Fo contours as strings of syllabic pulses

Björn Lindblom
Dept of Linguistics
Stockholm University

Sources of inspiration

Where do Swedish tonal accents come from?

- The 'stress clash' theory
- Danish stød, the 'curl'

(Tomas Riad)

How is Fo lowered?

- The AES model of laryngeal mechanisms.
- Fo lowering, creaky voice, glottal stop, Danish stød

(Jan Gauffin)

Learn more about Fo contours in 2010 CS!
(present project)

Outline

1. Revisit the classics

2. Present a numerical model that generates F0 contours from a string of syllabic pulses.

3. It works but what does it mean?

Stetson's chest pulse theory

- Every syllable is characterized by a '*ballistic chest pulse*'.

(based on measurements of rib cage movements, tracheal and pulmonary pressure, some EMG);

Stetson R H (1951): Motor Phonetics

Ladefoged's criticism

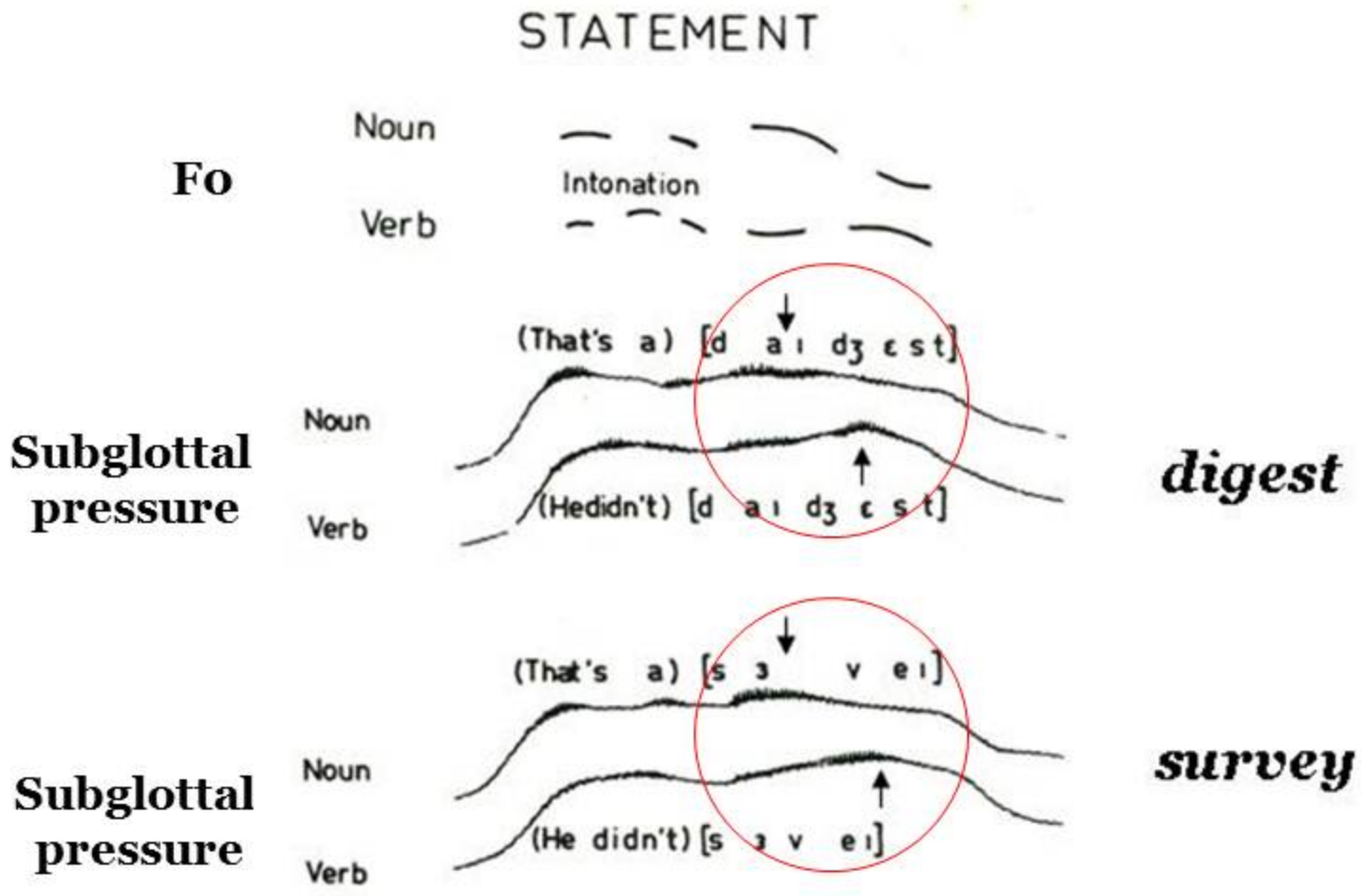
- Stetson's work not technically reliable.

"In our opinion there is certainly insufficient basis for a chest pulse theory of the syllable in normal speech"...

Ladefoged P (1967): Three areas of experimental phonetics

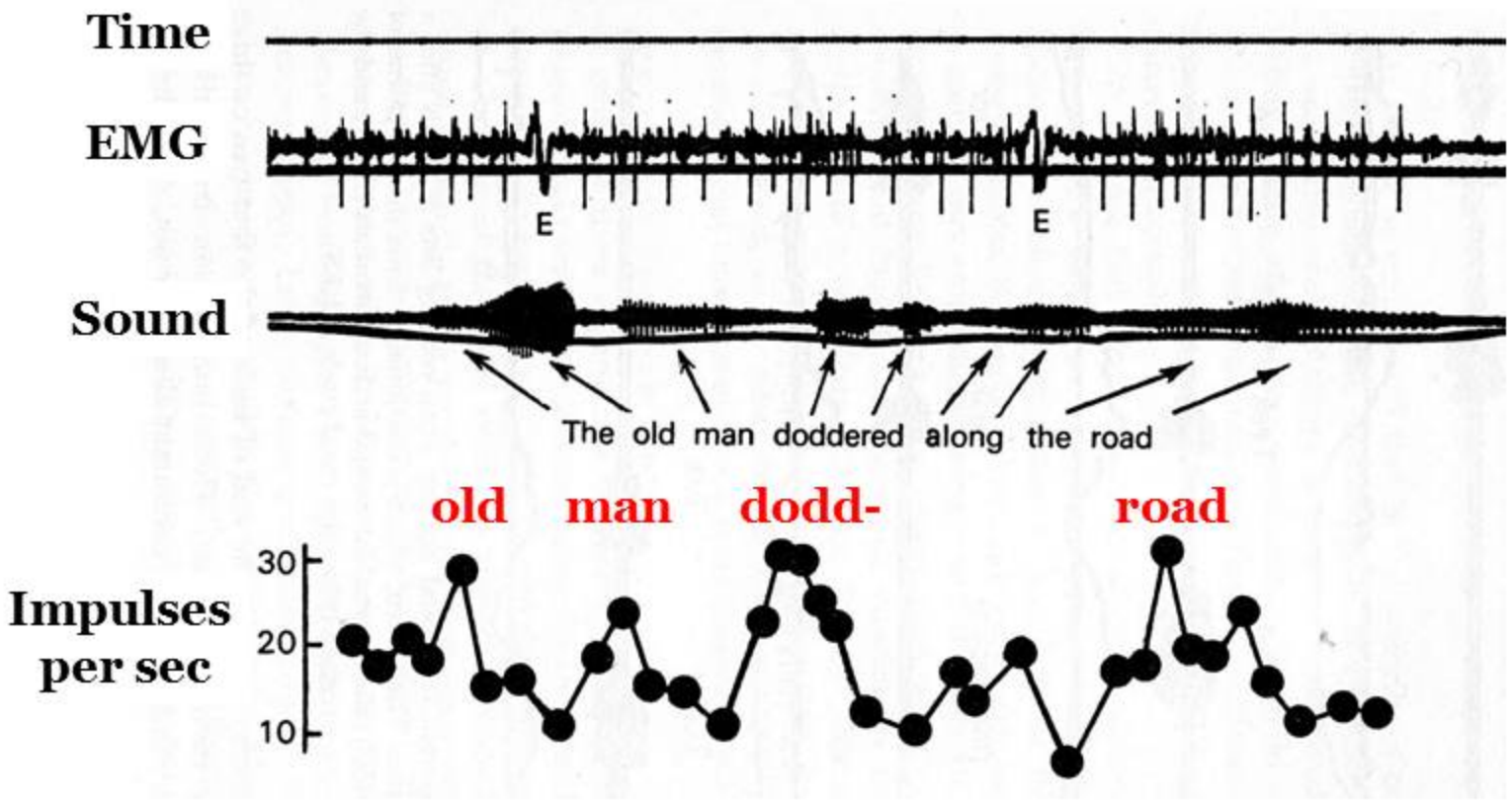
Subglottal pressure & stress

Ladefoged (1967)



EMG from expiratory muscle fiber

Ladefoged (1967)

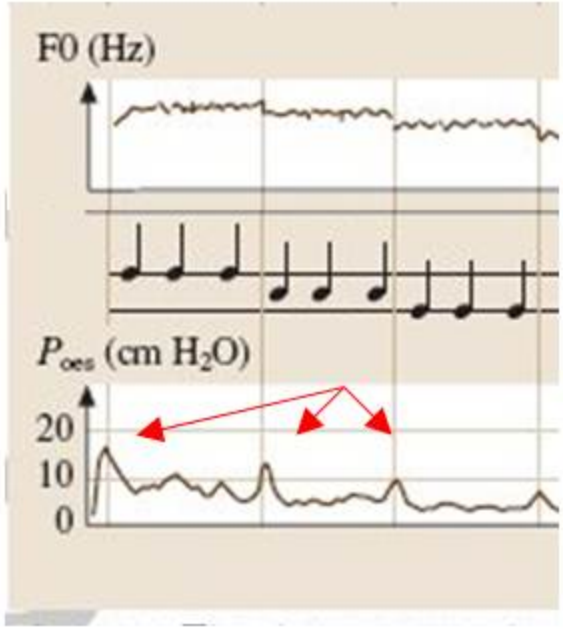


Clues from singing & stage speech

(Sundberg 1995, 2007)

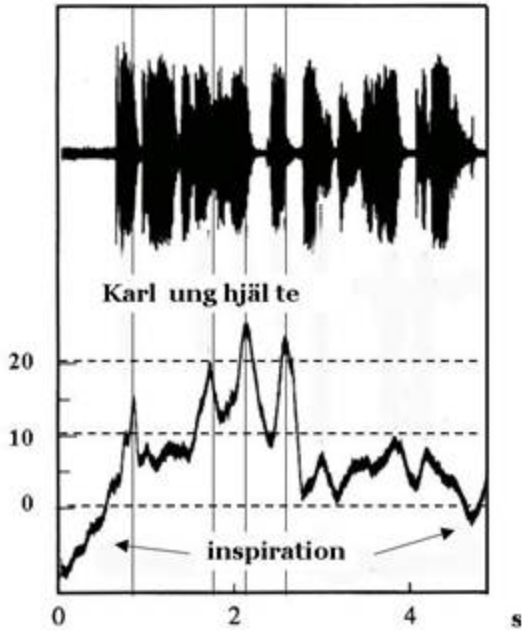
Barytone singing

Ps higher
1st beat of bar



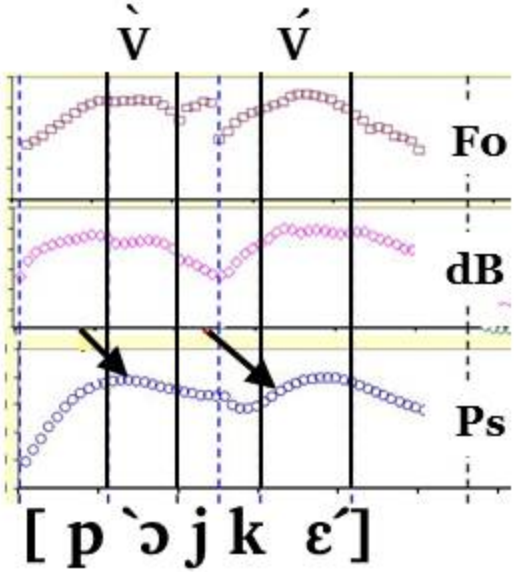
Stage speech

Ps higher
stressed syll's



Reading

Ps higher
on V's A2 word



Ladefoged (2005)

”Concluding …, we note that ***stressed syllables*** may or may not have a greater intensity or a higher pitch, and they may or may not have bursts of internal intercostals activity.

However, they ***always use greater respiratory energy.”***

Ladefoged P (2005): ”Speculations on the control of speech”, ***A figure of speech***, LEA : New Jersey.

A compromise

'Stetson light'

Let us abandon chest pulses for stress pulses

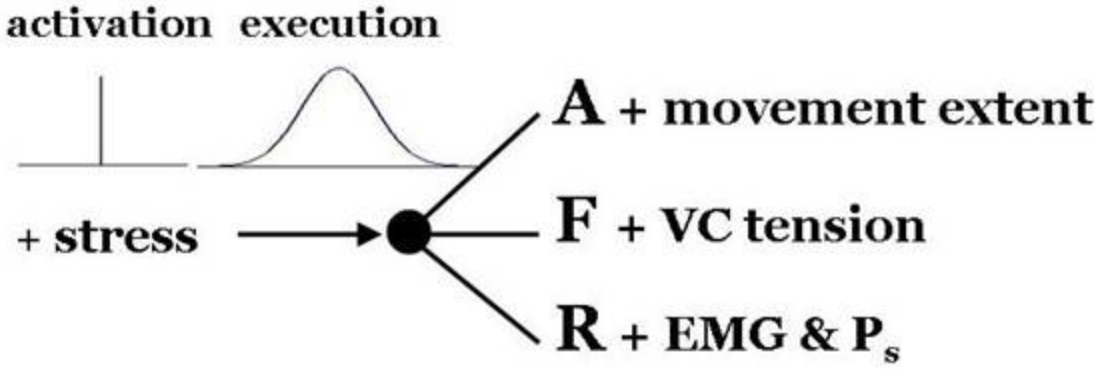
Stress

Jespersen's view

"Akzent (Druck) ist Energie, intensive Muskeltätigkeit, die nicht an ein einzelnes Organ gebunden ist, sondern der gesamten Artikulation ihr Gepräge gibt."

"Druck als Gesamtenergie"

Jespersen O (1926): ***Lehrbuch der Phonetik***, Teubner:Leipzig.

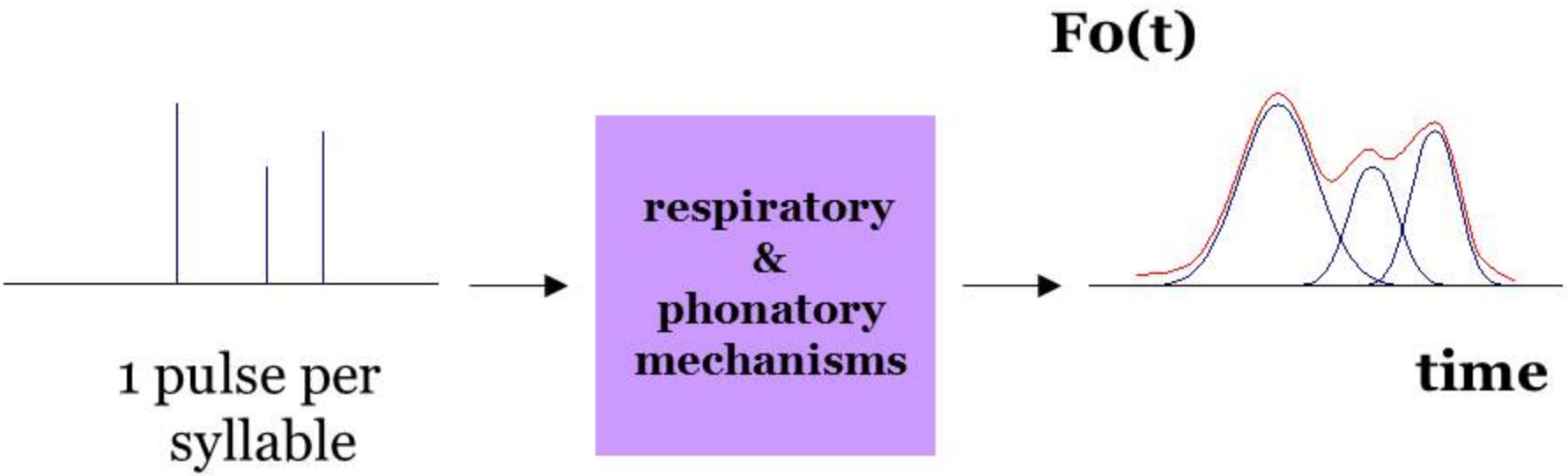


Modeling Fo contours using stress pulses

Assumptions

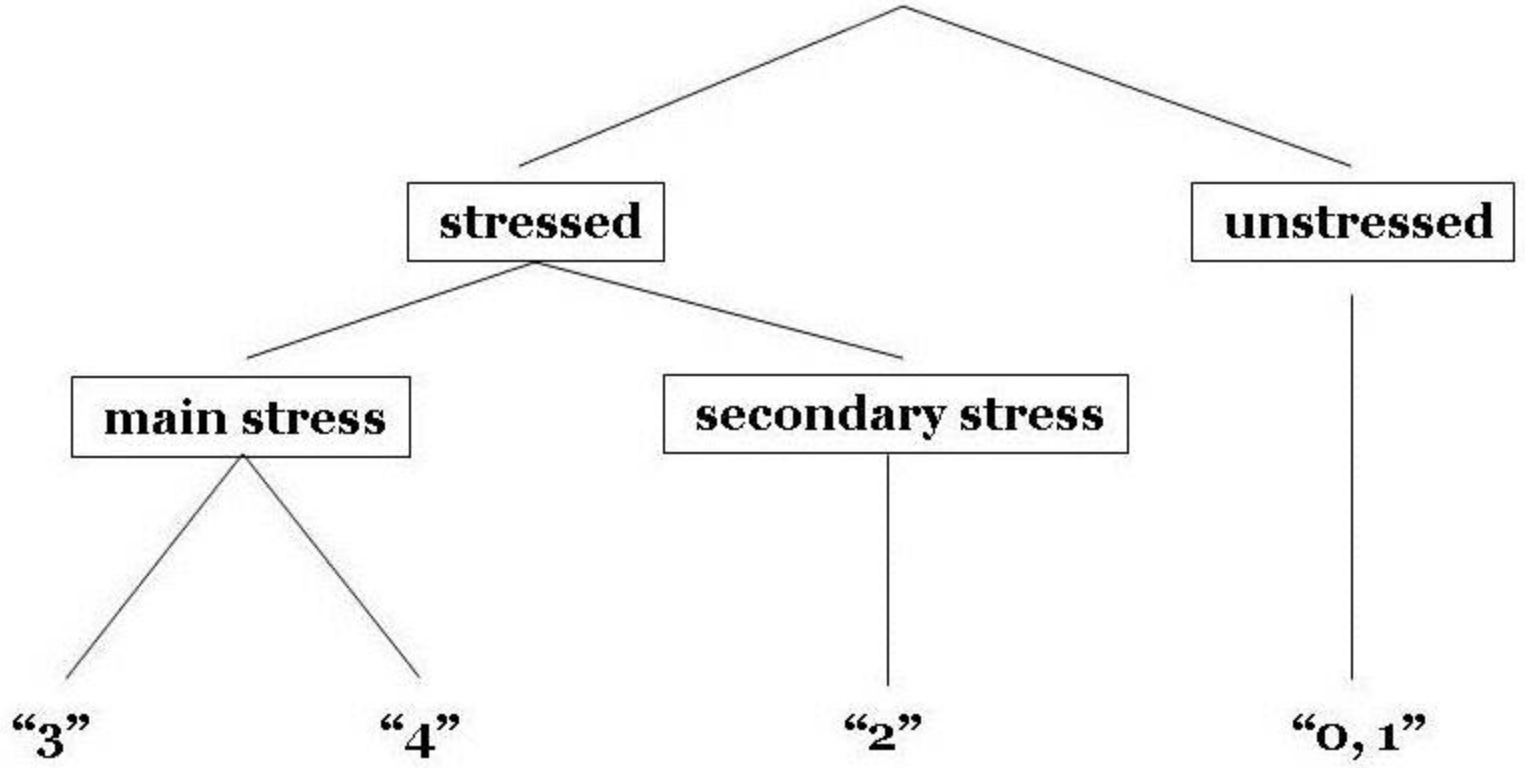
- *Stress* involves a strengthening of all the physiological activities (respiratory, phonatory and articulatory) that the production of a *syllable* gives rise to;
- Every syllable is produced with a certain physiological '*gain*', in other words a *quantum of energy* injected into the motoric systems of respiration phonation and articulation
- The input is idealized as an instantaneous event, a *pulse*.
- The output of is a smooth *bell-shaped curve* reflecting sluggish response characteristics.

Can Fo contours be analyzed as a summation of a string of syllabic pulses?

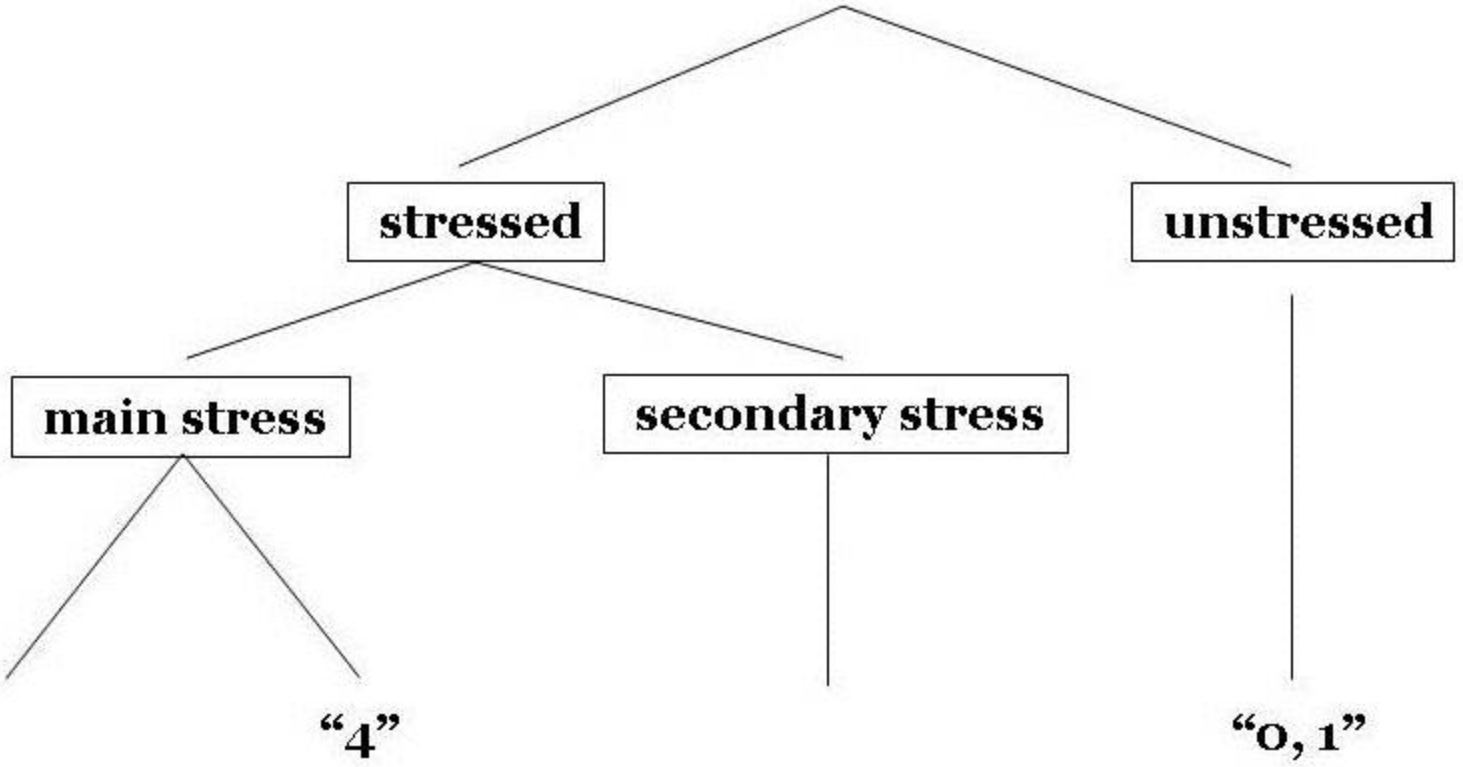


Speech samples and measurements

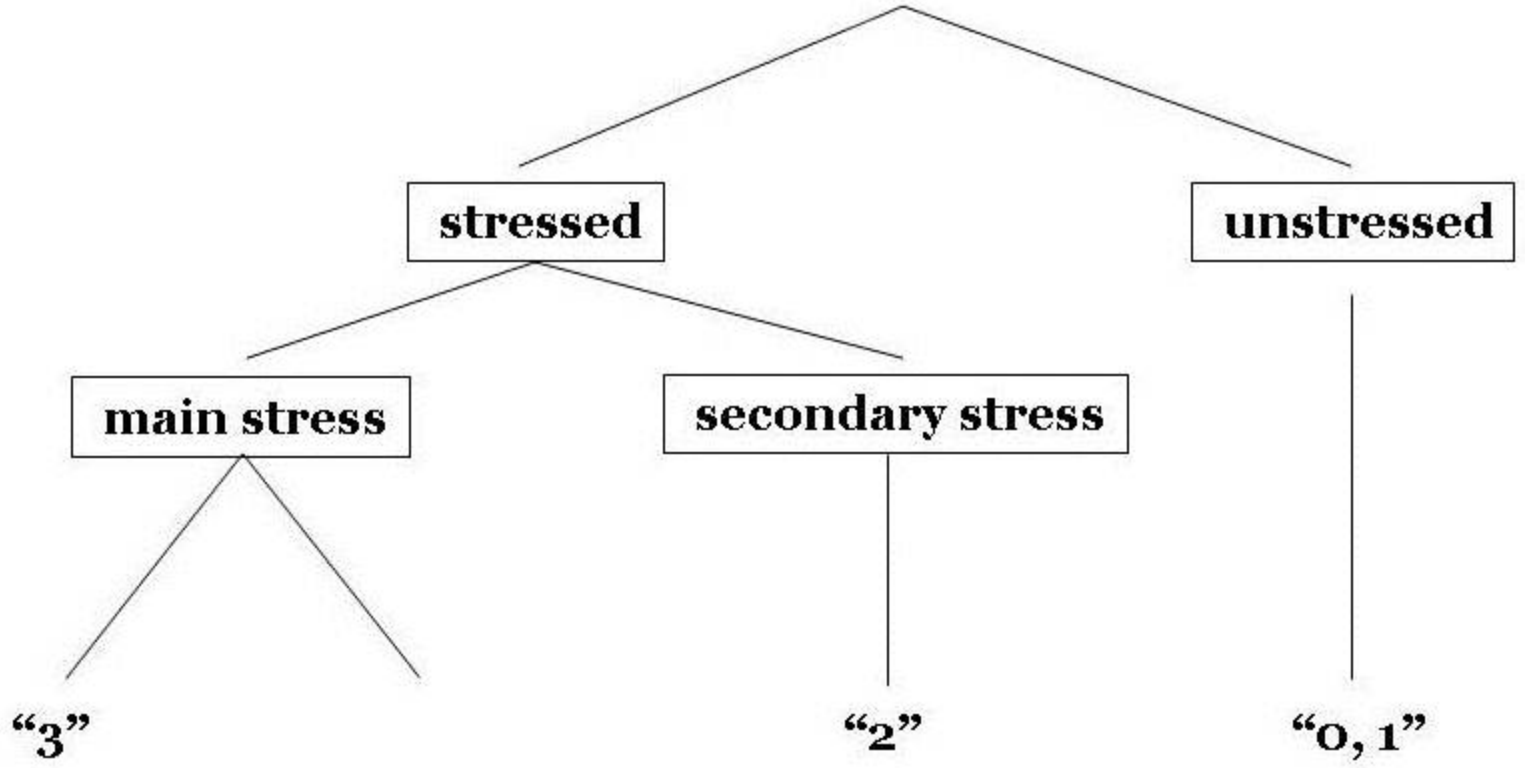
Prosodic syllable types (SAOB)



Words with accent I



Words with accent II



in compounds

Some words & phrases

4	sup
40	fjäder
04	pannå
400	kritiker
040	kaninen
004	logistik
4000	serierna
0400	botaniker
0040	intressera
0004	professionell
44	röd bil
444	Bengt går fort
440	Ulf läser
404	bra musik
044	Katrin vet
4004	Ted muckar gräl
40004	Per leker polis

32	kråkspark
320	brevlåda
302	motorbåt
032	miljövård
3200	busgrabbarna
3020	Londongatan
3002	väderprognos
0320	rabattkortet
0302	polispiket
0032	pyramidform

Speech samples

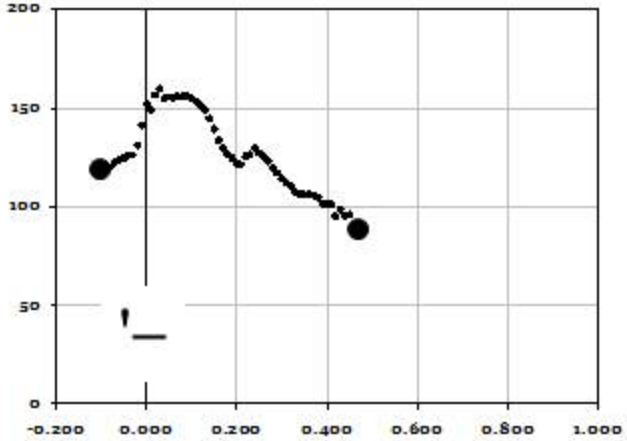
- **Words** 1-3 syllables long
- **Accent 1** & **Accent 2**
- **2^{ary} stress**
- **Unstressed** syllables
- **Phrases** 2-6 syllables long by combining the word forms.

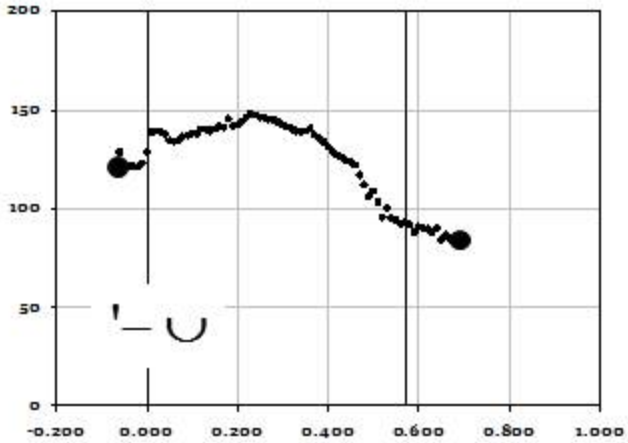
Fo data

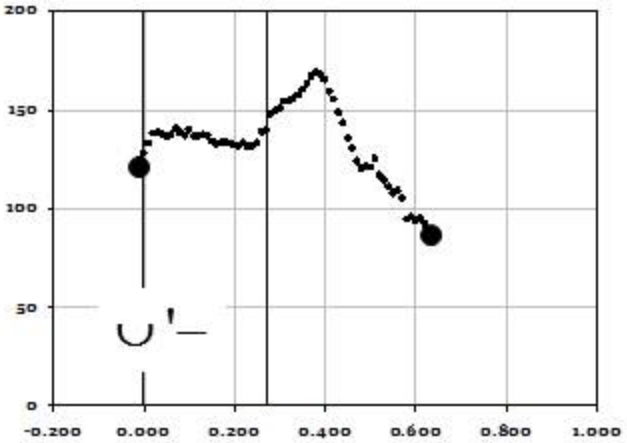
- Male speakers
- Central Swedish (Stockholm)
- 3-5 rep's of each test item

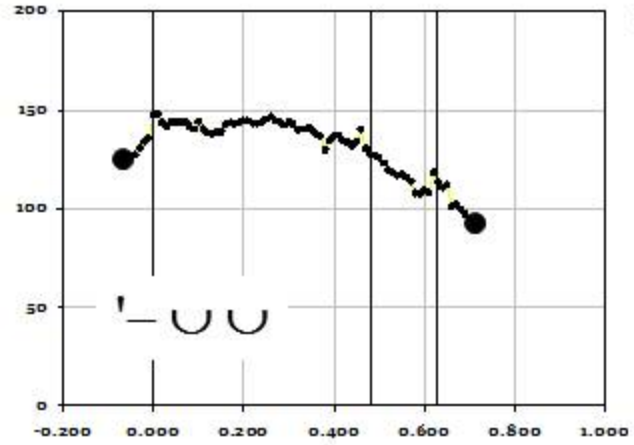
- Wavesurfer pitch tracker
- Fo curves synchronized @ vowel onsets and averaged.

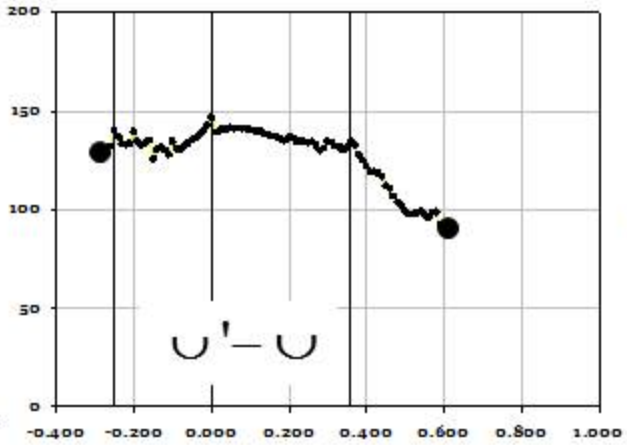
Accent I

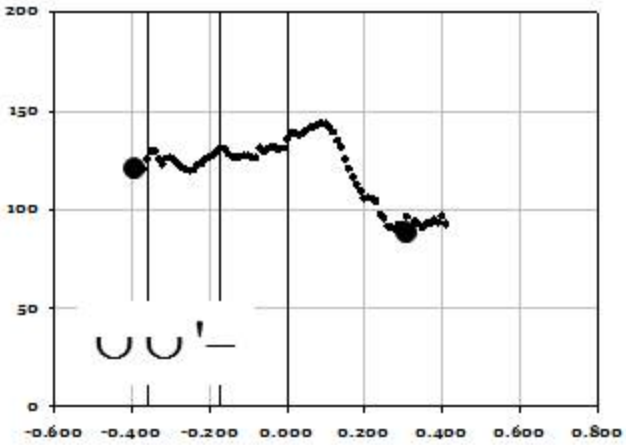

200
150
100
50
0
-0.200
0.000
0.200
0.400
0.600
0.800
1.000
'—


200
150
100
50
0
-0.200
0.000
0.200
0.400
0.600
0.800
1.000
'— ∪

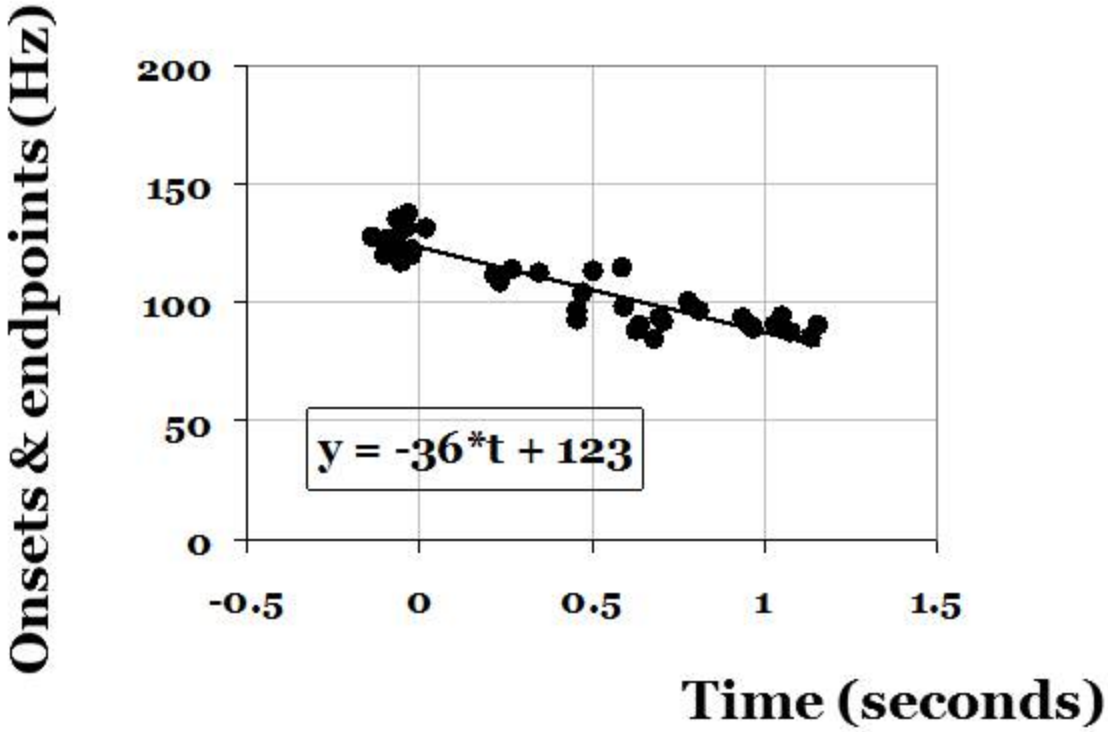

200
150
100
50
0
-0.200
0.000
0.200
0.400
0.600
0.800
1.000
∪ '—


200
150
100
50
0
-0.200
0.000
0.200
0.400
0.600
0.800
1.000
'— ∪ ∪

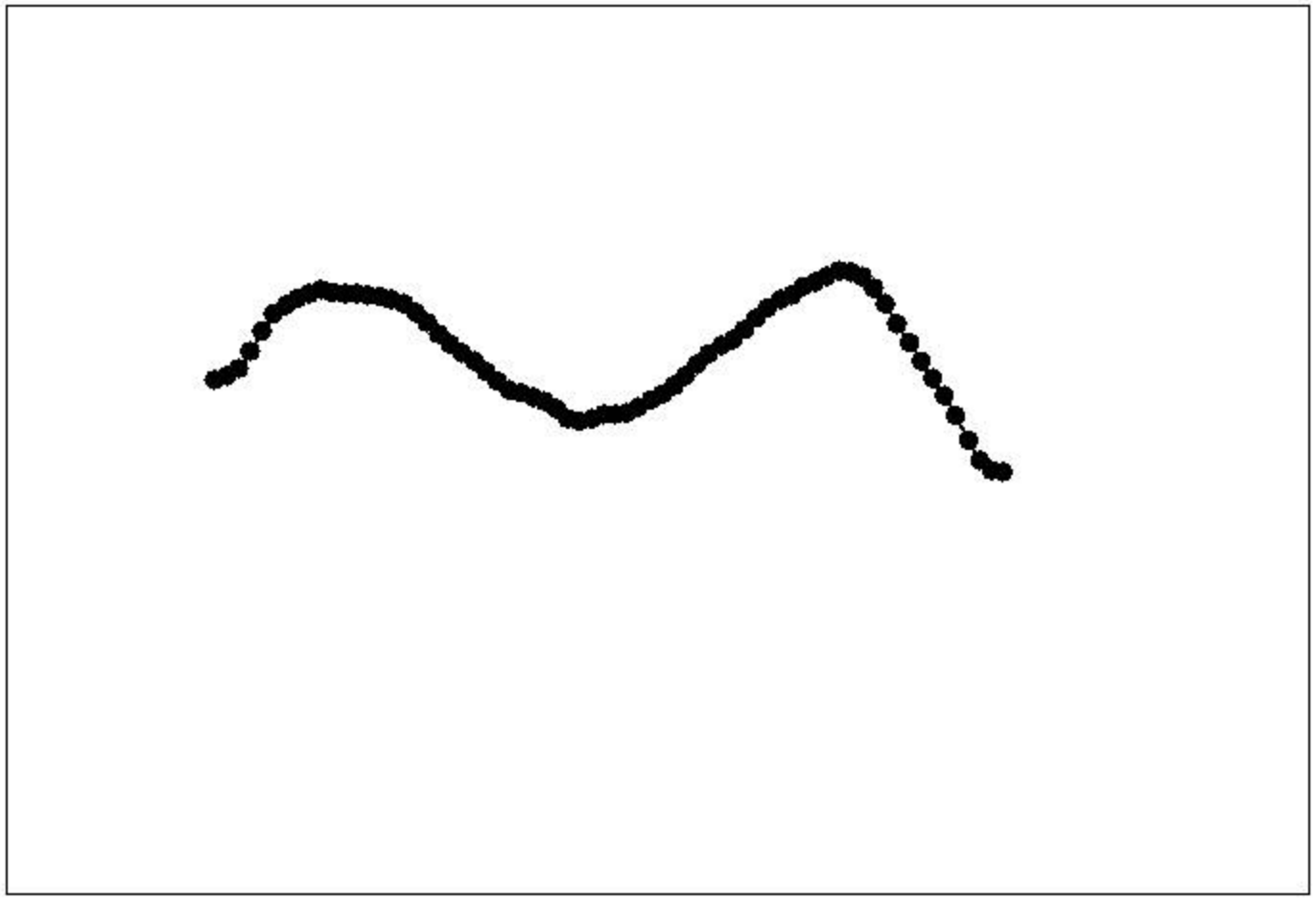

200
150
100
50
0
-0.400
-0.200
0.000
0.200
0.400
0.600
0.800
1.000
∪ '— ∪


200
150
100
50
0
-0.600
-0.400
-0.200
0.000
0.200
0.400
0.600
0.800
∪ ∪ '—

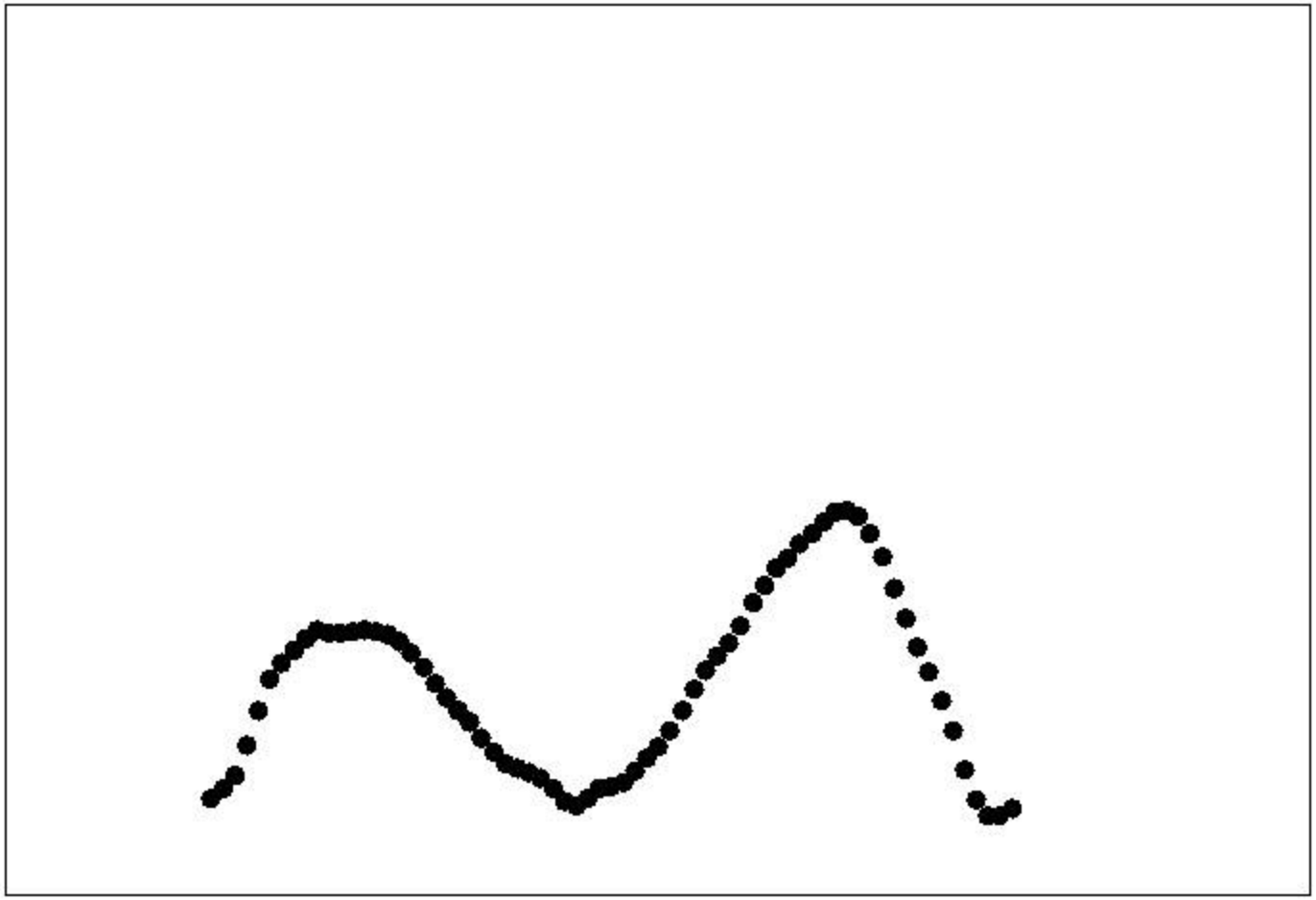
The declination component



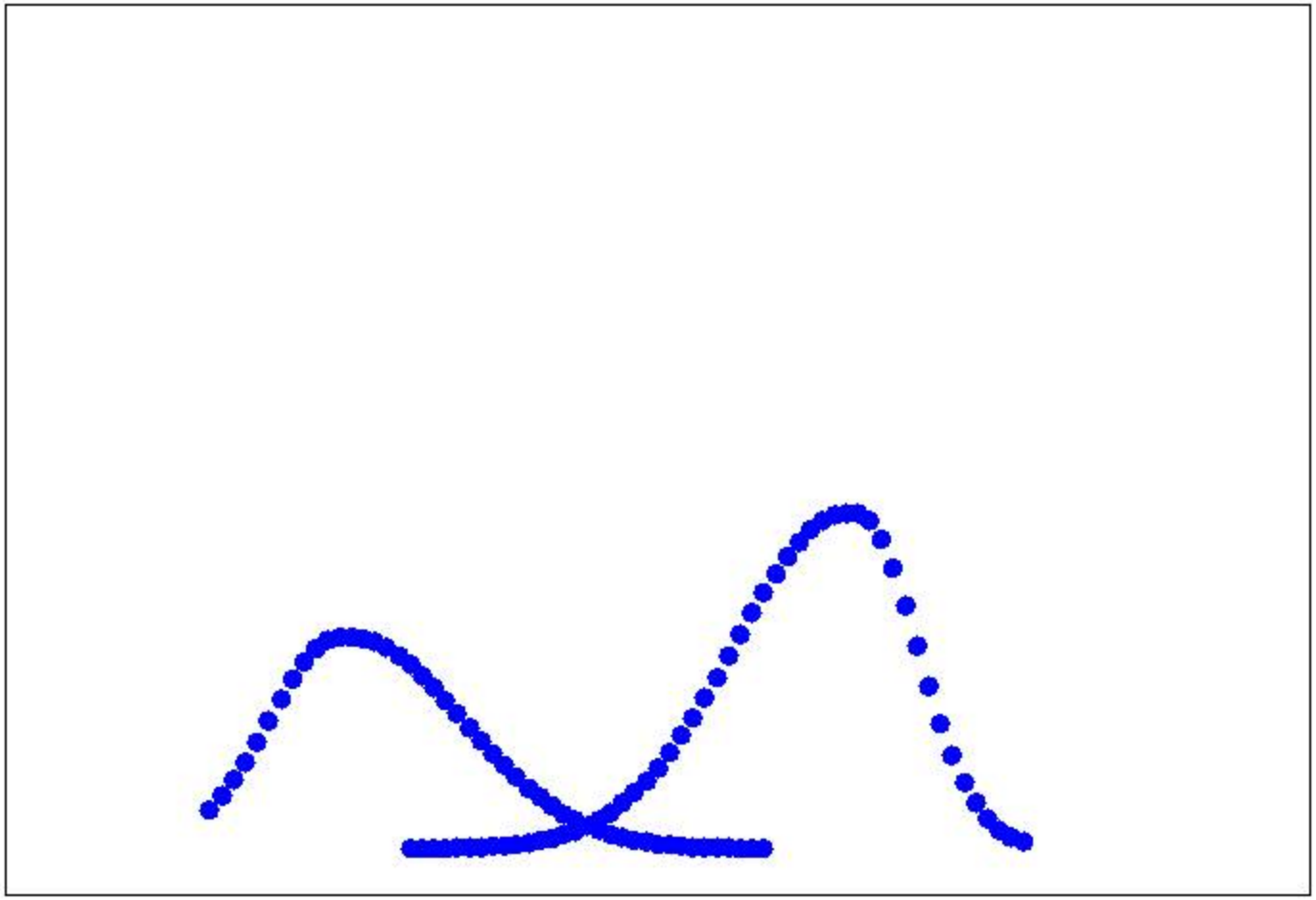
Raw data



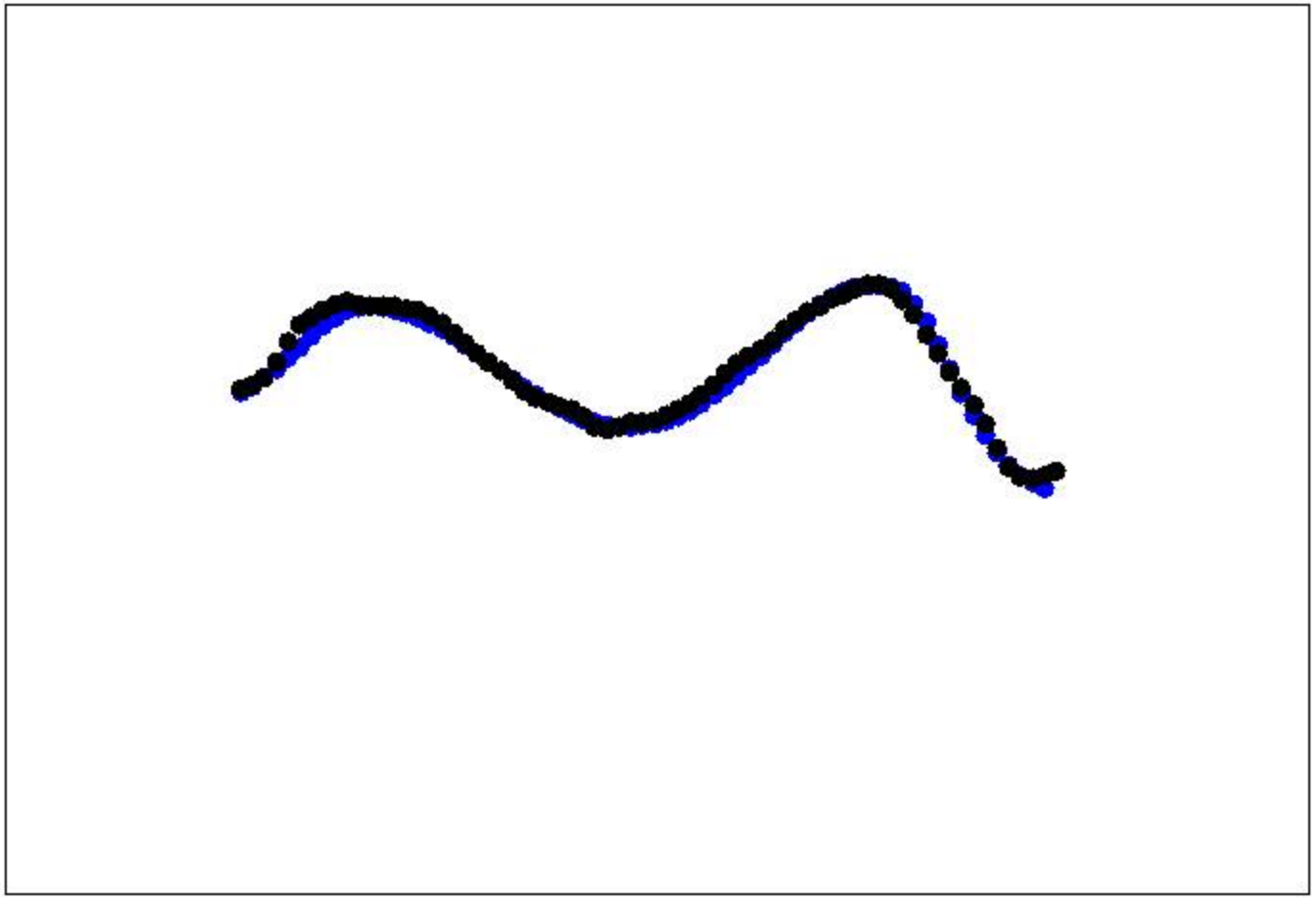
Rotate & translate!



Pulse analysis



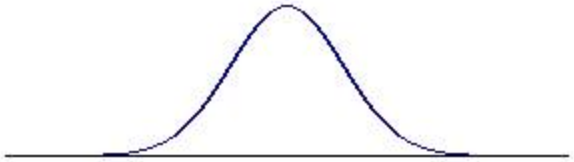
Sum the pulses, Undo translation & rotation!



Ground rules

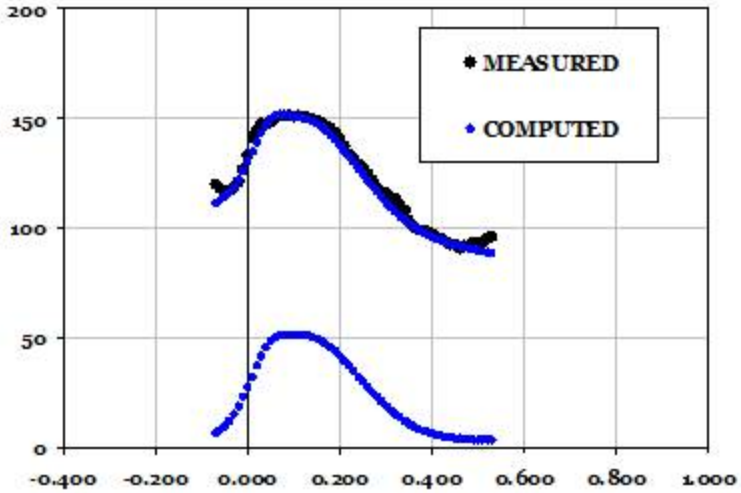
- Shape of pulse: Gaussian

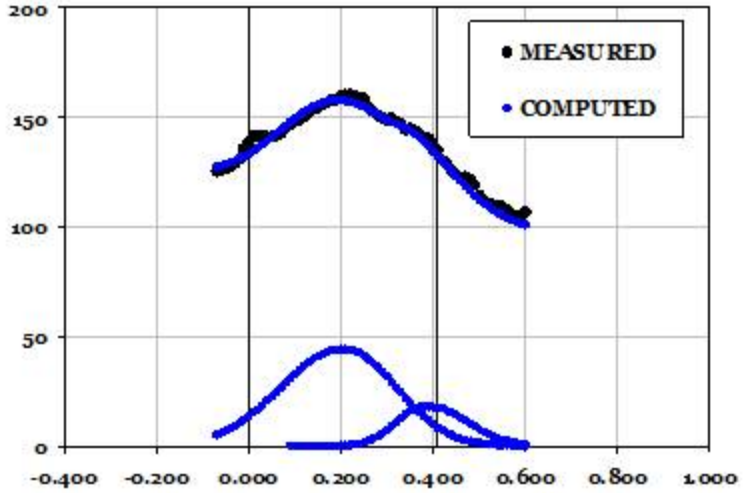
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

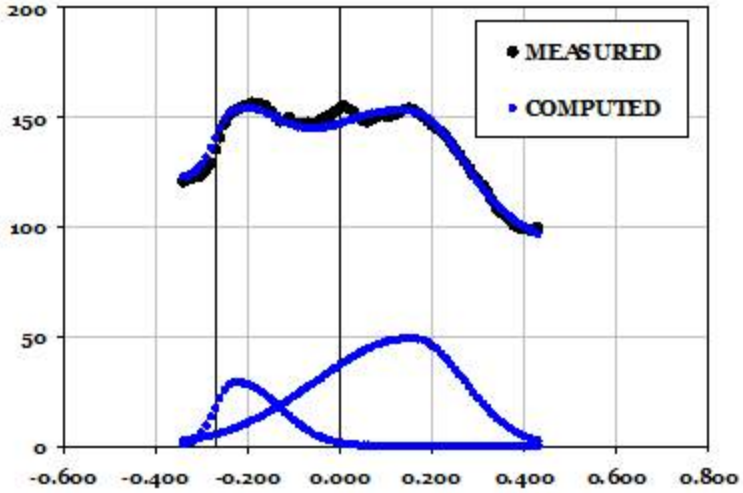


- Parameters:
  - ❖Time of peak
  - ❖Amplitude
  - ❖Fall & rise of slopes

Accent I words

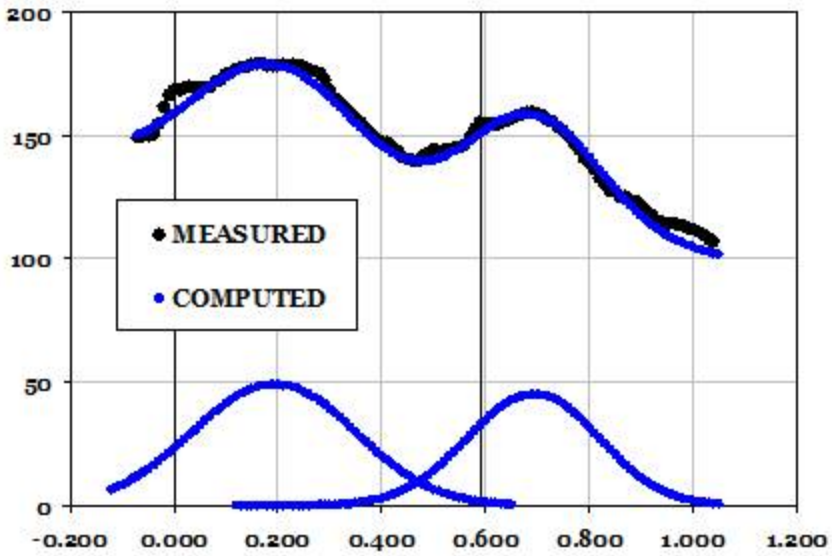




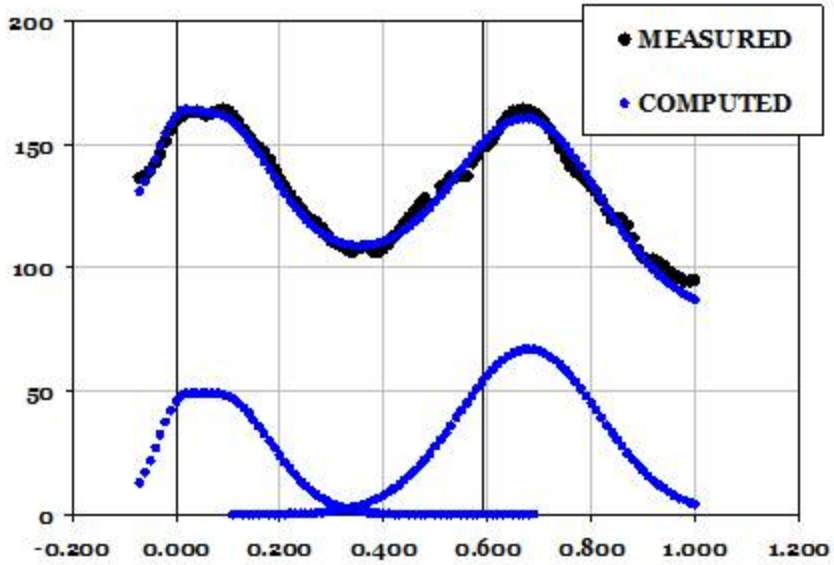


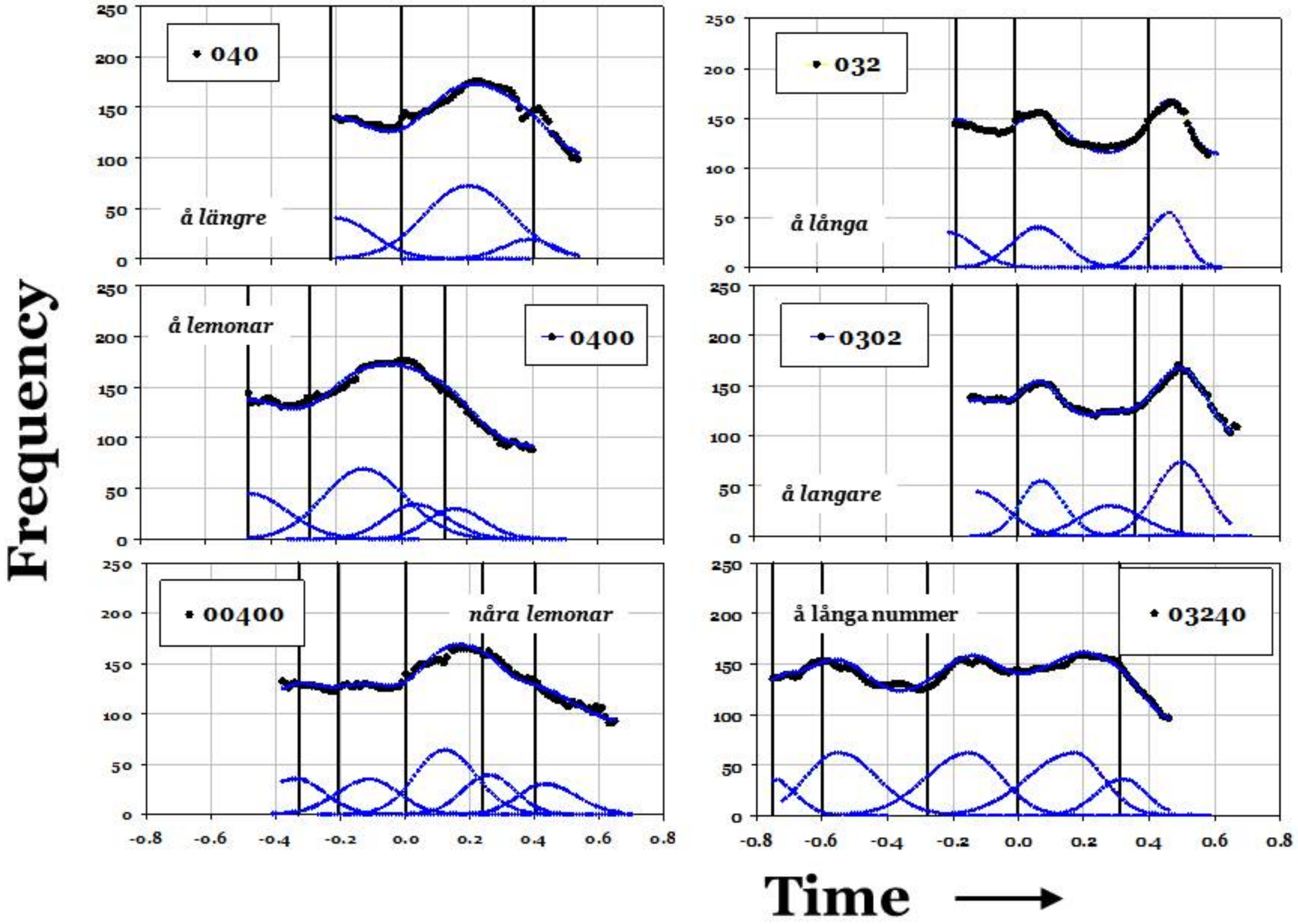
Comparing accent I and II

“grå ben”



“gråben”




Frequency
040
å längre
032
å långa
å lemonar
0400
0302
å langare
00400
nåra lemonar
å långa nummer
03240
Time

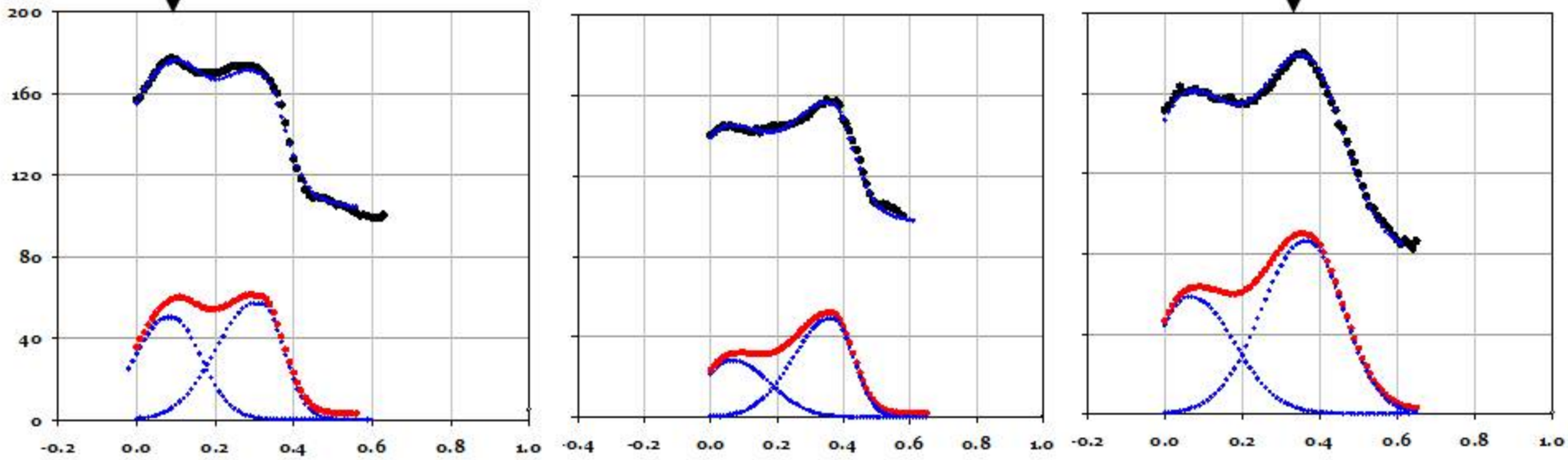
Emphatic stress

"Per går"

Emphasis on 1 st syllable

equal stress

Emphasis on 2 nd syllable

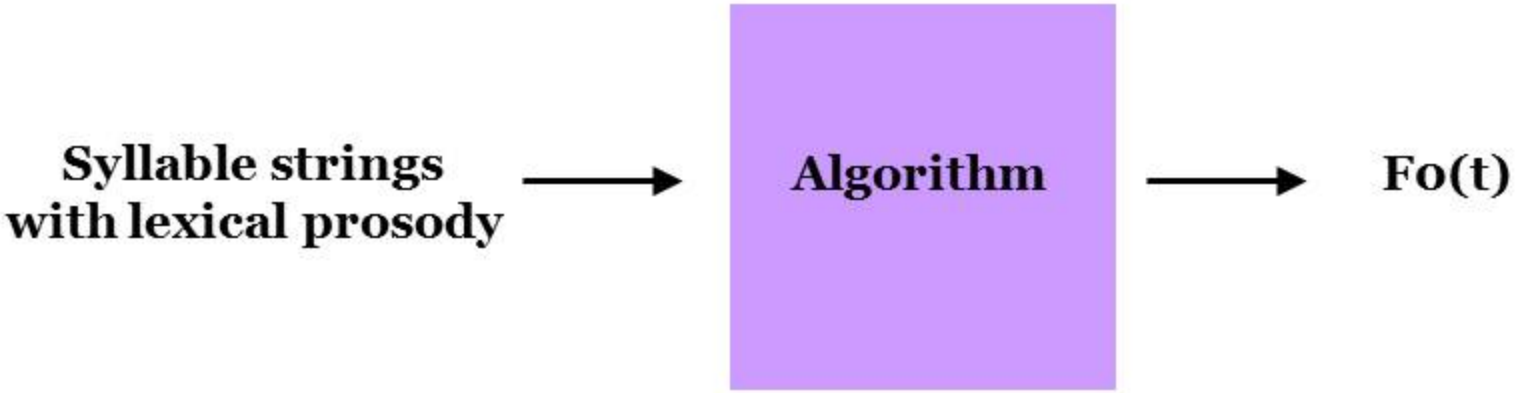


Observations

- Numerically accurate descriptions
- 1 pulse per syllable
- Pulse amplitudes correlate with 2 degrees of lexical stress
- Distinct representations of Accent I & II

Mere curve fitting?

A model for speech tech?



The explanatory claim

F0 contours of Central Swedish tonal accents

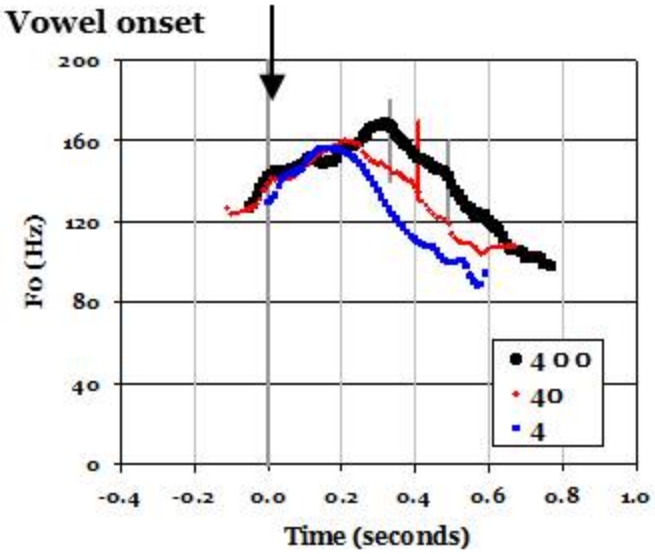
... are the way they are because they are close to *the unmarked (default) response of the pitch system to the Swedish form of syllabic stress*.

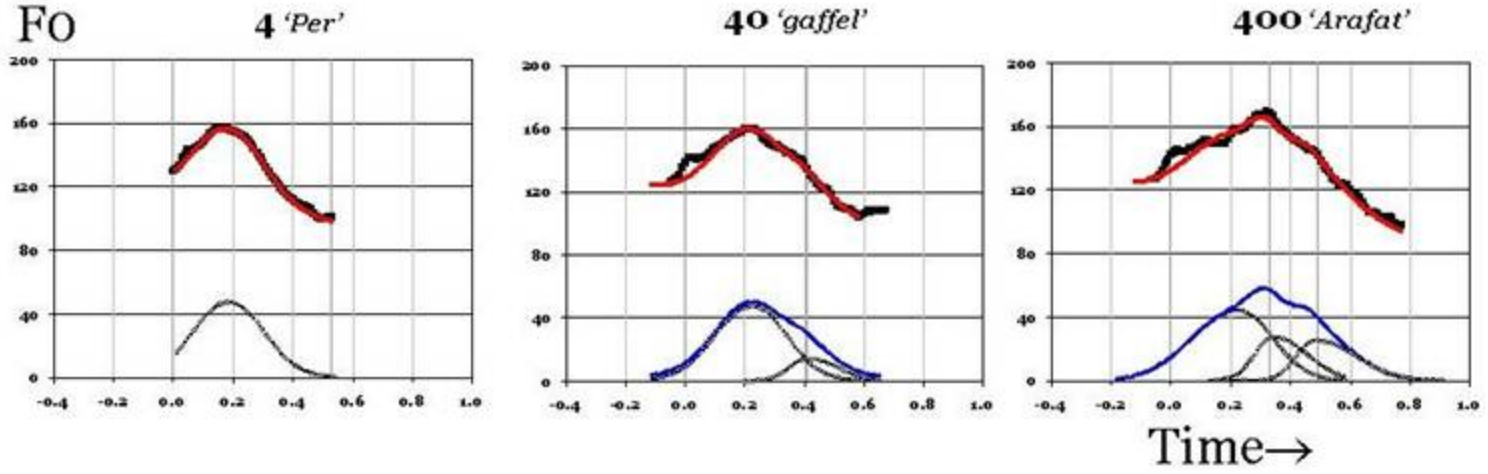
The pulse theory account

- Fo properties of SC tonal accents
  - Tonal accents occur only in stressed syllables
  - They derive from stress pulses
  - They form bell-shaped curves
- Fo properties of unstressed syllables
  - Also derive from stress pulses (but with lower amplitudes)
- Fo time properties (digital to analog conversion)
  - The details of their temporal integration arise from summation of pulses.

Any problems?

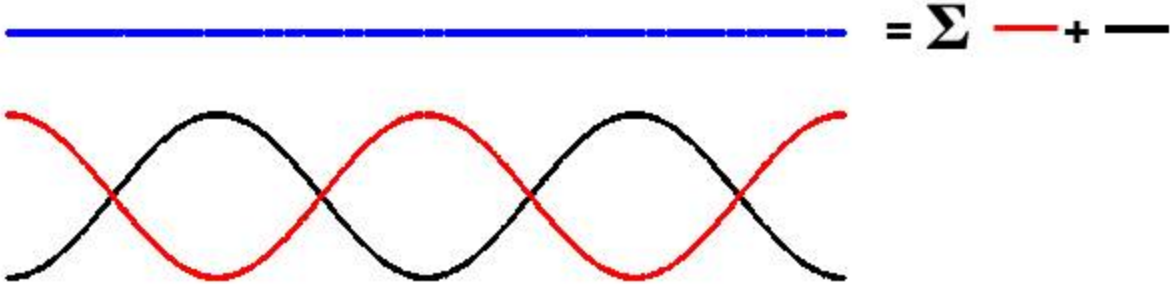
What about 'peak delay' effects?





Smooth and seamless yet pulsed?

Q: How come a continuous curve like Fo can be credibly be parsed into a series of discrete pulses?



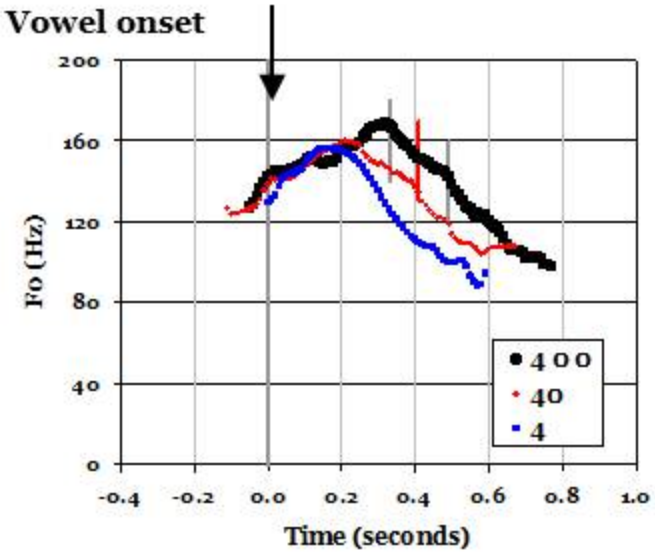
Clue: 2 sine waves added = horizontal line

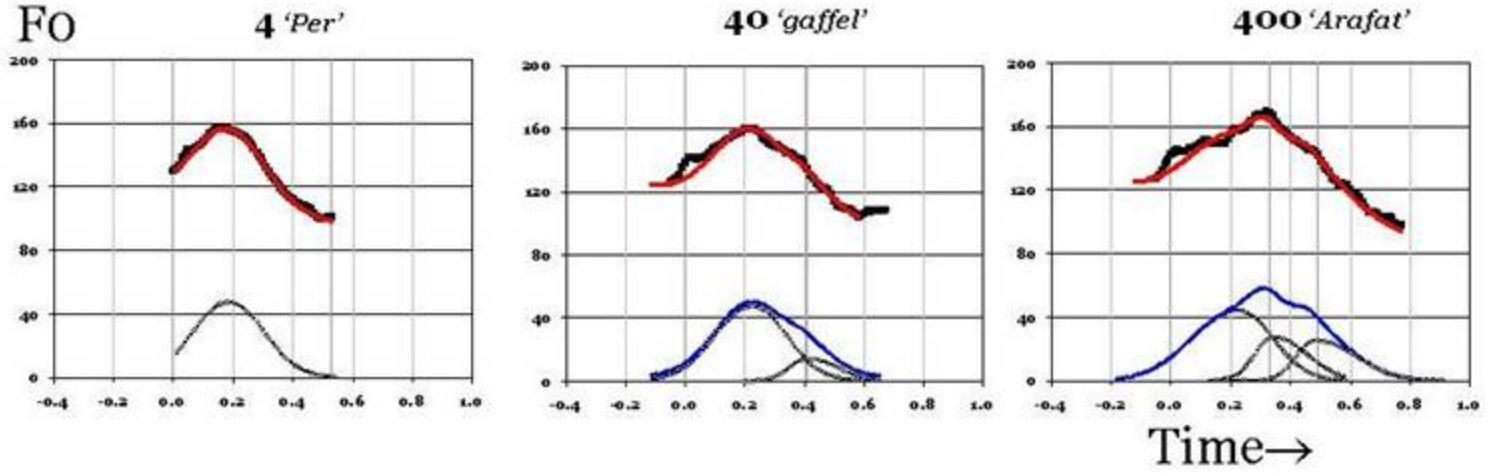
Stetson lives again!

There is more than meets the eye!

Any problems?

What about 'peak delay' effects?





What about laryngeal pitch control?

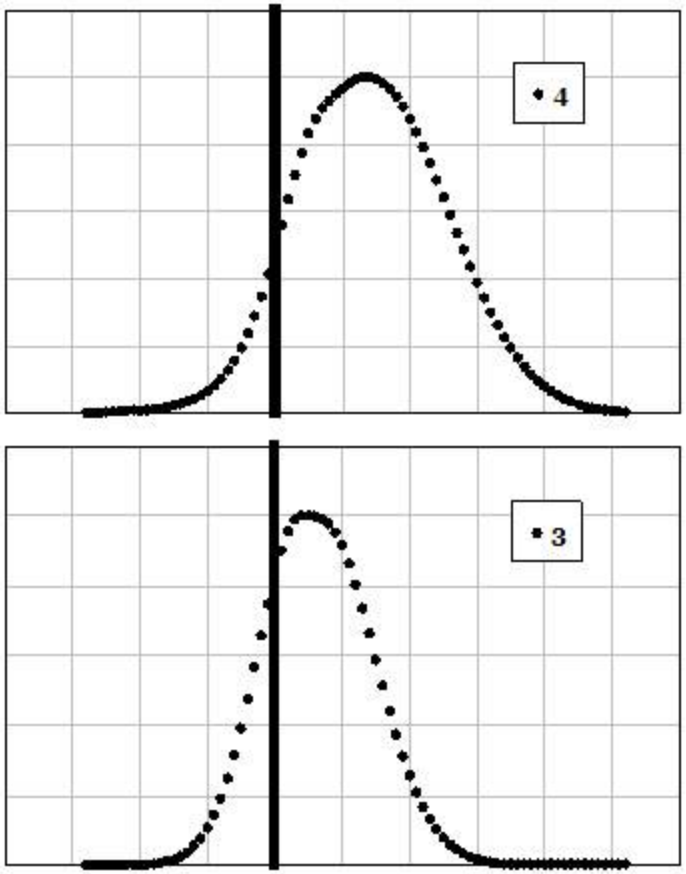
- Default case: If the tonal accent contours get their properties from "*the default response of the pitch system to syllabic stress*", we should expect a tight time locking between pulses and the segmental events (vowel onsets).

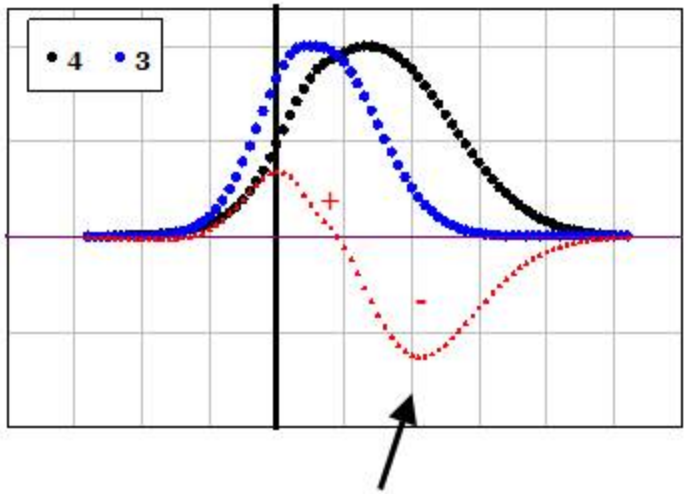
- In fact, although the F0 mechanism is here assumed to get such a contribution from syllable production, it must also be expected to be able to work independently on top of that input.

- The thing to be surprised about in the case of Swedish tonal accents is this: the default model seems to do pretty well without such contributions. No extra curlicues and wiggles!

Canonical pulse shapes

Comparing 4's and 3's

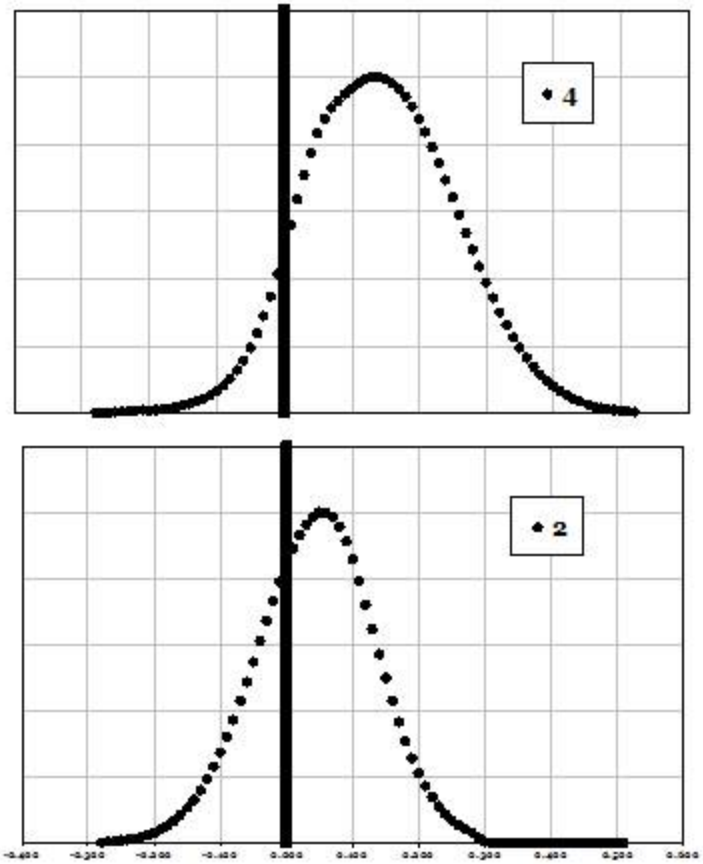


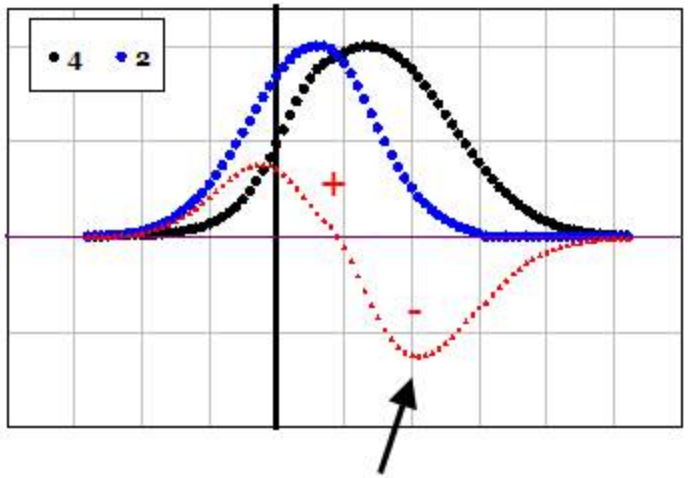


Difference curve
low tone
for accent II?

Canonical pulse shapes

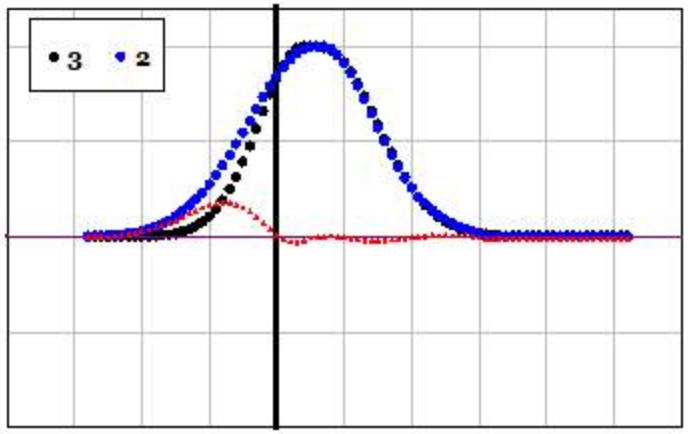
Comparing (non-final) 4's with (final) 2's





Difference curve
low tone
for terminal juncure?

3 and final 2 are very similar



Hardly any difference!

Can they be switched
with impunity?

Some samples 

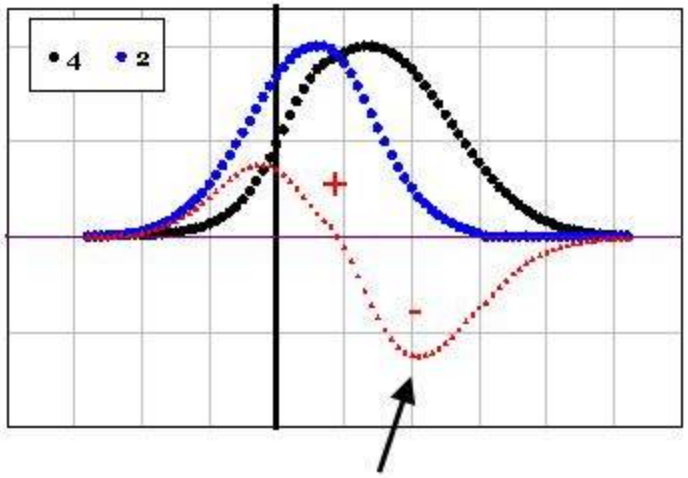
Recorded (32 pattern)	**After tape splicing**
”Lundberg”	*”Berglund”*
”matpersonal”	*”personalmat”*
”sportbil”	*”bilsport”*
”miljöakademi”	*”akademimiljö”*
“blågrå”	*“gråblå”*

Samples of ‘stress clash’

(Tomas Riad)

Old Norse

sammansättning	ˈland‚búi
	ˈsól-‚hvarf
avledning[9]	ˈsjúk-‚dómr
	ˈklók-‚skapr
	ˈmann-‚likr
	ˈvík‚ingr
stamsuffix	*ˈtung-‚o:
	*ˈdrup-‚an
böjning	*ˈherði-‚o:z
	*ˈdo:mi-‚de:



Difference curve
low tone
for terminal juncure?

Progress report

Where do Swedish tonal accents come from?

- Accent 2 HL pattern may arise from **re-use** of laryngeal gesture already in place in final main stress syllables (terminal juncture).

How is Fo lowered?

- The AES laryngeal mechanism of Fo lowering [and creaky voice, glottal stop, Danish stød]

Game plan and challenges

- Pitch patterns as deviations from stress-based default contours.
- Apply to historical and typological data
- Extend beyond Swedish

Thank you!