

Maailman ensimmäinen mikroprosessoripohjainen sääntösynteesillä varustettu kannettava puhesyntetisaattori SYNTE2 julkistettiin Matti Karjalaisen väitöskirjassa "An Approach to Hierarchical Information Processes with an Application to Speech Synthesis by Rule" TTKK:ssa 20.10.1978 ja Suomen Akatemian VAMKOM-tutkimushankkeen yhteenvetoraportissa 15.11.1978, joka hyväksyttiin asiantuntijalausunnoin Suomen Akatemiassa 1.2.1979. Kehittäjinä tehtäviä erittelemättä olivat Matti Karjalainen, Unto Laine, Seppo Lavonen ja Raimo Toivonen VTT/SAI:ssa. SYNTE2:ta kehittyneemmän version SYNTE3:n kehitys alkoi Matti Karjalaisen, Unto Laineen ja Raimo Toivosen toimesta TTKK:n elektroniikan laitoksella "Puhetulostin" tutkimushankkeen jälkeen 1980. Siinä käytettiin Unto Laineen kehittämää Euroka Oy:n PARCAS-ääniväyläkorttia ja tarvittavien korttien määrä SYNTE3:ssa oli vain kaksi. SYNTE3-hankkeen jälkeen Toivonen alkoi yksin kehittää 1981-1982 korkean tason SPL1-puhsynteesikieltä ja sen kääntäjää. Työn tuloksena Toivonen kehitti yksin 1982-1983 PARCAS-ääniväyläkortin pohjalta SPL1-tutkimuspuhsyntetisaattorin ja parametripuhsynteesiohjelmiston akateemiseen käyttöön EXORset30-tietokoneeseen. SPL1-syntetisaattoreita käytettiin Helsingin yliopistossa ja Tampereen yliopistossa. Niillä tehtiin opinnäytetöitä ja akateemisia tieteellisiä julkaisuja 1983-2004. Viimeisin niistä on Kari Leinosen väitöskirja "Finlands-svenskt sje-, tje- och s-ljud i kontrastiv belysning. Jyväskylä universitet. 2004."

80-luvulla Helsingin yliopistossa fonetiikkaa opiskelevat perehtyivät SPL1-tutkimuspuhsyntetisaattoriin puhsynteesikurssilla. Teoriaa ja työskentelyä koneen ääressä sisältäneestä opiskelusta antaa käsitystä kurssin vetäjän Anna-Liisa Niemisen laatima oppimateriaali.

Anna-Liisa Nieminen (nykyisin Lehessaari)

MUUTA

- ((SEGMENTIN NIMI) (RETURN)
 - KIRJOITTA A NIMEN KURSORIN OSOITTAMALLE SEGMENTILLE
- (O (RETURN)
 - PYYHKII POIS SEGMENTIN NIMEN
- F9-NÄPPÄIN
 - TEKEE KURSORIN OSOITTAMASTA SEGMENTISTÄ KOPION SEN ALLE
- F14-NÄPPÄIN
 - LISÄÄ UUDEN SEGMENTIN KURSORIN OSOITTAMAN SEGMENTIN ALLE
- F15-NÄPPÄIN
 - TUHOAA KURSORIN OSOITTAMAN SEGMENTIN
- L (RETURN)
 - LISTAA ESILLÄ OLEVAN TIEDOSTON PAPERILLE, KUNHAN MUISTAA PISTÄÄ RIVIKIRJOITTIMEN PÄÄLLE

SYNTEN PARAMETRIT

F0	A0	F1	Q1	F2	Q4	F5	AVL	AVH	AFL	AFH	QN
138.	34.	0.3	5.	1.39	3.5	5.11	63.5	63.5	7.	3.5	0.
384.	24.	24.	24.	24.	2.	4.	24.	24.	24.	24.	24.
9.		L2 *									12

DUR ↑ SEGMENTIN NIMI T-ARVOJEN RIVI SEGMENTIN JÄRJESTYS-
NUMERO TIEDOSTOSSA ↑

- F0 - PERUSSÄVELEN TAAJUUS
- A0 - KOKONAISINTENSITEETIN TASO
- F1, F2, ..., F5 - FORMANTTIEN TAAJUUDET
- Q1, Q2, Q3, Q4 - FORMANTTIEN VAIMENNUS: MITÄ SUUREMPI ARVO, SEN VÄHEMMÄN ENERGIAA
- AVL - SPEKTRIN ALAOSAN SOINTIKOMPONENTTI
- AVH - SPEKTRIN YLÄOSAN SOINTIKOMPONENTTI
- AFL - SPEKTRIN ALAOSAN HÄLYKOMPONENTTI
- AFH - SPEKTRIN YLÄOSAN HÄLYKOMPONENTTI
- QN - NASAALIFORMANTIN (250 Hz) VAIMENNUS
- T-ARVOT - LIUKUMA-AIKA ELI AIKA, JONKA MÄÄRÄTTYYN PARAMETRI-ARVOON SIIRTYMINEN VIE
- DUR - SEGMENTIN KESTO

PARAMETRIARVOJEN VAIHTELUVÄLI JA ASKELLUS

F0: 0 - 279.4, 2.2
F1: 0 - 1.524, 12
F2, F3, F4: 0 - 5.207, 41
F5: 4.64 - 10.101, 43
A- JA Q-ARVOT: 0 - 63.5, 0.5
T-ARVOT: 3 - 384, 2ⁿ × 3
DUR: 0 - 762

HUOM! ASKELLUKSEN (JA MUIDEN KIEROUKSIEN) TAKIA PARAMETRILLE NÄPPÄILEMÄSI ARVO EI AINA ILMESTY SEGMENTTIRIVIIN JUURI SELLAISENAAN VAAN YLEENSÄ VÄHÄN PIENEMPÄNÄ

ARTIKULAATIO - AKUSTIIKKA (FANT 1973)

- | | |
|--|---|
| 1. KIELI EDESSÄ | F2 - F1 SUURI |
| 2. KIELI TAKANA | F2 - F1 PIENI |
| 3. SUUN AVAUMA KAPEA | F1 MATALA |
| 4. HUULET LÄHEKKÄIN
JA SUPUSSA | F1, F2, F3 MATA-
LAMMALLA KUIN
CET. PAT. |
| 5. ALVEOLAARINEN AR-
TIKULAATIO | F4 MATALA, LÄ-
HELLÄ F3:STA |
| PALATAALINEN AR-
TIKULAATIO | F3 MATALA, LÄ-
HELLÄ F2:STA |
| 6. BILABIAALINEN TAI
LABIODENTAALINEN
SULKEUMA | F2 500 - 1500 Hz,
HÄLY MATALALLA
JA HEIKKO |
| 8. DENTAALINEN TAI
PREALVEOLAARINEN
ARTIKULAATIO | F2 1400 - 1800 Hz,
F3 KORKEA, HÄLY
VOIMAKAS |
| 9. PALAT. RETROFLEK-
SINEN ART. | F3 MATALA, VOIMA-
KAS HÄLY |
| PALAT. BI-RETRO-
FLEKSINEN ART. | F2, F3 JA VOIMA-
KAS HÄLY KORKEALLA |

MITÄ PITEMPI VÄYLÄ, SEN MATALAMMAT FOR-
MANTTITAAJUUDET; NAISILLA N. 20 % KORKE-
AMMAT KUIN MIEHILLÄ, LAPSILLA VIELÄ KOR-
KEAMMAT (FANT 1973)

(LAINE 1979)

dB
0 a o ö ä e u y i
-10 m n n l v j
-20 s h r b d g
-30
-40 f
k p t

ÄÄNTEIDEN KARKEA LUO-
KITUS KESKIMÄÄRÄISEN
INTENSITEETIN PERUS-
TEELLA

nopeat

k p t b d g
m n n l
s h f r
a e i o u y ä ö
j v

ÄÄNTEIDEN KARKEA LUO-
KITUS SIIRTYMÄNOPEUK-
SIEN PERUSTEELLA

ERI ÄÄNTEITTEN TUOTTAMISESTA SYNTYLLÄ

VOKAALIT

- F1, F2, F3 (& F4): VOKAALIN LAATU, F5 (& F4): PUHUJAN LAATU
- TÄYSI SOINTI, VÄHÄ HÄLY VOI OLLA HYVÄKSI
- REDUSOIMINEN FORMANTTITAAJUUKSIA MUUTTAMALLA

RESONANTIT

- PITKÄT SIIRTYMÄT, FORMANTTIEN, ETENKIN YLINTEN, VAIMENTELU
 - LAINE 1979: /j/ ~ /i/, F1↓, MUUT F↑, SIIRTYMÄT LÄHES ÄÄN-
TEEN ALUSTA LÄHTIEN
/l/: F1↓, F2↓ KUN VIERESSÄ /e, i/, ≡ /y, ä, ö/,
↑ /a, o, u/, SPEKTRIN YLÄOSA VAIMENEE
/v/ ~ /u/, SPEKTRIN YLÄOSAN VOIMAKAS VAIMENEMINEN
- NASAALIT: "PÖRRASMAINEN RAKENNE": F1:N JÄLKEEN
20 dB↓, F2:N JÄLKEEN 20 - 30 dB↓; F1 ~ 200 Hz,
F2 ~ 1100 Hz, F3: /m/ 1800 - 2000 Hz, /n/ 2500 -
2700 Hz, /ŋ/ 2000 - 2200 Hz

FRIKATIIVIT

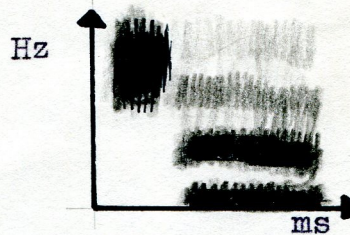
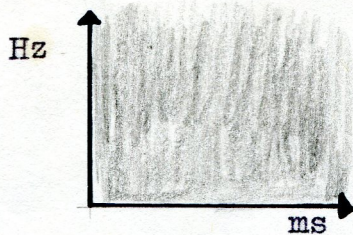
- HÄLY, FORMANTTIEN VAIMENTELU JA LEPATUS (~~~~~)
- VAIHTO SOINNISTA HÄLYYN GRADUAALISESTI: [SNEN] STON [SNEN]
- /h/ ~ VIEREISEN ÄÄNTEEN SOINNITON PÄTKÄ
LAINE 1979: F1↓ JA VAIMENEE, F2↑ JA VOIMISTUU, F3↓ LÄHELLE
F2:STA; /h/: VIEREISEN VOKAALIN PÄTKÄ, JOLLA KOKONAISINTENSI-
TEETTI 10 dB

KLUSIILIT

- LYHYET SIIRTYMÄT
- JOS KUULUU LOKSUMIA, KANNATTAA KOKEILLA T-ARVOJEN MUUTTAMISTA

PUHESYNTESISISTÄ

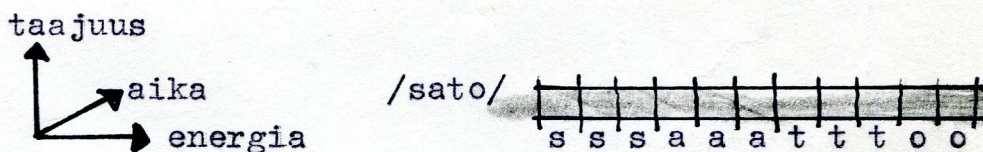
- Puhe: toisistaan erottuvien merkkien tuottamista.



valkoinen häly: energia tasaisesti jakaantunut; todennäköinen ja informaatioton tilanne

puhe: energia järjestelytiettyihin paikkoihin; epätodennäköinen ja informaatiollinen tilanne

- Puheen syntetisoiminen: oikean ihmisen puhetuotoksen korvikkeen rakentaminen.
- Voidaan joko jäljitellä itse puheentuottomekanismia (väyläanalogia), esimerkkinä tästä Kempelenin ja Bellin poikien askartelut, tai vain sen tuotosta (terminaalianalogia), kuten Syntellämme tehdään.
- Puheen syntetisoiminen edellyttää aina puheen analysoimista. Mitä perusteellisemmin niin akustinen, artikulatorinen kuin perseptuaalinen taso tunnetaan, sen helpompaa on päästä haluttuun lopputulokseen.
- Puhetta ei voi matkia kokonaisuutena, yhtenä äänenä, vaan sen tekijät on eriteltävä. Esimerkiksi akustisella tasolla:



Puhejatkuma pilkotaan aikasuunnassa jaotussegmentteihin ja kunkin segmentin spektraalinen laatu täsmennetään tarvittavien muuttujien avulla

- Puhesynteesi ei kuitenkaan ole luonnollisen puheen kopioimista. Kopioiminen edellyttäisi suunnatonta määrää säädettäviä parametrejä ja yhtä suunnatonta säännöstöä niiden säätelyyn, eikä olisi kuitenkaan takeita siitä, että fyysisiltä, mitattavilta ominaisuuksiltaan oikealta näyttävä tulos kuitenkaan kuulostaisi siltä mitä pitää. Passiivisen kopioimisen sijasta on aktiivisesti valittava oleelliset muuttujat ja oleelliset säätötoimet. Vertaa ekspressionismiin kuvataiteessa:

"Aion huomispäivänä lähettää sinulle muutamia kankaita. 'Louhoksensuusta' minä melkein pidän. [---] Tumma vihreä sopii mielestäni hyvin yhteen okrankeltaisten sävyjen kanssa. Siinä on jotain surullista - tervettä - ja siksi se ei kyllästyttä minua. Sama ehkä pätee myös 'Vuoreen'. Ihmiset tulevat sanomaan etteivät vuoret ole sen näköisiä ja että niissä on sormenlevyiset mustat ääriviivat.

'Oliivipuita'-maalauksessa on valkoinen pilvi ja taustalla vuoria, sekä 'Nouseva kuu' ja yötunnelmaa ovat sommittelultaan liioiteltuja; niiden viivat ovat kieroja kuin ikimetsissä. Oliivipuut vastaavat paremmin todellisuutta; yritin siinä kuvata sitä hetkeä päivästä jolloin näkee kultakuoriaisia ja kaskaita poukkoilemassa helteessä. Panen toisen kangaskäärön mukaan yhden kukkaharjoitelman - se ei ole paljon mistään kotoisin, mutta en sentään halua sitä repiäkään.

En pidä hyvinä muita kuin 'Viljapeltoa', 'Vuorta', 'Hedelmätarhaa', 'Oliivipuita' ja sen sinisiä kukkuloita, muotokuva ja 'Louhoksensuuta'. Muut eivät sano minulle mitään, koska niiden viivoista puuttuu yksilöllisyys ja tunne.

Missä nämä viivat ovat tiiviitä ja harkittuja, siinä on oikean maalauksen alkua, liioiteltunakin. Näin Gauguin ja Bernard tuntevat; he eivät vaadi puolta lainkaan oikeaa muotoa, heistä on vain osoitettava onko muoto pyöreä vai kulmikas, ja rehellisesti sanoen oikeassa he ovat. Eräiden tyyppien valokuvamainen ja onntto täydellisyys raivostuttaa heitä. He eivät vaadi vuorilta oikeaa värisävyä, vaan sanovat: Herran tähden, vuorethan olivat sinisiä! Iske siis kankaalle sinistä äläkä selitä että se oli enemmän tätä kuin tuota sinistä. Sehän oli vain sinistä. Hyvä - tee vuorista siniset ja se riittää!" (van Gogh 1889 (1981: Kirjeitä veljelleni (toim. I. Stone), s. 542, 543. Otava: Keuruu.))

- Tietysti jos tarkoituksena on jonkin tietyn akustisen ilmiön ja havainnon välisten suhteiden tutkiminen, on tämä akustinen ilmiö kopioitava sinällään.
- Säättöparametrien määrän lisääminen parantaa yleensä tuotoksen laatua, mutta tämä edellyttää, että uusia parametreja osataan myös käsitellä. Esimerkiksi F_4 - ja F_5 -parametri voivat tuottaa vaikeuksia, jos käytettävissä ei ole tietoa vastaavien formanttien käyttäytymisestä luonnollisessa puheessa.
- Puhetuotoksen ajallisessa pilkkomisessa on myös harkittava ihanteellisinta erottelutarkkuutta; mitä ohuemmiksi siivuiksi tuotos viipaloidaan, sen tarkemmin äännesek-

venssien sisäiset tapahtumatkin saadaan esitettyksi, mutta samalla työmäärä kasvaa. On otettava huomioon, että tietyn rajan jälkeen akustinen tarkentuminen ei enää erityisemmin paranna kuulovaikutelman luonnollisuutta.

- Puheen syntetisoimisessa on pidettävä silmällä/korvalla: äännesegmenttejä - niiden tunnistettavuutta ja sopivuutta kulloiseenkin ympäristöönsä

prosodiikkaa - äänneiden absoluuttisia kestoja (u vs. y), kvantiteettioppositioiden toteutumista; äännelaadun, äännekontekstin, sanarakenteen ja lauserakenteen vaikutuksia perussäveleen ja tuotoksen kokonaisintensiteettiin; lisäksi on säädettävä yleinen puhenopeus, perussävelen taajuusalue ja yleinen äänenvoimakkuustaso sekä näiden vaihtelu ilmauksen kuluessa

äänenlaatua - energian sijoittumista spektrin eri osaluueille, sointi- ja hälyenergian suhteita toisiinsa ynnä perussävelen, äänenvoimakkuuden (ja puhenopeuden?) hienovariaatiota; esimerkiksi perussävelen taajuudessa voi esiintyä jatkuvaa pientä heilahtelua:



Kirjallisuudesta ks. esim Fant, G. 1973: Speech sounds and features. Cambridge, Mass.: The MIT Press; Laver, J. 1981: The phonetic description of voice quality. Cambridge: Cambridge University Press.

- "Speech may be said to consist of movements made audible." (Abercrombie 1967:24.) - Puheessa heijastuvat paitsi aiotut aivojen ja puheentuottoelimistön liikkeet myös tottumukset ja rajoitukset näiden liikkeiden suorittamisessa, eli puhujan menneisyys ja nykyinen tila (Spede Pasanen, Viivi Luik, Amenofis IV, Knalli ja saateenvarjo).
- On muistettava synteessin päämäärät: onko etusijalla akustinen vai perseptuaalinen luonnonmukaisuus, onko riittävää, että tuotos on kielellisesti ymmärrettävää vai halutaanko puhujan lisäksi myös kuulostavan inhimilliseltä ja jopa jossain tunnistettavassa tunnetilassa olevalta.
- Jos tavoitellaan perseptuaalista luonnonmukaisuutta, on oikeiden parametriarvojen saavuttamisen lisäksi muistettava niiden laiminlyöminen ja niistä poikkeilu, se, ettei

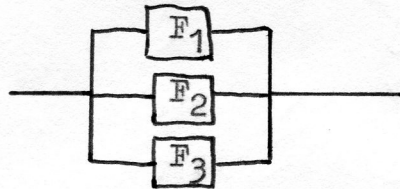
va)

luonnollisessa puheessa mikään oikein pysy paikoil-
laan.

- Parametrisynteesiksi syntetisaattorille syötetään pa-
rametriarvoja, sääntösynteesissä symbolimerkkejä, joista
kukin laukaisee esiin tietyt, ennalta suunnitellut para-
metriarvot. Puhesynteesin käytännön sovelluksissa on
pääsääntöisesti kyse sääntösynteesivempaimista.
- Muutamasta termistä:



sarjaankytketyt suodattimet (kaskadi)



rinnankytketyt suodattimet (paralleeli)



van Gogh: Auringonlasku (vrt. esim. Ingres!)