
Lasten yleistävän kielenkäytön kehityksen väitöksenjälkeinen tutkimus

A Data Management Plan created using DMPTuuli

Creator: Maija Surakka

Affiliation: University of Eastern Finland

Funder: Suomen Kulttuurirahasto

Template: Suomen Kulttuurirahasto 2018

ORCID ID: 0000-0002-0128-1073

Project abstract:

Tämä fennistiikan alaan kuuluva väitöksenjälkeinen tutkimushanke koostuu kolmesta tutkimuksesta, jotka tuottavat tietoa 2½–8-vuotiaiden lasten yleistämistarkoitteisista kielellisistä ilmauksista ja niiden kehityksestä. Hanke on suunniteltu kolmevuotiseksi. Hankkeen osatutkimusten tuloksissa painottuvat kielen syntaktisten rakenteiden ja käsitteistön kehitysnäkökulma sekä kielellisten ilmausten yhteys niillä tuotettuihin tilanteisiin merkityksiin. Koko tutkimushankkeen päätavoitteet ovat 1) osaltaan juurruttaa ja kumuloida nousevaa kielen ja ajattelun suhteen tutkimusalaa Suomessa sekä vakiinnuttaa lapsenkielen tutkimusta osaksi sitä sekä 2) tuottaa uutta perustutkimusta suomea äidinkielenään puhuvien lasten kielenomaksumisesta. Tällaista tietoa tarvitaan monien koulutus- ja kasvatusalojen toiminnan suunnittelun lähtökohtana. Tutkimusten aineistoja ovat arkisessa vapaa-ajan vuorovaikutuksessa kerättävät puhetallenteet. Aineistonkeruumenetelmiä ovat päiväkirja- ja äänitemenetelmät. Tutkimus keskittyy kieliopillisen persoonan, ajanilmausten ja muiden sanavalintojen avulla rakennettavaan geneerisyyteen, kontekstuaalisen merkitystulkinnan kysymyksiin sekä tiedon soveltamisen kehittymiseen tutkittavalla ikävälillä. Tutkimusten keskeisiä teemoja ovat geneerisyys, kielen ja käsitteellisen ajattelun kehityksen kumppanuus sekä vuorovaikutus. Tutkimuksen teoreettisen viitekehyksen muodostavat kognitiivinen kielitiede ja käyttöpohjainen kielenomaksumisen teoria. Tutkimustuloksia tarkastellaan rinnan psykologian alalla saavutetun tiedon kanssa. Psykologinen tutkimus lasten geneerisen ajattelun kehityksestä kytkee tutkimuksen kielitieteen rajat ylittävään, laajempaan tietokontekstiin lasten kielen ja käsitteellisen ajattelun kehityksen suhteista.

ID: 11647

Last modified: 28-04-2021

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Lasten yleistävän kielenkäytön kehityksen väitöksenjälkeinen tutkimus

1. Aineiston yleiskuvaus

1.1 Millaiseen aineistoon tutkimuksesi perustuu? Millaista aineistoa kerätään, tuotetaan tai käytetään uudelleen? Missä tiedostomuodoissa aineisto on? Arvioi myös karkealla tasolla, kuinka paljon levytilaa aineistosi lopulta tarvitsee?

Hankkeen osatutkimuksissa 1–3 käytetään kolmea aineistoa, joihin viitataan tunnuksilla A–C.

Tutkimusteni puhetalenteet on koottu; litterointityö jatkuu hankkeen edetessä. Kussakin osatutkimuksessa on oma ikärajauskensa. Rajaukset perustuvat aineistonkeruun kokemuksen ja aiempien tutkimusten tarjoamaan tietoon siitä, minkä ikävaihe on olennainen minkäkin kielellisen ilmiön kehityksessä.

Osatutkimuksen 1 (aineisto A) lähdeaineistona on tutkijan kahdentoista vuoden aikana tuottama puhetalennekokoelma, joka koostuu 269 äänitteestä. Äänitteiden yhteiskesto on 69 h. Äänitteet sisältävät 1–9-vuotiaiden lasten ja heidän läheistensä vuorovaikutuspuhetta. Litteroin äänitteistä osatutkimusta 1 varten selektiivisesti 2½–5-vuotiaiden perustelevia kielellisiä ilmauksia. Aineiston A koko on 393 esiintymää. Äänitelitteraatioiden lisäksi aineistossa A on mukana päiväkirjamerkintöjä.

Osatutkimuksen 2 lähdeaineistona on väitöskirja-aineistoni, joka koostuu 1 144 ajanilmauksesta (2½–8-vuotiaat). Väitöskirja-aineisto on koottu äänite-, päiväkirja- ja haastattelumenetelmillä. Eriytän tästä aineistosta yleistävät ilmaukset omaksi aineistokseen (**aineistoksi B**).

Osatutkimuksen 3 aineisto (aineisto C) koostuu kolmesta osasta: 1) aineiston B 6–8-vuotiaiden osuudesta, 2) kahden 10-vuotiaan lapsen haastattelulitteraatiosta sekä 3) puhetalennekokoelmasta selektiivisesti litteroitavista 6–9-vuotiaiden lasten äänitteistä. Haastattelulitteraatiot olen saanut käyttööni kollegaltani, joka on tuottanut aineistot omaa tutkimustaan varten. Vastaavanlaiset haastatteluaineistot ovat osa aineistoa B.

Aineistoni on äänitteistä litteroitua tai päiväkirjamerkinnöistä puhtaaksi kirjoitettua tekstiaineistoa, jonka levytilavaatimukset ovat pienet (< 5 MT). Hankkeessa käytettäviä puheäänitteitä sen sijaan on koottu 15.1.2021 mennessä lähes 300 (N = 278). Näiden tiedostojen levytilavaatimus on < 10 GT.

1.2 Miten aineiston yhtenäisyys ja laatu varmistetaan?

Aineistot A ja B on jo kerätty; aineiston B käsittely jatkuu yleistävien ajanilmausten eriyttämisellä. Aineistojen keräämisessä noudatettiin fennistiikan alalla vakiintuneita tapoja kerätä lapsenkielen aineistoja äänite- ja päiväkirjamenetelmällä. Tutkijalla on käytössään myös omat aineistonkeruukokemuksista nousevat laatuvaatimukset (Surakka 2019: 33). Aineiston B lähdeaineisto on tutkijan väitöskirja-aineisto, joten sen laatu on käynyt hyväksytyksi läpi tieteellisen arvioinnin; aineiston A pohjalta kirjoitettu tutkimus on vertaisarvioitu kansainvälisesti. Aineiston C keräämistä ohjaavat alan eettiset ohjeet (TENK 2019), fennistiikan alalla vakiintuneet keruumenetelmät sekä tutkijan yli kymmenen vuoden aineistonkeruukokemus.

2. Eettisten periaatteiden ja lainsäädännön noudattaminen

2.1 Mitä eettisiä seikkoja aineistosi hallintaan liittyy (esim. arkaluonteisten tietojen käsittely, tutkittavien identiteetin suojaaminen ja tietojen jakamista koskevan suostumuksen hankkiminen)?

Huolehdin tutkimukseni eettisyydestä voimassa olevan EU:n tietosuojasetuksen mukaisesti: jokaisen tutkimukseen osallistuvan lapsen huoltajalta on pyydetty kirjallinen suostumus puhennyttäjien keräämiseen, niiden käyttöön tieteellisessä tutkimuksessa ja tiedettä yleistajuistavissa kirjoituksissa sekä aineiston arkistointiin viiden vuoden määräajaksi.

Puhetallennetiedostoja säilytetään tutkimusprosessin ajan Amazonilta ostetulla palvelimella, joka on EU:n tietosuojasetusten mukaisesti suojattu, sekä Itä-Suomen yliopiston OneDrive-verkkolevyllä. Amazon-levytila on tärkeä siksi, että tutkija on ollut viime vuosina eri yliopistojen palveluksessa, jolloin aineistoihin pääsy riippuu kulloistenkin yliopistokohtaisten käyttöoikeuksien voimassa olosta. Aineistolitteraatioihin lasten tiedoista amerkitään ainoastaan ikä (muodossa vuosi;kuukausi). Litteraatioissa esiintyvät nimet pseudonymisoidaan. Aineistoon litteroidaan vain sellaisia puhejaksoja, jotka sisältävät siten yleisluontoisia aihetta, että ne voisivat kuulua kenen tahansa lapsen tai perheen

arkeen. Aineistoon ei litteroida sellaisia sisältöjä, jotka ovat arkaluontoisia tai jotka johtavat yksittäisten lasten tunnistamiseen.

Viimeistään tutkimusprosessin päätyttyä puhetallennetiedostot ja litteraatiot poistetaan yliopiston palvelimelta ja jätetään arkistoajaksi Amazonin palvelimelle. Suostumuslomakkeilla kerätyistä tiedoista (lapsen ja vanhemman nimitiedot) muodostuvasta henkilötietorekisteristä laaditaan seloste vastuuhenkilötietoineen (rekisterinpitäjä on Maija Surakka). Henkilörekisteriä säilytetään lukollisessa paloturvallisessa salkussa tutkijan kotona.

2.2 Miten aineiston omistajuuteen, tekijänoikeuksiin ja immateriaalioikeuksiin liittyviä asioita hallintaan? Estävätkö tekijänoikeudet, käyttöoikeudet tai muut rajoitukset aineiston käyttämisen tai jakamisen?

Aineisto on tutkijan keräämää, jolloin tutkija on yksin aineiston omistaja. Aineisto on kerätty alaikäisten lasten huoltajien suostumuksella. Huoltajat ovat saaneet tietoonsa oikeutensa vetää lastensa aineisto pois tutkimushankkeesta missä tahansa prosessin vaiheessa. Jo käsitelty tieto jää näissä tapauksissa hankkeen käyttöön (esimerkiksi jos tietoja pyydetään vedettäväksi pois vaiheessa, jossa tutkimusartikkeli on ehditty lähettää arvioitavaksi). Suostumuslomakkeella lasten huoltajille annetut tiedot aineiston käytöstä sitovat tutkijaa: tutkimusaineistoa saa käyttää vain tieteelliseen tutkimukseen ja sen perusteella laadittaviin yleistajuisiin julkaisuihin. Aineiston julkaisemisesta esimerkiksi Kielipankissa on alustavasti keskusteltu vanhempien kanssa, mutta aineiston julkaisemiseen liittyvä harkinta on vielä kesken.

3. Dokumentointi ja metatiedot

3.1 Miten dokumentoit aineistosi, jotta se on löydettävissä, saavutettavissa, yhteentoimivaa ja uudelleen käytettävissä sekä itseäsi että muita varten? Mitä metatietostandardeja, README-tiedostoja ja muuta dokumentaatiota käytät, jotta muut voivat ymmärtää ja käyttää aineistoasi?

Amazonin palvelimella ja yliopiston OneDrivessa sijaitsevat puhetallenteet on nimetty yhtenäisesti (esim. H_2;3(2)). Alkukirjan viittaa lapseen, jonka puhetta äänitteellä on. Nimen alkukirjainta seuraa lapsen ikä tallennushetkellä sekä tieto siitä, monesko tallenne saman lapsen samasta iästä on kyseessä.

Tallenteet on listattu ikävuosittain erilliseen Excel-tiedostoon, jossa seuraan myös äänitteiden kokonaiskestoja. Äänitteet litteroidaan Excel-taulukkoon, jossa on esiintymän lisäksi sarake lähdetiedostolle, kulloisenkin esiintymän kontekstitiedoille sekä aineiston luokittelussa käytettäville muuttujille, jotka selviävät tutkimustyön edetessä.

Aineistojen avaaminen muiden tutkijoiden käyttöön on vasta suunnitteilla, mutta yllä kuvattu tiedostokokonaisuus metatietoineen avautuu tarvittaessa myös muille tutkijoille.

4. Tallentaminen ja varmuuskopiointi tutkimushankkeen aikana

4.1 Minne aineistosi tallennetaan ja miten se varmuuskopioidaan?

Puhetallennetiedostot ovat tallennettuina Amazon Web Services -palvelimella, joka noudattaa EU:n tietosuojasetuksia. Varmuuskopiointi on osa Amazonin palvelua. Tallenteista on kopiot tutkimuksenaikaista käyttöä varten Itä-Suomen yliopiston OneDrive-palvelimella.

Äänitelitteraatiot tallennetaan tutkimusprosessin aikana Itä-Suomen yliopiston OneDrive-palvelimelle, josta tiedostot siirretään säilytykseen Amazonin palvelimelle.

4.2 Kuka valvoo pääsyä aineistoon ja miten suojattua pääsyä aineistoon valvotaan?

Aineiston ainoat tallennus- ja käyttösjainnit ovat Amazon Web Services ja Itä-Suomen yliopiston OneDrive, jotka ovat käyttäjätunnuksella ja salasanan suojattuja palvelimia. Tutkija saattaa tallentaa aineiston kulloinkin käsiteltäviä osia muistitikulle tai omalle tietokoneelleen, mutta tiedot siirretään näistä sijainneista viikoittain Itä-Suomen yliopiston OneDriveen. Tutkijan tietokone on suojattu salasanalla.

5. Aineiston avaaminen, julkaiseminen ja arkistointi tutkimushankkeen päätyttyä

5.1 Mikä osa aineistosta voidaan asettaa avoimesti saataville tai julkaista? Missä ja milloin aineisto tai siihen liittyvät metatiedot asetetaan saataville?

Aineiston avaamiseen liittyvät kysymykset ovat vasta suunnitteilla.

5.2 Mihin pitkällä aikavälillä arvokkaat tiedot arkistoidaan ja kuinka pitkäksi ajaksi?

Suostumuslomakkeella lasten huoltajien kanssa on sovittu, että aineisto arkistoidaan viideksi vuodeksi tutkimushankkeen päättymisen jälkeen. Tänä aikana tutkijalla on oikeus käyttää aineistoa tieteellisessä tutkimuksessa ja yleistajuisissa julkaisuissa.

6. Aineistohallinnan vastuut ja resurssit

6.1 Ketkä vastaavat aineistohallintaan liittyvistä tehtävistä tutkimusprojektin elinkaaren aikana? Arvioi myös aineistohallintaan tarvittavat resurssit (esim. taloudelliset, ajalliset, työmäärään liittyvät).

Kaikista aineistohallinnan tehtävistä vastaa tutkija Maija Surakka. Amazon Web Services -palvelusta on vähäisiä kustannuksia. Puhetallenteet on pääpiirteittäin kerätty, järjestetty ja kuvattu, joten niiden käsittelyyn ei kulu aikaa tutkimusprosessin edetessä. Litterointityö sen sijaan vie aikaa arviolta kaksi kuukautta.