

Digitaalinen ääni

Vielä viime vuosikymmenen puolella äänityksiä tehtiin lähinnä analogisille tallennusvälineille, esimerkiksi kelanauhoille ja C-kaseteille tai vinyylilevyille. Nykyisin äänitteet tallennetaan usein suoraan digitaaliseen muotoon.

Miten digitaalinen äänitiedosto syntyy?

Äänen tallentamiseen tarvitaan mikrofoni, joka pystyy muuttamaan pienet ilmanpaineen muutokset audiokaapelia pitkin kulkevaksi jännitevaihteluksi. Mikrofoneja on monenlaisia, mutta mikrofonin tarkoituksena ei yleensä ole digitoida ääntä vaan ainoastaan kerätä sitä eli toimia "anturina".

Kun ääntä halutaan tallentaa esimerkiksi mobiililaitteilla tai tietokoneella käsiteltävään muotoon, ääni pitää ensin kertaalleen digitoida. Äänen digitointi tarkoittaa, että mikrofonilla kerätystä ja sähköiseen muotoon muunnetusta äänestä otetaan näytteitä (eli tehdään mittauksia) hyvin lyhyin, mutta täsmällisin väliajoin. Näytteet poimii laite nimeltä A/D-muunnin (*Analog to Digital*). A/D-muunnin voi sijaita esimerkiksi äänitallentimessa tai mobiililaitteen äänipiirissä tai tietokoneen äänikortissa.

Jos äänite on tarkoitettu tutkimuskäyttöön, kannattaa digitointivaiheessa käyttää mahdollisimman hyvälaatuisia laitteita, jotta äänen näytteistäminen tapahtuisi luotettavasti. Esimerkiksi tavalliseen toimistotyöhön tarkoitettujen kannettavien tietokoneiden äänitysominaisuudet ovat yleensä huomattavasti heikommat kuin ammattikäyttöön tarkoitetuissa digitaalisissa tallentimissa tai tietokoneeseen liitettävissä ns. ulkoisissa audiorajapinnoissa (audio interface).

Digitoitu ääni tallentuu äänitiedostoksi, jota voidaan kopioida, muuntaa eri formaatteihin ja toistaa esimerkiksi tietokoneella tai muilla digitaalisilla soittimilla. Alkuperäinen äänitiedosto ja kaikki sen kopiot pysyvät sisällöltään täsmälleen samanlaisina, ellei tiedostoa muunneta toiselle näytetarkkuudelle tai pakata häviöllisesti (ja ellei tietokoneen kiintolevy, muistitikku tai muu tallennusväline vaurioidu, jolloin tiedostoista tietenkin tulee usein kertaheitolla lukukelvottomia). Toisin sanoen A/D-muuntimen ominaisuuksilla ei enää ole väliä sen jälkeen, kun ääni on kertaalleen digitoitu eli äänitiedosto on muodostettu.

Äänen näytteistäminen

Näytteitä voidaan ääntä digitoidessa ottaa esimerkiksi 22050 kertaa sekunnissa, jolloin sanotaan, että tallennettavan äänitiedoston näytteenottotaajuus on 22050 Hz (hertsiä), eli 22,05 kHz (kilohertsiä). Jokainen tällainen näyte esitetään tietokoneessa yhdellä luvulla, joka puolestaan esitetään bitteinä. Tämä näytteen koko eli ns. bittisyvyys tai bittileveys kertoo amplitudin tarkkuuden. Äänisignaalin aaltomuoto voidaan ikään kuin rakentaa uudelleen näiden lukujen muodostamasta jonosta.

Näytteenottotaajuus vastaa siis digitaalisen äänitteen taajuusresoluutiota: mitä korkeampi näytteenottotaajuus, sitä tiheämmin toistuvia ääniaaltoja ts. sitä korkeampia ääniä saadaan mukaan tallenteeseen. Suurin kuvattavissa oleva taajuus (ns. *Nyquistin taajuus*) on aina puolet äänen näytteenottotaajuudesta. Jos näytteenottotaajuus siis on 22050 Hz, digitoituun ääneen voi tulla mukaan korkeintaan 11025 Hz taajuisia ääniä. Kun tallennetaan puhetta

tutkimuskäyttöön, kannattaa heti alussa käyttää vähintään 22 050 Hz:n, mutta mieluiten 44100 Hz:n näytteistystä vähintään 16-bittisenä. Vielä parempi laatu valitaan tarpeen mukaan.

Häviöllinen ja häviötön pakkaus

Äänitiedostot vievät paljon tilaa varsinkin, mikäli alkuperäinen ääni halutaan kuvata tarkasti ts. tallennettavia näytepisteitä on kerätty hyvin lyhyin aikavälein. Tästä syystä on kehitetty äänen pakkausformaatteja, joista tunnetuin on MP3.

Pakkaamattomia äänentallennusformaatteja ovat esimerkiksi Microsoftin kehittämä WAV ja Applen kehittämä AIFF. Näistä kahdesta WAV alkaa olla laajimmalle levinnyt ja siksi suosittelavin. AIFF on jäänyt jo pitkälti pois käytöstä. Tutkimuskäytössä on yleensä suositeltavaa säilyttää koko aineistosta ainakin yksi pakkaamaton kopio esimerkiksi WAV-muodossa.

MP3-tiedostot on suunniteltu lähinnä musiikin tallennukseen ja kuunteluun kannettavissa laitteissa, ja ne ovat häviöllisesti pakattuja (engl. *lossy compression*). MP3-pakkauksen yhteydessä äänestä hävitetään tarkoituksella sellaista informaatiota, joka ei merkittävästi muuta musiikista syntyvää kuulovaikutelmaa. Puheentutkimuksen kannalta häviöllinen pakkaus on ongelma, sillä tutkijalla ei ole mahdollisuutta selvittää tai päätellä, mitä kaikkea alkuperäisestä äänisignaalista on pakkaamisen myötä muuttunut tai jäänyt pois. MP3 on lisäksi kaupallinen formaatti. Avoimen lähdekoodin Praat-ohjelmalla MP3-tiedostoja ei voi tuottaa, mutta niitä voi kyllä avata käsiteltäväksi ja kuunnella.

On olemassa myös häviöttömästi pakattuja äänitiedostoformaatteja (*lossless compression*), mm. FLAC. FLAC-tiedostojen pitäisi yleensä aueta myös Praatilla. Vastaan on kuitenkin tullut kenties vanhemmilla laitteilla tallennettuja FLAC-tiedostoja, joita ei ole saanut auki muuten kuin alkuperäisellä FLAC-dekooderilla. Etenkin isommissa puheaineistoissa FLAC-muoto voi joka tapauksessa tarjota mahdollisuuden tilansäästöön.

Stereo vai mono?

Stereotallenteita käyttämällä voidaan tarjota kuulijalle tilavaikutelma, joka parantaa esimerkiksi musiikkiesitysten tunnelmaa. Kun kaiuttimista tai kuulokkeiden kautta kunnellaan äänitettä, jossa vasemman ja oikean ääniraidan välillä on pieniä ajoituseroja, nämä synnyttävät vaikutelman kolmiulotteisesta tilasta. Stereovaikutelma voidaan yksinkertaisimmillaan saada aikaan tallentamalla samaa esitystä tai äänilähdettä kahdella eri mikrofonilla eri kanaville. Ääntä voidaan toki myös tallentaa monikanavaisena, miksata ja jälkikäsitellä, tai koko äänite voidaan tuottaa alusta lähtien digitaalisesti. Kuulijan havaintoon voidaan siis vaikuttaa eri tavoin.

Tallentimet, mobiililaitteet ja äänitysohjelmat saattavat tallentaa oletusarvoisesti stereoääntä. Se ei välttämättä haittaa, mutta on hyvä muistaa, että äänitiedostot kuluttavat levytilaa. Yksi minuutti CD-laatuista ääntä vie noin 10 MB (megatavua). Kun pakkaamaton stereoäänitiedosto muutetaan yksikanavaiseksi eli monoksi, sen tilantarve putoaa puoleen.

Puheentutkimuksessa tilavaikutelmasta ei välttämättä ole sanottavaa hyötyä – pikemminkin tietyn puhujan tuotos halutaan tallentaa niin suoraan ja "puhtaasti" kuin mahdollista. Jos puhetta voidaan tai halutaan äänittää vain yhdellä mikrofonilla, riittää yksikanavainen eli

monoäänite. Stereoäänitettä tai monikanavaista äänitettä kannattaa kuitenkin käyttää, jos puhujia on useita. Kun jokainen keskustelun osallistuja tallennetaan omalla mikrofonillaan ja omalle ääniraidalleen, kunkin ääntä voidaan tarpeen mukaan kuunnella joko erikseen tai yhdessä muiden kanssa, jolloin mm. aineiston litterointi helpottuu huomattavasti ja myös yksittäisen puhujan äänen akustinen analyysi onnistuu paremmin.