

MITÄ SYNTAKTISESTI KOODATTU KORPUS VOI KERTOA SANAJÄRJESTYKSESTÄ? PREDIKAATTI JA SUBJEKTI EDISTYNEIDEN SUOMENOPPIJOIDEN EKSISTENTIAALISISSA LAUSEISSA

Ilmari Ivaska

Abstrakti

Eksistentiaalilause kuuluu lausetyyppinä kaikkiin suomen kielen lausetyyppi-jaotteluihin. Yksi lauseen leimallisia piirteitä on sen sanajärjestys. Yleisesti suomen kielen lauseissa subjekti edeltää verbiä (SV-sanajärjestys), mutta prototyyppisessä eksistentiaalilauseessa lauseen subjektiksi kutsuttu jäsen on verbinjälkeisessä asemassa (VS-sanajärjestys). Tällainen ilmiö on hyvin kiinnostava kielen oppimisen näkökulmasta, ja sanajärjestyksen kaltaista kahden muuttujan välistä suhdetta on hyvin mielekästä tarkastella korpustutkimuksen keinoin. Tutkin tässä artikkelissa Edistyneiden suomenoppijoiden korpuksen (LAS2) avulla eksistentiaalisten lauseiden sanajärjestystä subjektin ja predikaatin osalta ja pohdin näiden varianttien tyypillisiä esiintymisympäristöjä. Vaikuttaa siltä, että SV-sanajärjestys on suhteellisesti todennäköisempi silloin, kun eksistentiaalinen lause on kielteinen. Huomio pätee niin edistyneiden suomenoppijoiden tutkimusaineistossa kuin

ensikielisten vertailuaineistossakin. On mahdollista, että sana-järjestyksen ja lauseen kielteisyyden välillä on edistyneiden suomenoppijoiden kielessä tiettyjä valintapreferenssejä nimenomaan eksistentialisten lauseiden kontekstissa, sillä kielteisissä lauseissa on suhteellisesti paljon useammin lausetyypille epäprototyyppinen SV-sanajärjestys. Vertailuaineistossa inversiolle on aina osoitettavissa jokin muu motivoija, kun taas tutkimusaineistossa tällaista motivoijaa ei aina näytä olevan.

Avainsanat: suomi toisena kielenä, korpuslingvistiikka, sana-järjestys, kiello, syntaksi¹

1. Johdanto

Eksistentialilauseeksi kutsuttu lausetyyppi on kuulunut suomen kielen syntaksin lausetyyppilähtöiseen tarkasteluun halkan historian. Suomen kielen eksistentialilauseella tarkoitetaan olemassaoloa ilmaisevaa rakennetta, jossa on prototyyppisesti VS-sanajärjestys, jonka subjektiksi kutsuttu jäsen² voi olla joko nominatiivissa tai partitiivissa ja jossa subjektin ja yksikön kolmannen persoonan näköisen predikaatin välillä ei ole kongruenssia (Hakanen 1972: 42–43). Usein eksistentialilauseessa on myös paikkaa ilmaiseva lokatiivadverbiaali (ISK: 851). Eksistentialilause eroaa tyyppipiirteidensä vuoksi suomen kielen ns. normaalilauseista niin sanajärjestyksen, kongruoimattomuuden kuin subjektin sijamerkinänkin osalta.

¹ Kiitän Helle Metslangia hänen esittämistään huomioista ja korjausehdotuksista.

² Kyseisen jäsenen subjektimaisuus on usein kyseenalaistettu. Esimerkiksi Marja-Liisa Helasvuo ja Tuomas Huumo ehdottavat jäsenen kutsumista neutraalisti E-NP:ksi (Helasvuo, Huumo 2010: 191–192).

Tästä syystä sen idiomaattinen käyttö vaatii monipuolista kielitaitoa, ja lausetyypin tarkasteleminen oppijansuomen viitekehyksessä on luontevaa. Tarkastelen tässä artikkelissa edistyneiden suomenoppijoiden eksistentiaalisten lauseiden sanajärjestystä ja pohdin, mitä syntaktisesti koodattu korpusaineisto voisi tästä kertoa. Tutkimus perustuu osin pro gradu -tutkielmaani (Ivaska 2009) ja se on osa Turun yliopistossa käynnissä olevaa Edistyneiden suomenoppijoiden korpus - tutkimushanketta (jatkossa LAS2, ks. esim. Ivaska, Siitonen 2009).

2. Eksistentiaalilauseen sanajärjestys ja toisen kielen korpustutkimus

Eksistentiaalilauseesta ja sen suhteesta muihin suomen kielen lauseisiin puhutaan useimmiten, kun tarkastellaan kielen syntaktisia lausetyyppejä. Kuhunkin lausetyyppiin liittyy kuitenkin aina niiden nimiä myöten presuppositioita myös semanttisesta sisällöstä. (Suomen eksistentiaalilauseesta käydystä keskustelusta ks. esim. Tiainen 1997; eksistentiaalilauseen luonteesta ja sen suhteesta muihin lausetyyppeihin ks. esim. Ivaska 2010: tulossa.) Käsite lainattiin alun perin Otto Jesperseniltä, joka käytti termiä *existential sentence* kuvaamaan sellaisia olemassaoloa ilmaisevia rakenteita, jotka vastaavat englannin *there is/there are* -tyyppisiä ilmaisuja ja joiden tarkoitteet ovat tyyppillisesti epämääräisiä (Jespersen 1924: 154–156). Partitiivisubjektin ja olemassolon ilmaisemisen välisen yhteyden nosti Suomessa ensi kerran esiin Martti Airila todetessaan näiden olevan lauseita, joiden predikaatin toiminnallisuus on korkeintaan vain taka-alalla (Airila 1924: 17–18). Termiä *eksistentiaalilause* käytti ensi kerran Osmo Ikola pohtiessaan niin ikään partitiivisubjektin käyttöä (Ikola 1954: 226).

Suomen eksistentiaalilauseen VS-sanajärjestyksen voidaan katsoa juontavan lausetyypin normaalilauseista poikkeavasta informaatorakenteesta: kun normaalilauseen subjekti on oletusarvoisesti lauseen alussa – sen teemapaikassa – on eksistentiaalilauseen teemana neutraalisti paikallissijainen adverbiaali, ja verbinjälkeinen subjektijäsen käsitetään remaattiseksi (ISK: 1310, 1313). Tässä artikkelissa keskeistä on kuitenkin lausetyypin sisäinen variaatio ja sen tarkasteleminen korpus-tutkimuksen metodein, kuten erilaisten sanajärjestyksvarianttien konkreettisia esiintymisympäristöjä tutkimalla.

Eksistentiaalilauseesta puhuttaessa on mielestäni perusteltua erottaa kaksi rinnakkaista termiä. *Eksistentiaalilauseella* tarkoitan prototyyppistä ja abstraktia lausetyyppiä ja *eksistentiaalisella lauseella* tarkoitan todellisessa kielenkäytössä tuotettua tekstilauseetta, joka ainakin joiltakin osin vastaa eksistentiaalilauseelle asetettujen muotopiirteiden kriteereitä.

Kun tietyn rakenteen ilmenemistä oppijankielessä tarkastellaan korpuksen avulla, on yksi luonteva tarkastelutapa frekvenssianalyysi. Tällä tavoin voidaan tavoittaa oppijankielelle ominainen sisäinen variaatio ja jatkuva kehittyminen ja asettaa tilannesidonnainen variaatio osaksi koko käytetyn kielen kirjoa (Cancino jt 1978: 209–210; Ellis, Barkhuizen 2005: 94). Näin oppijan välikieltä (engl. *interlanguage*, ks. tarkemmin Selinker 1972) voidaan tarkastella omana kokonaisuutenaan ja välttää välikielen ja kohdekielen välisestä tarkastelusta syntyvä vertailuharha (engl. *comparative fallacy*, ks. tarkemmin Bley-Vroman 1983: 2).

Tämän artikkelin aineisto ei yksin riitä kvantitatiiviseen tarkasteluun, mutta tulokset voi nähdä pilottikokeina, ja ne voivat avata uusia näkökulmia korpuksella tehtävään oppijankielen tutkimukseen. Suomen eksistentiaalilause tuskin on omaa, muista poikkeavaa kognitiivista alkuperää, vaan kaikki lausetyyppeihin vaikuttavat tekijät ovat ennemminkin jokaiseen

tekstilauseeseen vaikuttava rihmasto. Eksistentiaalilauseen tunnuspiirteet – subjektin sijamerkintä, subjektin ja predikaatin välinen inkongruenssi ja VS-sanajärjestys – ovat kaikki osia syntaktisesti hyvin perustavanlaatuisista mekanismeista. Eksistentiaalilauseessa ne kuitenkin poikkeavat joiltakin osin kunkin muotopiirteiden prototyyppisestä käytöstä, ja esimerkiksi eksistentiaalilauseen subjektin käyttöä määrittävät samat tekijät kuin passiivilauseen objektia (Hakanen 1972: 44). Analogia eli rakenteellinen samankaltaisuus on yksi kielen omaksumisen keskeisiä mekanismeja (ks. esim. Itkonen 2005: 93). Opittaessa uusia kielellisiä muotoja ne suhteutetaan jo tunnettujen rakenteiden suhdeverkostoon. Näin kielen osajärjestelmät vaikuttavat toisiinsa, vaikka niillä ei olisikaan suoraa yhteyttä (Goldberg 1995: 71–72).

Tässä mielessä oppijansuomen eksistentiaalisten lauseiden sanajärjestyksiä tarkasteltaessa on kiinnostavaa tarkastella eri sanajärjestysvalintojen myötäesiintymiä eli tekijöitä, joiden esiintyminen lisää jonkin toisen ilmiön tai elementin esiintymisen todennäköisyyttä. Erityisesti lauseiden sisäiset kolligaatiot eli leksikaalis-kieliopilliset suhteet ja näiden synnyttämät valintapreferenssit ovat seikkoja, joihin on syytä kiinnittää huomiota aineistoa analysoitaessa. (Myötäesiintymä, kolligaatio ja valintapreferenssi, ks. Jantunen 2004: 15, 29). Kuten Jarmo Jantunen toteaa, myötäesiintymien kaltaisen kontekstuaaliset epätyypillisyydet ovat piirteitä, joita voi tavoittaa välikielestä vain tarkastelemalla suuria aineistomassoja (Jantunen 2008: 83). Keskityn tässä artikkelissa yhteen eksistentiaalilauseelle ominaisena pidettyyn piirteeseen, sen sanajärjestykseen ja siinä esiintyvään variaatioon. Käsiteltävät esiintymisympäristöt ja myötäesiintymät ovat kuitenkin puhtaasti aineiston esiinnostamia kielenpiirteitä. Tarkastelussani on siis piirteitä niin korpuspohjaisesta kuin korpusvetoisestakin analyysistä (vrt. esim. Tognini-Bonelli 2001: 65–68, 84–87).

3. Aineiston eksistentiaaliset lauseet

Aineistoni 2575 virkeriviä jakautuvat yhteensä 4028 lauseeseen, ja se on kokonaisuudessaan osa LAS2-korpuksen tenttivastausaineistoa. Vertailuaineistona on 1310 virkeriviä ensikielisten suomenpuhujien tenttivastauksia yhteensä 27 informantilta (VA). Aineisto on edistyneiden suomenoppijoiden tenteissä kirjoittamaa suomea, eikä aineistoa ole valikoitu muin tavoin. Aineisto koostuu korpukseen vuonna 2007 tallennetuista tenttivastauksista. Tenttivastauksia pidetään oppijan kielen korpustutkimukseen hyvin soveltuvana aineistona, koska tällöin tuottamistilanne on hyvin samankaltainen keruuajankohdasta riippumatta. Tenttitapahtumaa voidaan myös pitää kohtalaisen autenttisenä kielentuottamistilanteena, sillä siinä informantti tuottaa kieltä jostain muusta syystä kuin kielen tutkimuksen aineistoksi. (Ks. esim. Granger 2002: 8–9.) Virkerivillä tarkoitetaan virkettä tai virkkeen kaltaista tekstuaalista kokonaisuutta.

Tätä artikkelia kirjoitettaessa korpuksen laaja syntaktinen ja morfologinen koodaus (ks. esim. Ivaska, Siitonen 2009) ei vielä ole käytettävissä. Analyysi onkin tehty digitalisoidusta raakatekstistä, mutta olen valinnut sovelletut metodit niin, että tarvittavat haut voitaisiin toteuttaa koodauksen avulla. Aineiston yhdeksän informanttia (jatkossa LAS2-1...LAS2-9) ovat hyvin heterogeeninen ryhmä³. Ensikieliset ovat englanti (LAS2-8), komi (LAS2-6), liettua (LAS2-5, LAS2-9), ruotsi (LAS2-3), tšekki (LAS2-2), venäjä (LAS2-1, LAS2-7) ja viro (LAS2-4).

Korpuksen avulla tehtävä tutkimus vaatii määriteltävissä olevat hakuehdot, tässä tapauksessa eksistentiaalisten lauseiden muotopiirteisiin perustuvat kriteerit, joita vasten variaatiota

³ Informanteista tarkemmin ks. esim. Ivaska 2010: tulossa.

voidaan sitten tarkastella (ks. tarkemmin Ivaska 2010: tulossa). Pidän muodoltaan prototyypinäkseen eksistentialilauseena Hakasen (1972) määrittämää mallia. Prototyypinäkseen lauseen kriteereitä – kutakin vuorollaan – väljentämällä voidaan edelleen tavoittaa niissä esiintyvä variaatio ja Tuomas Huumon ja Jari Perkon hahmottelema lauseiden eksistentialisuuden jatkumo (Huumo, Perko 1993: 389–391). Esimerkit (1) ja (2) ovat aineistoni sellaisia eksistentialisia lauseita, jotka täyttävät kaikki eksistentialilauseen tunnusmerkit.

(1) suomalais-ugrilaiset kielet ovat agglutinoivia, *niistä suurimmassa osassa on SOV-sanajärjestys*⁴ (subjekti-objekti-verbi). (las2-4tt01te04lo02⁵)

(2) *Saamessa ei ole soinnillisia konsonantteja.* (las2-7tt01te04lo03)

Näiden kriteerien perusteella tutkimusaineistossa on 449 sellaista eksistentialiseksi luokiteltavaa lausetta, joissa voidaan tarkastella subjektin ja predikaatin välistä sanajärjestystä. Prosenttiosuuksina ilmaisten 11 %:ssa tutkimusaineiston lauseista on eksistentialisia ja 17 %:ssa virkeriveistä on eksistentialinen lause. Eksistentialisten lauseiden joukosta 407 lauseessa (10 % ; 90 %)⁶ verbi edeltää subjektia. 42 (1 % ; 10 %)

⁴ Olen kursivoivut eksistentialiseksi tulkitsemäni lauseen ja lihavoinut sen subjektiksi käsittämäni jäsenen. Sama merkintätapa on käytössä läpi artikkelin.

⁵ Kunkin esimerkin lopussa on LAS2-korpukseen viittaava tunnistetieto, joka kertoo informantin (las2-4), tekstilajin (tt01 eli tenttivastaus, tässä artikkelissa ei muita tekstilajeja), tekstikokonaisuuden numeron (te06) ja tekstikokonaisuuden sijoittumisen kyseisen informantin tarkastelujaksolle kuukausina jakson alusta (lo02).

⁶ Sulkumerkinnän ensimmäinen luku (10 %) kertoo prosenttiosuuden kaikista aineiston lauseista ja jälkimmäinen (90 %) prosenttiosuuden eksistentialisiksi luokiteltujen lauseiden joukosta. Sama merkintä on käytössä läpi koko artikkelin.

tapauksessa subjekti edeltää verbiä. Edelleen, näistä SV-tapa-
uksista 16 lausetta (0,4 % ; 4 %) on ns. kvanttorilauseita (ks.
esim. Vähämäki 1987, 185–186; ISK, 848), joissa subjektijäsen
on jo prototyypillisesti verbinetisessä asemassa. Vertailuaineis-
tossa puolestaan on 139 sellaista eksistentiaaliseksi luokitel-
tavaa lausetta, joista voidaan tarkastella subjektin ja predikaatin
sanajärjestystä. Näistä lauseista 121:ssä on prototyypinen VS-
sanajärjestys ja 18:ssa SV-sanajärjestys. Eksistentiaalisten lausei-
den jakaumaa sanajärjestyksen kannalta havainnollistaa
asetelma 1.

Asetelma 1. Tutkimusaineiston eksistentiaalisten lauseiden jakauma
sanajärjestyksen kannalta. Sulkeiden ensimmäinen luku kertoo pro-
senttiosuuden kaikista lauseista ja jälkimmäinen prosenttiosuuden
eksistentiaalisiksi luokitelluista lauseista.

	LAS2	VA
VS	407 (10 % ; 90 %)	121 (87 %)
myönteisiä	87 %	91 %
kielteisiä	13 %	9 %
SV	42 (1 % ; 10 %)	12 (13 %)
myönteisiä	62 %	64 %
kielteisiä	38 %	46 %
joista alisteisia sivulauseita	4 (0,1 % ; 1 %)	7 (5 %)
kvanttorilauseita	16 (0,4 % ; 3,9 %)	8 (6 %)
ed. lauseiden viittaussuhteiden määrittämiä	2 (0,05 % ; 0,5 %)	2 (1 %)
Kysymyssana motivoi sanajärjestyksen <i>on olemassa</i> -lauseita	2 (0,05 % ; 0,5 %)	
Sanajärjestykselle ei motivoijaa	6 (0,15 % ; 1,5 %)	1 (1 %)
Yht.	449 (11 % ; 100 %)	139 (100 %)

Voidaan sanoa, että tutkimus- ja vertailuaineiston jakaumat ovat verraten lähellä toisiaan, ja että sanajärjestyksen kannalta eksistentiaalisten lauseiden käyttö on edistyneillä suomenoppijoilla kohtalaisen idiomaattista. Korpusvetoisen lähestymistavan kannalta on kuitenkin mielenkiintoista tarkastella niiden eksistentiaalisten lauseiden esiintymisympäristöjä, joissa on epäprototyypinen SV-sanajärjestys.

On luonnollisesti mahdotonta osoittaa yksiselitteisesti, mitkä tekijät ovat laukaisseet epäprototyypisen SV-sanajärjestyksen niihin tutkimusaineiston lauseisiin, joista ei ole löydettävissä mitään kohdekielen mukaista inversion motivoijaa. Se, että lauseiden sanajärjestys on kohdekielen kannalta epäidomaattinen, voi kuitenkin olla osaltaan myötäesiintymien ja kolligaatiosuhteiden vaikutusta. Jantunen toteaa tällaisten ko(n)tekstuaalisten epätyypillisyyksien olevan sellaisia välikielen piirteitä, jotka on mahdollista tavoittaa vain suurilla aineistomassoilla tarkastelemalla (Jantunen 2008: 83). Jokin ilmiö saattaa analogisen omaksumisen myötä saada sille kohdekielen kannalta kuulumattomia koodausominaisuuksia, mikäli kyseisen ilmiön todennäköinen esiintymisympäristö kohdekielessä tukee tällaisen analogian syntyä. Tämän artikkelin aineisto ei riitä luotettavien päätelmien tekemiseen, mutta nämä huomiot saattavat silti antaa viitteitä tarkempaa käsittelyä kaipaavista ilmiöistä.

Yksi mahdollinen kolligaatiosuhde niissä eksistentiaalisissa lauseissa, joissa subjekti edeltää predikaattia, on negaatio. Merkillepantavaa on ennen kaikkea se, että SV-sanajärjestys on suhteellisesti huomattavasti todennäköisempi kielteisissä lauseissa kuin myönteisissä lauseissa. Tämä huomio pitää paikkansa niin tutkimus- kuin vertailuaineistossakin. ISK ei mainitse sanajärjestyksestä käsitellessään kiellon ilmaissuhteita (ISK, 1534–1535). Kieltoa ja sen vaikutusalaa ilmaistaan muodon kannalta pääverbin erityisellä taiputusmuodolla sekä objektin,

eksistentiaalilauseen subjektin ja osman partitiivisijaisuudella. Näiden lisäksi kiellolla voi olla ns. saattoilmiöinä kieltokausia sananmuotoja ja sanoja. (Ks. mt., 1536.) Kieltolause voi kuitenkin usein olla luonteeltaan reaktiivinen, minkä vuoksi konteksti saattaa motivoida sanajärjestyksen muutoksen. Näin ollen negaatio ei itsessään motivoi epäprototyypistä sanajärjестystä mutta niiden yhteisesiintyminen on odotuksenmukaista. (Mt., 1326.)

Negaatio on epätyypillisen sanajärjестyksen yleinen myötäesiintymä, ja kieltolause onkin epätyypilliselle sanajärjестykselle suhteellisesti todennäköisempi esiintymisympäristö kuin myönteinen väitelause. Tällainen yhteisesiintyminen saattaa analogisen oppimisen myötä synnyttää oppijan välikieleen kohdekielen vastaisen koodausominaisuuden. Päätelmä ei ole tilastollisesti pätevä, mutta se esittää yhden mahdollisuuden kohdekielen myötäesiintymien potentiaalisesta vaikutuksesta oppijan välikieleen. Esimerkit (5) – (7) ovat tapauksia, joissa kieltolauseessa on eksistentiaalilauseelle epäprototyypillinen SV-sanajärjестys ilman kontekstuaalista motivoijaa tai sellaista reaktiivisuutta, että ratkaisu tuntuisi idiomaattiselta kohdekielen kannalta.

- (3) *Näissä sanoissa **astevaihtelua** ei ole*, eli geminaattoja esiintyy, koska ne kuuluvat sanoihin (sanojen vartaloihin) sellaisenaan. (las2-5t01te02lo02)
- (4) Suomeen on kehittynyt *-is(i) konditionaali, **-is(i) konditionaalina** ei ollut sgr. kantakielessä*. (las2-6tt01te03lo05)
- (5) linna: ei voi olla ikivanha sana, ***nn-geminaattaa** ei ollut* (las2-6tt01te02lo04)
- (6) Viime vuosisadalla eteläkarjalan murretta puhuttiin myös Valdain alueella, *mutta nykyään siellä **karjalaisia kyliä** ei ole*. (las2-7tt01te05loxx)

- (7) *Diftongia* kantakielessä ei ollut vaan oli syntynyt pitkistä vokaalista ē, ö ja myös esim. pohjoissaamessa pitkiä kantakielen vokaalia on jäänyt mutta suomessa on diftongutoineet {esimerkki} kieli~keel, nuori~noor. (las2-8tt01te02lo00)

Kiellon ja sanajärjestyksen vaikutusta toisiinsa nimenomaan eksistentiaalisissa lauseissa puoltaa se, että yleiskielessä näiden ilmiöiden välistä suhdetta ei näytä ainakaan pilottitarkastelun perusteella olevan. Tutkin sanajärjestyksen ja lauseen kielteisyyden välisen suhteen Lauseopin X-arkiston (LaX) yleiskielen aineistosta tietokirjallisuuden kielestä, tietosanakirjojen alaryhmästä. Aineistossa on 854 lausetta, joista 508:sta voidaan määrittää subjektin ja predikaatin välinen sanajärjestys (LaX-1). Valitsin aineiston, koska se on tekstilajina kohtalaisen lähellä tutkimusaineistona olevia tenttivas-
tauksia. Hakutulokset osoittavat, että tässä aineistossa sanajärjestyksellä ja kiellolla ei vaikuta olevan mitään vaikutusta toisiinsa, sillä kielteisten väitelauseiden suhteellinen määrä on niin SV- kuin VS-sanajärjestyksissäkin lauseissa 4 % kaikista väitelauseista (LaX-2...LaX-5). Asetelma 2 havainnollistaa jakaumaa tarkemmin.

Asetelma 2. LaX-arkiston yleiskielen aineiston tietokirjallisuuden kielen ja tietosanakirjojen alaryhmän sellaisten lauseiden esiintymät, joista voidaan määrittellä subjektin ja predikaatin välinen järjestys.

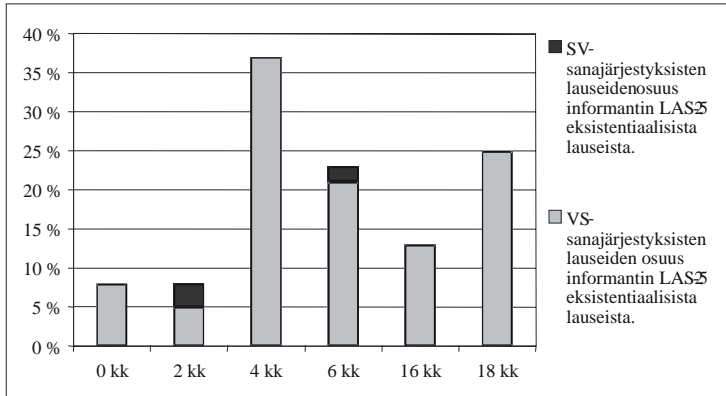
SV	376 kpl	74 %
myönteisiä (LaX-2)	361 kpl	96 % (SV:istä)
kielteisiä (LaX-3)	15 kpl	4 % (SV:istä)
VS	132 kpl	26 %
myönteisiä (LaX-4)	127 kpl	96 % (VS:istä)
kielteisiä (LaX-5)	5 kpl	4 % (VS:istä)
Yht.	508 kpl	100 %

Lauseen kielteisyys voidaan tulkita sanajärjestykseltään epä-prototyyppisen eksistentiaalisen lauseen myötäesiintymäksi sekä tutkimusaineistossa että vertailuaineistossa. Sanajärjestykseltään prototyyppisten eksistentiaalisten lauseiden jakauma tutkimusaineistossa vastaa likimain Hakulisen, Karlssonin ja Vilkunin esittämää myönteisten ja kielteisten lauseiden jakaumaa suomen kielessä yleisesti. Kun tämän tutkielman aineiston sanajärjestykseltään prototyyppisistä eksistentiaalisista lauseista 87 % on myönteisiä, on Hakulisen, Karlssonin ja Vilkunin saama luku 91 %. (Hakulinen, Karlsson, Vilkuna 1980: 120.) Kuten aiemmin todettiin, ei kielteisyys yksinään riitä selittämään odotuksenvastaista sanajärjestyksiä. Ilmiöiden yhteisesiintyminen on kuitenkin odotuksenmukaista. Analogisen oppimisen myötä myötäesiintymästä voi nousta yksi tekijä, joka on edistänyt näiden tutkimusaineiston tekstilauseiden subjektijäsenten järjestyksistä verbineitiseen asemaan juuri eksistentiaalisissa lauseissa.

4. Pitkittäismuutokset kahden informantin aineistossa

Aineiston koko ja tarkastelujakson pituus mahdollistaa pitkittäismuutosten tarkastelun informanttien LAS2-5 ja LAS2-8 aineistosta – molemmilla tarkastelujakso on 18 kuukautta. Tarkastelen eksistentiaalisten lauseiden esiintymistä niiden sanajärjestyksen kannalta tarkastelujakson eri vaiheissa.

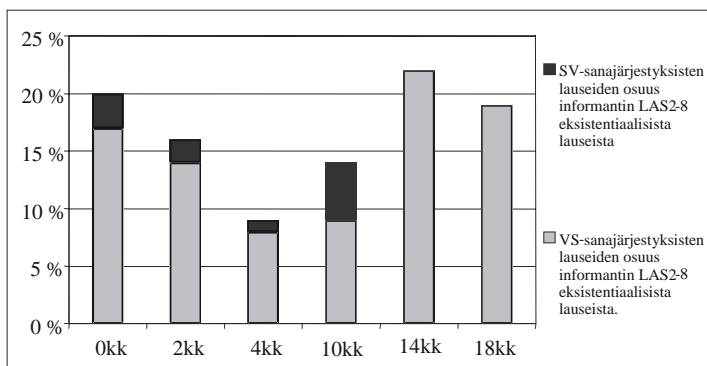
Informantin LAS2-5 aineiston 66 eksistentiaalista lausetta ovat lähes poikkeuksetta sanajärjestykseltään prototyyppisiä VS-lauseita. Tendenssi pysyy samanlaisena koko tarkastelujakson ajan. Kuvio 1 havainnollistaa eksistentiaalisten lauseiden esiintymistä ja niiden sanajärjestyksiä tarkastelujakson aikana.



Kuvio 1. Eksistentialisten lauseiden osuus ja sanajärjestys subjektin ja verbin osalta informantin LAS2-5 aineistossa. Prosenttiluku kertoo eksistentialisten lauseiden osuuden saman informantin virkeriveistä kunakin tarkasteluajankohtana.

Eksistentialisten lauseiden osuus kaikista lauseista kasvaa ajan kuluessa. Tarkastelujakson alussa ja kahden kuukauden kohdalla eksistentialisten lauseiden määrä (8 %:ssa virkeriveistä on eksistentialinen lause) on sekä vertailuaineiston että tutkimusaineiston keskiarvoja pienempi ja 99 virkerivissä on yhteensä 8 eksistentialista lausetta. Neljän kuukauden kohdalla eksistentialisia lauseita on 37 %:ssa virkerivejä eli 13 lausetta 35 virkerivissä. Määrän nopea nousu voisi viitata tekstin sisällölliseen poikkeuksellisuuteen, mutta aineistojakso ei tunnu poikkeavan oleellisesti sitä edeltävistä tai seuraavista jaksoista. Tämän jälkeen eksistentialisten lauseiden suhteellinen määrä laskee hieman. Tarkastelujakson lopussa eksistentialisia lauseita on 44 virkerivissä 11 kappaletta. Neljän kuukauden kohdalla tapahtuneesta määrän kasvusta alkaen määrä on lähes kaksi kertaa niin suuri kuin vertailuaineistossa. Eksistentialiset lauseet ovat siis selvästi yliedustuneina informantin LAS2-5 aineistossa. Luvut eivät ole tilastollisesti merkitseviä, mutta niiden osoittama tendenssi on selvä.

Informantin LAS2-8 tarkastelujakso on niin ikään 18 kuukautta. Tarkastelujakson eksistentiaalisten lauseiden suhteellinen määrä ja jakauma sanajärjestyksen kannalta on osin yllättävä. Kaikkiaan 16 %:ssa virkeriveistä on eksistentiaalinen lause, ja määrä on siis 6 prosenttiyksikköä vertailuaineiston määrää suurempi. Tarkastelujakson alussa eksistentiaalisia lauseita on 20 % virkeriveistä ja jakson lopussa 19 %. Ääripäiden välissä eksistentiaalisten lauseiden määrä kuitenkin laskee jopa alle 10 %:iin virkeriveistä. Tarkastelujakson alkuvaiheessa informantin eksistentiaalisissa lauseissa on sekä VS-että SV-sanajärjystä, mutta lopussa eksistentiaalilauseelle prototyyppinen VS-sanajärjestys muuttuu yksinomaiseksi. Kuvio 2 havainnollistaa eksistentiaalisten lauseiden jakaumaa informantin LAS2-8 aineistossa tarkastelujakson eri vaiheissa.



Kuvio 2. Eksistentiaalisten lauseiden osuus ja sanajärjestys subjektin ja verbin osalta informantin LAS2-8 aineistossa. Prosenttiluku kertoo eksistentiaalisten lauseiden osuuden saman informantin virkeriveistä kunakin tarkastelujakohtana.

Sekä lauseiden väliaikainen väheneminen että VS-sanajärjestyksen muuttuminen yksinomaiseksi saattavat kuvastaa informantin välikielen muutosta. Informantti on käsittänyt

VS-sanajärjestyksen eksistentiaalilauseelle prototyypiksi ja saattaa pyrkiä välttämään virheitä. Voidaan olettaa, että kyse on eksistentiaalilauseen ja sitä määrittävien funktioiden uudelleenjärjestymisestä oppijan välikielessä. Uuden sanajärjestysvaihtoehdon tultua kieleen se alkaa saada yksiselitteisen muodon ja funktion välisen tehtävän. (Vrt. Ellis 1985: 95–96.)

5. Päätelmät

Edistyneiden suomenoppijoiden eksistentiaalisten lauseiden sanajärjestystä koskevat tulokset osoittavat, että lausetyypille epäprototyypinen SV-sanajärjestys on todennäköisempi silloin, kun lause on kielteinen. Tämä huomio pätee niin tutkimus- kuin vertailuaineistossakin. Vertailuaineiston kaikissa eksistentiaalisten lauseiden SV-tapauksissa inversiolla on kuitenkin osoitettavissa jokin muu motivoija, kun taas tutkimusaineistossa tällaista motivoijaa ei aina ole. Syynä saattaa olla pyrkimys osoittaa reaktiivisuutta edeltävään kontekstiin, jolloin lauseen informaatorakenne, ja sitä myötä sanajärjestys, voi muuttua. Korpusvetoisen tutkimusotteen kannalta inversion ja kieltolauseiden yhteisesiintyminen on niin ikään kiinnostava tulkintavaihtoehto. Informanttien välikielten kieliopeissa kielteisyyden voi analogian myötä olla myös itsessään epäprototyypisen sanajärjestyksen motivoiva tekijä, tai ainakin kielteisyyden voi jo itsessään lisätä inversion mahdollisuutta. Ajatus sopii tuloksiin, jotka Maija Ikonen on saanut pro gradu-tutkielmassaan tarkastellessaan vieraskielisten lukiolaisten kieltolauseita suomenkielisessä kirjoituksessa (Ikonen 2006). Ikonen mukaan kielteisissä lauseissa ainoa sanajärjestyksen kannalta poikkeamaksi tulkittava ero on tarpeeton inversio (mt., 70). Yleisimmin vieras sanajärjestys liittyy subjektin paikkaan lauseessa, mutta tämäkin on verraten harvinaista

(mt., 66).⁷ Analogian syntymistä puoltaa se huomio, että eksistentiaalisten lauseiden SV-sanajärjestys on suhteellisesti todennäköisempi kielteisissä eksistentiaalisissa lauseissa sekä tutkimus- että vertailuaineistossa. Ottamatta kantaa lauseiden informaattiorakenteisiin ja suomen kielen inversiota laukaiseviin mekanismeihin eksistentiaalisen lauseen epäprototyypin SV-sanajärjestyksen ja kiellon yhteisesiintymisen on siis tyypillistä ja odotuksenmukaista myös kohdekielessä ja siitä saatavassa syötöksessä. Kieltä opittaessa oppija muodostaa oppimastaan analogioita, joita soveltamalla hän käyttää kieltä. Nämä analogiat voivat syntyä syötöksessä havaituista tendensseistä, osin riippumatta siitä, noudattaako analogia kohdekielen kieliopin malleja. Ellis toteaa vertikaalista vaihtelua käsitellessään, että tuotokseen vaikuttaa aina myös se, missä määrin kielenoppija osaa soveltaa kieli-taitoaan (Ellis 1988: 168–171). On ilmeistä, että prototyyppisyyden vähetessä myös tämä soveltaminen vaikeutuu, mikä voi kasvattaa suomen kielessä yleisemmän SV-sanajärjestyksen esiintymistodennäköisyyttä.

Tulosten valossa analogisen oppimisen vaikutus välikielen kielioppijärjestelmään vaikuttaa suurelta. Tämä puoltaa tilastollisesti todennäköisten myötäesiintymien tutkimista ja vertailua ensikielisen suomen ja oppijansuomen välillä myös syntaksin osalta. Leksikaalis-kieliopillisten ja kieliopillis-kieliopillisten kolligaatiosuhteiden laajamittainen tarkastelu vaatii mittavia aineistoja ja valistuneisiin hypoteeseihin perustuvaa kohteen tarkkaa rajaamista. Lauseen kielteisyys ja kontekstiin viittaavat elementit ovat sellaisia tekijöitä, joiden vaikutuksen tutkiminen edistyneiden suomenoppijoiden sana-

⁷ Ikosen aineistosta oleva lause *jolloin ei lapsi ajattele paljon asioita* kuvaa tarpeetonta inversiota, jossa subjekti on siirtynyt kieltoverbin jälkeiseen asemaan (Ikonen 2006: 66).

järjestykseen vaikuttaisi mielekkäältä. Nimenomaan korpusvetoisen tutkimusotteen soveltaminen syntaksin tutkimukseen tuntuu mielenkiintoiselta. Tämä tarkoittaa esimerkiksi edellä kuvattujen kaltaisten, tilastollisesti todennäköisten yhteisesiintymien tarkastelua ja niiden vaikutusten arvioimista oppijankieleen. Tässä artikkelissa osoitettujen tendenssien yksityiskohtaisempi tarkastelu suurella korpusaineistolla on kiintoisa jatkotutkimuksen kohde. Kieliopillisten ilmiöiden esiintymisympäristöjen tarkastelu on mahdollista järkevimmin juuri syntaktisesti koodatun korpuksen avulla.

Kirjallisuus

Airila, Martti 1924. Kielellisten ilmiöiden perustelemisesta. – Kieli- ja kansatieteellisiä tutkielmia. Juhlakirja professori E. N. Setälän kuusikymmenvuotispäiväksi 27.2.1924 (Suomalais-ugrilaisen seuran toimituksia 52). Helsinki: Suomalais-ugrilainen seura, 15–26.

Bley-Vroman, Robert 1983. The comparative fallacy in interlanguage studies: The case of systemacity. – *Language Learning* 33:1, 1–17.

Cancino, Herlinda, Ellen J. Rosansky, John H. Scumann 1978. The acquisition of English negatives and interrogatives by native Spanish speakers. – *Second Language Acquisition: A Book of Readings* / Ed. by Evelyn Marcussen Hatch. Newbury House, Rowley (Massachusetts), 207–230.

Ellis, Rod 1985. *Understanding Second language Acquisition*. Oxford: Oxford University Press.

Ellis, Rod 1988. *Classroom Second Language Development*. London: Prentice Hall International (UK) Ltd.

Ellis, Rod, Gary Barkhuizen 2005. *Analysing Learner Language*. Oxford: Oxford University Press.

Golberg, Adele E. 1995. *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.

Granger, Sylviane 2002. A Bird's-eye view of learner corpus research. – Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching / Ed. by Sylviane Granger, Joseph Hung, Stephanie Petch-Tyson. Amsterdam: John Benjamins Publ. Co, 3–33.

Hakanen, Aimo 1972. Normaalilause ja eksistentiaalilause. – Suomen kielen seuran vuosijulkaisu 14, 36–76.

Hakulinen, Auli, Fred Karlsson, Maria Vilkuna 1980. Suomen tekstilauseiden piirteitä: kvantitatiivinen tutkimus (Department of General Linguistics Publications No. 6). Helsinki: University of Helsinki.

Helasvuo, Marja-Liisa, Tuomas Huumo 2010. Mikä subjekti on? – Virittäjä, 165–194.

Huumo, Tuomas, Jari Perko 1993. Eksistentiaalilause lokaalisuuden ilmaisijana. – Virittäjä, 380–399.

Ikola, Osmo 1954. Suomen lauseopin ongelmia I–III. – Virittäjä, 209–245.

Ikonen, Maija 2006. Kieltolauseiden kiellonalaiset poikkeamat vieras-kielisten lukiolaisten suomenkielisessä kirjoituksessa. Pro gradu -tutkielma. Turun yliopiston suomalaisen ja yleisen kielitieteen laitos, suomen kieli.

ISK 2004 = Hakulinen, Auli, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen, Irja Alho 2004. Iso suomen kielioppi (Suomalaisen Kirjallisuuden Seuran Toimituksia 950). Helsinki: SKS.

Itkonen, Esa 2005. Analogy as structure and process. Approaches in linguistic, cognitive psychology and philosophy of science. Amsterdam: John Benjamins Publ. Co.

Ivaska, Ilmari 2009. Eksistentiaalilauseen ilmiöedistyneiden suomenoppijoiden kirjoituksessa. Pro gradu -tutkielma. Turun yliopiston suomalaisen ja yleisen kielitieteen laitos, suomen kieli.

Ivaska, Ilmari 2010 (tulossa). Eksistentiaalilauseen subjektijäsen edistyneiden suomenoppijoiden kirjoituksessa. – Suomen kielen seuran vuosijulkaisu 52.

Ivaska, Ilmari, Kirsti Siitonen 2009. Syntaktisesti koodattu oppijan-kielen korpus: mahdollisuuksia ja ongelmia. – Korpusuuringute metodoloogia ja märgendamise probleemid / Toim. Pille Eslon, Katre Õim (Tallinna Ülikooli eesti keele ja kultuuri Instituudi toimetised 11). Tallinn: Tallinna Ülikooli kirjastus, 54–71.

Jantunen, Jarmo 2004. Synonymia ja käännösuomi: korpus-näkökulma samamerkityksisyyden kontekstuaalisuuteen ja käännös-kielen leksikaalisiin erityispiirteisiin (Joensuun yliopiston huma-nistisia julkaisuja). Joensuu.

Jantunen, Jarmo 2008. Haasteita oppijankielen analyysille: oppijan-kielen universaalit. – Öppijakeele analüüs: Võimalused, probleemid, vajadused / Toim. Pille Eslon (Eesti filoloogia osakonna toimetised 10). Tallinn: Tallinna Ülikooli kirjastus, 67–91.

Jespersen, Otto 1924. *The Philosophy of Grammar*. London: George Allen & Unwin Ltd.

Selinker, Larry 1972. Interlanguage. – *International Review of Applied Linguistics* 10, 209–231.

Tiainen, Outi 1997. Suomen eksistentiaalilause – päättymätön tarina. – *Virittäjä*, 563–571.

Tognini-Bonelli, Elena 2001. *Corpus Linguistics at Work* (Studies in Corpus Linguistics 6). Amsterdam: John Benjamins Publ. Co.

Vähämäki, Börje 1987. On defining basic sentences in Finnish. – *Fennistica festiva in honorem Göran Karlsson* / Toim. Mauno Koski, Eeva Lähdemäki, Kaisa Häkkinen. Turku: Åbo Akademis förlag, 174–188.

Aineslähteet

LAS2 = Edistyneiden suomenoppijoiden korpus. Suomen kielen oppiaine, kieli- ja käännöstieteiden laitos, Turun yliopisto. Turku. Aineistona on korpukseen vuonna 2007 tallennetut tenttivastaukset.

- LaX = Lauseopin X-arkisto. Kotimaisten kielten tutkimuskeskus – Turun yliopiston kieli- ja käännöstieteiden laitos, Lauseopin arkisto. Turku
- LaX-1 = Aineisto: Yleiskieli/Tietokirjallisuuden kieli/Tietosanakirjat
XPath: //cl[w[@fun='npsubj'] and w[@fun='pred']]
[Viitattu 02.07.2010.]
- LaX-2 = Aineisto: Yleiskieli/Tietokirjallisuuden kieli/Tietosanakirjat
XPath: //cl[@type='affdecl' and w[@fun='npsubj'] and following-sibling::w[@fun='pred']]
[Viitattu 02.07.2010.]
- LaX-3 = Aineisto: Yleiskieli/Tietokirjallisuuden kieli/Tietosanakirjat
XPath: //cl[@type='negdecl' and w[@fun='npsubj'] and following-sibling::w[@fun='pred']]
[Viitattu 02.07.2010.]
- LaX-4 = Aineisto: Yleiskieli/Tietokirjallisuuden kieli/Tietosanakirjat
XPath: //cl[@type='affdecl' and w[@fun='pred' and following-sibling::w[@fun='npsubj']]]
[Viitattu 02.07.2010.]
- LaX-5 = Aineisto: Yleiskieli/Tietokirjallisuuden kieli/Tietosanakirjat
XPath: //cl[@type='negdecl' and w[@fun='pred' and following-sibling::w[@fun='npsubj']]]
[Viitattu 02.07.2010.]
- VA = tekijän hallussa oleva vertailuaineisto

What Can a Syntactically Encoded Corpus Reveal about Word Order?

Predicate and Subject in the Existential Sentences of Advanced Learners of Finnish

Ilmari Ivaska

Summary

Existential sentence is a clause type that is included in all clause-type classifications of Finnish. A prototypical existential sentence differs from so-called normal clauses of Finnish in the word order, case marking of the subject constituent and the lack of congruence between subject and verb. This paper discusses the word order in the existential sentences of advanced learners of Finnish. The study is part of the Advanced Finnish Learners' Corpus -project (LAS2) at the University of Turku.

The results show that the word order is VS in 90 % of the existential sentences and SV in 10 % of the existential sentences. The division is roughly the same in the comparison material. The analysis of the co-occurrences reveals that SV word order, the non-prototypical variant in the existential sentences, is more probable in the negative sentences both in the study material and in the comparison material. This lexicogrammatical co-occurrence works only in the context of the existential sentences as, in general, the negation does not impact on the word order of Finnish.

However, while something invariably motivated the SV word order in the comparison material, those motivators could not always be pinpointed in the study material. This might

suggest that given the analogy with the target language, negation may have become a motivator of word order inversion in existential sentences in the inter-languages of the informants.

The results show that analogical learning appears to be of great significance in the grammatical system of the inter-language. This encourages us to research the quantitatively probable co-occurrences in the learner language also from the syntactical point of view. A more thorough study on the tendencies pointed out in this article would hopefully tell us more about the grammatical and lexico-grammatical collocations.

Keywords: Finnish as a Second Language, Corpus Linguistics, Word Order, Negation, Syntax

Esittely

FM Ilmari Ivaska on työskennellyt Edistyneiden suomenoppijoiden korpus -tutkimushankkeessa sen alusta (2007) alkaen. Ivaska kirjoitti pro gradu -tutkielmansa edistyneiden suomenoppijoiden eksistentiaalisista lauseista. Tällä hetkellä hän tekee korpusvetoisena tutkimuksena toteutettavaa, edistyneiden suomenoppijoiden lausetyyppejä käsittelevää väitöskirjaa.

ilmari.ivaska@utu.fi