

Foneettinen tutkimusmatka häätäpuheluiden prosodiikkaan

Lauri Tavi: *Prosodic cues of speech under stress. Phonetic exploration of Finnish emergency calls*. Joensuu: Itä-Suomen yliopisto 2020. Johdanto 59 s. + artikkelit 40 s. ISBN 978-952-61-3402-4. Saatavilla verkossa osoitteessa <http://urn.fi/URN:ISBN:978-952-61-3403-1>.

Lauri Tavin väitöstutkimuksessa tarkastellaan hätäkeskukseen tulleita puheluita ja niissä kuultavaa stressinalaista puhetta eli emotionaalisesti kuormittuneiden ihmisten puhetta. Tavin tutkimus on ensimmäinen suomen kielen tunneilmaisuksen prosodiikkaa – puheen intonaatio-, voimakkuus- ja rytmikaavoja – koskeva työ. Se kuuluu kansainvälisestikin harvoin tutkittuihin, joissa on vertailtu aidosti stressinalaisen ja neutraalin puheen prosodisia eroja. Siten tutkimuksen tulokset ovat informaatioarvoltaan erityisen merkittäviä.

Tavin tutkimus on myös monella tavalla ajankohtainen. Ensinnäkin ihmisten hätä on tiedotusvälineiden mukaan kasvanut koronavirusepidemian aiheuttamien rajoitusten takia. Samaan aikaan on kuitenkin uutisoitu, että hätäkeskuksiin tulevasta soitoista jopa kolmasosa on luokiteltu epäasiallisiksi. Osa näistä asiattomista yhteydenotoista on mitä ilmeisimmin kiusantekoa. Tällaiset soitot sitovat viranomaisten resursseja, ja apu saattaa jäädä saamatta niiltä, jotka sitä oikeasti tarvitsisivat.

Tutkimus on ajankohtainen myös siksi, että se kulkee puheteknologian ja tekoälytutkimuksen menetelmien kehityksen aallonharjalla. Olisi helppoa ajatella, että jos ihminen on niin stressaantunut, että hän soittaa hätäkeskukseen, se kuuluisi äänestä selvästi ja erottuisi pilasoitosta. Monet tekijät, kuten taustahäly, puhujan outo aksentti ja puhelinlaitteen taajuussuodatus, kuitenkin haittaavat

aidon hädän tunnistamista. Aiempi tutkimus ei ole tuottanut yhtä ainoaa mittaria, joka paljastaisi stressinalaisen puheen. Tavin työssä on hyödynnetty useita aiemminkin käytettyjä mittareita mutta lisäksi kehitetty aivan uusia keinoja tunnistaa stressinalaisen puheen prosodisia piirteitä. Tekijä on tuonut perinteisten akustis-foneettisten menetelmien rinnalle perustietoa soveltavat, koneoppimiseen perustuvat työkalut.

Koneoppimisen kannalta on myös tärkeä selvittää, miten prosodiset piirteet vaihtelevat autenttisessa puheessa erilaisen kontekstuaalisten tekijöiden vuoksi. Tässä tutkimuksessa merkittäväksi tekijäksi havaittiin erityisesti hätäkeskukseen soittaneiden henkilöiden ja puheluihin vastanneiden virkailijoiden vuorovaikutus. Avun pyytäminen näyttää vaikuttavan virkailijan puheen prosodiikkaan, ja avun tarjoaminen muuttaa soittajan puhetta. Näin läheltä ei minkään auttamisinterventtion vaikuttavuutta liene tutkittu.

Tavin tutkimuksen tavoitteet kytkeytyvät selkeästi näihin ajankohtaisiin teemoihin. Artikkelimuotoisen väitöskirjan neljän osatutkimuksen avulla on pyritty 1) luomaan malleja, joilla olisi mahdollista luokitella stressinalaisen puheen laatua foneettisin perustein, 2) kehittämään keinoja erottaa oikeat hätäsoitot pilasoitoista ja muista asiaan kuulumattomista puheluista sekä 3) edesauttamaan ääneen perustuvien psykologisten arviointimallien kehittämistä.

Monipuoliset menetelmät

Tavin tutkimuksen erityisansio on sen metodologinen hienostuneisuus. Väitöskirjan yhteenveto-osan eri luvuissa on esimerkillisellä tavalla nostettu esiin tut-

kimuksen keskeiset käsitteet ja kuvattu niiden operationaaliset määritelmät matemaattisine kaavoineen, koneoppimisen periaatteiden esittelyä unohtamatta. Tarkastelu on kriittistä, eli Tavi tuo esiin eri menetelmien heikkouksia ja vahvuuksia ja perustelee erinomaisesti omat valintansa. Tutkimus olisikin sen ansiosta hyvin toistettavissa, ellei alkuperäisaineiston jatkokäyttöä olisi jouduttu epäämään sen arkaluontoisuuden vuoksi. Yhteenvedo- osan metodiluvut toimisivat hyvin puheentutkimuksen oppimateriaalina; luki- jalta toki vaaditaan vähintään lyhyttä fy- siikan oppimäärää kaikkien yksityiskoh- tien ymmärtämiseksi.

Tutkimuksen aineisto on peräisin Tavin itsensä kokoamista autenttista hätäkeskuspuheluista (*Emergency Call Corpus*) ja toisaalta kahdesta muusta puhekorpuksista (*Aalto University DSP Course Conversation Corpus* ja *C-PROM*). Osatutkimuksissa eri korpuksista on muodostettu stressinalaisten ja neutraa- lien puhujien ryhmät, joiden puheen prosodisia piirteitä tutkimuksessa ver- taillaan. Näiden vertailuryhmien avulla saadaan vastaukset tutkimuskysymyksiin luotettavasti, vaikka puhujien määrä on rajallinen ($N = 85$). Hätäpuheluaineiston käyttöoikeuden saaminen on ollut vaativa tehtävä, ja tutkimuksen eettiset näkökoh- dat onkin huolellisesti otettu huomioon ja kirjattu omaan alalukuunsa.

Aineisto ei kuitenkaan ole sama kuin tutkittavien henkilöiden tai näyttöiden antajien määrä. Itse analysoitavaa dataa Taville karttuikin kaiken kaikkiaan pal- jon: satoja lausumia, tuhansia vokaaleita ja kymmeniätuhansia tavuja. Aineiston muokkaus on mitä ilmeisimmin vaati- nut valtavasti käsityötä: kuuntelemista eli oman korvan ja kielitajun käyttöä, ma- nuaalista mittaussegmenttien merkintää, puhesignaalin vahvistamista ja niin edel- leen, ennen kuin on päästy tuottamaan täsmällisiä mittaus tuloksia tilastollista analyysia varten.

Tällainen kuulunee eksploratiivisen tutkimuksen luonteeseen; ihmiskorvaa ja ihmisen aivoja tarvitaan uuden etsin- nässä, ennen kuin automaattinen puheen prosodisten piirteiden tunnistus ja luokittelu ovat kehittyneet luotettaviksi. Tie- dämme kuitenkin, että kuulohavaintoon- kaan perustuvat arvioinnit ja luokitukset eivät ole täysin luotettavia, ja juuri siksi kuulonvaraisten luokitusten heikkouksia poistamaan tarvitaan akustis-foneettisia mittauksia. Kahdessa osatutkimuksessa puhetta tutkitaan vokaalien perusteella (Tavi 2017, 2019), ja kahdessa repertuaaria laajennetaan tavuihin ja lausumiin (Tavi, Alumäe & Werner 2019; Tavi & Werner 2020). Vaikka osatutkimusten lähtö- aineisto vaihtelee, mitattavat prosodi- set muuttujat ovat pääosin samoja, mikä tietysti on järkevää tiedon kumuloitumi- sen ja varsinkin mittareiden kehittelyn kannalta.

Vaikuttavia tuloksia

Tutkimuksen tavoitteet toteutuvat katta- vasti, sillä uutta tietoa suomenkielisissä hätäpuheluissa kuultavan stressinalaisen puheen prosodiasta saatiin kaikissa osa- tutkimuksissa. Akustis-foneettisen ana- lyysin ja useiden laskennallisten mal- lien haasteet käyvät selviksi, ja kone- oppimisen menetelmien mahdollisuudet ja rajoitukset tulevat hyvin esille eri osa- tutkimuksissa ja yhteenvedon erinomai- sissa selvityksissä.

Yhteenvedo-osan luvussa 8 Tavi tiivis- tää tutkimuksensa tulokset. Päätulos on, että emotionaalisesti kuormittunut pu- huja on mahdollista tunnistaa äänen laa- dun (mm. kireys ja hengästyneisyys) ja muiden prosodisten piirteiden perus- teella. Kaiken kaikkiaan Lauri Tavin tut- kimus sekä vahvistaa aiempien stressin- alaista puhetta koskevien tutkimusten tu- loksia (mm. korkea äänen perustaajuus) että tuottaa uutta tietoa prosodiikasta. Esimerkiksi *i*-vokaalin formanttien sent-

ralisoituminen selittyy stressin vaikutuksilla puhemotoriikkaan; stressi kangastaa myös kielen. Perinteistä koneoppimista hyödyntäen havaittiin, että stressinalainen ja neutraali ääni erottuvat toisistaan parhaiten, kun opetettava aineisto on homogeenista. Työssä luotiin myös hermoverkkomallinnuksen avulla uusi narinaäänien tunnistamisohjelma, joka kertoo puheäänien rentouden asteesta ja siten mahdollisesti myös puhujan stressin (vähäisestä) määrästä.

Viimeisimpänä vaan ei vähäisimpänä tuloksena neljännessä osatutkimuksessa havaittiin, että itsemurhaa pohtivat tai yrittävät soittajat tuntuivat rauhoittuvan saatuaan neuvoja tai muuta apua virkailijalta (Tavi & Werner 2020). Se on välitöntä vaikuttavuutta, jota on interventiotutkimuksissa usein vaikea tavoittaa. Tässä osatutkimuksessa kehitettiin uusi työkalu, prosodisen prominenssin mittari (*Syllabic Prosody Index*), jonka skriptissä yhdistyvät kolmen prosodisen piirteen mittaustulokset: äänen korkeuden, tavun keston ja energian jakautumisen. Erityisen selvästi avun tarjoamisen vaikutus ilmeni esimerkiksi soittajan puheen monotonisuuden vähenemisenä ja puherytmin hidastumisena. Soittajan ja työntekijän puheen prosodian ”tuunautuminen” näkyi äänen perustaajuuden korrelaationa.

Laadukas lukukokemus

Tavin tutkimus virittää lukijan pohtimaan monia kiinnostavia seikkoja, esimerkiksi emootioiden määritelmiä. Tavi pohtii niitä itsekin yhteenveto-osan luvussa 3 ryhtymättä kuitenkaan psykologiksi psykologin paikalle. Stressi on määritelty esimerkiksi emotionaaliseksi kuormituneisuudeksi, tilaksi, jossa ihminen kokee itsensä kohdistuvaa uhkaa. Samaisen luvun kuvauksessa ”tunneavaruudesta” (s. 26) stressi asettuu korkealle aktiivisuusasteikolla ja valenssiltaan melko ne-

gatiiviseksi. Neljännessä osatutkimuksessa (Tavi & Werner 2020) käsitellään toisenlaista emotionaalista kuormitusta, nimittäin suisdaalisuutta ja siihen usein liittyvää masentuneisuutta. Kun ihminen on aikeissa vahingoittaa itseään – kun hän uhkaa omaa henkeään – ja soittaa sen takia hätäkeskukseen, hänen luulisi olevan kovasti emotionaalisesti kuormittunut eli stressaantunut. Masennus asetuu yhteenveto-osan sivun 26 kuviossa kuitenkin aktiivisuusasteikon alapäähän, toki negatiiviseksi valenssiltaan, mutta aika kauas stressistä. Tämän tutkimuksen perusteella näyttääkin olevan niin, että korkea-aktiivisen stressinalaisen ja matala-aktiivisen depressiivisen äänen piirteet eroavat merkittävästi toisistaan. Jatkotutkimuksen selvitettäväksi jää, miten nämä vastakohtat täydentävät toisiaan puhujien psykologisen profiiloinnin kannalta.

Tavin väitöstyön pääpaino on stressinalaisten puhujien äänen laadussa ja muissa prosodisissa piirteissä. Tiedossani ei ole, että hätäkeskuspuheluiden vuorovaikutusta olisi tutkittu aiemmin. Voisin kuitenkin kuvitella, että tämän tutkimuksen tuloksilla olisi käyttöä esimerkiksi hätäkeskusten työntekijöiden kouluttamisessa ja heidän työkykynsä tukemisessa. Automaattinen stressinalaisen puheen tunnistaminenhan ei vielä poista stressaantuneen soittajan hätää.

Tavin käyttämistä perinteisistä akustisfoneettisista menetelmistä moni on tuttu myös patologisen puheen tutkijoille; logopedian piirissä on jo kauan mitattu esimerkiksi puhenopeuksia ja tutkittu vokaalien formanttikarttoja. Akustisfoneettisten menetelmien käyttö on olennainen osa myös äänentutkimusta. Olisi mitä suotavinta – sekä fonetiikan että logopedian kannalta – lyödä viisaat päät yhteen ja jakaa metodologinen osaaminen eri alojen tutkijoitten kesken. Tavin tutkimus osoittaa myös sen, että insinööritieteiden ja humanististen tie-

teiden tutkimusten tavoitteet eivät välttämättä ole kovinkaan kaukana toisistaan.

Tavin väitöskirjan artikkelit ja yhteenveto-osa muodostavat yhtenäisen ja johdonmukaisen kokonaisuuden akustis-foneettisesta tunneilmaisujen prosodiikan tutkimuksesta. Yhteenveto-osa perustelee kaikkien neljän osatutkimuksen aiheet ja menetelmien valinnan sekä kokoaa näiden artikkelien havainnot yhteen paikkaan. Siinä ei ole erillistä kirjallisuuskatsausta, vaan teksti keskusteluttaa eri aluvuissa aiempaa tutkimuskirjallisuutta osatutkimuksissa tutkittujen ilmiöiden, menetelmällisten ratkaisujen ja saatujen tulosten kanssa perinpohjaisella mutta helposti ymmärrettävällä tavalla. Tämä on hyvin raikas ratkaisu.

ANNA-MAIJA KORPIJAAKKO-HUUHKA
etunimi.sukunimi@tuni.fi

Kirjoittaja on Tampereen yliopiston
logopedian emeritaprofessori.
Hän toimi Tavin vastaväittäjänä.

Kirjallisuus

- TAVI, LAURI 2017: Acoustic correlates of female speech under stress based on /i/-vowel measurements. – *International Journal of Speech, Language and the Law* 24 s. 227–241. <https://doi.org/10.1558/ijssl.32506>.
- 2019: Classifying females' stressed and neutral voices using acoustic-phonetic analysis of vowels. An exploratory investigation with emergency calls. – *International Journal of Speech Technology* 22 s. 511–520. <https://doi.org/10.1007/s10772-018-09574-6>.
- TAVI, LAURI – ALUMÄE, TANEL – WERNER, STEFAN 2019: Recognition of creaky voice from emergency calls. – *Proceedings INTERSPEECH 2019* s. 1990–1994.
- TAVI, LAURI – WERNER, STEFAN 2020: A phonetic case study on prosodic variation of suicidal emergency calls. – *International Journal of Speech, Language and the Law* 27 s. 59–74. <https://doi.org/10.1558/ijssl.39667>.

Väitöskirja viron kielen partitiivin ja illatiivin variaatiosta

Ann Siiman: *Vormikasutuse varieerumine ning selle põhjused osastava ja sisseütleva käände näitel.* Dissertationes Philologiae Estonicae Universitatis Tartuensis 45. Tartu: Tartu Ülikooli Kirjastus 2019. ISBN 978-9949-03-124-5. Saatavilla verkossa osoitteessa <http://hdl.handle.net/10062/64787>.

Viron kielen illatiivin ja partitiivin päätevariaatio on hyvin monipuolista. Millaisia eri varianttien suhteet ovat nykyvirossa, ja missä määrin ne vastaavat sitä, mitä päätevaihtelusta on esitetty kieliopeissa? Mitkä

tekijät vaikuttavat varianttien esiintymiseen, ja mitkä niistä ovat tärkeimpiä? Tällaisia kysymyksiä tutkii Ann Siiman tuoreessa Tarton yliopistossa hyväksytyssä väitöskirjassaan. Väitöskirjan aihe on kiinnostava myös suomen kielen kannalta, sillä niin viron kuin suomenkin kielessä partitiivi ja illatiivi ovat päätevarianttien runsauden takia antoisia sijamuotoja.¹

1. Toimin Ann Siimanin vastaväittäjänä Tartossa ja olen julkaissut väitöskirjasta esittelyn myös *Keel ja Kirjandus* -lehdessä (Laalo 2020).