Kannatamme* yksityistä omistusoikeutta ja *parannamme* yrittäjyyden menestymisen edellytyksiä koko maassa. *Tarvitsemme* lisäksi osuustoimintaa ja valtion yrityksiä. *Suosimme* suomalaista omistusta, *kannustamme* suomalaisia omistajuuteen ja *rohkaisemme* yritysten sijoittumista Suomeen. (Kesk)

Silti alkanut vuosituhat *on* toivoa täynnä. Paitsi suuria riskejä, ihmiskunnalla *on* myös suuria mahdollisuuksia. Meillä *on* tietoa ja varallisuutta vastata ajan haasteisiin. Laadukas ja maksuton koulutus takaa, että yhteiskunta *on* tulevaisuudessakin yhdenvertainen ja sosiaalisesti oikeudenmukainen. Maailmanlaajuinen kansalaisyhteiskunta vahvistuu tuoden tahtoa ja keinoja poliittisten ratkaisujen löytämiseksi. (Vihr)

Työnteon ja työn teettämisen on aina *kannatettava* joutenoloa paremmin. Tarpeeton työttömien kurssittaminen työttömyydestä jo ennestään kärsiville aloille *on lopetettava*. Työvoimakoulutus *on suunnattava* esimerkiksi oppisopimuksella suoraan työllistävän yrityksen tarpeisiin. Valtion ja kuntien *on lopetettava* aiheeton pätkätöiden teettäminen ja niiden ketjuttaminen. Työn käsitettä *on laajennettava* huomioimalla kotityön ja omaishoidon merkitys. (PS)

Toisessa ääripäässä KD:n ohjelmassa on eniten suurkäyttöarvoja: substantiivit, monikolliset sanat, genetiivissä taivutetut sanat, peruslukusanat, intensifikaattorit ja verbien konditionaalimuodot. Erityisesti konditionaaleja KD käyttää moninkertaisesti muihin puolueisiin verrattuna. KD:lla on 101 konditionaalimuotoa, kun Kok:lla ja SDP:llä ei ole yhtään, RKP:lla yksi, Kesk:lla ja PS:llä kaksi ja Vas:lla neljä. KD:n konditionaaleista melkoinen osa on lievennettyjä direktiivejä (mm. *tulisi* 26 kertaa):

Jokaisen *tulisi* pyrkiä elämän heikoimpien mukaisesti. – – Inflaatio *tulisi* kyetä pitämään alle kahdessa prosentissa. – – Työviihtyvyyden *tulisi* lisääntyä. – – Yli sata puuhakkeella tai pelleteillä (lisäpolttoaineena turve) toimivaa 10–20 MW lämpö-sähkölaitosta *tulisi* perustaa maaseututaajamien yhteyteen. (KD)

Taulukko 5. Eräiden keskeisten kielipokkategorioiden esiintyminen.

	KD n	%	Kesk	Kok	Perus	RKP	SDP	Vas	Vähr	HS
virkkeiden määrä	1206		180	264	77	241	114	293	563	75
kieltoverbi ( <i>ei</i> )	93	0,08	6	12	0,05	8	0,03	26	0,09	12
passiivi ( <i>vätetään</i> )	592	0,49	20	77	0,29	42	0,17	105	0,36	29
verbi	2524	2,09	293	451	1,71	510	2,12	447	1,53	137
<i>olla</i> -verbi	636	0,53	68	166	0,63	126	0,52	176	0,60	44
A-infinitiivi ( <i>sanoa</i> )	629	0,52	68	102	0,39	7	0,09	62	0,21	9
imperfekti ( <i>sanoi</i> )	14	0,01	2	10	0,04	0	0,00	5	0,02	3
konditionaali ( <i>-isi</i> )	101	0,08	2	0	0,00	1	0,00	4	0,01	6
substantiivi	6108	5,06	827	1224	4,64	284	3,69	1325	4,52	366
adjektiivi	1680	1,39	279	330	1,25	83	1,08	430	1,47	61
komparatiivi ( <i>-mpi</i> )	73	0,06	3	10	0,04	19	0,08	8	0,03	5
adverbi	752	0,62	70	83	0,31	111	0,46	105	0,36	38
pronomini	491	0,41	93	131	0,50	21	0,27	143	0,49	42
peruslukušana	68	0,06	1	4	0,02	4	0,02	14	0,05	6
intensifikaattori	51	0,04	0	2	0,01	6	0,02	2	0,01	6
monikko ( <i>-t, -i</i> )	2894	2,40	257	443	1,68	138	1,79	452	1,54	85
mon. 1. p. poss.suff.						557	2,31	223	1,96	1,13
<i>-mme</i>	24	0,02	19	21	0,08	29	0,12	8	0,03	11
genetiivi ( <i>-n, -den</i> )	2793	2,32	350	558	2,11	121	1,57	577	2,0	106
nominatiivi	2420	2,01	342	490	1,86	154	2,00	660	2,3	201
elatiivi ( <i>-sta ~ -stä</i> )	416	0,34	38	81	0,31	4	0,05	75	0,26	29
partitiivi ( <i>-a, -ä</i> )	1541	1,28	231	310	1,17	301	1,25	321	1,10	71

Viimeinen äskeisistä virkkeistä havainnollistaa myös KD:n substantiivi-, monikko- ja lukusanapitoisuutta, yksityiskohtaisuutta sekä rakenteen mutkikkuutta. Konditionaaleista moni toimii epämääräistäjänä:

Puolisoista vain jompikumpi *voisi* saada vähennyksen verotuksessaan. – – Tällaiset yritykset *voisivat* tarjota jopa 160 000 työpaikkaa lisää. – – 10 laitosta *voisi* korvata 30. – – Työnantaja *voisi* hakea päivärahaetuuden Kelalta. – – Oikea paikka *olisi* terveyskeskuksen vuodeosastolla. – – Sen [valtakunnallisen joukkoviestinnän eettisen neuvottelukunnan] tehtävänä *olisi* käsitellä periaatteelliselta kannalta mediaan liittyviä eettisiä kysymyksiä. (KD)

Konditionaaleja on runsaanpuoleisesti myös vihreillä (30 kpl). Heillä on lisäksi viisi suurkanäyttöarvoa: *olla*-verbin esiintymät, adjektiivit, komparatiivit, adverbit ja partitiivissa taivutetut sanat.

SDP:llä on neljä suurkanäyttöarvoa, joista kiintoisin on imperfekti. SDP käyttää menneen ajan verbinmuotoja selkeästi enemmän kuin muut:

Työväenluokan järjestäytyminen poliittiseksi ja ammatilliseksi työväenliikkeeksi *antoi* Euroopan kehitykselle uuden suunnan. Suomen työväenpuolue *perustettiin* Turussa 1899. Se *asetti* tavoitteekseen työväestön taloudellisen ja yhteiskunnallisen vapauttamisen. Forssan kokouksessa 1903 puolue *muuttui* sosialidemokratiseksi puolueeksi. Se *hyväksyi* kansainvälisen työväenliikkeen sosialistisen ohjelman ja *omaksui* kansanvaltaiset menettelytavat. (SDP)

### *Päätelmä*

Puolueiden periaateohjelmat eroavat huomattavasti toisistaan morfosyntaktiselta kompleksisuudeltaan. Ne jakaantuvat tältä osin neljään ryhmään: rakenteellisesti yksinkertaisimmat eli Kesk ja PS, vähän mutkikkaampina Vas ja Kok, edelleen mutkikkaampina Vihr ja SDP ja vihdoinkin mutkikkaimpien ryhmä RKP ja KD, joista jälkimmäinen poikkeaa huomattavasti kaikista muista. Poliittiseen uutiskieleen verrattuna (HS) viimeksi mainittujen neljän puolueen tekstit ovat rakenteeltaan selvästi mutkikkaampia.

PS:n ohjelma on retorisesti iskevin. Siihen myötävaikuttavat paitsi yleinen vähäsanaisuus ja selkeästi lyhyimmät virkkeet myös runsaat kiellot ja passiivissa olevat toimintasuositukset. Monisanaisimmilla ohjelmilla (Vihr ja erityisesti KD) on taipumus myös rakenteelliseen kompleksisuuteen, samoin RKP:llä. SDP:n ohjelma on lyhyt mutta luettelomainen. Se sisältää myös eniten menneen ajan aikamuotoja.

Ei ole uskottavaa, että puolueet laatisivat ohjelmiaan erityisenä silmämääränään toisten puolueiden ohjelmat: puolueet eivät kommunikoi ensisijaisesti toisille puolueille vaan omille jäsenilleen ja kannattajilleen. Puolueet yrittävät mitä ilmeisimmin tehdä ohjelmistaan itsensä näköisiä. Analyysimme paljastaakin oleellisia eroja puolueiden kesken.

## LÄHTEET

*Aineistot*

<http://www.fsd.uta.fi/> (Yhteiskuntatieteellinen tietoaarkisto)  
<http://www.fsd.uta.fi/pohtiva> (puolueiden periaateohjelmien sähköiset versiot)

*Kirjallisuus*

- BORG, OLAVI 1965: *Suomen puolueet ja puolueohjelmat 1880–1964*. Helsinki: WSOY.
- HAKULINEN, AULI – VILKUNA, MARIA – KORHONEN, RIITTA – KOIVISTO, VESA – HEINONEN, TARJA RIITTA – ALHO, IRJA 2004: *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- KARLSSON, FRED 1990: Constraint Grammar as a Framework for Parsing Unrestricted Text. H. KARLGRÉN (toim.): *Proceedings of the 13th International Conference of Computational Linguistics, Vol. 3*, 168–173. Helsinki.
- KARLSSON, FRED – WIBERG, MATTI 2010: Puolueilla on eroja! Periaateohjelma-analyysi. *Politiikka* 52:1, 54–66.
- PAASTECLA, JUKKA – PALOHEIMO, HEIKKI (toim.) 2006: *Suomen puolueiden periaateohjelmat: ohjelmat, niiden taustaa ja tarkastelua*, moniste. Tampere: Tampereen yliopiston politiikan tutkimuksen laitos.

FRED KARLSSON and MATTI WIBERG: *Grammatical complexity of Finnish political party programs*

We carried out an analysis of the grammatical complexity of the party programs of all eight Finnish parliamentary parties. The texts were subjected to a computerized morphological analysis yielding the base forms (lexical words) and grammatical properties of inflected words. The party programs contain a total of 31,658 running words, comprising 9828 lexical words. There are considerable differences between the programs with respect to distribution, length, redundancy, specificity, complexity, readability, the presence or absence of catchphrases typical of a particular party, what the parties promise to do, and how they go about doing it by using various directive expressions.

Fred Karlsson: [fgk@ling.helsinki.fi](mailto:fgk@ling.helsinki.fi)  
 Nykykielten laitos / yleinen kielitiede, PL 24, 00014 Helsingin yliopisto

Matti Wiberg: [wiberg@utu.fi](mailto:wiberg@utu.fi)  
 Valtio-oppi, 20014 Turun yliopisto