



TAMPEREEN YLIOPISTO
Puheopin laitos

VOKOLOGIAA, PUHEVIESTINTÄÄ JA
MUUTA PUHEENTUTKIMUSTA

Juhlakirja Timo Leinolle

toim. Anne-Maria Laukkanen
ja Tuula-Riitta Välikoski

KUULONMALLI: UUDEN VUOSITUHANNEN PUHEENTUTKIMUSTA

Anna-Liisa Lehessaari
Helsingin yliopisto
Helsinki

Raimo Toivonen
Pitchsystems Oy.
Tampere

"Mutta en voi ymmärtää muuta, kuin että kielitiede tulee jatkuvasti löytämään problematiikkansa reaalisten tosiasioiden maailmasta. /---/ Uudet tulokset tuovat aina mukanaan uusia kysymyksiä, eikä tutkija koskaan tiedä, mihin tie hänet lopulta vie. Hänen täytyy olla ilman minkäänlaista apriorismia valmistautunut mihin tahansa, ja jos hän näin on asennoitunut, niin hän hyvin vastenmielisesti puhuu tieteen lopullisista päämääristä ja tarkoituseristä." (Ravila 1941/ 1967:25)

1900-luku - foneettisen frustraation vuosisata

Kiinnostus foneettisia tutkimusmenetelmiä kohtaan oli 1900-luvun kielentutkijoiden keskuudessa huomattavaa, ja esimerkiksi oman maamme fennisteillä tavallista. Ymmärrettävistäkin - olihan suomen murteiden ja sukukielten tarjoamista aineistoista ammentava kielentutkimus vahvasti nimenomaan kielellisen vaihtelun ja muutoksen tutkimusta. Kielen sosiaalista vaihtelua ja kielikontakteja tarkasteltaessa ovat foneettiset muutokset aina keskeisiä. Siitä, että tutkijat olivat hyvin perehtyneitä käyttämiinsä menetelmiin, kertoo myös heidän tyytymättömyytensä niitä kohtaan:

"Viime vuosikymmenien kokeellisen fonetiikan ei ole onnistunut luoda mainittavaa valaistusta painotuksen ongelmaan ja sen kielihistorialliseen asemaan. Tämä, ensi hetkessä ehkä oudostuttava tilanne (onhan fysiikalla näet kutakuinkin selvä käsitys äänen voimasta) johtuu suurimmaksi osaksi kielen intensiivisten suureiden - niistähän painotuksessa on kysymys - tavattomasta kompleksisuudesta, joka estää täsmällistä tietoa tuovien fysikaalisten mittausten käyttämisen /---/. Eräitä menetelmiä on silti tosin kyetty kehittämään, mutta toistaiseksi ne eivät ole osoittautuneet riittävässä määrässä kielitieteellisen tutkimuksen tarkoituksia vastaaviksi." (Penttilä 1926: 3).

Valitettavasti Penttilän ynnä monen muun haaveilemia menetelmiä on saatu odottaa niin kauan, etteivät he itse ehtineet niitä enää kokeilemaan. Toki uusia välineitä saatiin, ja aikanaan niistä jokaisen on täytynyt tuntua mullistavalta parannukselta tutkimuksen

edellytyksissä. Menetelmillä tehtiin se mitä kyettiin (esim. Penttilä 1958, 1959). Nykytutkijan on vaikea kuvitella sitä sinnikkyyden määrää, jota tarvittiin bensiiniä polttaen nokemalla ja lakkaamalla valmistettuihin analyysiliukoihin raapeutuneiden käyrien tarkastelussa, savuttavien spektrogrammilyllyjen pyörittelyssä tai sottaavien mustepiirtureiden säätämässä erikseen kunkin puhujan äänialan mukaan.

Tietokoneajan sarastus – ja päivännousu

Vasta 1900-luvun lopun tutkijasukupolvet ovat päässeet kokemaan sen mahdollisuuksien mullistuksen, jota tietokoneaika puheentutkimukselle on merkinnyt. Tämäkään mullistus ei ole tapahtunut yhtäkkiä.

Ensiksikin - ja alkuvaiheessa ehkä päällimmäisenä - merkitsi tietokoneiden tulo tutkimuksenteon rutiinien kevenemistä ja siistiyymistä. Vanhoilla, paljon mekaanista puurtamista vaativilla laitteilla tehtyjen, tulkinnassaan viivaimia ja laskutikkuja edellyttävien analyysikirjausten tilalle saatiin monitorille siististi piirtyviä kuvia, joista kone myös suoraan kykeni lukemaan tulokseksi valmiita lukuarvoja. Tosin mittaustarkeudessa ja -luotettavuudessa jäi usein toivomisen varaa.

Sitä mukaa kun tietokoneitten laskentateho on kasvanut ja käyttäjäystävällisyys parantunut, on myös saatu nopeutta analyysien suorittamiseen. Työskentelynopeuden lisääntyessä päästään tarkastelemaan suurempia aineistoja kuin mitä perinteisillä menetelmillä oltaisiin kohtuullisen ihmisiän puitteissa koskaan kyetty tuottamaan.

Ja toivon mukaan kehityksen myötä yhä useammin tullaan myös pääsemään käsiksi ilmiöihin, joita vanhoilla menetelmillä ei yksinkertaisesti ollut lainkaan mahdollista tavoittaa. Uusien havaintojen myötä saattaa edessä olla monien vanhojen totuuksien uudelleenarviointi. Esimerkiksi tietessämme niin keskeistä käsitystä toteuman ja toteutettavan suhteesta voi olla syytä tarkistaa, jos uusia menetelmiä sovellettaessa akustinen puhe-signaali yllättäen näyttääkin kuin itsestään jaottuvan fonologisia äänneyksiköitä varsin hyvin vastaaviksi segmenteiksi.

Uusia analyysimuotoja

Kun akustisen äänisignaalin ominaisuuksia rekisteröidään fysiikan ja yleisen akustiikan näkökulmasta, käsitellyssä ovat suureina sekunnit, desibelit ja hertsit. Puheen tarkastelussa on sen sijaan psykoakustinen näkökulma osoittautunut fysikaalisia

asteikoita tarkoituksenmukaisemmaksi. Psykoakustinen muunnos pyrki esittämään tietyn akustisen syötteen sellaisena kuin miksi se havaintoprosessimme käsittelyssä muovautuu. Tällä tapaa kuulonmallilla tehdyt muunnokset ovat osoittautuneet puheentutkimuksessa tarpeellisiksi ja tuloksekkaita.

Esittelemme seuraavassa uusia tällaisia analyysimuotoja **äännekkyyssäyrän**, **äännekkyyssuutoskäyrän** ja **spektrierokäyrän**, jotka ovat osa Raimo Toivosen kehittämää Intelligent Speech Analyser -puheentutkimusohjelmaa (ks. Toivonen 1986 ja kotisivu). Nämä analyysimuodot perustuvat kuulonmallilla laskettuihin auditoriseen spektriin ja äännekkyysspektriin. Auditorinen spektri on kuulonmallilla laskettu spektri. Spektrissä on 48 arvoa 0.5 Barkin välein ja pystysteikko on dB-asteikko. Kutakin spektriä varten ääntä tarkastellaan 20-40 ms:n aikaikkunalla. Äännekkyysspektri on kuulonmallilla (auditorinen spektri) laskettu spektri. Spektrissä on 48 pistettä 0.5 Barkin välein ja pystysteikko on soni-asteikko. Kutakin spektriä varten ääntä tarkastellaan 20-40 ms:n aikaikkunalla.

Äännekkyyssäyra on kuulonmallilla laskettujen äännekkyysspektrien kokonaisäännekkyyksien (48 spektripisteen äännekkyyksien summan) muodostama käyrä. Kutakin käyrän pistettä varten ääntä tarkastellaan 20-40 ms:n aikaikkunalla. Käyrän yksikkö on soni. Tässä käyrässä kokonaisäännekkyyden ratkaisee. Kahdella erilaiselta kuulostavalla äänneellä - joilla siis on keskenään erilainen spektri - voi olla sama kokonaisäännekkyyden. Sen sijaan tauon ja äänneen äännekkyysero on suuri.

Äännekkyyssuutoskäyra on äännekkyyssäyrästä 25 ms:n aikavälillä laskettu muutoskäyra. Käyrän yksikkö on soni. Käyra kuvaa äännekkyyssäyrän muutosvoimakkuutta. Äännekkyyden muutos näkyy pohjatasolta nousevana huippuna. Kun äännekkyyden on vakiota muutoskäyra palaa nolliatasolle. Käyrän paikallinen huippu edustaa suurinta paikallista äännekkyyssuutosta. Tämä käyra näyttää toimivan erittäin hyvin vierekkäisten äänneiden välisen rajan määrittämisessä.

Spektrierokäyra on kuulonmallilla 25 ms:n aikaerolla laskettujen äännekkyysspektrien spektriero. Kutakin spektriä laskettaessa tarkastellaan 20-40 ms:n aikavälillä. Käyra kuvaa äännekkyysspektrien muutosta ajassa ja äännekkyysspektrien erot näkyvät spektrierokäyrässä. Käyra ei niinkään kuvaa kokonaisäännekkyyden muutosta, vaan taajuuksien jakauman muutosta. Kahdella erilaiselta kuulostavalla äänneellä on aina erilainen spektri. Spektrin muutos näkyy pohjatasolta nousevana huippuna. Kun spektri ei muutu, palaa erokäyra nolliatasolle. Käyrän paikallinen huippu edustaa suurinta paikallista spektrimuutosta. Kun äänne vaihtuu, näkyy muutos huippuna käyrässä. Mitä

pienempi äänteiden laadun ero on, sen matalammaksi huippu jää. Esimerkiksi äänteiden [o] ja [u] välillä ovat käyrän muutokset pienet.

Uusien menetelmien merkitys

Uusien analyysimuotojen alustavassa testauksessa näyttää siltä, että niitä käyttäen saadaan puheen äännesegmenttien rajat määritettyä lähes poikkeuksetta tarkasti ja yksiselitteisesti. Verrattuna aiemmin käytettävissä olleisiin menetelmiin on ero varsin tuntuva. Esimerkiksi perinteinen FFT-spektrogrammi tai auditorinen spektrogrammi ovat foneettisesti harjaantumattomalle tarkastelijalle vaikeaselkoisia. Harjaantuneillakaan foneetikoille ei aina ole yksimielisyyttä siitä, mihin tarkalleen ottaen spektrogrammissa segmenttirajojen merkkejä asetetaan.

Ero toteutettavan ja toteuman välillä on yksi kielen- ja puheentutkimuksen keskeisiä teesejä - tai ainakin se on opittu sellaiseksi mieltämään. Fonetikassa puhesignaalin ominaisuuksissa todettava vaihtelu, joka johtuu jostain muusta kuin aiotusta viestin merkityssisällön vaihtelusta, on enemmän sääntö kuin poikkeus. Olisiko syytä miettiä uudenlaista maailmankuvaa siltä varalta, että kaikesta huolimatta funktionaalisten yksiköiden ja fyysikaalisten ilmiöitten väliltä löydetäänkin luultua selvempiä yksi-yhteen suhteita?

Kiintoisaa on tietysti ensinnäkin se, millaisia tuloksia uusilla menetelmillä saavutetaan. Uudet analyysimuodot ovat kokeiluissa vaikuttaneet lupaavilta esimerkiksi seuraavien ongelmien kohdalla: Onko puheen taukojen automaattinen paikantaminen mahdollista? Pystytäänkö suomen kielen diftongien sisältä osoittamaan selvä vokaalienvälinen raja? Miksi sananrajoilla tarvitaan glottaaliklusiilia? Voiko kvantiteetiltaan lyhyt äänne olla toteumaltaan monimutkaisempi kuin vastaava kvantiteetiltaan pitkä äänne?

Lisäksi tuntuu ajankohtaiselta pohtia, mitä tutkimuksen kannalta merkitsee tällainen menetelmien älykkyyden kasvaminen. Uudet menetelmät kykenevät yhä tarkemmin osoittamaan tutkijalle, mihin hänen tulee huomionsa kohdentaa. Tällöin kykenee tutkija saavuttamaan tuloksia paljon vähemmällä puheen akustisten ja artikulatoristen ilmiöiden vastaavuuksiin perehtymisellä, kuin mitä vaadittiin aiemmilta instrumentaalisen fonetiikan harrastajilta. Uudet menetelmät tarjoavat siis työskentelypohjan aiempaa laajemmalle joukolle puheen ilmiöistä kiinnostuneita tutkijoita.

Toisaalta voidaan myös kysyä, onko tällaisten, älykkään menetelmän ohjauksen varassa työskentelevien tutkijoiden saavuttama tutkimusaihetta koskeva ymmärrys ehkä myös vähemmän syvällistä, kuin mitä toivoisi. Tuleeko tieteen ammattilaisten joukko jakautumaan entistä selvemmin spesialisteihin, jotka rakentavat älykkäitä koneita, ja soveltajiin, jotka näitä koneita käyttävät kukin omien erityisalojensa ongelmien käsittelyyn? Olisiko tällainen muutos tutkimuksen koko kenttää arvioiden hyvästä vai pahasta? Ja ylipäänsä: olemmeko riittävästi "valmistautuneet mihin tahansa" kuten Ravila tieteen tekijöiltä kysyi edellä lainatussa *Totuus ja metodi* -teoksessaan, jonka sanoma tuntuu olevan aina ajankohtainen.

LÄHTEET

Penttilä, A. 1926. Suomen ja sen lähimpien sukukielten painotusoppia. Kielipsykologinen tutkimuskoe. Turun suomalaisen yliopiston julkaisuja sarja B, osa III, 2. Turku.

Penttilä, A. 1958. Intonaatiotutkimuksia (1 ja 2). *Virittäjä* 62, 1 - 18.

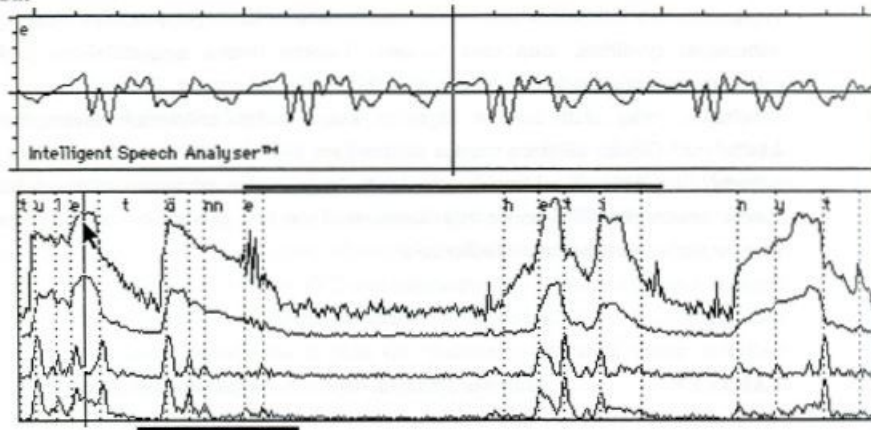
Penttilä, A. 1959. Intonaatiotutkimuksia 3. *Verba docent*. Juhlakirja Lauri Hakulisen 60-vuotispäiväksi 6.10.1959, 442 - 448. Suomalaisen kirjallisuuden seuran toimituksia 263. Helsinki: SKS.

Ravila, P. 1967. *Totuus ja metodi* WSOY Porvoo. Kirja on koottu vuosina 1941 - 1966 ilmestyneistä kirjoituksista.

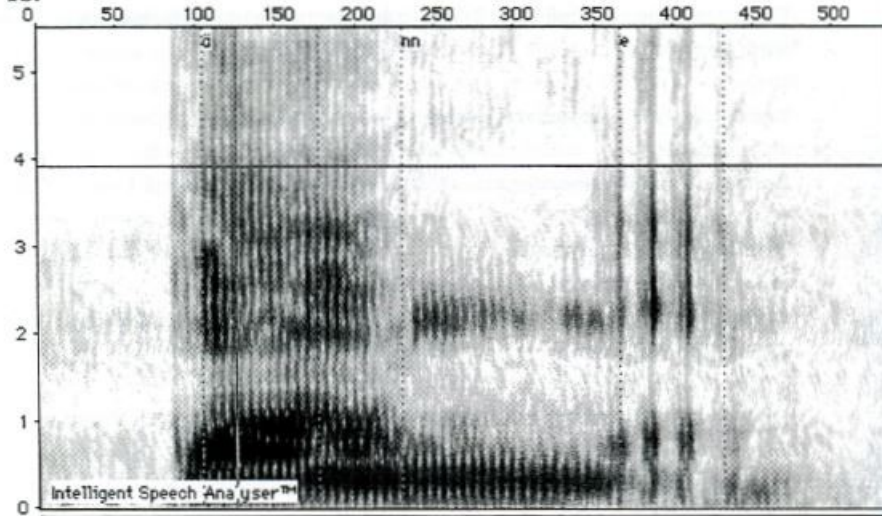
Toivonen, R. 1986. ISA-puheenkäsittelyohjelma. 14.12.1986. Tampere. Moniste.

Toivonen, R. Intelligent Speech Analyser -kotisivu: <http://www.sci.fi/~pitchsys>

1a.

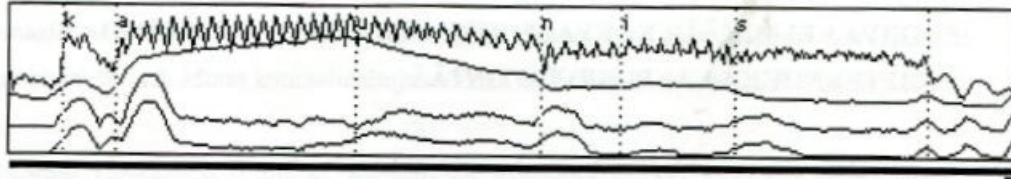


1b.

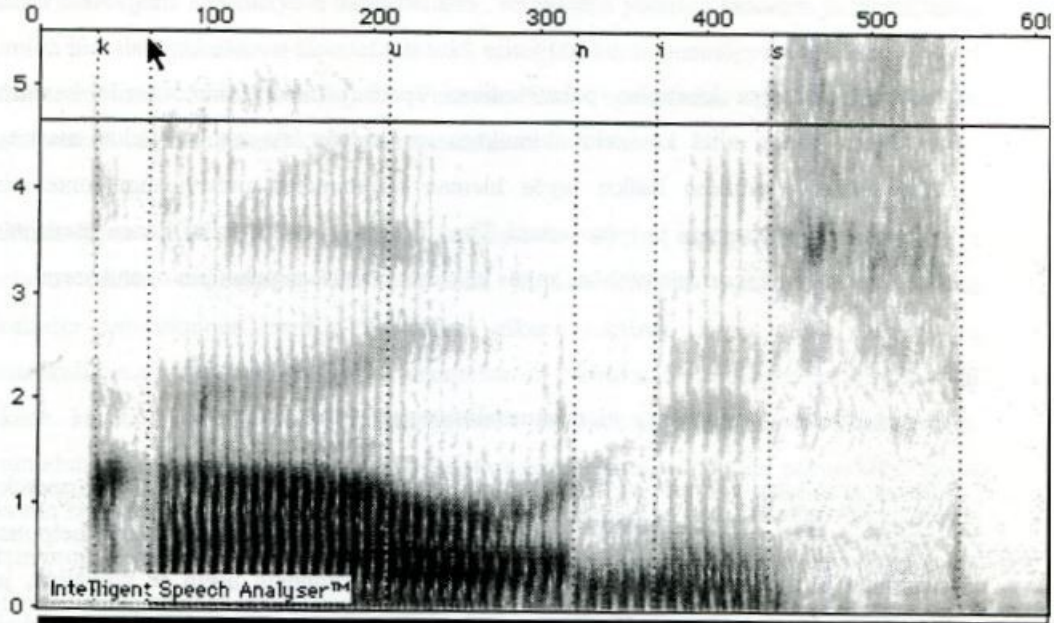


KUVAT 1a ja 1b. Käsittelyssä ilmaus "Tule tänne! Heti, nyt!" Kuvassa 1a ylinnä analyysin lähtökohtana oleva signaalikäyrä, tässä *tule*- sanan /e/-vokaalista (nuolen osoittama kohta alemmassa kuvassa). Signaalikäyrän alla olevassa kuvassa nähdään verhoikäyrä (ylin käyrä), äänekkyyssäyrä, äänekkyyssuutoskäyrä ja spektrinerokäyrä (alin käyrä). Kuvassa mustalla palkilla rajattu *tänne*-sana nähdään kuvassa 1b FFT-spektrogrammissa. Segmenttirajat ovat kuvissa 1a ja 1b samoilla paikoilla.

2a.



2b.



KUVAT 2a ja 2b. Sana *kaunis* miespuhujan ilmauksesta "Ai ai, kun on kaunis päivä."
Kuvassa 2a nähdään jälleen verhoikäyrä (ylin käyrä), äänekkyyskäyrä,
äänekkyysmuutoskäyrä ja spektrierokäyrä (alin käyrä). Kuvassa 2b *kaunis*-sana
nähdään FFT-spektrogrammissa. Segmenttirajat ovat kuvissa samoilla paikoilla.