

Lauluäänen digitoinnin haasteet ja mahdollisuudet musiikin opetuksessa

Veli Keitel

Opinnäytetyö
Joulukuu 2020
Kulttuuriala
Musiikkipedagogi (AMK)

Tekijä(t) Keitel, Veli	Julkaisun laji Opinnäytetyö, AMK	Päivämäärä Marraskuu 2020
	Sivumäärä 38	Julkaisun kieli Suomi
		Verkojulkaisulupa myönnetty: x
Työn nimi Lauluäänen digitoimisen haasteet ja mahdollisuudet musiikin opetuksessa		
Tutkinto-ohjelma Musiikkipedagogi (AMK)		
Työn ohjaaja(t) Sami Sallinen		
Toimeksiantaja(t)		
Tiivistelmä <p>Digitaalisen informaation aikana lauluäänen digitoiminen ja saattaminen digitaaliseksi instrumentiksi on ilmeinen tutkimuskohde. Tutkimuksen tehtävä oli kartoittaa tällaisen digitaalisen lauluinstrumentin valmistamisen haasteita ja mahdollisuuksia, sekä sen käyttökelpoisuutta musiikinopetuksen, säveltämisen, sovittamisen ja studiomusiikin työkaluna. Vertailukohtana oli useita vastaavia kaupallisia sovelluksia.</p> <p>Toteutustapana oli pääasiassa empiirinen tutkimus, jossa kahden laulajan ja yhden kuoron ääntä pilkottiin osiin (vokaalit ja konsonantit), ja tutkittiin, miten nämä lauluäänen osat voidaan rekonstruoida digitaalisesti MIDI-instrumentin muodossa, siten että sillä toistettu musiikki mahdollisimman autenttisesti kuulostaa samankaltaiselta kuin alkuperäinen ääni. Instrumentin rakentamisessa käytettiin sampling-teknologiaa, ohjelmoina Reaper (digitaalinen audiotyöasema, DAW) sekä TX16Wx-sampleriohjelma. Välineistönä oli desktop-tietokone, kaksi MIDI-koskettimistoa, laulumikrofoni sekä Zoom -äänityslaitte.</p> <p>Tutkimustyön tuloksena syntyi useita eri versioita kuvatuslaisesta MIDI-instrumentista sekä digitaalinen alusta (template), johon voidaan istuttaa uusi laulaja tai kuoro, hyödyntäen jo tehtyä pohjatyötä. Saatiin palautteen kautta näkemyksiä digitaalisten lauluinstrumenttien hyödyllisyydestä musiikinopetuksessa. Samalla saatiin käytännön tutkimuksen kautta monenlaista tietoa siitä, millaisia työvaiheita ihmisäänen digitaaliseen rekonstruktioon MIDI-instrumentin muodossa sisältyy.</p> <p>Johtopäätöksenä voidaan todeta, että ihmisäänen saattaminen sample-instrumentiksi on työläs ja vaikea prosessi, varsinkin jos halutaan instrumentin kuulostavan autenttiselta. Työn pohjalta voidaan kuitenkin todeta, että tällaiselle teknologialle on käyttöä, ja sen mahdollisuuksia musiikinopetuksessa on hyvä selvittää tarkemmin.</p>		
Avainsanat (asiasanat) DAW, digitoiminen, ihmisääni, instrumentti, kuoro, laulu, MIDI, musiikki, reaper, sample, samplekirjasto, sampling, simulointi, ääni		
Muut tiedot (Salassa pidettävät liitteet)		

Author(s) Keitel, Veli	Type of publication Bachelor's thesis	Date November 2020 Language of publication: Finnish
	Number of pages 38	Permission for web publication: x
Title of publication The possibilities and challenges of digital singing voice in music pedagogy		
Degree programme Music pedagogy AMK		
Supervisor(s) Sallinen, Sami		
Assigned by		
<p>Abstract</p> <p>In the age of digital information making a digital instrument out of the singing voice is an obvious subject to study. The aim was to examine the challenges and possibilities associated with the construction of such an instrument and its practical uses in music pedagogy, composition, arranging and as a tool in music studio. Some similar commercial applications were used as a base for comparison.</p> <p>The study was done through empirical research, where the voices of two soloists and one choir were sliced into individual vowels and consonants. The goal was to find out how these components could be reconstructed into a digital MIDI instrument, so that the reproduced music is as authentic as possible, as compared to the original voice. Sampling technology was applied, using the Reaper DAW software (digital audio workstation) and the TX16Wx sampler software. The equipment included a desktop computer, two MIDI keyboards, a singing microphone and a Zoom recorder.</p> <p>As a result of the study several versions of such a MIDI instrument were constructed, including a digital template on which a new singer or choir can be applied, utilizing the work that was done with other voices. Feedback was received to help evaluating the usability of this kind of instruments in music teaching. Through the practical research work a wide understanding of the process of reconstructing a human voice in the form of a digital instrument was gained.</p> <p>In conclusion it can be stated that transforming a singer's voice into a sampled instrument is a cumbersome and challenging process, especially if the instrument needs to sound authentic. Based on the study it can yet be said that this kind of technology can be useful and its applications in music pedagogy should be further examined.</p>		
Keywords/tags (subjects) choir, DAW, digitizing, instrument, MIDI, music, reaper, sample, sample library, sampling, simulation, singing, sound, voice		
Miscellaneous (Confidential information)		

Sisältö

1	Johdanto.....	3
1.1	Tutkimuskysymyksen määrittely ja tutkimusmenetelmä.....	3
1.2	Mihin digitaalista laulajaa tarvitaan	5
2	Sampling-teknologia ihmisäänen simuloimisessa	6
3	Tietoperusta	8
4	Fonetiikan rooli lauluäänen digitoimisessa	9
5	Digitoinnin tekninen toteutus	11
6	Tutkittavat osa-alueet ja tutkimuksen tulokset	16
7	Johtopäätökset.....	29
8	Pohdinta.....	31
8.1	Projektin pedagogiset mahdollisuudet	31
8.2	Mahdollisuudet säveltäjille ja sovittajille	33
8.3	Instrumentin kehitys jatkossa	35
	Lähteet.....	37
	Liitteet	38
	Liite 1. Sampler vocabulary (käsitteistö), Excel-taulukko	38
	Liite 2. Tutkimusta havainnollistava video: youtu.be/G5kIBQ5j3KE	38

Kuviot

Kuvio 1.	Yhden koskettimiston kartta.....	23
Kuvio 2.	CC-viestit ohjaamassa vokaalin valintaa	24
Kuvio 3.	Konsonanttiversion valintaa ohjaava matriisi	27
Kuvio 4.	Konsonanttisamplen eri vokaaleita vastaavat versiot.....	27

Taulukot

Taulukko 1. Tekemäni samplekirjasto verrattuna kaupallisiin vastineisiin	8
Taulukko 2. Simuloitavien ominaisuuksien tuottaminen	15

1 Johdanto

Nykyään mitä hyvänsä instrumenttia voidaan simuloida sähköisesti digitaalisen teknologian avulla. Yleensä nämä simulaatiot on tuotettu sample-teknologian avulla, mikä käytännössä tarkoittaa sitä, että halutun instrumentin tuottamat äänet taltioidaan digitaalisesti ja toistetaan sen jälkeen tarkoitukseen sopivan käyttöliittymän, yleensä koskettimiston, avulla (sample tarkoittaa tässä yhteydessä äänestä tai musiikista otettua digitaalista näytettä). Jotkin soittimet voidaan simuloida realistisesti myös syntetisoimalla; yleensä sampling-menetelmä tuottaa kuitenkin autenttisemmän kuuluisen äänen.

Melkeinpä ainoan poikkeuksen tässä muodostaa ihmisääni, erityisesti sanoitettu laulu. Samplekirjastoja (*sample library*, yleisesti käytetty termi instrumentille, joka vaatii lukuisia sampleja) löytyy, mutta ne kattavat aina vain tietyn osa-alueen ihmisäänestä instrumenttina. Laulun simuloiminen digitaalisella instrumentilla on erityisen ongelmallista, koska sampleja tarvitaan valtava määrä; äänen "soundin" ja dynamiikan lisäksi on simuloitava eri laulutekniikat, kaikki vokaalit ja konsonantit sekä niiden yhdistelmät ja glissandot eri sävelkorkeuksien välillä. Toinen ongelma on soittotekninen: Kuinka sanoitettua laulua voidaan soittaa reaaliaikaisesti - vai voiko sitä? Tarvitseeko sitä?

1.1 Tutkimuskysymyksen määrittely ja tutkimusmenetelmä

Tutkimukseni pyrkii selvittämään lauluäänen simuloimiseen liittyviä ongelmia ja etsimään niihin erilaisia ratkaisuja: Onko realistista tavoitella täysin luonnollisen kuuluisista digitaalista lauluinstrumenttia? Jos ei, onko tällainen instrumentti puutteellisenakin riittävän hyödyllinen musiikinopettajille, kuoroammattilaisille ja muusikoille yleensä; varsinkin, kun otetaan huomioon instrumentin valmistamiseen vaadittava työmäärä.

Koska olen jo vuosia ennen tähän opinnäytetyöhön ryhtymistä paneutunut lauluäänien digitoimiseen oman kokeilemisen kautta, minulle oli luontevaa valita tutkimusmenetelmäksi empiirinen tutkimus, jossa saadaan tekemällä konkreettisia havaintoja tutkimuskohteesta ja analysoimalla ja mittaamalla sitä. Tutkimuskohteena tässä tapauksessa on ihmisääni, sekä laulua että puhetta tuottavana instrumenttina. Analysoitavana ovat vokaalien ja konsonanttien ominaisuudet, niiden yhdistäminen toisiinsa luontevasti ja monet muut ääneen liittyvät ilmiöt. Mitattavana ovat äänen taajuuudet, äänen tuottamisen eri vaiheiden kestot, sävelkorkeudet jne.

Monilta osin tutkimusmenetelmän alalajina on *kokeellinen tutkimus*. Kokeellisessa tutkimuksessa pyritään havainnoimaan ilmiöiden vaikutuksia ja syy-seuraus-suhteita kontrolloimalla ja ottaen huomioon kaikki ilmiöön liittyviä tekijöitä. Tämän nim omaisen tutkimuksen kohdalla näitä tekijöitä ovat ihmisäänen ja sen digitaaliseen tallentamiseen liittyvät lukuisat osa-alueet. Tutkimuksen kohteena on digitoitu ääni, kontrolliryhmänä alkuperäinen autenttinen lauluääni.

Tutkimuskohde on erittäin laaja, joten sitä on selkeästi rajattava. Ensinnäkään en ole lähtenyt hakemaan laulunäytteitä kattavasti erityyppisiltä laulajilta, vaan keskittynyt pääasiassa omaan baritoniääneeni ja apulaisena toimineen Ana Pilatin altoääneen. Tarkoitukseni ei ole selittää puheen foneettisia yksityiskohtia vaan keskittyä sanoitetun laulun digitoimiseen liittyviin haasteisiin – ei niin, että vastaan kaikkiin kysymyksiin, vaan esitän niihin erilaisia ratkaisukeinoja sekä ratkaisujen jäädessä puuttumaan esitän uusia kysymyksiä, joita mahdollisesti voidaan käsitellä jatkotutkimuksessa.

Tarkoitus ei ollut myöskään valmistaa loppuun asti viimeisteltyä sovellusta, vaan tutkia, millaisia haasteita ja työvaiheita tällaisen sovelluksen valmistaminen vaatii sekä arvioida sovelluksen hyötyjä suhteessa sen tuottamiseen vaadittavaan työmäärään.

Työni tuloksena on paitsi tämä tutkimusraportti, myös konkreettinen esimerkki digitaalisesta lauluinstrumentista. Tätä demonstroin työhöni liitettyllä videolla (liite 2). Videon sisältämien ääninäytteiden kautta jokainen tutkimustani arvioiva henkilö voi itse päätellä, kuinka hyvin esimerkiksi omaa ääntäni jäljittelevä digitaalinen instrumentti vastaa kontrolliryhmänä toimivaa omaa akustisesti tallennettua ääntäni.

1.2 Mihin digitaalista laulajaa tarvitaan

Mikä yleensä on digitaalisen lauluinstrumentin käyttötarkoitus? On hyvin käytännöllistä, että pienen pop-yhtyeen kosketinsoittaja voi soittaa vaikuttavan trumpettisoolon. Mutta onko käytännöllistä tarvetta sille, että sama kosketinsoittaja voi olla myös laulusolisti - "laulamalla" koskettimillaan? Useimpien mielestä tämä ei varmaankaan ole tarpeellista. Mutta kenties kuitenkin olisi kätevää, jos hän voisi soittaa yhtyeen laulusolistia säestävää "taustakuoroa".

Lienee kuitenkin selvää, että lauluäänen simuloiminen on hyödyllistä studiomusiikoille. Säveltäjä saattaa esimerkiksi haluta elokuvan soundtrackille mukaan kuoron, mutta elokuvan budjetti ei riitä elävän kuoron käyttämiseen. Tällöin on kätevää, jos uskottavan kuuluisen kuoron voi simuloida digitaalisesti. Tai laulusolisti voi haluta demonauhalle taustalaulajan, joka on esimerkiksi eri sukupuolta. Jos tämä laulaja on mahdollista saada nauhalle mukaan kotistudiossa, säästyy vaivaa ja rahaa.

Reaaliaikaisen lauluäänen tuottaminen digitaalisella soittimella on toki mahdollista siinä missä minkä hyvänsä muun instrumentin, jos unohdetaan lyriikka - tai minimoidaan se. Haluttu vokaali voidaan vaihtaa esim. *keyswitchillä* eli painamalla soivan alueen ulkopuolella olevaa kosketinta tai MIDI-kontrollerin painiketta (MIDI=*musical instruments digital interface* eli musiikki-instrumenttien digitaalinen käyttöliittymä). Jos konsonantitkin on äänitetty ja prosessoitu, keyswitcheille voidaan ohjelmoida myös kokonaisia lauseita tavu kerrallaan. Tällöin soittaja voi valita haluamansa lauseen ja sen jälkeen koskettimen (sävelkorkeuden) vaihtaminen johtaa seuraavaan tavuun. Eri asia on sitten se, kuinka aidolta tämä digitaalinen laulu kuulostaa.

Ihmisen jäljittely teknologian keinoin koetaan monesti jollain tavoin uhkaavana, jäljittelläänpä sitten ihmisen ulkonäköä, käyttäytymistä, tunteita – tai ääntä. Kirjassaan *Singing the body electric: The human voice and sound technology* Miriama Young to-

teaa, että teknologian hyödyntäminen voidaan nähdä yksinkertaisesti keinona laajentaa ihmisäänen luontaista kapasiteettia (Young 2015). Digitaalisesti äänitetty ja toistettu ihmisääni ei siis korvaa eikä uhkaa todellista ihmistä!

2 Sampling-teknologia ihmisäänen simuloimisessa

Mitä on sampling eli suomeksi sämpläys? Englannin sana 'sample' tarkoittaa yksinkertaisesti näytettä. Musiikin tapauksessa sämpläminen on näytteiden ottamista musiikista tai äänestä. Mark Katz määrittelee tuo asiaan kuitenkin pragmaattisemman ulottuvuuden täsmentämällä, että kyseessä on tietynlainen tietokonesynteesi, jossa ääni muutetaan ensin dataksi, joka samalla on se informaatio, jonka pohjalta samainen ääni rekonstruoidaan takaisin ääneksi (Katz 2010, 115).

Periaatteessa kaikki digitaalisesti tallennettu ja toistettu musiikki muodostuu digitaalisista sampleista. Digitaalisesta musiikintoistosta tuli arkipäivää 1980-luvun alussa, kun ensimmäiset CD-levyt julkaistiin (Salo 2020, 11). Sample-instrumentti on hieman eri asia. Daniel Duffel määrittelee sen yhden MIDI-kanavan ohjaamaksi samplejen koelmaksi (Duffel 2005, 8).

Kun lähdetään tekemään samplesta digitaalista instrumenttia, MIDI-järjestelmä tulee väistämättä vastaan. MIDI mahdollistaa syntetisaattoreiden, samplerien, tietokoneiden ja muiden laitteiden välisen kommunikoinnin (Salo 2020, 24). Jos esim. sämpläystä lauluäänestä rakennettua digitaalista instrumenttia halutaan soittaa koskettimistolla, käytössä täytyy olla ohjelmisto, jonka avulla koskettimistosta tietyltä MIDI-kanavalta tuleva signaali voidaan ohjata sampler-ohjelmaan, joka puolestaan poimii soitettua kosketinta vastaavan samplen digitaalisesta muistista. Juuri tällä periaatteella olen omaa MIDI-instrumenttianikin lähtenyt rakentamaan, käyttäen TX16Wx-nimisen ilmaisen samplerin maksullista Pro-versiota ja Reaper-nimistä DAW-ohjelmaa (*digital audio workstation* eli digitaalinen audiotyöasema - näitä esittelen videollani: youtu.be/G5kIBQ5j3KE).

Tarjolla on nykyään lukuisia erilaisia samplekirjastoja, jotka on tehty lauluäänestä. Nämä ovat yleensä saatavina joko itsenäisinä ohjelmina, joita varten on tehty erikseen graafinen käyttöliittymä, tai usein VST-liitännäisinä, joita voi käyttää mistä hyvänsä DAW-ohjelmasta käsin. Hyvin usein molemmat vaihtoehdot ovat tarjolla.

Kuoroja on samplattu enemmän kuin solisteja; löytyy sekakuoroja, mies- ja naiskuoroja, poikakuoroja ja erikoisuuksia kuten gregoriaanisia munkkikuoroja. Monet kuoroista tehdyt samplekirjastot sisältävät myös solisteja. Laulusolistin simuloiminen digitaalisesti on selvästi kuoroa vaikeampaa. Näkisin, että tämä johtuu erityisesti siitä, että kuorossa monet äänet sekoittuvat toisiinsa, eikä yksittäisen lauluäänen mahdollisia virheitä pysty joukosta erottamaan. Lisäksi kuoro kuulostaa eri sävelkorkeuksilta melko lailla samalta, kun taas yksittäisen laulajan äänessä voidaan erottaa hienoiset vivahte-erot äänen siirtyessä ylös tai alas päin.

Jos ääninäytteitä otetaan riittävästi, voidaan tuottaa eri instrumentit jokaisesta vokaalista kaikilla mahdollisilla laulutavoilla, äänenväreillä ja voimakkuuksilla. Sen sijaan, että kukin laulutyyli rakennetaan eri instrumentiksi, voidaan tehdä myös dynaaminen yksittäinen instrumentti, jossa soittovoimakkuutta tai erilaisia MIDI-kontrollereita käyttäen voidaan vaikuttaa siihen, mikä sample kulloinkin soi. Näin olen ainakin teoriassa mahdollista tallentaa yksittäisen solistin lauluääni hyvinkin kattavasti. Vielä kattavampi instrumentti saadaan, jos ääni tallennetaan samanaikaisesti useilla mikrofoneilla, jotka on sijoitettu eri etäisyyksille ja eri kulmiin.

Seuraavassa taulukossa olen esitellyt oman menetelmäni verrattuna joihinkin kaupallisiin tuotteisiin, joita olen vuosien varrella hankkinut, ja jotka edustavat alan kärkeä maailmalla. Esitän taulukossa myös arvion siihen, millaisiin ulottuvuuksiin menetelmäni on realistisesti laajennettavissa tulevaisuudessa.

Taulukko 1. Tekemäni samplekirjasto verrattuna kaupallisiin vastineisiin

VOCAL SAMPLE - kirjastot									
Tuote	Kuorot	Solistit	True legato -vokaalit	Sustain -vokaalit	mallinnettu legato -vokaalit	marcato/staccato samplet	dynaamisiet samplet	sanat/tavut sampleina	vapaa sanojen rakennus
Oma menetelmäni nyt	1	2	1	10	10	ei	ei	ei	kyllä
- potentiaalisesti	4+	4+	?	12+	12+	4+	kyllä	ehkä	kyllä
Symphonic choirs (Eastwest)	5	3	ei	5	5	ei	kyllä	ei	kyllä
Olympus symphonic choir (Soundiron)	2	8	5	20+	ei	kyllä	kyllä	kyllä	ei
Voices of Gaia (Soundiron)	ei	4	4-5	6	ei	ei	kyllä	kyllä	ei
Miroslav philharmonik (IK multimedia)	2	ei	ei	3	ei	ei	ei	ei	ei
Realivox blue (Realitone)	ei	1	8	8		ei	ei	ei	kyllä

3 Tietoperusta

Tutkimuksen teoreettisena viitekehyksenä toimivat digitaalista äänieditointia ja sampling-teknologiaa käsittelevä tietous ja lähdekirjallisuus, tutkimus fonetiikasta ja siihen liittyvät lähteet sekä oma kokemukseni kuoro- ja soololaulajana, laulunopiskelijana (olen suorittanut sivuinstrumenttitutkinnot sekä klassisessa että pop/jazz-laulussa) sekä studiomuusikkona.

Työssäni käyttämiäni lähteitä ovat mm. seuraavat näitä alueita käsittelevät teokset, jotka olen löytänyt JAMK:in kirjaston valikoimista ja käymällä läpi useita kirjallisuuden ja lehtiartikkeleiden tietokantoja.

Fonetiikka

Fonetiikan perusteet (Kalevi Wiik, WSOY 1999)

Tämä fonetiikan perusteos tarjoaa hyvän pohjan äänteiden (konsonanit, vokaalit, puolivokaalit) ominaisuuksien ymmärtämiseen.

Fonetiikan peruskäsitteitä (toim. Antti Iivonen & Reijo Aulanko, Yliopistopaino 2001)

Tässä kirjassa on laajasti määritelty fonetiikan, fonologian ja prosodii-
kan peruskäsitteiden ja -ilmiöitä. Oman työni kannalta merkityksellisiä
ovat nimenomaan fonetiikan ilmiöt, joita olen empiirisesti oman ääneni
kautta selvittänyt.

Ihmisäänen ja teknologian suhde

Singing the body electric : the human voice and sound technology (Miriam Young,
Farnham, Surrey, UK: Routledge. 2015)

Ihmisäänen digitaalinen prosessointi ja erityisesti lauluäänen sämplää-
minen ja yhdistäminen MIDI-teknologian kanssa ovat aiheita, joista on
niukasti teoksia saatavana. Lehtiartikkeleita toki löytyy. Tämä kirja käsit-
telee laajasti aihetta ja on siten hyvä lähde teos tätä tutkimusta varten.

Sampleteknologia

Making Music with Samples (Daniel Duffell, Backbeat Books, 2005)

Duffell käsittelee kattavasti sampling-teknologiaa, lähtien äänilähteiden
valinnasta ja käsitteiden selittämistä edeten sample-instrumenttien ra-
kentamiseen eri tyyppisillä laitteilla ja ohjelmilla. Myös erilaiset tehos-
teet ja äänenmuokkausmenetelmät käydään kirjassa läpi.

Teoksen pohjalta olen laatinut myös tähän teknologiaan liittyvän sanas-
ton (liite 1).

4 Fonetiikan rooli lauluäänen digitoimisessa

Ihmisen lauluääni on ainoa puhuva musiikki-instrumentti. Jos siis laulajan ääni halu-
taan digitoida lyriikoineen, kyse ei ole pelkästään musiikin vaan myös puhutun kielen
digitoimisesta. Digitaalisesti tuotettu puhe on useimmille tuttu ilmiö esim. automaat-

tisista puhelinvastaajista tai autojen navigaattoreista. Kun siirrytään puheesta lauluun, mukaan tulee melodian (sävelkorkeuden vaihtelu) lisäksi monenlaisia lauluteknisiä asioita.

Fonetiikkaa kiinnostaa toisaalta kielellisten muotojen ja rakenteiden konversio puheääneksi (puhujan suoritus) ja toisaalta puheäänien konversio kuulijan havainnoksi ja kielelliseksi tulkinnaksi (kuulijan suoritus). (Iivonen & Aulanko 2001, 2). Projektini puitteissa olen joutunut paljon selvittämään nimenomaan kuulijan suoritusta: millaiset tekijät määrittelevät sen, että toisiaan seuraavat konsonantit ja vokaalit nivoutuvat kuulohavainnoissa luontevasti ja uskottavasti toisiinsa.

Fonetiikkaa kiinnostavat myös esim. sukupuolten ja yksilöiden väliset äänteelliset erot (mts. 3). Kun lauluääntä lähdetään muuttamaan digitaalisesti instrumentiksi, kyse ei ole pelkästään instrumentin, vaan *yksilön* digitoimisesta. Kun kuuntelemme minkä hyvänsä muun soittimen ääntä, emme juurikaan erota soittimen yksilöllisiä piirteitä, niin että osaisimme tunnistaa, esimerkiksi millä nimenomaisella viululla jokin kappale on soitettu. Laulaja on kuitenkin aina yksilö, jonka äänen määrittelevät mm. ikä, sukupuoli sekä henkilökohtainen äänenväri ja äänenlaatu. Näiden pohjalta jokainen yksittäinen laulaja on tunnistettavissa.

Äänenväri eli teknisesti ilmaistuna *spektri* määräytyy ensisijaisesti laulajan iän, sukupuolen ja henkilökohtaisten ominaisuuksien perusteella, mutta on myös jonkin verran laulajan muunneltavissa. Äänen spektri tarkoittaa *ääneen sisältyvien osavärihenteiden kuvausta taajuuden ja amplitudin asteikoilla* (mts. 36).

Jotta yksittäisen laulajan erilaiset laulutavat saataisiin taltioitua hänen ääntään jäljittelevään digitaaliseen instrumenttiin, kaikki äänenlaadut pitäisi tallentaa erikseen (kaikki vokaalit ja soivat konsonantit kullakin laadulla). Iivonen ja Aulanko luettelevat lukuisia äänenlaatuja: käheä (breathy), karhea (coarse), henkäyssoinnillinen (breathy voiced; suomeksi usein "vuotoinen") ääni, falsetti (falsetto), kuiskaus (whisper), narina (creak, vocal fry), tiukka (tense voice) ja rento (lax voice) ääni (mts. 4). Asian tekee vielä haastavammaksi se, että ihminen instrumenttina ei ole kone, joka kullakin äänenlaadulla laulaessaan tuottaa kunkin äänteen aina samalla tavalla. Toisaalta siis

on vaikea saada äänitettyä kaikki ääniteitä tarkalleen samalla äänenlaadulla lauletuina, toisaalta digitaalisesti imitoitu laulu ei kuulosta täysin luonnolliselta, jos ääniteet toistuvat systemaattisesti täysin samanlaisina kunkin laulutekniikan sisällä – inhimillinen luovuus ja erehtyväisyys jäävät puuttumaan.

Fonetiikkaan kytkeytyy toinen läheistä sukua oleva tutkimuskenttä, prosodia: *prosodia ominaisuuksia ovat keston (ajoitukseen), painotukseen, sävelkulkuun ja -korkeuteen sekä rytmiin liittyvät ilmiöt* (mts. 4). Prosodian myötä puheeseen tulevat siis mukaan rytmi ja melodia; fonetiikka ja prosodia yhdessä tutkivat ja määrittelevät laulua. Tämän vuoksi lauluäänien analysointiin ja digitaaliseen reproduktioon liittyvä tutkimus onkin luoneelta prosodista.

Ihmisiä äänen digitoimisessa vokaalien ja konsonanttien äänittäminen ovat kaksi selkeästi toisistaan poikkeavaa osa-aluetta, mm. siksi, että vokaalit kytkeytyvät selkeästi sävelkorkeuteen ja siten musiikkiin, kun taas konsonantit enemmän laulun kielelliseen sisältöön. Iivonen ja Aulanko määrittelevät vokaalin seuraavasti: Äänne, joka tuotetaan ääntöväylän avaamalla (ääntiö). Perusvokaalien ominaisuuksia kuvataan kielen aseman suhteellisen korkeuden ja etisyysasteen sekä huuliaukon muodon avulla. (mts. 15). Konsonantti on äänne, joka muodostetaan ääntöväylän sulkeuman tai supistuman avulla (mts. 10).

5 Digitoinnin tekninen toteutus

Kattavan digitoinnin kannalta laulettuun äänen tekniset ominaisuudet voidaan luokitella seuraavasti:

1. Fonetiikka: vokaalit ja konsonantit ja siirtymät niiden kesken
Tämä osa-alue ei liity musiikkiin, vaan puheeseen; sanoitettu laulu on ”musiikkipuhetta”.
2. Sävelkorkeus: sustain (paikallaan pysyvä vokaali) ja legato / portamento (äänne korkeuden asteittainen muutos kahden sävelen välillä)

3. Dynamiikka: laulun eri voimakkuudet
4. Laulutyyli: esim. huokoinen, metallinen, nasaali jne.

Kunkin osa-alueen realistinen toteuttaminen digitaalisesti on mahdollista kolmella periaatteella:

1. Kukin variaatio (vokaali, konsonantti, sävel, voimakkuus, portamento, tyyli) sämplätään erikseen. Tämä vaikuttaa olevan yleisin tapa toteuttaa vokaalisia samplekirjastoja, perustuen hankkimieni kaupallisten kirjastojen analyysiin (esim. *Realivox blue* /Realitone ja *Voices of Gaia* /Soundiron).

Hyvää: Kaikki äänet ovat täysin autenttisia ja aidon kuuloisia.

Huonoa: Vaati valtavasti työtä, aikaa ja resursseja. Sampleja kertyy suunnaton määrä, jolloin samplekirjastosta tulee erittäin raskas. Vaikka yksittäiset samplet ovat autenttisia, ne voivat olla keskenään erilaisia, jolloin siirtyminen samplesta toiseen ei välttämättä ole luontevaa.

2. Samplet kerätään näytteinä tietyin välein ja puuttuvat samplet tuotetaan lennossa ohjelmallisesti.

Hyvää: Tarvittavien samplejen määrä on kohtuullinen. Samplekirjastosta tulee kevyt. Mahdollista päästä autenttisen kuuloiseen lopputulokseen.

Huonoa: Vaatii paljon tutkimustyötä, koodaamista ja reaaliaikaista äänen prosessointia, mikä vie paljon tietokoneen resursseja live-soittotilanteessa. Ei päästä välttämättä aivan yhtä hyvään autenttisuuteen kuin 1. menetelmällä.

3. Samplet kerätään näytteinä tietyin välein ja puuttuvat samplet mallinnetaan, pääasiassa käyttäen *pitch shift* (sävelkorkeuden muuttaminen digitaalisesti) -menetelmää. Pitch shift on mahdollista noin kvartti-intervallin puitteissa toteuttaa niin, että äänen spektri ("väri") sekä formantti (sukupuoli/ikä) säilyvät luontevasti.

Hyvää: Tarvittavien samplejen määrä on kohtuullinen. Kun mallinnusmenetelmät on luotu, niitä on helppo soveltaa uusiin sampleihin (esim. toinen laulaja). Mallinnuksen avulla voidaan vaikuttaa paljon samplejen keskinäiseen yhteensopivuuteen.

Huonoa: Mallintaminen vaati paljon työtä. Vaikka suurin osa sampleista tuotetaan mallintamalla, niitä kertyy paljon, jolloin samplekirjastosta tulee raskas.

Olen kokeillut jonkin verran kaikkia menetelmiä, mutta pääasiassa päätyneet menetelmään 3. Näin tein erityisesti siksi, että kyseinen lähestymistapa mahdollistaa uusien laulajien ja kuorojen ”istuttamisen” toteutettuun alustaan kaikkein vähimmällä vaivalla.

Kuitenkin esim. vibrato, reverb (akustiikan simulointi) ja dynamiikka on käytännöllistä tuottaa lennossa, käyttäen laadukkaita äänifilttereitä. Näin esimerkiksi soittovoimakkuus ei vaikuta pelkästään toistetun samplen volyyymiin vaan myös sen spektriin, vibraton määrään ja alukkeen terävyyteen / pehmeyteen.

Katsotaan vibratoa esimerkkinä reaaliaikaisesti tuotettavista efekteistä. Kun laulaja laulaa vibratoa tai tremoloa käyttäen, lauluäänien sävelkorkeus (vibrato) ja/tai voimakkuus (tremolo) vaihtelevat tietyllä välillä ja taajuudella. Vibrato ei yleensä ala heti, vaan sen syttymisellä on viive ja tietty nopeus (siirtyminen täysin tasaisesta äänestä maksimivibratoon). Jos tällainen laulutapa äänitetään, sample sisältää vain tietynlaisen vibraton, jossa kaikki nämä ominaisuudet ovat kiinteästi paikoillaan ja niitä on hyvin vaikea jälkikäteen varioida.

Alla oleva taulukko havainnollistaa asiaa:

Vibraton ominaisuudet

- Syttymisen viive: kuinka kauan kestää, ennen kuin vibrato alkaa kuulua
- Syttymisnopeus: siirtymäaika tasaisesta äänestä maksimivibratoon

- Äänenkorkeuden ja -voimakkuuden vaihtelun taajuus - keskimäärin 6 Hz
- Sattumisen viive: missä vaiheessa vibrato alkaa heikentyä
- Sattumisnopeus: siirtymäaika maksimivibratosta minimiin

(Sundberg 1994, 49-50.)

Äänen korkeuden ja voimakkuuden vaihtelut, samoin kuin niihin liittyvät nopeudet ja viiveet voidaan varsin hyvin lisätä staattiseen ääneen esiohjelmoiduilla tai reaaliaikaisilla efekteillä. Vibraton syvyyttä voidaan säätää esimerkiksi modulaatiopyörällä (modulation wheel), joka lähes kaikista MIDI-koskettimistoista löytyy ja nopeutta halutessa jollain muulla MIDI-säätimellä (<https://youtu.be/G5kIBQ5j3KE?t=476>).

Myöskään äänen dynaamisia variaatioita (esim. hiljaa, kohtuullisen voimakkaasti, tai voimakkaasti) ei välttämättä tarvitse äänittää erikseen. Sen sijaan voidaan analysoida mitä äänen spektrissä tapahtuu, kun laulaja muuttaa äänen voimakkuutta. Soittolanteessa äänen spektriä voidaan sitten muuttaa lennossa perustuen soittovoimakkuuteen.

Samalla tavoin voidaan muuttaa myös äänen väriä. Huomasin puolivahingossa, että käyttäessäni MIDI-koskettimiston pitch bend -pyörää samplen sävelkorkeuden nostamiseen tai laskemiseen, äänenväri samalla vaaleni tai tummui vastaavasti. Lasten ääni on yleensä vaaleampi kuin aikuisten, samoin tenorit ja sopraanot laulavat yleensä vaaleammalla äänellä kuin bassot ja altot (Schmidt 2019). Niinpä jos pitch bend on ohjelmoitu nostamaan / laskemaan ääntä esim. puolisävelaskelen verran, soittamalla vastaavasti tuon verran ylempää tai alemmää ja kääntämällä pitch bendiä vastakkaiseen suuntaan, saadaan näppärästi aikaan vaaleampi / tummempi äänenväri (<https://youtu.be/G5kIBQ5j3KE?t=390>).

Vuoden 2019 loppuun mennessä olin tuottanut kattavat samplekirjastot (sisältäen vokaalit ja konsonantit) omasta äänestä ja lisäksi naisäänestä (Ana Pilat, altto). Lisäksi olen tehnyt vokaali-instrumentit JAMK:in KUMUHAKU-harjoituskuoron naiskuorosta.

Alla oleva taulukko näyttää, mitkä lauluäänen ominaisuudet olen näissä tuottanut äänittämällä (samplaamalla), mitkä mallintamalla tai reaaliaikaisilla filtereillä.

Taulukko 2. Simuloitavien ominaisuuksien tuottaminen

Simuloitava ominaisuus	äänitetty	mallinnettu	filterit
Tasainen ääni (sustain); kukin vokaali erikseen	x		
Ääninäytteet terssin- kvartin välein	x		
Muut sävelkorkeudet pitch shifter -efektillä		x	
Sävelten välinen portamentoliuku (legato):		x	
Vibrato (sävelkorkeuden huojunta)			x
Tremolo (äänenvoimakkuuden huojunta)			x
Äänenväriin muutokset eri dynamiikoilla (Peak filter)			x
Konsonantit (äänitetty eri vokaalien kanssa)	x		
Siirtymisnopeus ja viive kons. ja vok. välillä			x

Työvaiheet

1. Aineiston kerääminen: näytteiden äänittäminen.
2. Äänitallenteiden pilkkominen ja käsittely.
3. Puuttuvien "välisamplejen" mallintaminen.
4. Samplejen kerääminen sampler-ohjelmaan ja asettelu koskettimistolle.
(<https://youtu.be/G5kIBQ5j3KE?t=208>)
5. Samplejen virittäminen ja muu editointi.
6. Ääninäytteiden soittaminen eri henkilöille / yleisöille ja palautteen kerääminen.
7. Materiaalin analyysi ja raportointi.

6 Tutkittavat osa-alueet ja tutkimuksen tulokset

Jotta sanoitettua lauluääntä jäljittelevä digitaalinen instrumentti kuulostaisi luonnolliselta ja aidolta, sen pitää toistaa kaikki lauluäänen eri osa-alueet mahdollisimman tarkasti. Kaikkein realistisin instrumentti saataisiin aikaan äänittämällä kaikki äänen ominaisuudet ja *mappaamalla* (engl. mapping) äänitetyt näytteet sampleri-ohjelmalla koskettimistolle. Tämä tarkoittaisi jokaisen vokaalin äänittämistä puolisävelaskelen välein useilla eri voimakkuuksilla, ilman vibratoa ja erityyppisten vibratoefektien kanssa, erilaisilla laulutekniikoilla (tumma ja vaalea, metallinen, huokoinen ym.) jne. Sekä äänittämisessä että ohjelmoinnissa olisi tällöin valtava työmaa ja lisäksi lopputuloksena syntyvässä samplekirjastossa olisi kymmeniä tuhansia sampleja, joiden toistaminen vaatisi tietokoneelta hyvin paljon resursseja.

Jotta turhalta työltä ja massiivisilta datamääriltä vältyttäisiin, on tutkittava, kuinka tarkasti kutkin äänen ominaisuudet on todella äänitettävä ja mitä ominaisuuksia voidaan tuottaa digitaalisesti mallintamalla tai simuloimalla. Esimerkiksi näytteenottotiheys (kuinka pitkin askelvälein eri korkeudella soivat vokaalit ja konsonantit täytyy äänittää) ja näytteiden pituus ovat ominaisuuksia, joita voidaan optimoida.

Jos kaikkia äänen ominaisuuksia ei äänitä, vaan osa tuotetaan mallintamalla, efekteilä tai eri sampleja yhdistämällä, on myös tutkittava mitä näissä simuloitavissa lauluäänen ominaisuuksissa teknisesti tapahtuu ja miten näitä ominaisuuksia voidaan jäljitellä.

Selvitettäviä osa-alueita

1. Tarvittavat intervallivälit vokaalien äänittämiseen
2. Tarvittava samplen kesto
3. Portamento-liu'un nopeus kahden sävelen välissä
4. Mitä tapahtuu äänelle portamentossa

5. Sustain- ja portamentosamplen saumaton yhdistäminen
6. Äänitapahtumat ja viiveet konsonanttien ja vokaalien välillä
7. Konsonanttien ja vokaalien ohjaus live-soitossa
8. Monenko vokaalin kanssa kukin konsonantti täytyy äänittää
9. Miten sampleri tietää soittaa konsonantista oikean version
10. Diftongien toteuttaminen
11. Äänen spektrin muutokset äänen voimakkuuden vaihdellessa

Käyn seuraavaksi läpi nämä kysymykset ja millaisiin vastauksiin ja käytännön ratkaisuihin olen päätenyt. Olen liittänyt mukaan linkkejä youtubessa olevaan listaamattomaan videoon, jossa esittelen tekemääni instrumenttia.

Tarvittavat intervallivälit vokaalien äänittämiseen

[Äänen] formantti on ääntöväylän siirtofunktiosta johtuva spektrin voimistunut alue, jolla yleensä yksi osasävel on voimakkain, mutta voimistuminen voi kohdistua useaan osasäveleeseen (Iivonen & Aulanko 2001, 14). Lauluäänen tapauksessa formantti määrittelee äänen karaktääriä; erityisesti sitä, kuulostaako laulu miehen, naisen vai lapsen ääneltä. Kun digitoitun ääninäytteen sävelkorkeutta muutetaan keino-tekoisesti jälkikäteen, myös formantin sävelkorkeus muuttuu. Tästä johtuu äänen nopeuttamisesta tuttu ilmiö, jossa sävelkorkeuden lisäksi äänen karaktääri muuttuu ”akuankkamaiseksi”.

Koska TX16Wx-sampleri ei mukauta äänen formanttia äänen korkeuteen, jo yhdellä kokosävelellä nostettu / laskettu sample kuulostaa luonnottomalta. Siksi sampleriin ääninäytteet täytyy tuoda kokosävelaskelen väleihin. Tämä ei kuitenkaan tarkoita, että kukin vokaali pitäisi äänittää kokosävelaskelin, sillä esim. Reaper-ohjelman mukana tuleva pitch shifter (sävelkorkeuden muutos) -efekti mahdollistaa myös formantin säilyttämisen sävelkorkeutta muutettaessa (<https://youtu.be/G5kIBQ5j3KE?t=235>).

Totesin, että ihanteellinen tallennusväli on suuri terssi. Sävelten laulaminen suuren terssin välein on kuitenkin haasteellista, koska tällöin syntyy ylinouseva kolmisointu, joka on vaikea laulaa puhtaasti. Siksi olen käyttänyt välillä askelvälinä suurta terssiä, välillä kvarttia. Koska pitch shifteriä käyttämällä ääninäytettä voi helposti nostaa tai laskea kokoaskelella, periaatteessa riittäisi, että vokaalit äänitetään kolmen kokoaskeleen välein (ylinouseva kvartti). Tämä ei kuitenkaan ole hyvä käytäntö, koska kolmen kokosävelaskeleen välillä äänenväri saattaa jo havaittavasti muuttua. Tasaisen siirtymän kannalta siis pienempi väli on parempi.

Tarvittava samplen kesto

Jotta ääni saataisiin soimaan jatkuvasti niin pitkään, kun kosketinta pidetään alhaalla, siihen täytyy asettaa luuppi (engl. *loop*, silmukka). Jos äänessä olisi täysin tasaisena soiva kohta, hyvinkin lyhyt luuppi riittäisi. Ihmisäänen ollessa kyseessä tällaista täysin tasaista kohtaa ei yleensä löydy. Tämän seurauksena luuppi voidaan kuulla pienenä aaltoiluna, vaikka luuppikohdat olisi asetettu hyvin tarkasti. Mitä pitempi luuppi, sitä aidommalta sample kuulostaa, mutta sitä pitempi ja samalla tiedostokooltaan suurempi sample tarvitaan. Jos samplen koko halutaan pitää mahdollisimman kohtuullisena, on ihanteellista sovittaa luupin taajuus samaan kokoluokkaan kuin luonnollinen lauluäänen vibrato. Tällöin samplen tarvitsee olla pituudeltaan vain sekunnin luokkaa.

Portamento-liu'un nopeus kahden sävelen välissä

Saadakseni selville, miten nopeasti lauluääni siirtyy sävelkorkeudelta toiselle, äänitin omaa lauluani eri intervalleja käyttäen – pyrkien laulamaan mahdollisimman luontevalla laulutyyllillä. Toin eri intervalleja sisältäviä ääninäytteitä Reaper-ohjelmaan ja poistin näytteestä tasaisena soivat alku- ja loppusävelet. Näin pystyin selvittämään, kuinka pitkä sävelten välinen liukuma oli kestoaltaan.

Portamentoliu'un kesto on toki riippuvainen siitä, kuinka legatossa (sitoen) lauletaan; siksi kokeilin muutamaa eri pituutta, kunnes päädyin mielestäni keskimääräisesti hyvän kuuloiseen nopeuteen, joka on karkeasti noin 0,2 sek.

Mitä tapahtuu äänelle portamentossa

Kun ääni liukuu pehmeästi sävelestä toiseen, pelkästään sen korkeus ei muutu. Oletin, että äänen formantin pitäisi muuttua samaan tahtiin kuin sävelkorkeuden. Totesin, että näin ei kuitenkaan ole. Saadakseni aikaan tasaisen sävelkorkeuden nousun ja laskun, asetin kaksi eri MIDI-parametria ohjaamaan äänen korkeutta ja formanttia. Ohjelmoin sitten nämä parametrit liikkumaan erilaisilla käyrillä ja kuuntelin lopputulosta. Jos asetin sävelkorkeuden ja formantin muuttumaan samaa tahtia, ääneen tuli luonnoton "kiekaisu". Lukuisten kokeilun jälkeen löysin sopivan viiveen formantin muutoksen alkamiselle suhteessa sävelkorkeuden muutokselle. Tämän taustalla lieinee jokin fysiologinen seikka, jonka ymmärtäminen ei tämän tutkimuksen puitteissa ole tarpeellista.

Korkeuden ja formantin lisäksi erityisesti suurien intervallien ollessa kyseessä myös äänen voimakkuudessa tapahtuu muutoksia; samoihin aikoihin kuin liuku tavoittaa päätepisteensä, äänessä kuuluu pieni pehmeähkö aksentti. Aksentin voimakkuus riippuu siitä, millä korkeusalueella liikutaan; täytyykö laulajan juuri tuossa kohtaa vaihtaa äänenmuodostustekniikkaa. Tämä tulee esille erityisesti, kun sävelväli on oktaavin kokoluokkaa. Mallinsin aksentit kuitenkin samanlaisiksi riippuen pelkästään intervallista, en itse absoluuttisesta sävelkorkeudesta

(<https://youtu.be/G5kIBQ5j3KE?t=1041>).

Aluksi ajattelin, että jokaiselle intervallille portamentoliuku pitää tehdä erikseen, jolloin niitä olisi tarvittu oktaavin sisällä (ylös ja alas) peräti 24. Mutta jälleen kokeiltuani erilaisia vaihtoehtoja, päädyin käyttämään huomattavasti pienempää liukujen lukumäärää kutakin säveltä kohti. Kun sitten ohjelmoin nämä äänitiedostot koskettimistolle, säädin kunkin alkamiskohdan aina sen mukaan, mikä intervalli on kyseessä (liukua ei soiteta alusta asti vaan "matkan varrelta").

Jokainen intervalli oli tästä huolimatta ohjelmitava erikseen puolisävelaskeleen välein. Kullekin vokaalille järjestelmässäni on yhteensä 15 samplea kokosävelaskelluksella (reilu kaksi oktaavia E - F#) ja kullekin 15 sävelelle vaaditaan periaatteessa 24 eri ohjelmaa, jotta saadaan katettua kaikki intervallit sekä ylös että alas päin. Todellinen

lukumäärä vähenee siksi, että matalilla sävelillä alas päin ei tarvita tietenkään koko oktaavia eikä korkeimmilla sävelillä ylös päin, koska intervallien pitää mahtua instrumentin kokonaisskaalaan. Toisaalta projektin myöhemmässä vaiheessa totesin, että oktaaviliukuja voi käyttää myös pienelle ja suurelle nooni-intervallille (oktaavi + puoliaskel ja oktaavi + kokoaskel, ylös ja alas), joten maksimissaan ohjelmia tarvitaan 24 sijasta 28.

Kaikkienensa tämä portamentoliukujen tuottaminen, sovittaminen koskettimistolle ja kaikenlainen hienosäätö oli kaikkein työläin vaihe koko projektia ja siihen kului valtavasti aikaa. Valmiissa legatot sisältävässä sample-instrumentissa on yhteensä 75 samplea (joista alle 10 on varsinaisesti äänitetty, muut tuotettu äänitettyjä sampleja modifioimalla) ja niiden ohjelmointi koskettimistolle vaatii yli 800 erillistä määritellyä. Instrumentin html+ koodissa on liki kaksi miljoonaa merkkiä ilman välilyöntejä. Tämä kaikki siis vaaditaan ainoastaan yhdelle vokaalille. Koodamisessa olen käyttänyt TX16Wx-samplerin lisäksi notepad++ ohjelmaa, ja samalla olen sivutuotteena oppinut hahmottamaan html+ ohjelmointikieltä. Luonnollisestikaan näin pitkiä koodeja ei kirjoiteta kokonaisuudessaan käsin, vaan perusrakenteiden koodaamisen jälkeen työskentely tapahtuu pitkälti copy-paste (kopioi-liitä)- ja find-replace (etsi-korvaa)-menetelmillä. Kun koko ohjelma on kertaalleen tehty, sen soveltaminen toiselle vokaalille on hyvin vaivatonta (<https://youtu.be/G5kIBQ5j3KE?t=849>).

Sustain ja portamentosamplen saumaton yhdistäminen

Aluksi tein portamentoliu'ut pitkän aikaa erillisinä, ilman sustain-osuutta. Idea oli, että sustain-sävelet tuotetaan eri sampleista kuin portamentot, jolloin samplet vaativat mahdollisimman vähän muistia – liu'ut ovat varsin lyhyitä kestoltaan. Jouduin kuitenkin kovasti kamppailemaan yrittäessäni välttää saumakohtien kuulumista. Loppujen lopuksi päädyin siihen, että on parempi sisällyttää sustain-osuus samaan äänitiedostoon kuin liuku. Tällöin ainoastaan edellisen sustain-sävelen ja uuden liu'un väliin jäi saumakohta, ja se oli paljon helpompi työstää kuulumattomaksi kuin loppupään sauma. Lisäetu oli se, että nyt kussakin neljässä portamentoäännessä sustain-osuus on

identtinen luuppikohtineen. Ainakin joissakin hankkimissani kaupallisissa sample instrumenteissa liukua seuraava sustain-sävel kuulostaa erilaiselta riippuen intervallista, mikä tuo soivaan lopputulokseen hallitsematonta epätasaisuutta.

Jäljelle jäävä saumakohta on riittävän kuulumaton, kun edellisen äänen lopussa on pieni häivytyks ja liukuosuuden alussa pieni nosto. Väliin jäävä lyhyt notkahdus äänen voimakkuudessa ei ole korvin havaittavissa.

Äänitapahtumat ja viiveet konsonanttien ja vokaalien välillä

Eri konsonantit ovat keskenään varsin erilaisia. Jotkin konsonantit, kuten [l] ja [m] ovat vokaalien tapaan *soinnillisia*; soinnillisuus on *äänihuulten värähtelyn aikaansaama äänteen foneettinen tai fonologinen piirre* (Iivonen & Aulanko 2001, 7). Toiset taas, kuten [k] ja [p] ovat *klusiileja*. Klusiili on äänne, joka syntyy ääntöväylän täydellisen sulkeuman seurauksena (mts. 10).

Klusiileilla on erilaisia *ääntövaiheita*: imploosio, okkluusio, eksploosio, soinnin alkamisaika (mts. 10). Imploosion aikana ääntöväylä sulkeutuu, okkluusion aikana se on täysin sulkeutunut ja syntynyt paine purkautuu eksploosion kautta ääntöväylän avautuessa. Juuri viime mainitun vaiheen aikaansaama ääni muodostaa konsonantin kuuluvan osuuden. Oleellista on kuitenkin juuri viimeinen ääntövaihe eli soinnin alkamisaika.

Tässä on huomioitava myös ilmiö, jota kutsutaan nimellä *aspiraatio*. Aspiraatio on suuväylän sulkeuman taakse syntyneen ilmanpaineen voimakkaan laukeaman aiheuttama *h*-mainen lisä-äänne klusiilissa (mts. 13). Tämä on hyvin vähäinen suomalaisissa klusiileissa (P, K, T), mutta selkeä näiden englanninkielisissä versioissa (vrt. esim. sanat "Pekka" ja "Peter" – jälkimmäisessä lyhyt *h* kuuluu P-äänteen jälkeen). Aspiraaation puuttuminen suomalaisista klusiileista tarkoittaa, että vokaali seuraa konsonanttia hieman nopeammin kuin aspiroivissa vieraskielisissä klusiileissa.

Siirtyminen esim. M:stä vokaaliin on täysin erilainen kuin P:stä. Soinnillista konsonanttia vokaali seuraa välittömästi, tosin pienen äänen voimakkuuden notkahduksen

saattelemana. Soinnittoman konsonantin jälkeen kuitenkin on pieni hiljainen tauko. Molemmissa tapauksissa konsonantin äänenvoimakkuus on huomattavasti pienempi kuin vokaalin. Vaikka miellämme T:n ja P:n terävinä ja vahvoina äänteinä, ilman vokaalia ne ovat vain kuiskauksia. Soinnillisten ja soinnittomien konsonanttien lisäksi on välimuotoja, kuten B ja D.

Frikatiivi eli hankausääänne on äänne, joka syntyy ääntöväylän supistumakohdalla il-mavirran aiheuttamasta hankaushälystä, esim. [f, x, s]. (mts. 11). Soinnittomat frikatiivit, kuten H ja S poikkeavat terävistä klusiileista T:stä ja P:stä siinä, että ne soivat pitempään. Saadakseni selville, miten kukin konsonantti toimii ja mitä tapahtuu konsonantin ja vokaalin välissä, minun piti analysoida monia erilaisia konsonantti-vokaalilyhdistelmiä tarkasti.

Vokaalin alkamisajankohta konsonantin alkamisen jälkeen vaihtelee merkittävästi (noin sekunnin kymmenesosan verran) konsonantista riippuen. Halusin, että samaa sample-instrumenttia voitaisiin käyttää ilman konsonantteja ja kaikkien konsonanttien kanssa; siksi ei ollut käytännöllistä lähteä koodaamaan vokaaleille vaihtelevia viiveitä, kullekin konsonantille sopivaksi. Paremmaksi keinoksi totesin viiveen asettamisen konsonantin alkamiselle. Vokaalille piti asettaa riittävä viive, jotta kestoltaan pisinkin konsonantti, [S], mahtuisi väliin. Tällöin S-konsonantti alkaa ilman viivettä heti kosketinta painettaessa ja esim. P-konsonantti alkaa selkeällä viiveellä. Jotta tästä johtuva vokaalia seuraava ylimääräinen viive ei kuuluisi liikaa, päätin käyttää melko reilua häivytystä (release) vokaaleissa (vokaali ei lakkaa soimasta välittömästi kosket-timen vapauttamisen jälkeen vaan häipyä vähitellen).

Konsonanttien ja vokaalien ohjaus live-soitossa

Kun vokaali- ja konsonanttisamplet ovat olemassa, täytyy olla myös keino, miten niitä ohjataan koskettimiston ja/tai MIDI-kontrollerin avulla. Itselläni on käytössä Yamaha MX-49, jossa on neljän oktaavin koskettimisto sekä Samson-merkkinen MIDI-kontrolleri, jossa niin ikään on neljä oktaavia ja lukuisia ohjelmoitavia painikkeita ja nuppeja.

Olen tehnyt erilaisia "koskettimistokarttoja", sekä yhdelle että kahdelle koskettimistolle. On kätevää, jos toinen koskettimisto on vain konsonanttien ja vokaalien valitsemista varten ja toinen soittamista varten. Kuitenkin järjestelmän täytyy olla käytettävissä myös yhdellä koskettimistolla. Kuvio 1 näyttää viimeisen rakentamani yhden koskettimiston asettelun, joka vaatii viisi oktaavia. Asettelu on kuitenkin käytettävissä neljän tai jopa kolmen oktaavin koskettimistolla, koska MIDI-koskettimiston oktaavialaa voidaan tilapäisesti nostaa ja laskea. Koskettimiston keltaiselta alueelta valitaan vokaali ja siniseltä alueelta vokaalia edeltävä konsonantti tai vokaalia seuraava toinen vokaali. Vihreällä merkityllä alueella vokaaleja ja konsonantteja soitetaan. (<https://youtu.be/G5kIBQ5j3KE?t=42>)



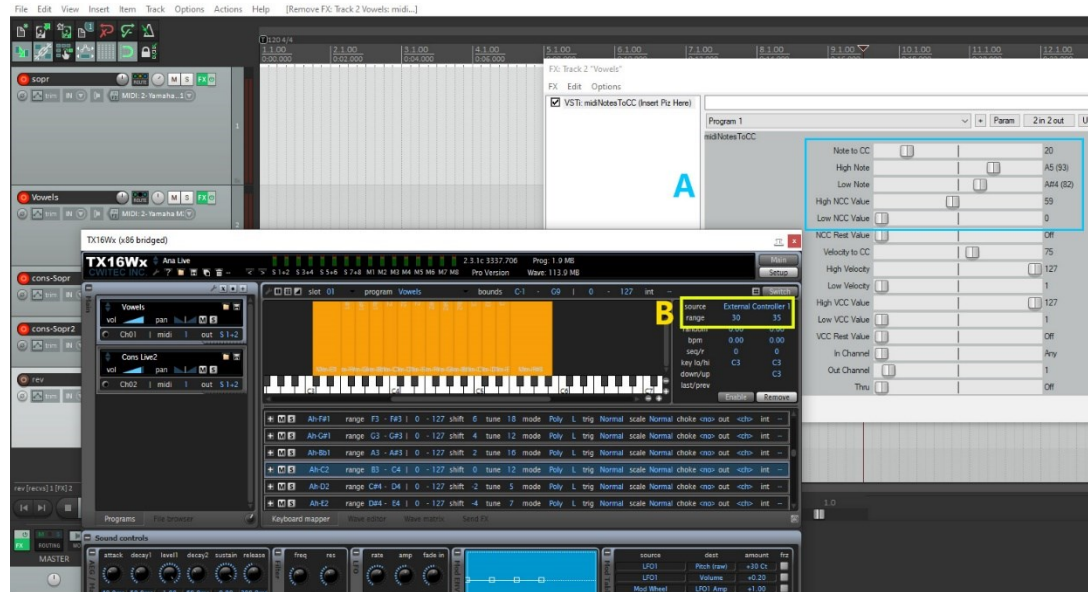
Kuvio 1. Yhden koskettimiston kartta

Tähän asetteluun olen valinnut tyypillisimmät konsonantit, joita lauletuissa tavuissa käytetään - ainakin Suomessa. Loput konsonantit olen sijoittanut seuraavaan alempaan oktaaviin, joten nekin voidaan hakea transponoimalla koskettimistoa alas päin. Sinisellä merkityt äänteet 'm' ja 'i' (vasemmalla) soivat vokaalin jälkeen. Punaisella X:llä merkitty C-sävel sammuttaa kaikki konsonantit, siten että pelkkä vokaali soi.

TX16Wx-samplerin maksullisessa versiossa on oma osionsa "switchejä" varten: Täällä voidaan määritellä, miten MIDI-viesteinä tulevat arvot ohjaavat sitä, mitkä samplet soivat milloinkin. Tätä menetelmää käyttäen olen määritellyt, mitkä koskettimet aktivoivat kunkin vokaalin vastaavat samplet (<https://youtu.be/G5kIBQ5j3KE?t=198>).

Olen käyttänyt ilmaisia VST-liitännäisiä, joilla MIDI-signaaleja voidaan kontrolloida ja ohjata monin eri tavoin. Esimerkiksi vokaalien vaihtaminen koskettimiston kautta toimii, siten, että koskettimen antama signaali käännetään toiseksi MIDI-

kontrolliviestiksi (continuous controller eli lyhyesti CC), joka on ensin sampleriohjelmasta käsin määriteltä. Esim. Vowels-niminen sampleriohjelma sisältää kaikki vokaalit yhdessä paketissa, mutta kukin vokaali soi vain silloin kun mainitun MIDI-viestin arvo on tietyn suuruinen.



Kuvio 2. CC-viestit ohjaamassa vokaalin valintaa

Kuviossa 2 näkyy "Vowels"-nimisellä raidalla oleva VST-pluginin "midiNotesToCC" (A), joka ottaa vastaan signaalit koskettimista A#4-A5 ja muuttaa ne CC-viestiksi nro 20 arvoilla 0 - 59. Kun esimerkiksi painetaan kosketinta E5, arvoksi tulee 30. Samplerissa on mahdollista ohjata haluttuja CC-viestejä "ulkoisiksi kontrollereiksi" (external controller), ja controllerille 1 on valittu CC nro 20. Kuvassa näkyy, että valittuna on ohjelma "Ah-C2" (sininen palkki) eli A-vokaalisample koskettimella C2, joka soi vain silloin kun External Controller 1:n arvo on välillä 30-35.

Samoin konsonantteja ohjaa CC-viesti, tässä tapauksessa *channel pressure*, joka on vapaana käytettäväksi.

Kuvion 1 mukaisella koskettimistokartalla on siis mahdollista hyvin nopeasti valita konsonantti vokaalin eteen ja myös perään joko 'm'-äänne tai vokaali 'i'. Näin voidaan siis soittaa hyvin monenlaisia erilaisia tavuja, kuten tyypilliset "duu duu", "laa laa", "pom pom", "jee jee" mutta myös hieman erikoisempia kuten "roi roi", "tam tam" tai vaikkapa "mei mei" (<https://youtu.be/G5kIBQ5j3KE?t=520>).

Konsonantin tai toisen vokaalin ohjaaminen soitettavaksi päävokaalin *jälkeen* on mahdollista valitsemalla samplerin ohjelman "trig"-arvoksi normaalin sijasta *Release* (vapautus), mikä tarkoittaa, että kyseinen sample soitetaan, kun määritelty kosketin päästetään vapaaksi. Niin pitkään, kun kosketinta pidetään pohjassa, ensimmäinen vokaali soi jatkuvasti. Kun kosketin vapautetaan, toinen konsonantti tai vokaali soite- taan niin pitkänä kuin se on ohjelmoitu. Samalla periaatteella voitaisiin luoda kosket- timistokartta, joka tuottaa toiseksi vokaaliksi minkä hyvänsä toisen vokaalin tai kon- sonantin.

Tämä menetelmä ei ole riittävä, silloin kun konsonantteja on useita peräkkäin, kuten konsonanttiyhtymässä, joka on samassa tavussa esiintyvä konsonanttijono; suo- messa esim. /kl/ klapi, /pr/ prutku (Iivonen & Aulanko 2001, 14). Konsonanttiyhty- mien ohjaus live-soitossa on alue, jota en toistaiseksi ole ratkaissut, mutta niiden tuottamiseen MIDI-tiedostosta käsin olen kokeillut erilaisia tapoja.

Monenko vokaalin kanssa kukin konsonantti täytyy äänittää

Voisi helposti ajatella, että soinnittomat vokaalit kuten P, T ja K ovat aina samanlaisia riippumatta siitä, mikä vokaali niitä seuraa ja miltä sävelkorkeudelta. Nämä äänneet- hän tuotetaan täysin ilman äänihuulia. Tosiasia on kuitenkin se, että *vokaali muodos- tuu suuontelossa*, ei äänihuulissa. Juuri siksi meidän on mahdollista puhua kuiskaa- malla, käyttämättä äänihuulia ja soivia vokaaleja.

Vaikka soinnittoman konsonantin ja vokaalin välinen viive on kestoaltaan vain sekun- nin kahdeskymmenesosan luokkaa, sen aikana kuuluu kuiskauksenomaisena se vo- kaali, joka pian syttyy kuuluvana värähtelemään. Tässä on kyseessä ns. *koartikulaa- tio*: Puhe-elinten toimintatavasta aiheutuva, puhetta helpottava ilmiö, jossa saman- kaltaistumista tapahtuu etenkin peräkkäin esiintyvien äänneiden kesken (Iivonen & Aulanko 2001, 8).

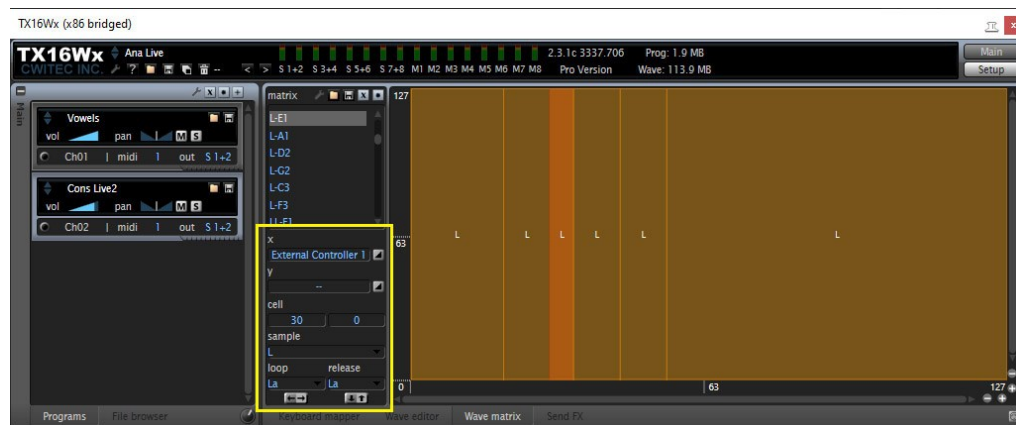
Jos tämä lyhyt kuiskattu vokaali ei täsmää soivan vokaalin kanssa, ihmiskorva ei kuule soivaa tavua luonnollisena (joskaan ei erota, mikä tarkkaan ottaen menee pieleen). Myös soivissa konsonanteissa, erityisesti L-äänneessä, itse konsonantin väri muuttuu

hieman sen mukaan, minkä vokaalin ääntämiseen suu on muotoutunut. Tämän vuoksi kukin konsonantti täytyy äänittää usean eri vokaalin kanssa yhdessä.

Kuinka monen vokaalin kanssa kukin konsonantti pitää äänittää? Totesin, että tämä vaihtelee. A- ja O-vokaalit ovat niin lähellä toisiaan, että niitä ei aina tarvitse huomioida erikseen. Monien konsonanttien kohdalla vain jompikumpi vokaaleista I tai E riittää. Kaikkein vaativimmiksi tässä suhteessa osoittautuivat L ja amerikkalainen R, joita seuraavan vokaalin aiheuttama koartikulaatio vaikuttaa erityisen voimakkaasti. Tämä johtunee siitä, että suuontelo mukautuu seuraavan vokaalin muotoon jo konsonantin alkaessa, ja vokaalin ääntöväylä on vain osittain suljettu (L:n tapauksessa keskikohta, amerikkalaisen R:n kohdalla päinvastoin reunat).

Miten sampleri tietää soittaa konsonantista oikean version

Koartikulaatiosta johtuen samplerin pitää osata toistaa oikea version konsonantista, siten että se täsmää konsonanttia seuraavan vokaalin. TX16Wx-samplerin professionaalisessa versiossa kullekin koskettimille voidaan ohjelmoida *matriisi*. Kuviossa 3 näkyy L-konsonantin matriisi. Koska L on soinnillinen konsonantti, sille täytyy vokaalien tavoin olla sampleja äänitettynä eri sävelkorkeuksilta. Olen valinnut näytteiden väliksi kvartin (matalin sample on E1, seuraava A1, sitten D2 jne). Kuvan tilanteessa valittuna on matriisi ohjelmalle "L-E1" ja sen kolmas lohko.



Kuvio 3. Konsonanttiversioiden valintaa ohjaava matriisi

Jos katsomme arvoja keltaisen laatikon sisällä, näemme että matriisin X-akselilla olevia parametreja ohjaa External controller 1. Kun arvo on välillä 30-35 (arvo 35 on seuraavan lohkon aloitusarvo), soitetaan samplesta 'L' viipale (slice) nimeltään 'La' (tällöin myös seuraavaksi soiva vokaali on A). Kuviossa 4 näkyy kyseinen sample, joka on jaettu lukuisiin viipaleisiin (engl. *slice*). Siinä on konsonantti L kuudelta eri sävelkorkeudelta kuuden eri vokaalin kanssa, siis yhteensä 36 versiota tästä yhdestä konsonantista. Kukin viipale on kestoaltaan hyvin lyhyt, eli tällaisenaankaan sample ei ole kuin vajaan 5 sekunnin mittainen.



Kuvio 4. Konsonanttisamplen eri vokaaleita vastaavat versiot

Sen sijaan, että olisin tallentanut 36 erillistä äänitiedostoa (mikä toki olisi myös mahdollista), olen yhdistänyt kaikki 36 variaatiota samasta konsonantista yhteen tiedostoon. Samoin olen menetellyt muiden konsonanttien kanssa, joskaan useimmille konsonanteille aivan näin monia versioita ei tarvita.

Diftongien toteuttaminen

Diftongi eli pariääntiö on kahdesta kvalitatiivisesti erilaisesta, samaan tavuun kuuluvasta vokaalielementistä koostuva äänneskevenssi, esim. tuo, työ, tie. Vokaaliyhtymä on vokaalijono, johon sisältyy tavunraja, esim. suomen /ta.e/ (17). (Iivonen & Aulanko 2001, 17.) Näissä vokaali muuntuu toiseksi pehmeän siirtymän kautta, ja tätä siirtymää ei sisälly erikseen digitoituihin yksittäisiin vokaaleihin.

Aluksi kokeilin, saisiko kahden vokaalin yhdistelmän kuulostamaan luontevalta siirtymällä yhdestä vokaalisamplesta toiseen ristihäivytyksen kautta. Totesin, että tämä ei toiminut tyydyttävästi. Aloin siis äänittämään vokaaliketjuja siten, että kaikki tarvittavat diftongit ovat näytteessä mukana.

Erottelin sitten näistä ketjuista kohdat, joissa vokaalien välinen siirtymä (morphing) tapahtuu ja tallensin nämä eri tiedostoiksi. Sen jälkeen piti kehitellä menetelmä, jolla diftongit saatiin valittua koskettimistolta käsin. Törmäsin tässä samaan ongelmaan kuin portamentoliukujen kanssa; saumakohtia oli vaikea saada kuulostamaan ehyiltä. Päädyin kokeilemaan ratkaisua, jossa diftongit ovat saatavilla sekä erillisinä siirtyminä että yhdessä siirtymän jälkeen soivan sustain-vokaalin kanssa. Tällä ratkaisulla sain melko hyvää jälkeä aikaiseksi, mutta kaikkien äänitiedostojen tuottaminen ja ohjaaminen oli erittäin työlästä.

Kun olin äänittänyt Ana Pilatin äänen niin ikään yleisimmät diftongit sisältäen, päädyin kuitenkin vielä tutkimaan mahdollisuutta pärjätä ilman erillisiä diftongisampleja. Kun modifioin kahden vokaalin välistä ristihäivytystä sopivan mittaiseksi, sain diftongin kuulostamaan kohtalaisen luontevalta. Päätin tehdä kompromissin ja luopua erittäin hankalasta menetelmästä, vaikka soiva lopputulos ei välttämättä ollut aivan yhtä luontevan kuuloinen.

Äänen spektrin muutokset äänen voimakkuuden vaihdellessa

Kun ihminen laulaa eri voimakkuuksilla, todellisuudessa pelkästään äänen volyyymi ei muutu. Hiljaa tai kovaa laulaminen vaatii erilaista laulutekniikkaa, jolloin myös äänen spektrissä tapahtuu selkeitä muutoksia. Jotta ihmisääntä jäljittelevä virtuaali-instrumentti mahdollistaisi myös laulun eri dynamiikat, tämä äänen spektrin muuttuminen pitää huomioida.

Tämä voidaan tehdä joko niin, että kukin vokaali äänitetään eri voimakkuuksilla laulettuna tai niin, että eri dynamiikat mallinnetaan keskivoimakkuudella laulettuna näytteestä. Tämä modifioitu versio voidaan joko tuottaa reaaliaikaisesti efektejä

käyttäen tai efektin asettamisen jälkeen tallentaa erillisenä äänitiedostona ja tuoda sampleriohjelmaan.

Käytin H-EQ-nimistä kaupallista ekvalisaattori-pluginia, joka näyttää reaaliaikaisen graafisen käyrän äänen spektristä sen soidessa. Äänitin omaa lauluani samalla vokaalilla hiljaisena ja voimakkaana ja säädin sitten nämä soimaan samalla volyyymilla. Sitteen vertailin äänen spektrin eroa näissä näytteissä. Totesin, että suurin ero voimakkaassa äänessä hiljaiseen verrattuna oli tietyn taajuusalueen korostuminen. Lisäksi toisella alueella oli pientä vaimentumista.

Kun olin tämän havainnut, totesin että käyttämässäni sampleriohjelmassa (kuten oletettavasti useimmissa muissakin) on suodin nimeltään *peak*, joka mahdollista tietyn taajuusalueen nostamisen tai laskemisen. Myös korostetun taajuusalueen laajuutta voidaan muuttaa. Asetin suotimen arvoksi löytämäni vahvistettavan taajuuden ja sen kontrolleriksi koskettimen alaspainonopeuden (*key velocity*). Käytännössä tämä tarkoittaa, että soitettaessa kovempaa kyseinen taajuusalue korostuu ja näin ollen jäljittelee tehokkaasti sitä, mitä tapahtuu äänessä luonnollisesti, kun sitä vahvistetaan! Mikä parasta, soitettaessa hiljaa, sama taajuusalue heikkenee alkuperäiseen sampleen nähden, jolloin yhdellä samplella saadaan tuotettua sekä hiljaiset, keskivoimakkaat että voimakkaat vokaaliäänet tyydyttävän realistisesti.

7 Tutkimuksen johtopäätökset

Olen käyttänyt satoja työtunteja saadakseni aikaan luonnolliselta kuulostavan digitaalisen *lyyrisen* (sanoilla laulavan) lauluinstrumentin. Työ on koostunut äänittämisestä ja äänitystilanteiden järjestämisestä, äänen käsittelystä, ääninäytteiden analysoinnista ja tutkimuksesta, lähdemateriaalin opiskelusta, sampleriohjelman käytöstä ja sen tekemien tiedostojen editoimisesta, muistiinpanojen ja suunnitelmien tekemisestä sekä joistakin muista satunnaisemmista elementeistä.

Kaiken tämän työn jälkeen voin todeta, että täysin autenttiselta kuulostavan digitaalisen laulavan instrumentin rakentaminen kotikonstein on valtavan työn takana ja käytännössä melko lailla mahdotonta. Tarvitseeko tällaiseen täydelliseen luonnollisuuden kuitenkaan pyrkiäkään? Tarkoitushan ei alunalkaenkaan ole korvata elävää laulajaa tai kuoroa digitaalisella vastineella, eikä laittaa sitä esiintymään konserttisaleissa heidän sijastaan. Sen sijaan tarkoitus on luoda riittävän tarkasti äänitettyä laulajaa tai kuoroa muistuttava instrumentti, jota voidaan käyttää työkaluna musiikin opettamisessa, säveltämisessä ja sovittamisen sekä studiomusiikissa taustaäänien tuottamisessa tilanteissa, joissa laulajien käyttäminen ei ole käytettävissä olevien resurssien puitteissa mahdollista.

Mitä tulee tuohon autenttisuuteen, olen havainnut, että suurin haaste liittyy vastoin odotuksiani vokaalien yhtenäisyyteen. Konsonanttien ja vokaalien yhdistäminen luontevasti on erittäin vaativa työmaa, mutta vaikka urakka on suuri, se on mahdollista toteuttaa. Tyypillisissä sessioissa käytin paljon aikaa yksittäisten konsonantti-vokaaliparien työstämiseen ja loppujen lopuksi yleensä pääsin itseäni tyydyttävään tulokseen. Pettymyksekseni huomasin, että digitaalinen laulajani ei silti kuulostanut aidolta. Ongelmana oli se, että vokaalien vaihdellessa nopeasti, niiden pienet sävy- ja voimakkuuserot tulevat esille, ja saavat laulun kuulostamaan luonnottomalta.

Ymmärsin tämän johtuvan siitä, että kukin vokaali on äänitetty ja jatkokäsitelty erikseen, jolloin ei ole mitään takeita siitä, että vokaalit keskenään kuulostavat yhteensoivalta. Tätä ei auta edes se, että Ana Pilatia äänittäessäni käytin äänityksen suunnittelemaani laulua, joka sisältää kaikki haluamani vokaalit ja vokaaliyhdistelmät; laulun eri kohdissa vokaaleissa kuitenkin esiintyy sattumanvaraisia variaatioita. Yksittäinen tavu saadaan kuulostamaan hyvin aidolta, mutta peräkkäisissä tavuissa ja diftongeissa haaste tulee vastaan.

Yllä oleva haaste koskee kuitenkin pääasiassa soololaulajan äänen digitointia, ja sama asia on tullut esille, kun olen kokeillut hankkimiani kaupallisia soololaulajien äänestä rakennettuja sample-instrumentteja – vaikka ensin en huomannutkaan, että niissä kuuluva hienoinen luonnottomuus johtuu pääasiassa tästä tekijästä.

Jos puhutaan pelkästään yhdellä vokaalilla kerrallaan ”laulavasta” instrumentista (mukaan lukien tavut, joissa on mukana konsonantteja), on sekä oman tutkimukseni pohjalta että vertailussa käyttämieni kaupallisten instrumenttien perusteella mielestäni selvää, että hyvin autenttiseen ja käyttökelpoiseen tulokseen voidaan päästä, verrattain pienellä työmäärällä. Koska tällaisella instrumentilla voidaan digitaalisesti toteuttaa ”Aa-kuoroja” (ja muita tavuja tai tavuyhdistelmiä, kuten esim. *dum, laa, badam, shalalaa* jne.), se on jo käytettävissä moniin tarkoituksiin niin pedagogisissa kuin muissakin sovelluksissa.

8 Pohdinta

Tutkimuksen taustalla on oma henkilökohtainen tavoitteeni luoda mahdollisimman laadukas laulajaa jäljittelevä MIDI-soitin; projekti, joka tulen viemään eteenpäin tämän tutkimuksen jälkeenkin. Vaikka täydellisen autenttisesti soivan valmiin instrumentin laatiminen ei ollut opinnäytetyön tavoite, olen sen myötä päässyt tässä henkilökohtaisessa tavoitteessani valtavasti eteenpäin. Samalla olen hankkinut paljon arvokasta tietoa ihmisäänien ja sampling-teknologian eri aspekteista, sekä konkreettisen tutkimustyön että lähdekirjallisuuden avulla.

Lisäksi tavoitteeni projektin osalta on hieman täsmentynyt: Sen sijaan, että haluan luoda täydellisesti aidon kuuloisen digitaalisen laulajan (kun tiedän, että omat resursit eivät siihen riitä), pyrin luomaan sellaisen alustan, jonka pohjalta kuka hyvänsä laulaja tai mikä hyvänsä kuoro voidaan verrattain helposti ja käyttökelpoisella laadulla saattaa MIDI-instrumentiksi.

8.1 Projektin pedagogiset mahdollisuudet

Miten lauluääntä simuloiva digitaalinen instrumentti hyödyttää musiikinopettajaa?
Mitä hyötyä siitä on musiikin opiskelijalle?

Selvästikin kyseinen instrumentti, tai pikemmin *työkalu*, kytkeytyy vahvimmin kahden perinteiseen instrumenttiin, jotka kumpikin ovat suosituimpien instrumenttien joukossa opiskelijoiden keskuudessa; laulu ja piano.

Tiettyä instrumenttia pääaineenaan opiskeleva henkilö hyötyy luonnollisestikin eniten oman instrumenttinsa soittamisesta ja sen ominaisuuksien tuntemuksesta. Mutta lienee melko selkeää, että minkä hyvänsä instrumentin soittaja hyötyy pianon ja laulun opiskelusta sivuaineena. Piano (ja kosketinsoittimet yleensä) poikkeaa useimmista muista instrumenteista siinä mielessä, että sillä voidaan tuottaa itsenäisesti kaikki musiikillisen ilmaisun elementit: melodia, rytmi, harmonia. Näiden elementtien tuntemus taas on jokaiselle muusikolle vähintäänkin hyödyllistä, jos ei suorastaan välttämätöntä!

Lauluääni toisaalta on instrumentti, jonka me kaikki omistamme – lukuun ottamatta yksiöitä, joiden äänielimistö on jollain tavoin vioittunut tai vajavainen. Omalla *sisäisellä* instrumentillamme voimme tuottaa samoja melodioita, joita opettelemme soittamaan *ulkoisella* työkalulla. Lauluääni resonoi omassa kehossamme, jolloin emme ainoastaan kuule näitä melodioita vaan myös tunnemme ne itsessämme syvämminkin.

Lauluäänemme on kuitenkin pianoon verrattuna sikäli puutteellinen, että emme voi yksin tuottaa harmonioita. Harmoniseen lauluun tarvitaan useita laulajia, mutta puhaiden harmonioiden laulaminen on toisaalta musiikillisesti vaativa taito, jonka opiskelemiseen ei kaikilla ole aikaa tai motivaatiota. Niinpä olemmekin löytäneet yhden käytännöllisen hyödyn; *digitaalinen lauluinstrumentti mahdollistaa harmonisen laulun ilman muiden kuorolaulutaitoisten ihmisten läsnäoloa*. Opettaja tai oppilas itse voi säestää laulamaansa melodiaa koskettimistolla, vieläpä mieleisellään äänellä.

Laulunopetuksessa tällainen työkalu löytää myös toisentyyppisen tarkoituksen; sillä voidaan havainnollistaa erilaisia laulutekniikoita ja sävyjä, joita opettaja ei voi omalla äänellään tuottaa. Tähän päästään toki myös käyttämällä äänitteitä, mutta tällöin ei aina voida löytää juuri tiettyä kappaletta laulettuna tietyltä korkeudelta. Jos esimerkiksi naispuolinen laulunopettaja opettaa bassolaulajaa, hän voi demonstroida laulua

myös bassoäänellä. Mikäli käytössä oleva samplekirjasto on riittävän laaja ja monipuolinen, hän voi käytettävän vokaalin lisäksi valita dynamiikan ja laulutyylin (esim. klassinen/pop, vibratolla tai ilman jne.) ja soittaa haluamansa kappaleen. Ja jos mukana ovat konsonantit ja diftongit (kahden vokaalin yhdistelmät), myös lyriikka voidaan ottaa mukaan.

Ehkä kaikkein selkeimmin tällainen työkalu osoittaa voimansa kuorojen ja lauluyhtyeiden käytössä. Sen avulla harjoitettava kappale voidaan tuottaa äänitteeksi, jossa kaikki stemmat soivat puhtaasti. Tyypillisesti stemmojen harjoitteluun valmistetaan MIDI- tai MP3-tiedostoja, joissa melodia soi pianon tai muun instrumentin äänellä. On aivan eri asia, jos harjoitettava stemma - ja kaikki muut stemmat - soivat kuoron äänellä, jolloin opiskelija voi harjoitella omaa stemmaansa muuta kuoroharmoniaa vasten - ja täten saa hyvinkin realistisen kokemuksen siitä, miten oma stemma sopii yhteen muun kuoron kanssa.

Jotta saataisiin selkeämpi kuva tällaisten instrumenttien hyödyllisyydestä musiikin opetuksessa, niitä pitäisi esitellä alan ammattilaisille ja kokeilla niiden käyttöä opetustilanteissa. Tällöin voitaisiin kerätä kokemukseräistä palautetta sekä opettajilta että oppilailta ja näin saada arvokkaita vinkkejä projektin kehittämiseen nimenomaan pedagogisia näkökulmia silmällä pitäen. Kuoronjohdon opiskelijoilta saatu palaute oli rohkaisevaa, ja heidän esittämänsä parannusehdotukset olivat enimmäkseen toteutettavissa. Pienellä viilauksella tekemäni instrumentit ovat jatkossa käytävissä tämän tyyppiseen kokeilukäyttöön, ja ainakin omassa työssäni musiikkipedagogina voin sitä toteuttaa.

8.2 Mahdollisuudet säveltäjille ja sovittajille

Nykyaikainen tietokoneteknologia tarjoaa huikeat mahdollisuudet musiikkikappaleiden orkestrointiin. Säveltäjän tai sovittajan ei enää tarvitse kuvitella mielessään, miltä kappale kuulostaa tietyillä instrumenteille soitettuna, koska hän voi työskennellä digitaalisesti simuloitujen (yleensä sämplättyjen) instrumenttien kanssa alusta asti. Hän voi kokeilla erilaisia soolosoittimia, erilaisia jousi- ja puhallinkombinaatioita jne.

Lauluääni on sovittajan kannalta yksi instrumentti muiden joukossa. Ideaalisessa tilanteessa myös laulajat voidaan simuloida, jotta kuullaan miten heidän osuutensa sopii yhteen muiden instrumenttien kanssa – tai *a cappella* -kuoron tapauksessa miten eri laulustemat soivat yhteen. Useimmiten kuitenkin on saatavilla vain digitaalinen "Aa-kuoro", joka laulaa vain yhtä vokaalia, yhdellä dynamiikalla.

Useimpien instrumenttien digitalisaation kanssa ei tarvitse miettiä soittajan roolia; kitaralla tai viululla soitettu staattinen ääni kuulostaa jokseenkin samalta riippumatta soittajasta. Ihmisäänen kohdalla näin ei ole. Meillä on kyky erottaa hyvin monen tyyppisiä lauluääniä, niinpä tietty vokaali ei kuulosta samalta eri ihmisten laulamana. Äänen ominaisuuksiin vaikuttavat iän ja sukupuolen lisäksi laulutyyli (klassinen, pop, kansanlaulu jne.) sekä laulajalle ominainen äänenväri.

Sovittajan pitäisi kyetä valitsemaan vähintäänkin haluamansa vokaali joko mies- tai naispuolisen solistin tai kuoron laulamana. Ja jos kappale sävelletään tai sovitetään tietyille laulajille tai tietylle kuorolle, olisi ihanteellista, jos sitä voitaisiin testata juuri kyseisten laulajien äänellä.

Tutkimukseni kautta kehitetty menetelmä poikkeaa kaupallisista vokaalisista samplekirjastoista sikäli, että sen avulla kukin kuoro (tai yksittäinen laulaja) voidaan kohtuullisella vaivalla saattaa digitaalisesti instrumentiksi. Riittää, että kaikki vokaalit ja soivat konsonantit äänitetään erikseen miehiltä ja naisilta noin viideltä eri korkeudelta, mielellään ainakin kahdella dynamiikalla (piano ja mezzoforte). Äänittämiseen riittää muutama tunti, mielellään akustisesti kuivassa, kaiuttomassa tilassa. Tarvittava äänieditointi voidaan helposti tehdä millä hyvänsä äänieditointisovelluksella. Sävelten välille tarvittavat portamento-liu'ut tuotetaan lähestulkoon automaattisesti kehittämälläni tekniikalla. Nämä tallennetaan sitten äänitiedostoina tiettyyn hakemistoon, jossa jo valmiiksi sijaitsevat niitä käyttävät sampleritiedostot.

Ainoa ominaisuus, joka tässä kohtaa ei toimi täysin automaattisesti, on vokaalien "luupaus" (looping), eli rajatun mittaisen ääninäytteen venyttäminen jatkuvaksi. Kuitenkin myöhemmin yhteistyössä osaavan ohjelmoijan kanssa voidaan kehittää sovellus, joka etsii myös luuppikohdat automaattisesti. Koska sampleja kertyy erittäin

paljon, olen pyrkinyt pitämään ne mahdollisimman lyhyinä; tällöin sample instrumentit eivät käytä kohtuuttomasti tietokoneen resursseja ja tiedostokoot pysyvät kohtuullisina. Luuppaamiseen riittää parin sekunnin mittainen pätkä staattista ääntä.

Lukuun ottamatta jo mainittuja soivia konsonantteja, yleensä ottaen konsonantteja ei tarvitse äänittää uudestaan jokaisen laulajan ja kuoron kohdalla. Tämä johtuu siitä, että emme juurikaan erottele laulajan äänenväriä konsonanttien vaan vokaalien pohjalta. Konsonantti soi laulettuun tavun alussa vain sekunnin murto-osan ja konsonantin ja vokaalin välissä on yleensä pieni tauko tai kevennys, minkä vuoksi ne soivat luontevasti yhteen, vaikka konsonantti olisi äänitetty eri laulajalta kuin vokaali. Kun kaikki konsonantit (mukaan lukien useimmat vieraskieliset konsonantit) on kerran perusteellisesti äänitetty ja prosessoitu, miehiltä ja naisilta erikseen, tätä työtä ei tarvitse enää tehdä uudestaan.

Sämplätyn lauluäänen ei tarvitse, eikä sen ole tarkoitus, olla täydellinen kopio jäljittelemästänsä ihmisäänestä. Vaikka se on kopiona epätäydellinen, se on kuitenkin omalla tavallaan instrumenttina täydellisempi; se ei koskaan väsy, ei koskaan kadota omaa tasoaan, tekee aina juuri sen mitä halutaan. Se, mitä Young toteaa synteettisestä ihmisäänestä, pätee pitkälti myös sämplättyyn ääneen: *Mekaaninen ääni on väistämättä puutteellinen kopio alkuperäisestä, ei koskaan sidoksissa inhimilliseen, fyysiseen alkuperään. Kuitenkin sen fyysisessä muodossa on tiettyä puhtautta, sillä soiva ääni ei ole lainkaan taipuvainen inhimillisiin poikkeamiin, virheisiin, sattumanvaraisuuteen tai hiipumiseen* (Young 2015, 77).

8.3 Instrumentin kehitys jatkossa

Laatimani MIDI-instrumenttien jatkotyöstöön kuuluu äänieditointiin liittyvän viimeistelyn lisäksi ehdottomasti käyttöliittymän kehittäminen. Olisi tarpeellista kehittää itsenäinen sovellus, joka ei vaadi sen enempää DAW-ohjelman kuin yleisen sample-riohjelmaankaan käyttöä. Toki, kuten yleisesti sample-instrumenttien kanssa tehdään, olisi hyvä tehdä myös plugin-versio (mielellään VST-muodossa). Tällaisessa sovelluksessa tulisi olla selkeät painikkeet ja säätimet vokaalien ja konsonanttien valintaan sekä muiden äänen ominaisuuksien kontrolloimiseen.

Uskoisin, että tekemäni tutkimus ja sen kautta löytämäni mahdollisuudet ja toisaalta myös ongelmakohdat voivat olla suureksi hyödyksi muille, jotka työskentelevät ihmisäänen digitoimisen parissa. Siksi olisi hyödyllistä saattaa tutkimukseni ja sen tulokset sellaiseen muotoon, joka olisi helposti saatavilla – joko kirjallisena tuotteena tai esimerkiksi videona. Itse tulen epäilemättä jatkamaan aiheen parissa työskentelyä ja käyttämään jo valmistamiani sample-instrumentteja omassa työssäni musiikkipedagogina, säveltäjänä, sovittajana ja studiomuusikkona.

Lähteet

Young, M. 2015. Singing the body electric : the human voice and sound technology. Farnham, Surrey, UK: Routledge.

Duffell, D. 2005. Making Music with Samples. Backbeat Books.

Iivonen A., Aulanko R. 2001. Fonetikan peruskäsitteitä. Helsinki: Yliopistopaino.

Salo, M. 2020. Tietokone äänenkäsittelyn työvälineenä. Helsinki: Planeetta 9.

Johnson, Keith, et al. 2011. Acoustic and Auditory Phonetics, Ebook. John Wiley & Sons, Incorporated.

Kefauver, A., Patschke, D. 2007. Fundamentals of Digital Audio. A-R Editions.

Katz, M. 2010. How Technology Has Changed Music. Capturing Sound, University of California Press.

Wiik, K. 1999. Fonetikan perusteet. Helsinki: WSOY.

Sundberg, J. 1994. Acoustic and psychoacoustic aspects of vocal vibrato. Music and Hearing Quarterly Progress and Status Report, vol. 36.

Schimdt, K. 2019: What is bright and warm singing? www.liveabout.com/bright-and-warm-singing-2994193

Liitteet

Liite 1. Sampler vocabulary (käsitteistö), Excel-taulukko

Liite 2. Tutkimusta havainnollistava video: youtu.be/G5kIBQ5j3KE