

# Maailman kielten ja kielitieteiden valintakoe 2023

## Osa 1: Aineistotehtävä (0-50 p)

Alla on 20 monivalintatehtävää, jotka perustuvat liitteenä oleviin katkelmiin Turo Vartiaisen ja Tanja Säilyn artikkelista "Historiallinen korpuslingvistiikka" (teoksessa M. Hamunen ym. (toim.) Kielentutkimuksen menetelmiä, 2020). Lue teksti ja vastaa sen perusteella tehtäviin. Valitse jokaisessa kohdassa se vaihtoehto, jota pidät oikeana. Jokaiseen tehtävään on vain yksi oikea vastaus. Oikeasta vastauksesta saa 2,5 pistettä, vastaamatta jättämisestä ja väärästä vastauksesta 0 pistettä. Voit valita vain yhden vaihtoehdon. Vastausvaihtoehdot ovat satunnaisessa järjestyksessä.

Aineisto: Historiallinen korpuslingvistiikka. *Aineisto on poistettu tekijänoikeudellista syistä.*

### 1.1. Mihin historiallinen korpuslingvistiikka perustuu?

- **Historiallisiin kieliaineistoihin, jotka ovat usein säilyneet sattumanvaraisesti**
- Kielikuntien oletettujen varhaisvaiheiden eli protokielen rekonstruktioihin
- Yksityiskohtaisiin kielioppeihin
- Kielenkäyttäjille tehtyihin kielenkäyttöä koskeviin kyselyihin

### 1.2. Mitä seuraavista ei ole mainittu historiallisen korpuslingvistiikan harjoittajien tutkimuksen kohteeksi?

- Yksittäisen kielellisen rakenteen vaihtelu ja muutos
- Kieliyhteisöjen sosiaalisten suhteiden vaikutus kielenmuutokseen
- Muuttuneen piirteen omaksumisen vauhti tietyssä ihmisryhmässä
- **Yksittäisen kielenpuhujan käsitykset kielenmuutoksesta**

### 1.3. Miksi historiallisessa korpuslingvistiikassa käytetään usein tilastollisia menetelmiä?

- Jotta voidaan arvioida onko kyseessä synkroninen vai diakroninen frekvenssi
- Jotta voidaan arvioida tarkasti ajankohta, jolloin kielenmuutos tapahtui
- **Jotta voidaan arvioida kertooko ilmiön esiintyminen todennäköisesti kielenmuutoksesta vai satunnaisesta vaihtelusta**
- Jotta korpusaineistoon saadaan sattumanvaraista vaihtelua

### 1.4. Mitä rajoitteita korpusaineistojen käytössä mainitaan olevan?

- Tutkijat eivät voi luottaa kielenpuhujien intuitioihin
- **Tutkijat eivät voi saada tietoa kielimuodon puhujilta suoraan**
- Tutkijat eivät voi selittää ilmiöitä sosiaalisten tai psykologisten prosessien kautta
- Tutkijat eivät voi soveltaa esimerkiksi nykyihmisen kognition tutkimuksia tuhansien vuosien takaisten ihmisten kognitioon

### 1.5. Mitä tarkoitetaan merkityksen subjektifikaatiolla?

- Fyysisen liikkeen tai muun konkreettisen tapahtuman olemista lähimpänä subjektin toimintaa
- Kielenkäytön kontekstien siirtymistä subjektin ilmaisuun piiriin
- **Kielenpuhujan taipumusta siirtää sanojen ja rakenteiden merkitystä lähemmäs itseään**
- Kieliopillisen subjektin taipumusta siirtä ilmaisemaan mielipiteitä ja olettamuksia

### 1.6. Minkälainen on edustava ja tasapainoinen yleiskorpus?

- **Sen voidaan ajatella kuvastavan yhteiskunnan kielenkäytön eri tilanteita tasapuolisesti**
- Epämuodollisia tekstejä, kuten sähköposteja, ei ole syytä sisällyttää edustavaan korpukseen
- Se sisältää tekstejä kaikista eri tekstilajeista, joissa kielenkäyttäjien ikä ja sukupuoli on tasaisesti jakautunut
- Kielenkäyttäjien sukupuolta ja ikää ei pidä ottaa mukaan, jotta tasa-arvoinen tutkimus on mahdollista.

### 1.7. Minkälainen korpus ei ole historiallisessa korpuslingvistiikassa tarkoituksenmukainen?

- Useita eri tekstilajeja tasaisesti sisältävä
- Vain yhtä ajanjaksoa koskeva
- Vain yhtä sosiaalista ryhmää edustava
- **Useita ajanjaksoja satunnaisesti kattava**

### 1.8. Miksi Brown-korpusperheessä on useita erillisiä korpuksia?

- Alkuperäisen korpuksen rinnalle on kerätty samoilla periaatteilla ajallisesti uudempia tekstejä historiallisen muutoksen seuraamiseksi sekä britti- että amerikanenglannissa
- Alkuperäisen korpuksen kokoaminen oli tehty epätasaisilla valintaperiaatteilla ja myöhemmät korpuksot paikkaavat sen aukkoja, mm. brittienglannin osalta
- **Alkuperäisen korpuksen rinnalle on kerätty samoilla periaatteilla sekä uusia että vanhempia tekstejä historiallisen muutoksen seuraamiseksi sekä britti- että amerikanenglannissa**
- Vuonna 1964 julkaistu alkuperäinen korpus ei enää ollut ajan tasalla joten sitä on pitänyt päivittää

### 1.9. Mitä ongelmia varhaisten kielivaiheiden tutkijat kohtaavat?

- **Aineistoa ei aina ole säilynyt siten, että se edustaisi tasapuolisesti eri tekstilajeja ja eri sosiaaliluokkien kielenkäyttöä**
- Aineistoa ei ole säilynyt juurikaan yhteiskunnallisesti tärkeiltä aloilta ja alueilta
- Vain yksityishenkilöiden kirjeistä ja päiväkirjamerkinnöistä ei voida muodostaa tutkimukseen sopivaa korpusta
- Valinnanvaraa tekstikorpuksen kokoamiseen historiallisten kielivaiheiden osalta on liian monipuolisesti

### 1.10. Miten murteiden rooli vaikuttaa tutkijan työhön historiallisisten kielikorpuksen käytössä?

- On tärkeää, että historiallisissa korpuksissa on eri murteita edustavaa aineistoa, jotta saadaan riittävän monipuolinen kuva yleiskielen kehityksestä diakronisesti
- **Tietyn historiallisen kielivaiheen aineisto voi keskittyä eri murrealueelle kuin sitä edeltävän tai seuraavan kielivaiheen aineisto, jolloin yleisen vaihtelun ja murteiden väliset erot pitää osata erottaa toisistaan.**
- Tietyn historiallisen kielivaiheen aineisto voi keskittyä eri murrealueelle kuin sitä edeltävän tai seuraavan kielivaiheen aineisto, jolloin aineistojen keskinäinen vertailu on erityisen hedelmällistä runsaan vaihtelun vuoksi.
- Tutkijan kannattaa valita ja vertailla keskenään toisistaan poikkeavien alueiden, esimerkiksi Turun ja Kuopion, murteiden eri historiallisia vaiheita.

1.11. Mikä seuraavista väitteistä *Helsinki Corpus of English Texts* -kokoelmasta (HC) ei pidä paikkaansa?

- **HC kattaa noin tuhannen vuoden ajanjakson ajalta englanninkielisiä tieteellisiä artikkeleita.**
- HC käsittää useita tarkkaan määriteltyjä tekstikategorioita noin tuhannen vuoden ajalta muinaisenglannista alkaen.
- HC on sanamäärältään pienempi kuin monet nykykorpukset mutta sen tarkat kuvailutiedot ja tekstilajien luokittelut auttavat tutkijoita.
- HC:n avulla voidaan tutkia vain suhteellisen yleisiä kielen ilmiöitä, koska se on kooltaan verrattain pieni.

1.12. Miten korpusaineistojen kehitystä nykypäivänä kuvaillaan?

- **Yhä suurempia aineistoja voidaan saada uusilla menetelmillä aikaiseksi, mutta edelleen pienillä, tarkasti kuvailluilla ja kehittyneillä annotointimenetelmillä rikastetuilla aineistoilla on tärkeä rooli tutkimuksessa.**
- Erittäin suurissa aineistoissa taustatiedot tekstien kirjoittajista, kuvat ja tekstin asetelut ovat yhä suuremmassa roolissa.
- Yhden tekstilajin aineistot tulevat menettämään asemansa, koska ne eivät voi vastata nykypäivän vaatimuksiin rikkaasta datasta digitaalisten ihmistieteiden alalla.
- Uusilla tarkoilla annotoinneilla rikastetut aineistot eivät voi samaan aikaan palvella useiden alojen tutkijoita kuten historioitsijoita, kirjallisuudentutkijoita tai kielentutkijoita.

1.13. Historiallisen korpuslingvistiikan menetelmillä voidaan tutkia esimerkiksi kieliopillistumista eli jonkin sanan tai rakenteen merkityksen muuttumista asteittain kieliopillisemmaksi. Kieliopillistuneena konstruktio

- viittaa konkreettisiin asioihin tai tapahtumiin.
- **esiintyy tiheämmin kuin se esiintyi ei-kieliopillistuneena.**
- abstrahoi aikamuotoja, aspektia tai lukuja.
- muuttaa kielen fonologiaa yksinkertaisemmaksi.

1.14. Mitä muinaisenglannin *willan*-verbin kieliopillistumisesta kerrotaan?

- Sillä ilmaistiin muinaisenglannissa tulevaisuuteen kohdistuvia ennustuksia.
- Kontrahoituneessa muodossaan (*'ll*) se merkitsi 'tahtoa' tai 'haluta'.
- **Satojen vuosien kuluessa siitä tuli futuuria ilmaiseva apuverbi *will*.**
- Muoto kontrahoitui ilmaisujen *sort of*, *kind of* ja *type of* -mallin mukaan.

1.15. Englannin kvanttorit *a lot (of)* ja *a bit (of)* ovat kieliopillistuneet konkreettisista substantiivilausekkeista. Mitä prosessissa on ensisijaisesti tapahtunut?

- Kontrahoituminen
- Asteadverbiaalistuminen
- Binominaalistuminen
- **Rakenteellinen uudelleentulkinta**

1.16. Englannin *be going to* -konstruktion kieliopillistumisprosessissa futuurin ilmaukseksi on korpustutkimuksissa havaittu seuraavia vaiheita:

- **Konkreettista liikettä tarkoittava ilmaisu sai subjektifikaation kautta puhujan aikomuksia korostavia merkityksiä ja sen jälkeen sen käyttö yleisty tulevien tapahtumien ennakoinnissa.**
- Ilmausta ryhdyttiin käyttämään kysymyslauseissa ja *go*-verbin fyysinen ulottuvuus muuttui objektiivisemmaksi.
- Ilmaus alkoi ajan saatossa viitata enenevässä määrin puhetilannetta välittömästi seuraaviin tapahtuviin tapahtumiin samalla kun subjektin merkitys korostui.
- Konstruktio viittasi subjektin fyysiseen liikkeeseen yleisesti myöhäisuusenglannissa, kun varhaisuusenglannissa sillä viitattiin aikomuksiin ja tapahtumiin.

1.17. Artikkelissa mainitaan näennäisaikahypoteesi. Mitä sillä tarkoitetaan?

- Sitä, että naiset ovat kielenmuutoksessa näennäisesti edelläkävijöitä eli käyttävät uusia muotoja aikaisemmin.
- **Sitä, että vanhemman ikäluokan kielenkäyttö edustaa varhaisemman ajanjakson kielenkäytön suuntauksia kuin samanaikainen nuoremman ikäluokan kielenkäyttö.**
- Sitä, että eri-ikäisten ryhmien kielenkäyttöä ei voida verrata historiallisessa korpuslingvistiikassa.
- Sitä, että kielenmuutokseen kuuluva aika on näennäisesti samanlaista eri ihmisryhmien välillä.

1.18. Artikkelissa kerrotaan englannin yksikön kolmannen persoonan presensin päätteiden, variaabelin *-s/-th* kehityksessä. Esimerkkinä toimivasta *eat*-verbistä mainitaan esiintyneen muotoja *eateth*, *eates* ja *eats*. Mitä voidaan sanoa varianttien esiintymisestä tekstin ja kuvan 3 perusteella?

- Persoonapäätte *-eth* levisi etelästä pohjoiseen 1500-luvun alkupuolella, joten päätte *-es* ei saanut vielä laajempaa suosiota.
- **1400-luvun loppupuolella oli ensimmäisiä merkkejä sibilanttiversioiden käytöstä pohjoisilla alueilla vaihtelun koskiessa vokaalisia variantteja *-es/-eth*.**
- Muodot tyyppiä *eates* ja *eats* levisivät laajempaan käyttöön, koska olivat lyhyempiä kuin tyyppi *eateth*.
- Painoton vokaali /e/ alkoi kadota 1400-luvulla ja 1500-luvulla *-s*-variantti syrjäytti sen.

1.19. Mikä seuraavista sosiaalisista tekijöistä ei vaikuttanut variaabelin *-s/-th* kehitykseen Nevalaisen ja Raumolin-Brunbergin kirjekorpuustutkimuksessa?

- Alue, jolta kirjoittaja oli kotoisin
- Kirjoittajan sukupuoli
- Kirjoittajan asema yhteiskunnassa
- **Vastaanottajan suhde kirjoittajaan**

1.20. Minkä vuoksi Nevalaisen ja Raumolin-Brunbergin tutkimuksessa hyödynnettiin monimuuttujamenetelmää?

- Variantin normalisoidun frekvenssin laskemiseksi
- Jotta voitiin muodostaa sosiaalisten kategorioiden alakorpuukset
- **Jotta voitiin arvioida eri muuttujien välistä suhteellista tärkeyttä muutoksessa**
- Jotta hypoteesia voitiin testata  $\chi^2$  -testillä

## Osa 2. Morfeemianalyysi ja lauseenmuodostus (0-50 p)

Tieteen termipankki määrittelee morfeemin seuraavasti: "Määritelmä: kielen pienin omaa merkitystä kantava rakenneyksikkö. Selite: Morfeemi on morfologian perusyksikkö. Sitä vastaava konkreettinen rakenneyksikkö on morfi. Morfeemeja ovat esim. sanavartalot, tunnukset, päätteet ja liitteet. Sekä leksikaalisilla morfeemeilla (esim. sanoilla talo tai minä) että kieliopillisia tehtäviä kantavilla kieliopillisilla morfeemeilla (esim. suomen sijapäätteellä -ssA) on omat merkityksensä."

### Morfeemianalyysi (0-30 p)

Liitteessä on kuvitteellisen meirapaun kielen aineistoa. Alla on kymmenen aineiston morfeemeihin liittyvää monivalintakysymystä, jossa kussakin on yksi oikea vastaus. Oikeasta vastauksesta saa 3 pistettä, vastaamatta jättämisestä ja väärästä vastauksesta 0 pistettä. *Aineisto on poistettu tekijänoikeudellista syistä.*

### Lauseenmuodostus (0-20 p)

Liitteessä on kuvitteellisen wikikin kielen aineistoa. Valitse oikea wikikinkielinen käännös. Oikeita vaihtoehtoja on vain yksi, ja siitä saa 4 pistettä. *Aineisto on poistettu tekijänoikeudellista syistä.*

Tehtävien ratkaisemiseksi kannattaa analysoida aineistosta niin monta morfeemia kuin mahdollista.

### Tehtävä 2.1 Morfeemianalyysi (0-30 p)

Liitteessä on kuvitteellisen meirapaun kielen aineistoa. Alla on kymmenen aineiston morfeemeihin liittyvää monivalintakysymystä, jossa kussakin on yksi oikea vastaus. Oikeasta vastauksesta saa 3 pistettä, vastaamatta jättämisestä ja väärästä vastauksesta 0 pistettä. Vastausvaihtoehdot ovat satunnaisessa järjestyksessä.

Huomioi, että yhtä merkitystä voidaan ilmaista useammalla kuin yhdellä muodolla, eli morfeemilla on tällöin allomorfeja. Huomaa myös, että jotain merkitystä voidaan ilmaista nollamorfeemilla, jolloin sillä ei ole ilmimuotoa. Tehtävän ratkaisemiseksi kannattaa analysoida aineistosta niin monta morfeemia kuin mahdollista. Huomioi, että aineiston kaikki substantiivit ovat määräisessä muodossa.

Seuraavassa on kuvitteellisen meirapaun kielen aineistoa.

Warohai han paw korow. 'Kameli näki minut.'

Warohao pan koron hao. 'Minä näin kamelin.'

Babalisko pen exul hao. 'Minä luen kirjeen.'

Babawalikko pen exul koye. 'Te luitte kirjeen.'

Rohakko han koye. 'Te näette minut.'

Babaliski pen exul se. 'Hän lukee kirjeen.'

Wapaammo pin rata moye. 'Me saimme kärryn.'

Babawaliis pen exul pis ratas soye. 'He lukivat kirjettä kärryssä.'

Paapi pan koron se. 'Hän saa kamelin.'

Warohammo pan koron pes exus moye. 'Me näimme kirjeessä kamelin.'

### 2.1.1. Subjektia verbin yksikön ensimmäisessä persoonassa ilmaisee morfeemi

- o
- i
- ko
- hao

### 2.1.2. Subjektia yksikön kolmannessa persoonassa ilmaisee verbiin liittyvä morfeemi

- o
- i
- se
- ye

### 2.1.3. Objektia ilmaisee mm. morfeemi

- n
- w
- s
- mmo

### 2.1.4. Subjektia verbin monikon ensimmäisessä persoonassa ilmaisee

- n
- w
- s
- **mmo**

### 2.1.5. Mennyttä aikaa ilmaisee morfeemi

- ye
- pan
- **wa**
- pen

### 2.1.6. Lokatiivia eli paikkaa ilmaisee morfeemi

- **s**
- n
- wa
- ba

### 2.1.7. Yksikön ensimmäisen persoonan pronomini nominatiivissa on

- **hao**
- han
- pan
- pen

### 2.1.8. Yksikön ensimmäisen persoonan muoto nykyajan aikamuodossa verbistä 'nähdä' voidaan päätellä aineiston perusteella. Se on

- rohai
- korow
- koron

- rohao

2.1.9. Määräinen artikkeli esiintyy subjektin sijamuodossaan aineistossa vain yhdessä substantiiviluokassa, jota edustaa sana *kameli*. Mikä olisi artikkelin subjektin sijamuoto *kärrysanan* suvussa, kun se on muodostettu saman säännön mukaan?

- pew
- **piw**
- pen
- pin

2.1.10. Voiko *kärry*-sanan nominatiivin (perusmuodon) päätellä aineiston perusteella?

- Voi, sen pääte on samanlainen kuin *kirje*-sanan nominatiivissa.
- Ei voi, sillä meillä ei ole aineistossa muita saman substantiiviluokan sanoja.
- Voi, sen pääte on samanlainen kuin *kameli*-sanan nominatiivissa.
- **Ei voi, sillä emme tiedä tarpeeksi nominatiivimuotojen muodostumisperiaatteista.**

## Tehtävä 2.2 Lauseenmuodostus (0-20 p)

Liitteessä on kuvitteellisen wikikin kielen aineistoa. Valitse oikea wikikin kielinen käännös.

Oikeita vaihtoehtoja on vain yksi, ja siitä saa 4 pistettä. Tehtävän ratkaisemiseksi kannattaa analysoida kaikki aineiston morfeemit.

Fringa ogana plix. 'Peippo söi siemenen.'

Varri ko lõn mane. 'Menemme metsään huomenna.'

Ofringa nachsen kaule sveixo. 'Näin peipon lentävän eilen.'

Pebelse sveiet ovarri mo okremama? 'Näettekö metsän ja vuoret?'

Matula osilvega krai brezex. 'Haukka jahtasi lintua tänään.'

Kremama lis lämexo. 'Saavuin vuorilta.'

Pematulala plisi outribaba? 'Syövätkö haukat hyönteisiä?'

Ganana kaule lämexi. 'Siemenet saapuivat eilen.'

Osilvegaga sveechsen krai man. 'Hän menee tänään katsomaan lintuja.'

Pebelse pliet lõn oganana? 'Syöttekö huomenna siemeniä?'

### 2.2.1. Haukat näkivät peippoja

- Ofringaga utriba plixi.
- **Matulala ofringaga sveixi.**
- Gana omatula plixi.
- Pematulala ofringaga sveixi.

### 2.2.2. Jahtaatteko lintua tänään?

- **Pebelse brezet krai osilvega?**
- Belse pebrezet krai osilvega?

- Pebelse silveget man ofring?
- Pesilveget brezex kraï man?

### 2.2.3. Linnut saapuvat metsästä.

- **Silvegaga varri lis lämsi.**
- Man ovarri lôn kaulexi.
- Gana ovarri lôn lisi.
- Silvega lis kremane lämsi.

### 2.2.4. Näen peipon ja haukan lentävän vuorelle.

- **Ofringa mo omatula krema ko nachsen sveio.**
- Osilvega mo outribaba krema ko kaule sveio.
- Pebelo ogana mo omatula kaulechsen lämö.
- Oplixa mo okrema nachsen fringo.

### 2.2.5. Saavumme syömään hyönteisiä.

- **Outribaba plichsen läme.**
- Petribaba plichsen lämexe.
- Oganana lämechsen lise.
- Pematulala kaulechsen pline.