Making music with voice
MENU:

A: The instrument
B: Getting heard
C: Expressivity
The instrument
Summary

Radiated Spectrum

Frequency

Level

Vocal Tract

Frequency curve

Formants

Voice Source

Spectrum

Level

Frequency

Waveform

Transglottal airflow

Time

Lungs

Trachea

Vocal folds

Velum
Demo

Variable sine wave traveling through a tube
Thus, spectrum peaks at formant frequencies
Demo

Pulsating airflow through tube

Result: vowel-like sound
Formant frequencies determined by vocal tract shape
Vocal tract shape of vowel /i/
Vocal tract shape of vowels
Demo:

Pulsating airflow through pinched tube produces vowels
Formant frequencies of vowels
Demo: Crusade in the $F_1$ & $F_2$ archipelago

Result: all vowels available by varying $F_1$ & $F_2$
Articulatory tools:

• Jaw opening
• Lip opening
• Tongue body shape
• Tongue tip
• Larynx position
Vocal tract length is also important

Short vocal tracts have higher resonance frequencies than longer vocal tracts
Listen to voice timbre difference produced by different vocal tract lengths!
Tuning formants

First formant:
Mostly jaw opening

Second formant:
Mostly tongue shape

Third formant:
Cavity behind lower incisors

Higher formants:
Vocal tract length, Larynx position

Determine vowel
CONCLUSIONS

Formants

• controlled by vocal tract shape (articulation)

• first two resonances determine vowel quality

• higher formants relevant to personal voice quality
MENU:

A: The instrument
B: Getting heard
C: Expressivity
Getting heard
The high-pitch case
Question:
Where is the fundamental and where is the first formant?
Is fundamental allowed to pass the first formant?

Listen!
Jaw opening is particularly efficient tool for raising first formant

Female singers tend to widen the jaw opening at high pitches!

Vowel [i]

Vowel [u]
Larynx height strategy in professional soprano

![Graph showing the relationship between larynx height and fundamental frequency.](image)
Jaw opening strategy in professional soprano

![Graph showing the relationship between jaw opening and fundamental frequency for spoken and sung sounds.](image)

Distinguished lecture, CIRMSTM Jan 2009, Copyright Johan Sundberg
In [a:] the jaw opening is widened when pitch frequency approaches first formant, but in [e:] some semitones higher.
Experiment
MRI analysis of professional soprano singing different vowels on a triad pattern covering her range
Measuring jaw opening
Measuring tongue dorsum height
Tongue dorsum first

Pitch

Jaw opening later

Fundamental = Normal F1

Pitch [semitones above start tone]

Change of articulator position [mm]

Time [image number]
So why not reduce tongue bulging also in /a/?
APEX, please
This formant strategy expands the dynamic range

![Graph showing the relationship between pitch frequency and SPL at 0.3 m. The graph compares Singers and Untrained with different dynamic levels: pppp and ffff.](image-url)
Considerable sound level gain:
Loud tones at low cost,
Vocal economy!
Which singers can profit from this strategy?
Formant frequencies for vowels

Classification

Bass

First formant [kHz]

Second formant [kHz]
Formant frequencies for vowels

Classification

Baritone
Formant frequencies for vowels

Classification

Tenor
Formant frequencies for vowels
Formant frequencies for vowels
Formant frequencies for vowels

Classification
Soprano
Alto
Tenor
Baritone
Bass

Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg
Singers singing in pitch ranges above normal value of first formant need to learn a pitch-dependent vowel articulation!!!!!!!
Don’t allow pitch frequency to pass the first formant!

Trick:
• reduce articulatory constriction
• widen jaw opening

Result:
• loud sound at minimum effort/vocal economy
Getting heard

The male case
Singer’s formant cluster

Also called singer’s spectrum peak

The fine art of clustering resonances;
Performed by male classically trained

• tenors
• baritones
• basses
Check spectrum
Singer’s formant cluster

Long-term-average spectrum of orchestra ± singer

Sound example
1. Noise corresponding to orchestral sound
2. Singer
   - Singers’ formant cluster
   + Singers’ formant cluster
3. Examples 1. & 2. together

<table>
<thead>
<tr>
<th>Frequency [kHz]</th>
<th>Mean sound level [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Orchestra</td>
</tr>
<tr>
<td></td>
<td>Singer + orchestra</td>
</tr>
</tbody>
</table>

Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg
Production of singer’s formant cluster

Good voice source
Wide pharynx/Low larynx
Clustering of formants 3, 4, 5
Formants generate spectrum peaks
Their levels determined by their frequencies

(Resonances are like good friends, proximity strengthens)
Creating singer’s formant cluster produced by clustering formants

Formant 5 lowered from 4500 to 2700 Hz

Distinguished lecture, CIRMNT Jan 2009, Copyright Johan Sundberg
Center frequency of singer’s formant cluster is perceptually relevant
Lowering third formant by 300 Hz

Frequency [Hz]

Average level [10 dB/division]

Manipulated

Original

Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg
Who possesses a singer’s formant cluster?
LTAS

Distinguished lecture, CIRMNT Jan
2009, Copyright Johan Sundberg
Is this a singer's formant?
Any peak at high frequency is not a singer’s formant cluster
Female singers don’t have a singer’s formant cluster.
Vocal loudness and spectrum tilt
Effect of vocal loudness on spectrum slope,
Check spectrum, when loudness is increased!

High overtones gain more than low, so singer’s formant cluster becomes more dominant, if vocal loudness is increased!

A bad strategy for achieving singer’s spectrum peak!
Summary

Singer’s formant cluster

• occurs near 3000 Hz in tenor, baritone, and bass voices
• center frequency varies with voice classification
• generated by clustering formants 3, 4, & 5
• can be achieved by a wide pharynx/lowered larynx
• helps male solo singer’s voice to cut through loud accompaniment

Another case of vocal economy
MENU:

A: The instrument
B: Getting heard
C: Expressivity
Expressivity
What is the code?
Where did we learn it?
The Analysis-by-Synthesis strategy

Music file → DIRECTOR MUSICES (Performance grammar) → Synthesis

Translators: Sundberg & Friberg

New/Modified Performance rules

Musician Lars Frydén

Synthesised performance

Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg
Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg

DIRECTOR MUSICES interface

Quantity

Performance rules

2009, Copyright Johan Sundberg
Three performance principles:

1. Marking the structure
2. Sharpening contrasts
3. Emphasising important notes
Principle 1:
Mark the structure!
Tempo change is an *expressor*
Similar expressor used in speech
Final lengthening in speech

Stressed syllable, initial position

Stressed syllable, final position
Origin of this expressor?
Velocity of hand movement

Change of hand position along straight line

Position

Start

Target

(according to Juslin, Friberg and Bresin, forthcoming)
Velocity of hand movement translated to tempo

(translated to tempo (according to Juslin, Friberg and Bresin, forthcoming))
Hand movement pattern translated to tone duration

Note position in phrase

Relative tone duration

Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg

(according to Juslin, Friberg and Bresin, forthcoming)
Phrase marking

Phrases level

1 + 2

Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg
Modelling Hagegård’s phrase marking with hand movement pattern

Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg
Final ritardando,
another expressor

Normalised tempo

Normalised time

2009, Copyright Johan Sundberg
Final ritardando & stopping running

- Normalised time
- Normalised velocity
- Normalised tempo

- Mean ritardando
- Runners' mean velocity

2009, Copyright Johan Sundberg
Do we understand tempo change expressors because they allude to experience of movement?
Principle 2:
Sharpen contrasts!
The case of pitch
Director musices example
Felix Mendelssohn Bartholdy: *Scherzo from Ein Sommernachtstraum*, op 61

Deviation from equally tempered tuning [cent]

K=0
K=5
K=10
K=-5

Distinguished lecture, CIRMMT Jan
2009, Copyright Johan Sundberg
Other examples
Sharpening contrasts in speech

Swedish example:

Long and short vowels: (ha:t = hatred; hatt: = hat)

Vowel duration contrast enhanced by formant frequency differences

Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg
Sharpening contrasts in speech

Approval directed to

Infant

Adult

Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg
Principle 3:
Emphasise important events
Lag of vowel onset relative to the accompaniment chord [ms]

Tone onset and vowel onset

Occurrence [%]

N = 552
Thus, sung tone start at the vowel onset
Expressor in singing:

Timing of tone onset/Tone duration
Tone durations in Hagegård’s material

Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg
Expressor in singing:
Emphasis by delayed arrival
Expressor in speech:

Syllable duration
Syllable duration in actor’s speech

Stressed syllable

![Graph showing syllable duration comparison](image)
Expressor in speech:
Emphasis by delayed arrival
Expressor in singing:
Amplitude of fundamental
How does it sound?
Physiological factors affecting voice timbre

Voice source
(Transglottal airflow)

Subglottal pressure
(Loudness)

Vocal fold length & tension
(Pitch)

Glottal adduction
(Phonation type)

Formants
Amplitude of voice source fundamental

- Adduction
  - Weak
  - Firm

Phonation type:
- Breathy
- Flowy
- Neutral
- Pressed

Fundamental
- strong
- weak

Distinguished lecture, CIRMMT Jan 2009, Copyright Johan Sundberg
Long-Term-Average Spectra of expressive and neural versions of examples

Red: Expressive, Gray: Neutral

**Agitated**

- a: Ford's Monologue
- b: Wie du auch

**Peaceful**

- e: Ich denke dein
- f: Du bist die Ruh
- c: Mein Vater
- d: Mahler Messer
- g: Es schwelen
- h: Wanderes Nachtlied

Frequency [10 dB / division]

Pitch range

Frequency [100 Hz / division]
Expressor in speech:
Amplitude of fundamental
Expressor in speech:

Amplitude of fundamental

Courtesy of Klaus Scherer, Geneva

Distinguished lecture, CIRMNT Jan 2009, Copyright Johan Sundberg
Summarising

Instrument:
Pulsating transglottal airflow, controlled by subglottal pressure, glottal adduction and vocal fold length and tension formant, controlled by articulation

Getting heard:
Use formants to reach audibility when accompaniment is loud

Expression:
Principle 1. Mark the structure
Principle 2. Enhance contrasts
Principle 3. Emphasise important events