



Turun puhekielen 2020-luvun tutkimushankkeen (Tupu2020) tietosuojaseloste

Olet ottamassa osaa Turun yliopistossa järjestettävään tieteelliseen tutkimukseen. Tässä tietosuojaselosteessa kerrotaan, miten henkilötietojasi käsitellään tutkimuksessa.

1. Rekisterinpitäjä

Turun yliopisto

Yhteyshenkilö tutkimusta koskevissa asioissa:

Nimi: Tommi Kurki / Suomen ja sen sukukielten arkisto

Puh.: 029 450 3326

E-mail: tommi.kurki@utu.fi

2. Tietosuojavastaavan yhteystiedot

Turun yliopiston tietosuojavastaava on tavoitettavissa sähköpostitse dpo@utu.fi ja puhelimitse +358294503009.

3. Kuvaus tutkimuksesta ja henkilötietojen käsittelystä

Turun puhekielen 2020-luvun tutkimushankkeessa tutkitaan sitä, millaista Turussa ja Turun lähiympäristössä puhuttu suomi on nykyään, miten se vaihtelee puhujan taustan ja tilanteen mukaan, miten se on muuttunut ja miten se on muuttumassa. Hankkeessa tutkitaan myös, millaisia kielikäsitteitä Turun seudulla asuvilla on.

Tutkimusta varten kerätään ja tallennetaan erilaisilta turkulaisilta ja Turun seutulaisilta puhuttua kieltä erilaisista kielenkäyttötilanteista. Hankkeeseen kerätään myös kyselyin tietoja siitä, millaisia käsityksiä heillä on muun muassa Turun puhekielestä, murteista, yleiskielestä ja ns. yleispuhekielestä. Jotta aineistoa saadaan erilaisilta tutkimushenkilöiltä ja jotta aineisto saadaan ryhmiteltyä systemaattisesti, kysytään tutkimukseen osallistujilta tiettyjä taustatietoja.

Aineiston käsittelyyn osallistuvat ensi vaiheessa vain tutkimusryhmän jäsenet. Ainoastaan tallenteiden litteroinnissa saatetaan käyttää ulkopuolisia toimijoita, jotka eivät pääse kuitenkaan käsiksi tutkimushenkilöiden henkilötietoihin. Tutkimushenkilöiden henkilötietoja ei luovuteta työryhmän ulkopuolelle.

Hankkeen päätyttyä aineisto annetaan pitkäaikaissäilytettäväksi Turun yliopiston Suomen ja sen sukukielten arkistoon sekä myöhemmin myös kansallisten tutkimusaineistojen pitkäaikaistallennukseen perustettuun Kielipankkiin. Sieltä aineistoa saavat käyttöönsä vain sellaiset tutkijat, jotka sitoutuvat sen eettiseen käyttöön ja jotka pystyvät esittämään uskottavan tutkimussuunnitelman, jossa kuvataan, miten ja mihin aineistoa käytetään. Ne, jotka saavat luvan käyttää aineistoa, saavat tietoonsa vain, minkälaisia puhujia (esimerkiksi millä vuosikymmenellä syntyneitä puhujia) kussakin osa-aineistossa on. Sellaisia henkilötietoja ei luovuteta, joista ilmenisi, keitä tallenteilla puhuvat henkilöt ovat.



4. Henkilöt, jotka osallistuvat henkilötietojen käsittelyyn

Henkilötietoja käsittelee hankkeen johtaja, Suomen ja sen sukukielten arkiston vastaava erikoistutkija tai erityisissä tapauksissa heidän nimeämäänsä hankkeessa tai arkistossa työskentelevä henkilö. Tutkimushenkilöitä koskevien kuvailevien tietojen (synnyinvuosikymmen, sukupuoli jne.) lukemiseksi pitää pystyä avaamaan Turun yliopiston tarjoamista ja varmentamista tallennuspaikoista kaksi erillistä tiedostoa sekä niiden yhdistämisen mahdollistava koodiavain. Tallennuspaikkoihin on pääsy yliopiston UTU-käyttäjätunnuksin vain em. johtajalla, erikoistutkijalla ja arkiston johtoryhmän jäsenillä, mutta yhdistävään koodiavaimeen vain johtajalla ja erikoistutkijalla.

Muut hankkeen aineistoja käsittelevät henkilöt pääsevät luvanvaraisesti käsiksi vain pseudonimisoituihin aineistoihin.

5. Henkilötietojen käsittelyn kesto

Aineisto säilytetään pysyvästi Turun yliopiston Suomen ja sen sukukielten arkistossa sekä tutkimushankkeen päätyttyä myös Kielipankissa. Kerätty aineisto edustaa 2000-luvun alun Turun ja Turun seudun puhekieltä erilaisine variaatioineen, ja se on ainoa tällaista puheaineistoa edustava tutkimusaineisto. Tärkeä osa murteentutkimusta ja sosiolingvististä tutkimusta on kielen muuttumisen tutkiminen, minkä vuoksi on tärkeää, että tätä aineistoa voidaan säilyttää. Tavoitteena on kerätä vastaavanlainen aineisto vuosikymmenten päästä: tällöin pystyttäisiin toteuttamaan pitkittäistutkimus, jossa olisi puhetta samoilta yksilöiltä eri vuosikymmeniltä. Tällaiset laajat tutkimusaineistot ovat kansainvälisestikin vielä verrattain harvinaisia. Jotta samoja tutkimushenkilöitä pystyttäisiin tavoittamaan seuraavalla tallennuskierroksella, on välttämätöntä säilyttää joitain perustietoja tutkitavista.

6. Henkilötietojen lainmukainen käsittelyperuste

Henkilötietojen käsittely perustuu tietosuoja-asetuksen 6(1) artiklan alakohtaan:

- ☒ yleistä etua koskeva tehtävä:
 - ☒ tieteellinen tai historiallinen tutkimus tai tilastolliset tarkoitukset
 - ☒ tieteellisten aineistojen tai kulttuuriperinnöllisten materiaalien arkistointi
- ☐ rekisteröidyn suostumus
- ☐ rekisterinpitäjän lakisääteisen velvoitteen noudattaminen
- ☐ rekisterinpitäjän tai kolmannen osapuolen oikeutettu etu

7. Tutkimusaineistoon sisältyvät henkilötiedot

Tutkimushenkilöiltä kysytään taustatiedoiksi heidän nimeään, synnyinvuottaan, sukupuoltaan, paikkakuntia, joilla he ovat eläneet pisimpään lapsena ja aikuisena sekä heidän nykyistä asuinpaikkakuntaansa. Lisäksi heiltä kysytään heidän ammattiaan yleisluonteisesti, koulutusastettaan sekä heidän äidinkieltään, mikäli heidän äidinkieltensä on muu kuin suomi. Näitä tietoja tarvitaan tutkimusaineiston järjestämisessä, jotta tutkittavat voidaan ryhmitellä kielentutkimusta varten eri puhujaryhmiin (ennen muuta ikäryhmiin ja murteellisiin ryhmiin). Ryhmittelyvaiheen jälkeen tarkka synnyinvuosi ja nimi näkyy vain kirjallisissa aineistonkäyttöluvista, ja luvat ovat säilytteillä vain tutkimushankkeen johtajan lukollisessa työhuoneessa sekä yliopiston verkkolevyllä tiedostossa, johon on pääsy vain johtajalla. Puhujaryhmien kuvailutietoihin jää näkyviin puhujaryhmien puhujien yhteinen synnyinvuosikymmen, sukupuoli, paikkakunta- tai seututiedot (jos paikkakunnalla on alle 10 000 asukasta, kerrotaan seutu, muuten paikkakunta), ammattitieto ylimalkaisesti, koulutusaste



sekä äidinkieli. Lisäksi aineistoon saattaa tallentua muitakin henkilötietoja, mikäli henkilö kertoo niistä haastattelussa.

8. Erityiset henkilötietoryhmät (arkaluontoiset henkilötiedot)

Tutkimuksessa ei käsitellä erityisiä henkilötietoryhmiä. Jos tutkimukseen osallistuvat kertovat oma-aloitteisesti sellaisia erityisiä henkilötietoja, jotka koskevat heidän terveyttään tai poliittisia mielipiteitään, poistetaan tällaiset kohdat aineistosta. Tällaiset kohdat voidaan kuitenkin säilyttää, jos tutkittava ilmaisee erikseen haluavansa ne säilytettäväksi osana aineistoa.

9. Henkilötietojen keräämisen lähteet

Tutkimushenkilöt ovat itse antaneet heitä koskevat tiedot joko suullisesti tai kirjallisesti ennen tallennustilannetta tai heti sen jälkeen tai kyselyn yhteydessä kirjallisesti.

10. Henkilötietojen suojatoimenpiteet

Tietojärjestelmissä käsiteltäviä tietoja suojataan seuraavilla tavoilla:

- ☒ käyttäjätunnus ja salasana
- ☒ käytön rekisteröinti/lokitus
- ☒ kulunvalvonta
- ☐ salaust/kryptaus
- ☐ kaksivaiheinen tunnistautuminen
- ☐ muu, mikä:

Manuaalista (esim. paperimuodossa tai muuten aineellisessa muodossa) olevaa aineistoa suojataan seuraavilla tavoilla

Suorien tunnistetietojen käsittely:

- ☐ Aineisto kerätään ilman suoria tunnistetietoja
- ☒ Suorat tunnistetiedot poistetaan analysointivaiheessa:

Aineiston ryhmittelyvaiheen jälkeen tarkka tutkimushenkilöiden synnyinvuosi ja nimi näkyy vain kirjallisissa aineistonkäyttöluvista, ja luvat ovat säilytteillä vain tutkimushankkeen johtajan lukollisessa työhuoneessa sekä yliopiston verkkolevyllä tiedostossa, johon pääsy vain johtajalla. Puhujaryhmien kuvailutietoihin jää näkyviin puhujaryhmien puhujien yhteinen synnyinvuosikymmen, sukupuoli, paikkakunta- tai seututiedot (jos paikkakunnalla on alle 10 000 asukasta, kerrotaan seutu, muuten paikkakunta), ammattitieto ylimalkaisesti, koulutusaste sekä äidinkieli. Lisäksi aineistoon saattaa tallentua muitakin henkilötietoja, mikäli henkilö kertoo niistä haastattelussa. Tallenteista tehtäviin litteraatteihin muutetaan henkilönimet.

- ☐ Aineisto analysoidaan suoraan tunnistetiedoin, koska:

11. Henkilötietojen siirtäminen ja jakaminen kolmansille osapuolille

Hankkeen aineisto arkistoidaan Turun yliopiston Suomen ja sen sukukielten arkistoon sekä hankkeen päätyttyä myös Kielipankkiin, josta tutkijat voivat Turun yliopiston tutkimusryhmän luvalla saada sen käyttöönsä tutkimustarkoituksiin.



Kielipankkiin luovutetaan hankkeen tallenteet, litteraatit sekä näiden pohjalta mahdollisesti tehtävät puheen rakenneominaisuuksia koskevat kieliopilliset annotaatiot. Puhujaryhmien koostumusta luonnehditaan aineiston luvanvaraisille käyttäjille vain seuraavilla tekijöillä: sukupuoli, synnyinvuosisikymmen, keruupaikkakunta, yleisluonteinen asuinhistoria (syntyperäinen paikkakuntalainen, paikkakunnalle muuttanut, paikkakunnalta muuttanut; jos muuttanut miltä/mille murrealueelle), ammatti yleisluonteisesti, koulutus sekä äidinkieli. Litteraateissa henkilönnimet muutetaan. Ne, joille myönnetään lupa käyttää aineistoa, sitoutuvat tutkimuksen yleisiin eettisiin periaatteisiin ja aineistokohtaisiin käyttörajoituksiin.

12. Henkilötietojen siirtäminen EU:n tai ETA:n ulkopuolelle

Henkilötietoja ei siirretä Euroopan unionin tai Euroopan talousalueen ulkopuolelle.

13. Henkilötietojen käsittely tutkimuksen päättymisen jälkeen

☐ Tutkimusaineisto poistetaan

☐ Tutkimusaineisto säilytetään tämän tutkimuksen tulosten luotettavuuden arvioimista varten:

☐ ilman suoria tunnistetietoja ☐ tunnistetiedoin

☒ Tutkimusaineisto säilytetään myöhempää, yhteensopivaa tieteellistä tutkimusta varten tietosuoja-asetuksen vaatimusten mukaisesti:

☒ ilman suoria tunnistetietoja ☐ tunnistetiedoin

Tutkimusaineisto säilytetään pysyvästi Turun yliopiston Suomen ja sen sukukielten arkistossa, ja valikoitua tutkimuksellisesti edustavaa osaa aineistosta säilytetään myös Kielipankissa.

Tutkimusaineiston säilytys perustuu tietosuoja-asetuksen 5 artiklan 1 kohdan b ja e alakohtiin. Tutkimusaineiston uudesta tutkimuskäytöstä lähetetään tutkittaville uusi tietosuojaseloste, paitsi jos rekisterinpitäjä ei enää pysty tunnistamaan tutkittavia tutkimusaineistosta.

14. Rekisteröidyn oikeudet sekä niihin tehtävät poikkeamat

Lisätietoa tietosuoja-asetuksen mukaisista oikeuksistasi antaa kohdassa 1 mainittu yhteyshenkilö.

Rekisteröidyn oikeudet

Tietosuoja-asetuksen mukaan rekisteröidyllä on oikeus:

- tarkistaa itseään koskevat tiedot (15 artikla)
- oikaista omia tietoja (16 artikla)
- poistaa tiedot ja tulla unohdetuksi (17 artikla)
- rajoittaa omien tietojen käsittelyä (18 artikla)
- siirtää tiedot rekisterinpitäjältä toiselle (20 artikla)
- vastustaa tietojen käsittelyä (21 artikla)
- olla joutumatta automaattisen päätöksenteon kohteeksi (22 artikla)

Rekisteröity ei kuitenkaan voi käyttää kaikkia oikeuksia kaikissa tilanteissa riippuen esimerkiksi henkilötietojen käsittelyperusteesta.



Rekisteröidyn oikeuksiin tehtävät poikkeamat

Tietosuojalainsäädäntö mahdollistaa rekisteröidyn oikeuksista poikkeamisen silloin, kun henkilötietoja käsitellään tieteellisessä tutkimuksessa ja oikeuksien toteuttaminen estäisi tai vaikeuttaisi suu-
resti käsittelyn tarkoitusten saavuttamista. Tarve tehdä poikkeamia rekisteröityjen oikeuksiin arvioi-
daan aina tapauskohtaisesti.

Tämän tietosuojaselosteen tutkimuksessa on todennäköisesti tarpeen tehdä poikkeamia seuraaviin rekisteröityjen oikeuksiin:

- ☐ Oikeus tarkistaa itseään koskevat tiedot (15 artikla)
- ☐ Oikeus oikaista omia tietoja (16 artikla)
- ☒ Oikeus poistaa tiedot ja tulla unohdetuksi (17 artikla)
- ☐ Oikeus rajoittaa omien tietojen käsittelyä (18 artikla)
- ☐ Oikeus vastustaa tietojen käsittelyä (21 artikla)

Jos henkilötietojen käsittely tutkimuksessa ei edellytä rekisteröidyn tunnistamista eikä rekisterinpi-
tāja pysty tunnistamaan rekisteröityä, niin oikeutta tietojen tarkastamiseen, oikaisuun, poistoon,
käsittelyn rajoittamiseen, ilmoitusvelvollisuuteen ja siirtämiseen ei sovelleta, ellei rekisteröity pysty
antamaan luotettavan tunnistamisen mahdollistavia lisätietoja (11 artikla).

Perusteet tehtäville poikkeamille sekä tehtävien poikkeamien ulottuvuus

Jos tutkimushenkilön puhekielen tallenteita tai kyselyvastauksia, tallenteista tehtyjä transkriptioita
ja annotaatioita on päätetty sisällyttää lopulliseen Tupu2020-tutkimusaineistoon (korpukseen), ei
näitä poisteta, koska tällöin ainutlaatuisen jäsennetyn tutkimuskokonaisuuden kielitieteellinen käyt-
tö ei olisi enää mahdollista: korpus on kokonaisuus, jossa sen jokainen osa on valittu edustamaan
tietyn puhujaryhmän kielenkäyttöä 2020-luvun Turun seudulla. Osa-aineistojen poisto tästä kor-
puksesta estäisi tai vähintäänkin vaikeuttaisi suuresti tämän ainutlaatuisen puhekorpuksen kielitie-
teellistä käyttöä.

Valitusoikeus

Sinulla on oikeus tehdä valitus tietosuojavaikuttetulle, jos koet, että henkilötietojasi on käsitelty
soveltuvan tietosuojalainsäädännön vastaisesti.

Tietosuojavaikuttetun yhteystiedot:

Tietosuojavaikuttetun toimisto
Käyntiosoite: Lintulahdenkuja 4, 00530 Helsinki
Postiosoite: PL 800, 00531 Helsinki
Puhelinvaihe: 029 566 6700
Sähköposti (kirjaamo): tietosuoja(at)om.fi