Acoustic synthesis: KlattGrid formant synthesis

David Weenink
What is an acoustic synthesizer?

An acoustic synthesizer is

a computer program to synthesize speech on the basis of a specification of acoustic parameters as a function of time.

The KlattGrid is a tier-based time-stamped speech synthesizer!

Example of KlattGrid specifications

Add pitch point... 0.1 120
Add voicing amplitude... 0 90
Add oral formant frequency point... 1 0.15 800
Add oral formant bandwidth point... 1 0.15 50
What is an acoustic synthesizer?

An acoustic synthesizer is a computer program to synthesize speech on the basis of a specification of acoustic parameters as a function of time.

The KlattGrid is a tier-based time-stamped speech synthesizer!

Example of KlattGrid specifications

Add pitch point... 0.1 120
Add voicing amplitude... 0 90
Add oral formant frequency point... 1 0.15 800
Add oral formant bandwidth point... 1 0.15 50
Formant synthesis

- Unit synthesis gives nearly perfect sounds
- Why do we need formant synthesis?
- To have more control over the final speech sound
- Contra: It is very difficult, especially for consonants
Formant synthesis

- Unit synthesis gives nearly perfect sounds
- Why do we need formant synthesis?
  - To have more control over the final speech sound
  - Contra: It is very difficult, especially for consonants
Formant synthesis

- Unit synthesis gives nearly perfect sounds
- Why do we need formant synthesis?
- To have more control over the final speech sound
- Contra: It is very difficult, especially for consonants
Formant synthesis

- Unit synthesis gives nearly perfect sounds
- Why do we need formant synthesis?
- To have more control over the final speech sound
- Contra: It is very difficult, especially for consonants
What is a formant?

- **Production**: a resonance of the vocal tract
- **Perception**: a local maximum in the spectrum
- **Acoustics**: a peak in the spectrum
How can we make it

**Second order recursive filter**

\[ y_n = ax_n + by_{n-1} + cy_{n-2}, \text{ (filter)} \]

Resonance for \( b = 2r \cos(2\pi FT) \) and \( c = -r^2 \), where \( r = e^{-\pi BT} \). \( F, B \) are formant frequency and bandwidth, \( T \) is sampling time (=1/samplingFrequency)

**How to do this in Praat**

```plaintext
f1 = 800
b1 = 80
fs = 44100
r = exp(-pi*b1/fs)
b = 2*r*cos(2*pi*f1/fs)
c = -r^2
Create Sound from formula... f Mono 0 0.2 fs if col = 1 then 1
...else self+b*self[col-1]+c*self[col-2] fi
```
How can we make it

**Second order recursive filter**

\[ y_n = ax_n + by_{n-1} + cy_{n-2}, \text{ (filter)} \]

Resonance for \( b = 2r \cos(2\pi FT) \) and \( c = -r^2 \), where \( r = e^{-\pi BT} \).

\( F, B \) are formant frequency and bandwidth, \( T \) is sampling time (=1/samplingFrequency)

**How to do this in Praat**

```plaintext
f1 = 800
b1 = 80
fs = 44100
r = exp(-pi*b1/fs)
b = 2*r*cos(2*pi*f1/fs)
c = -r^2
Create Sound from formula... f Mono 0 0.2 fs if col = 1 then 1
...else self+b*self[col-1]+c*self[col-2] fi
```
How does a formant look like?

As a damped sine:  \[ f(t) = Ae^{-\alpha t} \sin(2\pi Ft + \phi), \]
where \( A \) is the amplitude, \( F \) is the formant frequency and \( \phi \) is a phase.

To demo function_simulator_damped_sine.praat
1. Phonation 2. Vocal tract - phonation coupling
Phonation section

- Pitch tier
- Voicing amplitude tier
- Flutter tier
- Glottal flow function tiers (power1, power2, collision phase)
- Open phase tier
- Double pulsing tier
- Aspiration tier
- Breathiness tier
- Spectral tilt tier
Glottal flow function (top) and derivative (bottom).
Example of flow function (top) and derivative (bottom) with collision phase.
Breathiness

Glottal flow with breathiness during open phase.
Vocal tract & coupling section

- Vocal tract tiers
  - Oral formant tiers (frequencies & bandwidths)
  - Nasal formant tiers
  - Nasal antiformants (spectral dips)

- Coupling tiers
  - Tracheal formant tiers
  - Trachael antiformant tiers
  - Delta formant tiers (delta frequencies & bandwidths)
Example of formant coupling

Extreme coupling. First formant changes from 500 to 1000 Hz during open phase.
Frication section

- Noise amplitude tier
- Frication formants (frequencies, bandwidths & amplitudes)
- Bypass tier
Comparison with other implementations

- Time-stamped (instead of frame-based)
- Unlimited number of formants and antiformants (instead of 6F, 1 NF/AF, 1 TF/TAF)
- Frication and normal formants uncoupled (instead of equal)
- Glottal-flow function time-stamped (instead fixed)
- No limit on number of delta formants (instead of one)
- No parameter quantization (instead of quantization of amplitudes)
- Completely integrated in Praat (instead of separate tool)