On acoustic salience of vowels and consonants predicted from articulatory models

Many thanks to Shinji Maeda and students

Jacqueline Vaissière
Université Sorbonne Nouvelle - Paris
Laboratoire de Phonétique et de Phonologie

PPC2008, Pékin, Chine
Menu?

1. What is a model?
2. Why modelling?
3. Which models?
4. What is a useful model?
5. Demo (vowels consonants)
6. How done?
7. Other uses
   - speaker anatomy
   - speaker strategy
   - singing formant
   - inversion
   - etc

PPC2008, Pékin, Chine
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
vowels
consonants
Other uses
speaker anatomy
speaker strategy
singing formant
inversion etc;

The questions

PPC2008, Pékin, Chine
A model is a simplified description of a complex system

A model is not a goal per se, but a tool

1. to know what we know
2. and what we don’t know
3. And where to look for …
Why modelling?

Modeling is needed to apprehend complex phenomena

Speech is complex enough to need modelisation

Modeling allows to evaluate the separate contribution of individual parameters and the acoustic and perceptual consequences of large and small movements

Modelisation in turn makes speech looking less or sometimes more complex
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

Phonetics

• Prosodic modelling
• Source modelling
• Vocal tract
  • 2D articulatory
  • 3D articulatory
  – Articulatory modelling

Psycholinguistics

• perception and recognition
  • Phoneme
  • Word access
  • Sentences
  – Etc.

PPC2008, Pékin, Chine
What is a useful model?

Be compatible with the observations
Be testable
Describes well the data and to be able to reproduce them from

Few number of parameters
As motivated as possible
As independent as possible

And for AM, react as a human vocal tract
no stop without a burst, or no frication when humans do.

To draw a direct link between articulatory, acoustic and perceptual observations

PPC2008, Pékin, Chine
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?

– By Shinji Maeda
– From X-ray data for French (from Strasbourg)
– By statistical analysis
– The most widely used AM
– Used for pedagogical purpose in our lab and research in Grenoble, Nancy, UCLA, MIT, etc.

PPC2008, Pékin, Chine
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
vowels
consonants
Other uses
speaker anatomy
speaker strategy
singing formant
inversion

X-ray data
One speaker
10 sentences
>1000 images
VT shapes

First subtract the effect of the jaw

Lindblom & al, 1971

10 sentences
>1000 images
VT shapes

Guided PCA
7 parameters

PPC2008, Pékin, Chine
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo vowels:

1+2 = 58%, 56%

1+2+3 = 81%, 79%

1+2+3+4 = 88%, 86%

Other uses:

Jaw + tongue

4 first parameters

How done:

Demo vowels:

1+2 = 58%, 56%

1+2+3 = 81%, 79%

1+2+3+4 = 88%, 86%
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
vowels
consonants
Other uses

How done?

Demo
vowels
consonants

LIP
ht 5
pr 6
larynx
ht 7

15%
24%
nasal 8

3 more = Lip and larynx
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
vowels
consonants
Other uses

7 parameters = Speech organs

How done?

LIP
ht 5
pr 6
larynx
ht 7

15%
24%
nasal

7%
Surprisingly or not

The 7 statistically found parameters could be interpreted in phonetic terms:
2 for lips, 3 for tongue, 1 for jaw, 1 for larynx.
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

JAW

LIP

ht 5

pr 6

larynx

ht 7

nasal

15%

24%

7 commands

body

shape

u, k

apex

i, e, dentals

7 commands

2 43%

3 23%

4 7%
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant inversion etc;
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant inversion etc;

• 1) vowels

• 2) CVC sequences

Examples?

PPC2008, Pékin, Chine
1) vowels

VaiSSiÈRE, 2007

PPC2008, Pékin, Chine
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels
consonants

Other uses
speaker anatomy
speaker strategy
singing formant inversion etc;

PPC2008, Pékin, Chine
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

  vowels

  consonants

Other uses

  speaker anatomy
  speaker strategy
  singing formant
  inversion
  etc;

2) consonants

[b]

F1 371
F2 898
F3 2136
F4 3683
F5 4024

[a]

F1 712
F2 1145
F3 2229
F4 3745
F5 4024

[b]

closure

[i]

after release
Beans on a string synthesis

1) What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo
  vowels
  consonants

Other uses
  speaker anatomy
  speaker strategy
  singing formant inversion etc;

PPC2008, Pékin, Chine
• /a/ during 30 msec
  cosine transition for 50 ms
• /k/ closure for 100 ms
• /k/ release for 17 msec
  cosine transition for 17 ms
• /u/ for 150 msec
• Start by close VF for 50 msec
  Then cosine transition to open glottis (0.4) at 100 msec
  Stay open for 100 msec
  Then cosine transition to close glottis at 250.
  Stay close for 100 msec
  Open slightly for the last 50 msec. (prepausal opening)
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?

Demo
vowels
→
consonants

Other uses
speaker anatomy
speaker strategy
singing formant
inversion
etc;

Glottis and the three sources in the model

• A) closed
• voicing

• B) open
• If less open than VT: aspiration
• If VT less open than glottis: frication

PPC2008, Pékin, Chine
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo
vowels
consonants

Other uses
speaker anatomy
speaker strategy
singing formant
inversion
etc;

Invariant place and shape

Figure 2: Interpolation of target area function
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?

**Different degrees of sophistication**

1) **Simple concatenation**

/asa/ = 3 targets  
= /a/ + /s/ + /a/  
/ata/ = 4 targets  
= /a/ + /tclosure/ + /t release/ + /a/

+ Cosine transitions

2) **Contextual allophones**

Ease of articulation  
Necessary for velar

Demo
vowels
→ consonants

Other uses
speaker anatomy
speaker strategy
singing formant
inversion
etc;

PPC2008, Pékin, Chine
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?

Demo
vowels
consonants

Other uses
speaker anatomy
speaker strategy
singing formant
inversion
etc;

• /k/ and /g/ require the burst to be in special position to the ‘main pitch’ of the vowel
• Different /k/ allophones
• Main pitch /i/= 3000 Hz, /y/= 2000 Hz, etc.
• Not articulatory ease, but perceptual requirement
• It seems to be not a linguistic but a purely psychoacoustic phenomena (temporal masking)

PPC2008, Pékin, Chine
the reference to the vowel is not needed for discriminating [p] and [t].

Perceptual requirement?
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

1) Simple concatenation

2) Contextual allophones

Velar

3) Anticipation and carry over phenomena

Other uses

speaker anatomy
speaker strategy
singing formant inversion etc;

PPC2008, Pékin, Chine
**What is a model?**

**Why modelling?**

**Which models?**

**What is a useful model?**

**How done?**

**Demo**

- vowels
- consonants

**Other uses**

- speaker anatomy
- speaker strategy
- singing formant
- inversion
- etc;

**/aku/**

**Mouth already open**

**/t0/**

**/aku/**
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

Maximum opening for /a/

Jaw starts to rise again
<table>
<thead>
<tr>
<th>Other uses</th>
<th>Demo</th>
<th>How done?</th>
<th>What is a useful model?</th>
<th>Which models?</th>
<th>Why modelling?</th>
</tr>
</thead>
<tbody>
<tr>
<td>vowels</td>
<td>consonants</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Round gesture starts**

**/k/ closing gesture starts**

/Aku/
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

Other uses

singing formant

inversion

e tc
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
vowels
consonants
Other uses
Closure
Lip continue to round
<table>
<thead>
<tr>
<th>Why modelling?</th>
<th>What is a useful model?</th>
<th>How done?</th>
<th>Which models?</th>
<th>Demo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **vowels**
- **consonants**

Other uses:
- Speaker anatomy
- Speaker strategy
- Singing formant inversion

---

/aku/

K de aku
/a 5k u/
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo
vowels
consonants

Other uses
speaker anatomy

/a ku/

37
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

t7
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?

Demo
vowels
consonants

Other uses
speaker anatomy

/a ku/

/t8/
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
  vowels
  consonants
Other uses
  speaker anatomy
Maximum rounding
/aku/
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
vowels
consonants
Other uses
Velum starts to lower
Derounding starts
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo
  vowels
  consonants

Other uses
  speaker anatomy
  speaker strategy
  singing formant
  inversion
  etc;

Larynx maximally low
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
vowels
consonants
Other uses
speaker anatomy
Speaker strategy
Singing formant
Inversion

/aku/
**Different degrees of sophistication**

1) **Simple concatenation**

2) **Contextual allophones**
   - Velar

3) **Anticipation and carry over phenomena**

4) **Suprasegmental influences**
   - Position in the syllable (nasal in preparation, masking)
   - Position in the word or relative to stress (lenition and fortition) seen

---

Other uses:
- speaker anatomy
- speaker strategy
- singing formant
- inversion etc;

**Demo**
- vowels
- consonants
The points
1) Pedagogical point of view

Draw the link between
-articulatory data,
-acoustic data
-and perception

>> Understanding for linguists
the basics of
Acoustic Theory of Speech production
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

Vowels

Consonants

Other uses

Speaker anatomy

Singing formant inversion etc.

Production

Acoustics

Perception

1) DATA

linguist

Language teacher

phonetician

phonologist

engineer

Speech therapist

Medical students

2) Coherent teaching Research tool

Modelisation
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo
  vowels
  consonants

Other uses
  speaker anatomy
  speaker strategy
  singing formant
  inversion
  etc;

Modelisation
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant

inversion

e tc;
2) For what concerns the vowels

a) Need for reference vowels

Cardinal vowels

Sensibility to formant amplitude and bandwidth

AM could help
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant

inversion

etc.
1) Helps to represent in a comprehensive way the language-specific prototypes
1) Helps to represent in a comprehensive way
the language-specific prototypes
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

Vowels

Consonants

Other uses

Speaker anatomy

Speaker strategy

Singing formant inversion etc;

Modelisation
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant

inversion

etc;

(neutral + protrusion)

(\(F2 \downarrow \downarrow F3\))
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant

inversion

etc;

neutral

+ protrusion

and

+ compression

Lip height and protrusion

neutral

+ compression

spectral magnitude (dB)

Frq(Hz) Bw(Hz) A(dB)

F1 464 47 22
F2 1610 64 20
F3 2414 56 20
F4 3652 57 19
F5 4240 58 18

spectral magnitude (dB)

Frq(Hz) Bw(Hz) A(dB)

F1 433 37 21
F2 1455 46 19
F3 2352 47 18
F4 3590 46 18
F5 4210 34 19

spectral magnitude (dB)

Frq(Hz) Bw(Hz) A(dB)

F1 402 59 16
F2 1362 23 21
F3 2352 42 14
F4 3621 24 18
F5 4240 23 17

Modelisation
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant

inversion

e tc;

Lip height and protrusion

(neutral + compression + protrusion + compression + protrusion)

LIP + TONGUE

(F2 F3)
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
vowels
consonants
Other uses
speaker anatomy
speaker strategy
singing formant
inversion

Lip height and protrusion
neutral
+ compression
+ compression
+ protrusion (F2, F3)
LIP
LIP + TONGUE
<table>
<thead>
<tr>
<th>Lowest formant</th>
<th>Possible</th>
<th>Highest formant</th>
<th>Possible</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>F1</strong></td>
<td>anterior part of the VT</td>
<td><strong>F1</strong></td>
<td>posterior part of the VT</td>
</tr>
<tr>
<td><strong>F2</strong></td>
<td>velar region + lip rounding</td>
<td><strong>F2</strong></td>
<td>mid-palatal region + glottal region</td>
</tr>
<tr>
<td><strong>F3</strong></td>
<td>pharyngeal region + bunching of the tongue, retroflexion + lip rounding and lip protrusion</td>
<td><strong>F3</strong></td>
<td>apical and prepalatal regions + lip spreading + glottal region (larynx lowering)</td>
</tr>
</tbody>
</table>

Table 3: Tongue constriction position, lip configuration, tongue shape and F-pattern values: a summary.
2) For what concerns the vowels

a) Need for references vowels

b) The rôle of anatomical details
   - the laryngeal cavity for the vowels
   - rôle of pharyngeal volume, etc.
Role of anatomical details

What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
  vowels
  consonants
Other uses
  speaker anatomy
  speaker strategy
  singing formant inversion
  etc;

Modulation

2) For what concerns the vowels

a) Need for references vowels

b) The rôle of anatomical details

c) The acoustic properties largely Prevail over articulation
Perception prevails over articulation

About the same sound (except the focal and extreme sounds can be obtained by Fairly different VT configuration

Modelisation
2) For what concerns the vowels

a) Need for reference vowels

b) The rôle of anatomical details

c) The acoustic properties largely prevail over articulation

d) Different strategies can be modelled
Speaker different strategies
2) For what concerns the consonants

a) *Simple concatenation works better than expected*
2) For what concerns the consonants

a) Simple concatenation work than expected

b) The rôle of anatomical details: sublingual cavity
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo
  vowels
  consonants

Other uses
  speaker anatomy
  speaker strategy
  singing formant inversion etc;

sublingual cavity
Necessary
To recreate existing
Acoustic discontinuity
2) For what concerns the consonants

a) Simple concatenation work than expected

b) The rôle of anatomical details: sublingual cavity

c) The acoustic properties largely Prevail over articulation: the velar
2) For what concerns the consonants

a) *Simple concatenation work than expected*

b) *The rôle of anatomical details: sublingual cavity*

c) *The acoustic properties largely*  
*Prevail over articulation: the velars*

d) *Exotic consonants may be easy to model*
Non French consonants

- Start from articulatory model for the flanking vowels
- Tube or area function for the consonant
- Examples: posterior arabic consonants (Yeou’s thesis)
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant

inversion

e tc;

Glottis more open?

95%, 69*

100%, 76

100%, 72

100%, 79

10 listeners

10 repetitions

*Alwan, 1989
• What is a model?
• Why modelling?
• Which models?
• What is a useful model?
• How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant

inversion

etc;

Pharyngeal sourde

16.5

0.7 0.7 0.7 0.7 0.30

5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0

Pharyngeale sonore

16.5

0.7 0.7 0.7 0.7 0.35

5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0

[uvulaire sourde]

17

2.0 2.0 2.0 2.0 2.0 2.0 2.0 2.0 2.0 0.2

5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0

[uvulaire sonore]

17

2.0 2.0 2.0 2.0 2.0 2.0 2.0 2.0 2.0 0.35

5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0 5.0

Ag = 0 cm²

Ag = 0.20-0.25 cm²

+ timing
2) For what concerns the consonants

a) Simple concatenation work than expected

b) The rôle of anatomical details: sublingual cavity

c) The acoustic properties largely Prevail over articulation: the velars

d) Exotic consonants may be easy to model

e) Lenition and fortition not difficult to model
lenition

- Less strong constriction
- More close glottis
- Ratio between the two areas may change, and to the category of the consonant
- But the underlying F-pattern is the same
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant

inversion

etc;

Figure 4.4.2. Normal and two degrees of target undershoot of a voiced stop.
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
vowels
consonants

Other uses
speaker anatomy
speaker strategy
singing formant inversion etc;

Less strong closure in the middle

Vt more close than vf
Less vf opening
Voiceless > voiced
Noise created at the narrowest

Modelisation
2) For what concerns the consonants

a) Simple concatenation work than expected

b) The rôle of anatomical details: sublingual cavity

c) The acoustic properties largely Prevail over articulation: the velars

d) Exotic consonants may be easy to model

e) Lenition and fortition not difficult to model

f) Shape of the tongue counts more that exact place of constriction
What is a model?
Why modelling?
Which model?
What is a useful model?
How done?
Demo
vowels
consonants
Other uses
speaker anatomy
speaker strategy
singing formant
inversion
etc;

And shape of the tongue is influenced by the vowels

laminal
apical

F2
2) For what concerns the consonants

a) Simple concatenation work than expected

b) The rôle of anatomical details: sublingual cavity

c) The acoustic properties largely
Prevail over articulation: the velars

d) Exotic consonants may be easy to model

e) Lenition and fortition not difficult to model

f) Shape of the tongue counts
more that exact place of constriction

g) Suppressing the coarticulation effect allows
to distangle the perceptual effects C to V and V to V
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
vowels
consonants
Other uses
speaker anatomy
speaker strategy
singing formant
inversion

Speaker adaptation
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses:

speaker anatomy
speaker strategy
singing formant
inversion
etc;

VLAM (Maeda, 1994)

Variable Linear Articulatory Model

lip, jaw, tongue, larynx

The Larynx Height Index variation is introduced by modifying the longitudinal dimension of the vocal-tract according to 2 scaling factors and 3 zones:

(1) for the anterior part of the vocal-tract
(2) for the pharynx
(3) interpolating the zone in-between

Babies

Neanderthal monkey

Modelisation
Vowel prototypes for ontogenesis

Which models?

What is a useful model?

- speaker anatomy
- speaker strategy
- singing formant inversion etc;

Maeda + ICP-Grenoble

Modelisation

Babies Neanderthal monkey

Fo

Larynx

Adult

Newborn

Ontogenesis
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?

Man  Woman  Child (10y 4y)  Newborn

Ape
Newborn  Adult

Neandertal

La Ferrassie  La Chapelle-aux-Saints

Maeda + ICP-Grenoble

Babies  Neanderthal  monkey

Larynx Height Index

<table>
<thead>
<tr>
<th>1.00</th>
<th>0.95</th>
<th>0.88</th>
<th>0.79</th>
<th>0.61</th>
<th>0.55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Man</td>
<td>Woman</td>
<td>Child</td>
<td>Newborn</td>
<td>Ape</td>
<td>Neandertal</td>
</tr>
<tr>
<td>Newborn</td>
<td>Adult</td>
<td>La Ferrassie</td>
<td>La Chapelle-aux-Saints</td>
<td>Maeda + ICP-Grenoble</td>
<td></td>
</tr>
</tbody>
</table>
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?

Vowel prototypes (adult man)

/i/  L = 16.3 cm
/a/  L = 17.4 cm
/u/  L = 19.8 cm

Maeda + ICP-Grenoble

Babies Neanderthal monkey
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant

inversion

etc;

Maeda + ICP-Grenoble
Contra Lieberman & Crelin

Modelisation of the differences suggests that the Neandertals were no more vocally handicapped than children at birth are...

Maeda + ICP-Grenoble
Conclusions?

- AM is a very useful research and teaching tool
  - Essential as a platform for discussions between phoneticians and engineers

- As a complement of other instrumentation such as EMMA, echograph, IRM, etc.

- Still in progress (a third cavity)

- New thesis coming in fricatives in different languages (Toda/Maeda), MRI, ATR
Conclusions?

• **limits**

1. /l/ and many other sounds not yet done (as presented in this congress)
2. Register/Phonation types/laryngeal region area and Fo/quality interaction not included (could replace the actual 3 parameters)
3. Aerodynamic contrains increases when VT very close
4. mid-saggital profile conversion to area function not straightful process (3D)
5. New X-ray difficult to obtain (MRI and new instrumentation)
What is a model?
Why modelling?
Which models?
What is a useful model?
How done?
Demo
vowels
consonants
Other uses
speaker anatomy
speaker strategy
singing formant
inversion
e etc;

Thank you!
What is a model?

Why modelling?

Which models?

What is a useful model?

How done?

Demo

vowels

consonants

Other uses

speaker anatomy

speaker strategy

singing formant

inversion

etc;

Thank you!

PPC2008, Pékin, Chine