

Spektri ja spektrogrammi

Kerrataanpa ensin, mitä ääni on

Ääni on kuultavissa olevaa mekaanista aaltoliikettä. Äänen siirtymiseen tarvitaan jokin väliaine. Esimerkiksi ilmassa ääni etenee hyvin pieninä ilmanpaineen muutoksina.

Ihmisen korvaan päätyessään pienet ilmanpaineen vaihtelut heiluttavat tärykalvoa, joka puolestaan siirtää liikkeen välikorvan kuuloluille, joiden vahvistamana liike siirtyy sisäkorvaan. Sisäkorvan sisältämässä hyytelössä etenevä pikkuruinen paineaalto liikuttaa samalla sisäkorvan ns. basilaarikalvoa. Basilaarikalvon eri kohdissa puolestaan sijaitsevat eri taajuisille äänille herkäät kuuloaistinsolut, jotka lähettävät signaalin eteenpäin kuulohermoon ja aivojen tulkittavaksi.

Äänen nopeus ilmassa on noin 330 metriä sekunnissa. Kaikki aineet eivät kuitenkaan johda ääntä samalla nopeudella. Äänen etenemisnopeus riippuu myös aineen lämpötilasta. Kun ääniaalto kohtaa jonkin esteen, osa ääniaallon energiasta heijastuu takaisin päin. Kappaleen muoto vaikuttaa sen värähtelyominaisuuksiin.

Yksittäistä ääniaistimukseen tarvittavan aaltoliikkeen aiheuttajaa kutsutaan äänilähteeksi. Äänilähteenä voi toimia jokin värähtelevä kiinteä kappale, mutta myös esimerkiksi putken tai ontelon sisällä oleva ilmapatsas voi muodostaa värähtelevän massan. Luonnossa esimerkiksi kitka ja muut voimat vaikuttavat värähteleviin kappaleisiin (ja ilmahiukkasiin) ja vaimentavat niiden liikettä. Niinpä värähtely ja siitä syntyvä ääniaalto vaimenevat ja lakkaavat vähitellen, ellei jokin voima ylläpidä värähtelyä.

Yksinkertaisin mahdollinen ääni on *siniääni* joka koostuu *siniaalloista*. Nimitys ei liity väriin vaan se tulee *sinifunktiosta*, jolla siniaallon muoto voidaan matemaattisesti kuvata. Englanniksi siniäänestä käytetään myös termiä *pure tone*, ”puhdas ääni”. Siniäänellä on tietty taajuus eli **frekvenssi** (*frequency*) ja tietty **amplitudi** (*amplitude*) sekä **vaihe** (*phase*) eli se suhteellinen ajanhetki, jolla yksittäinen aalto tai värähdys alkaa ja päättyy. Luonnossa ei kuitenkaan käytännössä koskaan esiinny puhtaita siniääniä, vaan esimerkiksi puhe ja muut luonnolliset äänet ovat aina monimutkaisia eli **kompleksisia**. Voidaan siis ajatella, että jokaisella ajanhetkellä äänessä on mukana monia eri taajuisia ja amplitudiltaan erilaisia siniääniä.

Mikä tahansa kompleksinen ääni (*complex tone*) voidaan ns. **Fourier-muunnoksen** (*Fourier transform*) avulla laskennallisesti purkaa joukoksi yksinkertaisia osasäveliä eli osataajuuksia, joista kukin siis edustaa vain yhtä tiettyä taajuutta. Osataajuuksista käytetään joskus myös nimityksiä **äänes** tai **(taajuus)komponentti** (*frequency component, spectral component*). Osataajuudet siis ”soivat” yhtä aikaa, mutta suhteellisilta ”voimakkuuksiltaan” (energiämäärältään tai intensiteetiltään) erilaisina.

Kuulijan saama kokonaisvaikutelma jonkin äänen väristä ja luonteesta riippuu äänen osataajuuksien äänekkyyksien keskinäisistä suhteista sekä niiden ajallisista muutoksista. Kannattaa myös pitää mielessä, että digitaalista ääntä analysoidaan tietokoneella yleensä lyhyinä, ajallisesti osin päällekkäisinä pätkinä. Ihminen sen sijaan aistii ja havaitsee ääntä

jatkumona, hetkeksikään pysähtymättä. Havaittu ja tulkittu ääni vaikuttaa jatkuvasti siihen, mitä odotamme kuulevamme – ja samalla siihen, mitä lopulta oikeasti kuulemme!

Spektrianalyysi

Minkä tahansa äänipätkän sisältämät osataajuudet ja niiden keskinäiset voimakkuussuhteet voidaan siis selvittää Fourier-muunnokseen perustuvan spektrianalyysin avulla. Kaikki luonnosta äänitettävissä olevat äänet – myös puhe – muuttuvat kuitenkin käytännössä jatkuvasti. Tämän vuoksi myös spektrin muoto riippuu aina jossain määrin siitä, millainen ja kuinka pitkä pätkä ääninäyttestä on valittu spektrianalyysin kohteeksi. Spektrianalyysi onkin jossain määrin dilemma. Jos analysoimme liian pitkän pätkän ääntä kerrallaan, sen sisällä tapahtuneet ajalliset muutokset eivät erotu lopputuloksessa. Jos taas analysoimme liian lyhyen pätkän ääntä, emme ehkä pysty havaitsemaan sellaisia tapahtumia, jotka kestävät analysoidun näytteen kestoa kauemmin.

Spektrin tarkastelu Praatissa

Praatin ääni- tai annotaatioeditorissa voit laskea spektrin joko lyhyestä ikkunasta kursorin ympäriltä tai valitusta äänipätkästä, kun valitset valikosta Spectrum: View spectral slice. Spektrin vaaka-akseli kuvaa äänen (osa)taajuutta ja pystyakseli suhteellista intensiteettiä (yksikkönä desibeli, dB). Spektrikuvasta puuttuu siis kokonaan aika ja valitusta aikaikkunasta poimitun äänipätkän tiedot näkyvät yhtä aikaa samassa kuvassa. Spektrin asetuksia voi säätää editorissa samasta kohdasta kuin spektrogrammin asetuksia. Lisätietoa:

https://www.fon.hum.uva.nl/praat/manual/Intro_3_7_Configuring_the_spectral_slice.html

Spektrogrammi: paljon spektrejä “käsi kädessä”

Spektrogrammi on kuvaustapa, jolla voidaan havainnollistaa äänen taajuusrakenteen muutoksia ajassa. Spektrogrammista voidaan saada nopeasti kokonaiskäsitys äänisignaalin osataajuuksien jakautumisesta. Spektrogrammi perustuu spektrianalyysiin, jota toistetaan lukuisia kertoja peräkkäin siirtämällä määrätyn levyistä aikaikkunaa pienin askelin ääninäytteen alusta loppuun.

Spektrogrammeja voidaan piirtää monella tavalla. Praatin perinteisessä spektrogrammissa kuvataan äänen energiatiheyttä (keskimääräistä intensiteettiä) harmaasävyillä: mitä tummempi kohta spektrogrammissa, sitä suurempi keskimääräinen intensiteetti äänisignaalissa on todettu kyseisellä taajuusvälillä

(https://www.fon.hum.uva.nl/praat/manual/Intro_3_1_Viewing_a_spectrogram.html).

Joskus intensiteettiä kuvataan harmaasävyjen sijaan eri väreillä. Vastaaan voi tulla myös spektrogrammeja, joissa aika kulkeekin syvyys suunnassa, edestä taaksepäin. Spektrogrammin voi kuvata myös vaikkapa animaationa, kuten esimerkiksi Mark Newmanin videossa *What is sound?*:

<https://youtu.be/ZRZlZ81nXo4>

Leveä- ja kapeakaistaiset spektrogrammit

Eri tarkoituksiin on perinteisesti laskettu kahdentyyppisiä spektrogrammeja, ns. **leveäkaistaisia** ja **kapeakaistaisia**. Niiden mukaisia asetuksia käyttämällä voidaan

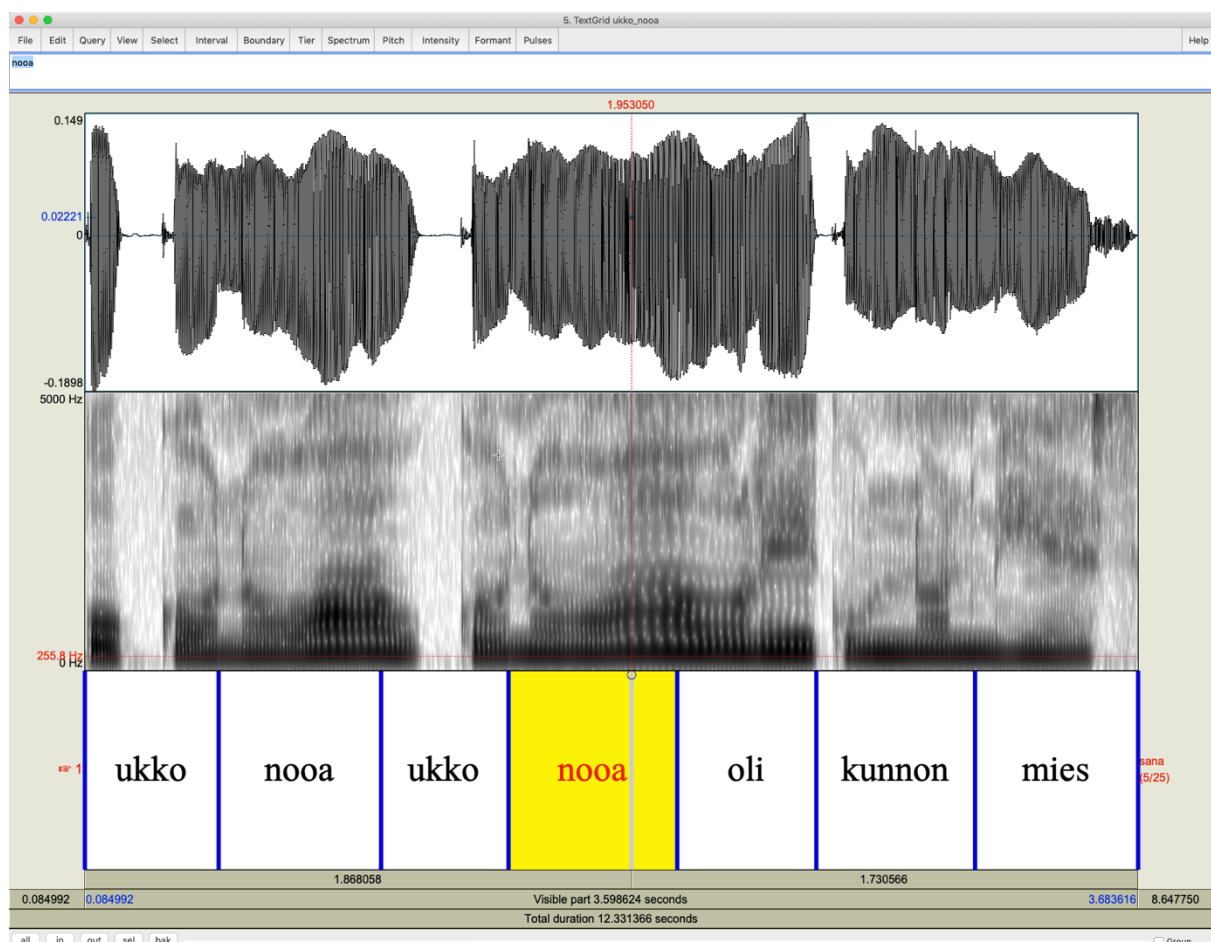
tarkastella puheäänien spektraalisia ominaisuuksia eri näkökulmista.

Leveäkaistainen spektrogrammi (analyysi-ikkunan pituus n. 5 ms; Praatissa oletuksena) näyttää tarkasti ajassa tapahtuvia muutoksia. Esimerkiksi glottisperiodit (äänihuulten tuottamat yksittäiset "paukaukset") näkyvät selvästi pystyraitoina.

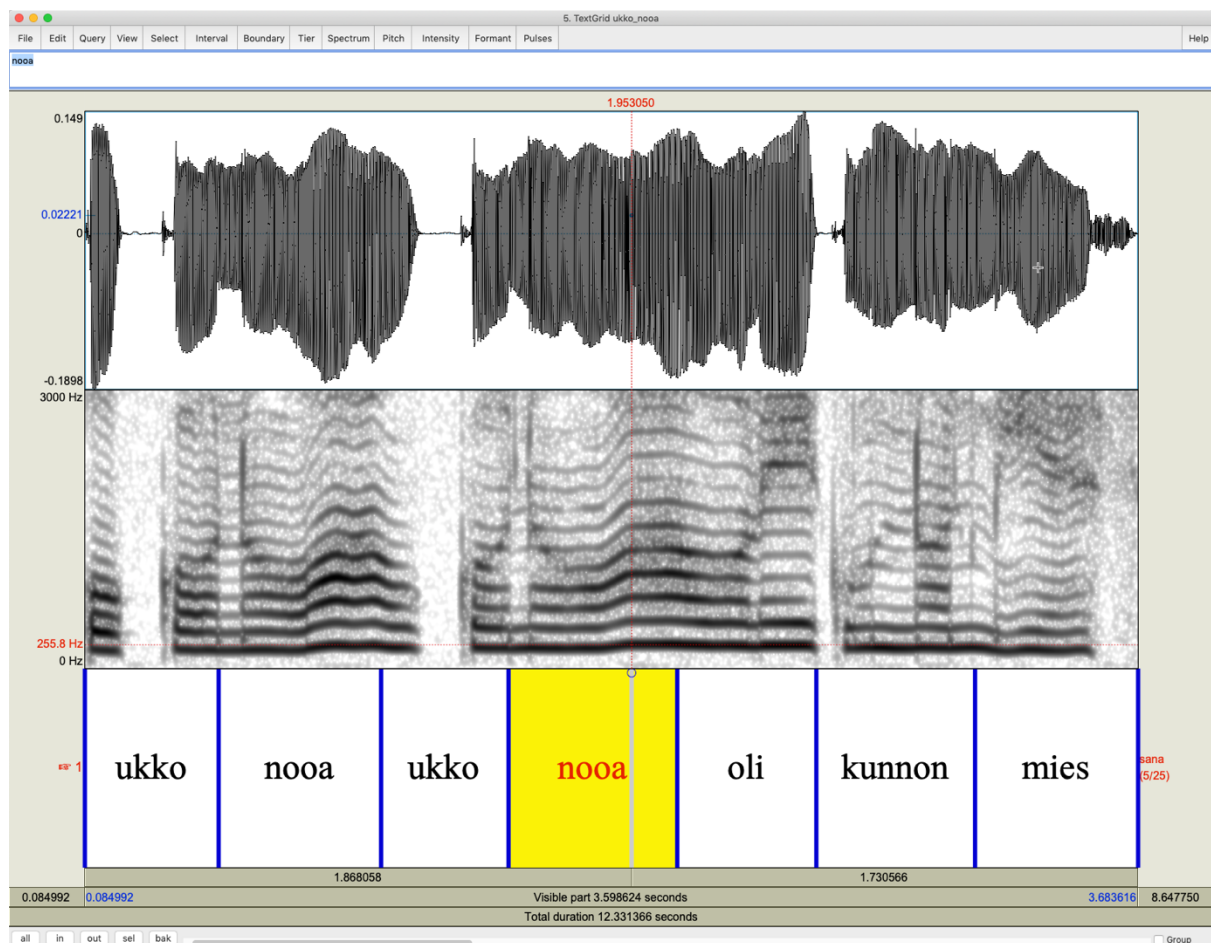
Leveäkaistainen spektrogrammi on kätevä esimerkiksi puhenäytteen segmentoinnissa ja nimikoinnissa, kun määrittelet äänteiden rajoja. Spektrin ajalliset muutokset näkyvät tarkasti ja onnistut ehkä löytämään nopeammin spektraalisia "käännekohtia", jotka ovat usein hyviä ehdokkaita segmenttirajoiksi. (Rajakohta on tietenkin varmistettava kuuntelemalla.)

Kapeakaistaisessa spektrogrammissa (analyysi-ikkunan pituus n. 30 ms) sekä formantit (näkyvät tummina vaakakaistaleina etenkin soinnillisissa äänteissä), äänen osataajuudet (ns. harmoniset, harmonics) että näiden ajassa tapahtuvat muutokset erottuvat parhaiten.

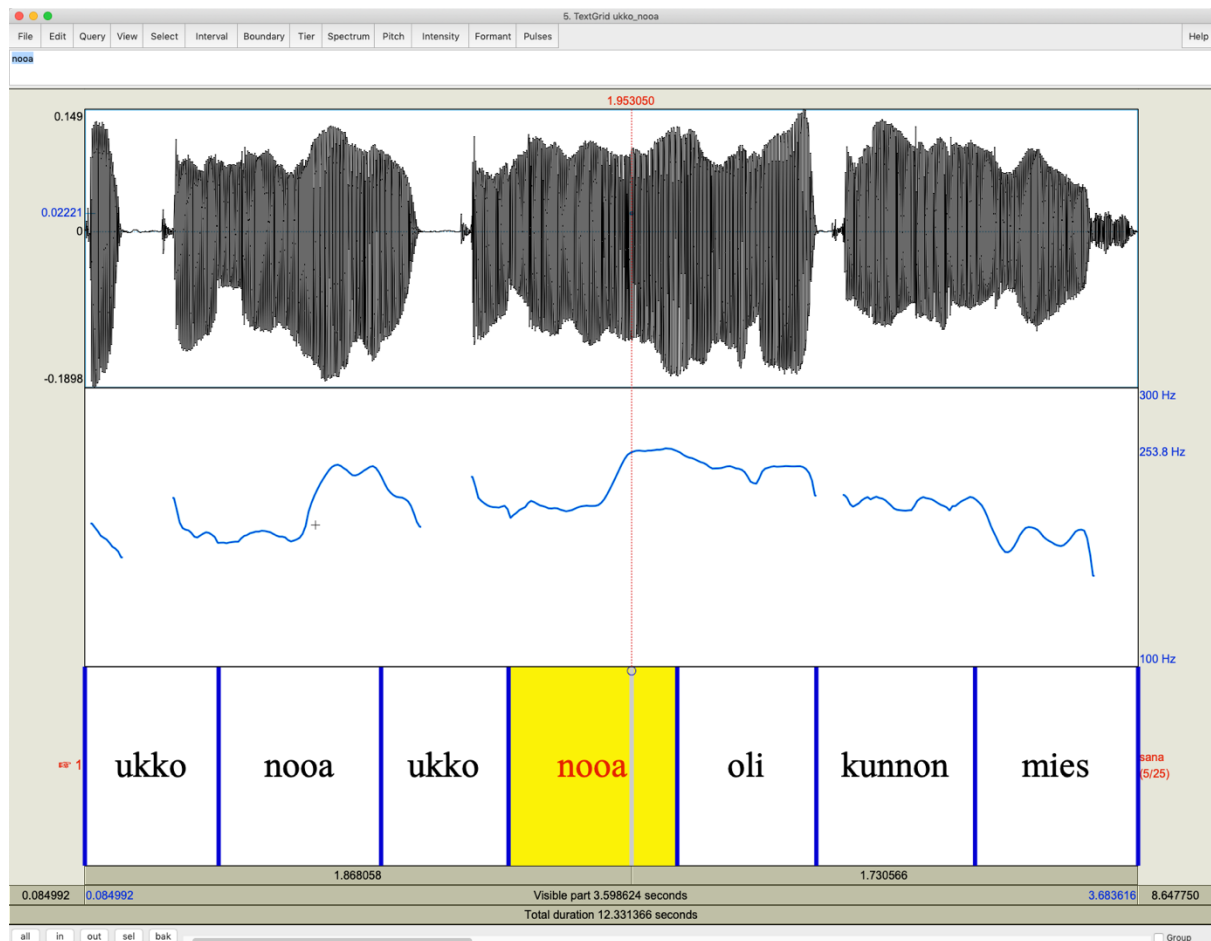
Kapeakaistainen spektrogrammi voi olla hyödyllinen esimerkiksi, kun haluat tarkastella F0- eli perustaajuusanalyysin oikeellisuutta. Tällöin editori-ikkunassa kannattaa määritellä näkyviin vain spektrogrammin alimmat taajuudet, niin että vain muutama alin osataajuus erottuu selvästi. Jos kytket samalla näkyviin myös perustaajuusanalyysin ja määrittelet sille saman näyttöasteikon kuin spektrogrammille, voit helposti nähdä, seuraako perustaajuskäyrä puheen alinta osasäveltä (kuten sen oikeastaan pitäisi).



Leveäkaistainen spektrogrammi naispuhujan laulamasta Ukko Nooa -laulun säkeestä. Analyysi-ikkunan pituudeksi on asetettu 0.005 s (= Praat-ohjelman oletusarvo).



Kapeakaistainen spektrogrammi naispuhujan laulamasta Ukko Nooa -laulun säkeestä. Analyysi-ikkunan pituudeksi on asetettu 0.03 s. Näkyvissä olevaksi taajuusalueeksi on säädetty 0-3000 Hz, jotta alimmat osasävelet erottuisivat paremmin. Hiiren osoittimen punainen pystyviiva on asetettu Nooa-sanin viimeisen tavun paikkeille aikapisteeseen 1,953050 s. Vaakaviiva taas on merkitty silmämääräisesti spektrogrammissa näkyvän alimman osasävelen keskikohdalle, jonka pitäisi vastata äänen perustaajuutta. Vaakasuoran punaisen viivan kohtaa vastaa spektrogrammikuvasa taajuus 255,8 Hz, joka näkyy ikkunan vasemmassa reunassa punaisella tekstillä. Huomaa, että tämä ei ole tarkka mittaustulos vaan spektrogrammiin perustuva arvio. Tätä silmämääräisesti saatua tietoa voi verrata seuraavassa kuvassa näkyvään, perustaajuusanalyysin antamaan tulokseen, sillä kursori on molemmissa kuvissa sijoitettu samaan aikapisteeseen ääninäytteessä.



Aaltomuoto ja perustaajuuskäyrä (Pitch) naispuhujan laulamasta Ukko Nooa -laulun säkeestä. Perustaajuusanalyysin alaraja-asetukseksi on määritetty Praat-ohjelman äänieditori-ikkunan Pitch settings -kohdasta 75 Hz ja ylärajaksi 500 Hz. Näkyvissä olevaksi taajuusalueeksi on säädetty Advanced pitch settings -kohdasta 100-300 Hz, jotta perustaajuuden liikkeet erottuisivat paremmin. Kursori (punainen pystyviiva) on samassa aikapisteessä kuin kahdessa edellisessä kuvassa. Kursorin kohdalta mitattu perustaajuusarvo 253,8 Hz näkyy ikkunan oikeassa reunassa sinisellä tekstillä.

Spektrogrammin tuottaminen Praatissa

Editori-ikkunassa

Spektrogrammi saadaan näkyviin äänieditorissa tai TextGrid-editorissa valitsemalla **Spectrum**-valikosta **Show spectrogram**. Tällöin spektrogrammin tärkeimpiä asetuksia voidaan tarvittaessa muuttaa **Spectrum**-valikon komennolla **Spectrogram settings...**

- Jos haluat laskea leveäkaistaisen spektrogrammin (oletus Praatissa), analyysi-ikkunan pituudeksi (window length) tulisi antaa noin 0.005 s eli 5 ms (tällöin ns. kaistanleveydeksi saadaan 260 Hz).
- Jos haluat laskea kapeakaistaisen spektrogrammin, analyysi-ikkunan pituudeksi tulisi antaa noin 0.03 s eli 30 ms (tällöin ns. kaistanleveydeksi saadaan 43 Hz). (Voit

verrata kapeakaistaisen spektrogrammin kuvaa myös samasta pienestä laulunäytteestä laskettuun perustaajuuskäyrään yllä olevassa kuvassa.)

Objekti-ikkunassa

Objekti-ikkunassa uusi spektrogrammiobjekti voidaan luoda valitsemalla ensin haluttu ääniobjekti ja painamalla dynaamisen valikon painiketta Analyse spectrum: To Spectrogram... Kohtaan Window length merkitään analyysi-ikkunan pituus kuten äänieditorin kautta laskettaessa.

Mihin spektrogrammia voi käyttää?

Spektrogrammit ovat vain yksi tapa kuvantaa äänen taajuusrakennetta ajan funktiona. Spektrogrammi ei sinänsä ole varsinainen tulos tai tilastollista aineistoa; se on vain yleiskuva tietyistä äänisignaaleista. Tilastolliseen tutkimukseen siis tarvitaan muita numeerisia mittauksia. Jos taas teet kvalitatiivista tutkimusta, spektrogrammi sinänsä ei riitä, vaan sitä on voitava tulkita tutkimuskysymykseen nähden perustellulla tavalla.

Oman kokemukseni mukaan spektrogrammi on paras väline tutkijalle itselleen. Sen avulla voi kahlata aineistoa läpi ja etsiä varsinaisia mitattavia tutkimuskohteita. Kokemuksen karttuessa foneetikko alkaa tunnistaa näytteistä erilaisia äänteitä ja ääniä jopa pelkästään spektrogrammia katsomalla. Yksittäisiä aineistosta poimittuja esimerkkejä voi toki tarvittaessa havainnollistaa spektrogrammeilla myös varsinaisissa julkaisuissa.