

# Phonetic modifications of AM categories

Tamara Rathcke

<http://www.kent.ac.uk/secl/ell/staff/rathcke.html>



***Aix Summer School on Prosody 2016***

**07/09/2016 - 16:30-18:30**

University of  
**Kent**

# There will be questions!

<https://b.socrative.com/>

STUDENT

PHONISTA

JOIN ROOM

Enter your name

geek\_chick

DONE

# Goal of this lecture

Discuss F0-modifications that can happen to

1. **scaling** (= pitch height) of **H & L** (phonological tone primitives)
2. **alignment** (= timing of pitch targets) of associated **H\* & L\*** - in relation to the accented syllable
  - ✓ Technical (erroneous)
  - ✓ Real but misleading
  - ✓ Systematic and predictable (and can also be misleading)

*Lead to problems in the phonological generalisation if these not accounted for*

# Pitch categories

## Core questions:

- *What kinds of pitch pattern can occur and serve linguistic functions in language X?*
- *How can they be systematically described and classified?*

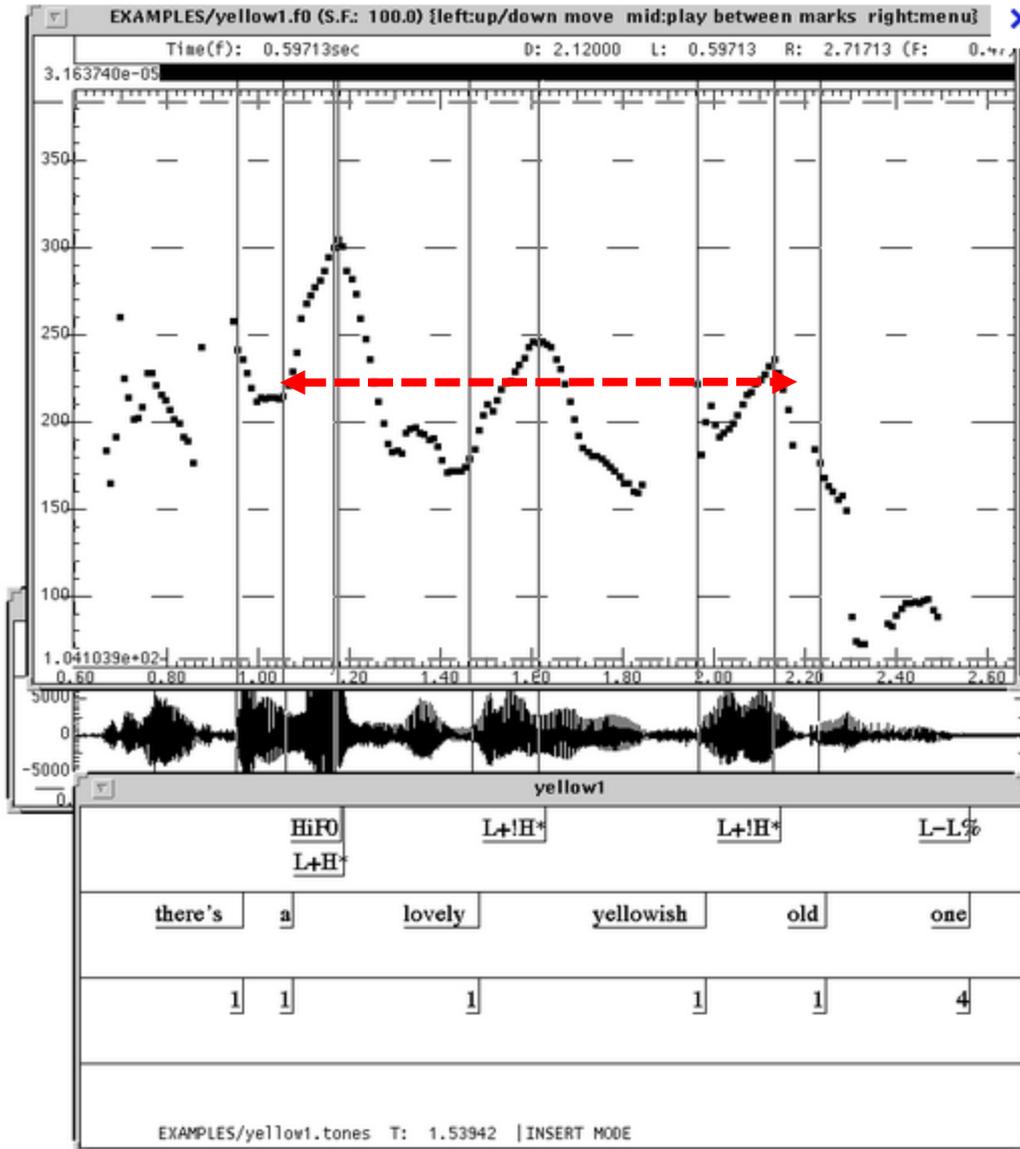
## Core evidence:

- *Acoustic: databases* (a list of sentences read by trained speakers, unconstrained corpora)
- *Perceptual: experiments* (systematically manipulated pitch tracks to derive listeners' responses)

# Intonational phonology: What is essential?

- In older schools of intonation, **frequency** is often ‘phonologised’
  - e.g. *low fall* vs. *high fall*, *high rise* vs. *low rise*
- In AM, frequency domain is only reflected in the H & L units (intonational primitives); **time** domain is ‘phonologised’
  - How do we decide on these categories?

# Intonational primitives: H and L (M rare, !H frequent)



*There is a lovely  
yellowish old one.*

L+H\*

L+!H\*

A long phrase:

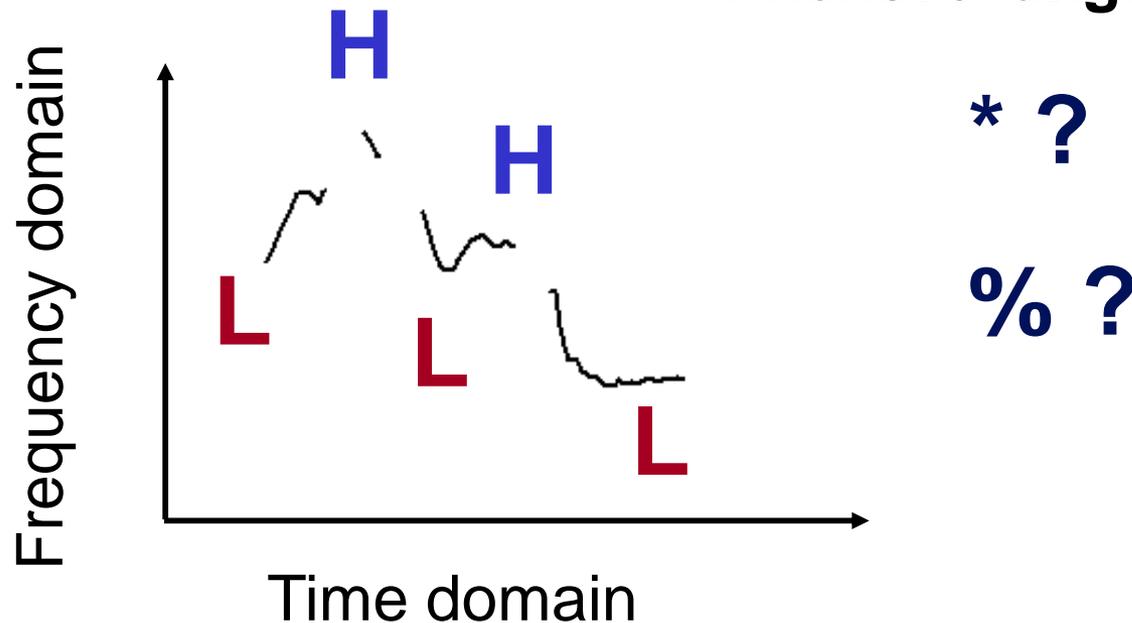
L at the beginning ≈  
H at the end

*Example from ToBI  
manual*

# Phonologised time domain

Temporal coordination between tonal primitives H & L and the accented syllable is essential:

- Phonological: **association** (\*)
- Phonetic: **alignment**



# AM categories

L+H\*

H\*

H+L\*

L\*+H

H-

L-

L-H%

etc.

A perception experiment

Socrative task #1



*She's gone to Malaga.*

11 repetitions

**1 2 3 4 5 6 7 8 9 10 11**

Listen to the melody.

Can you perceive any changes?

If yes, which repetition?

# SOCRATIVE TASK #1

# SSBE listeners from Cambridge

Two perception shifts (Kleber 2006):

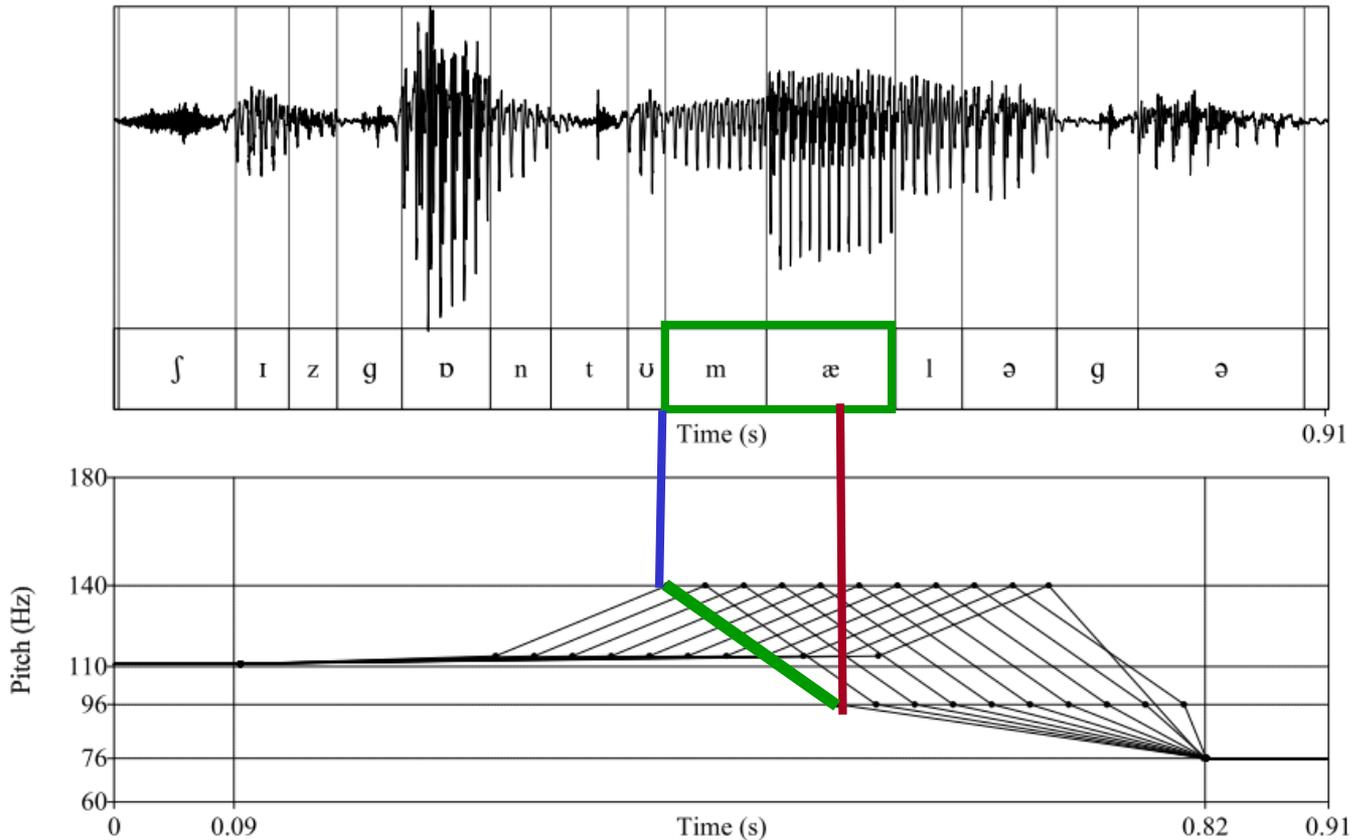
*She's gone to Malaga.*

1 2 3 4 5 6 7 8 9 10 11

# Exploiting time domain

H

L

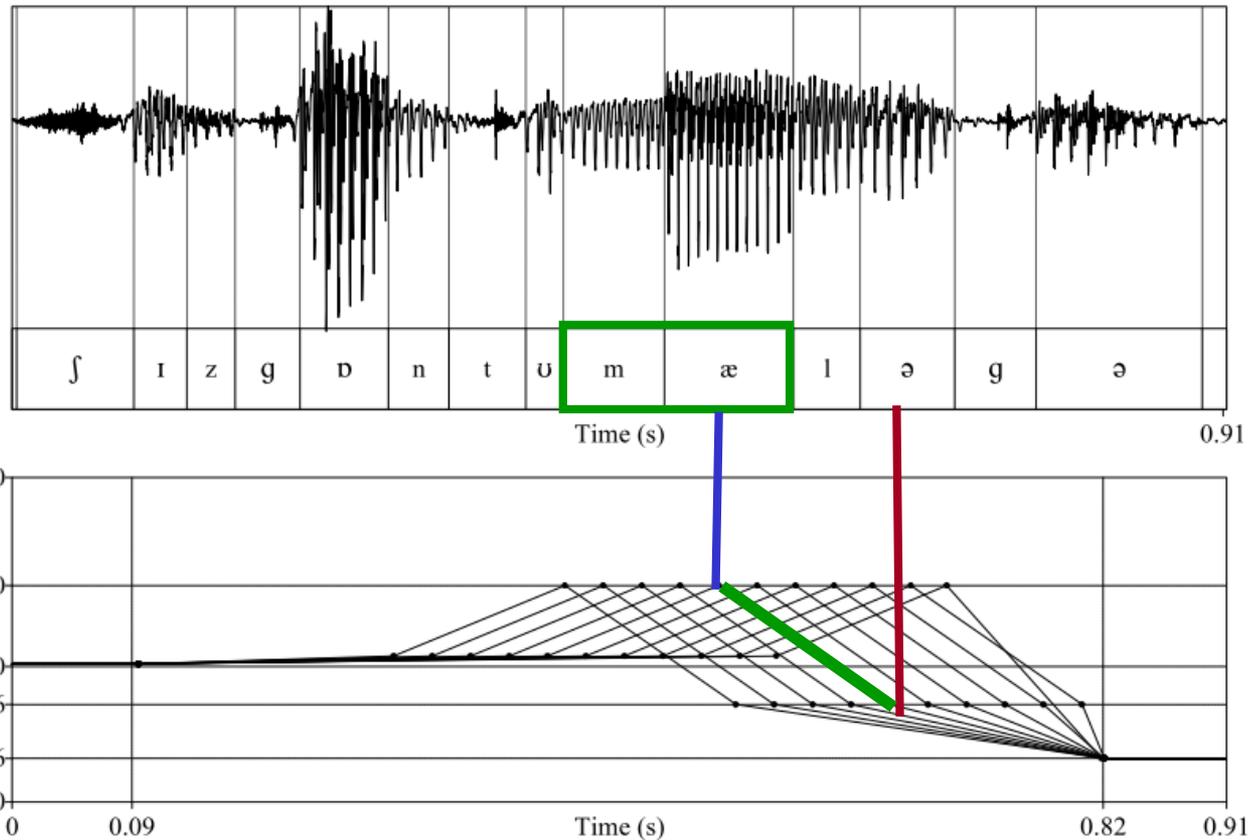


~ “early peak” : H+L\*

# Exploiting time domain

H

L

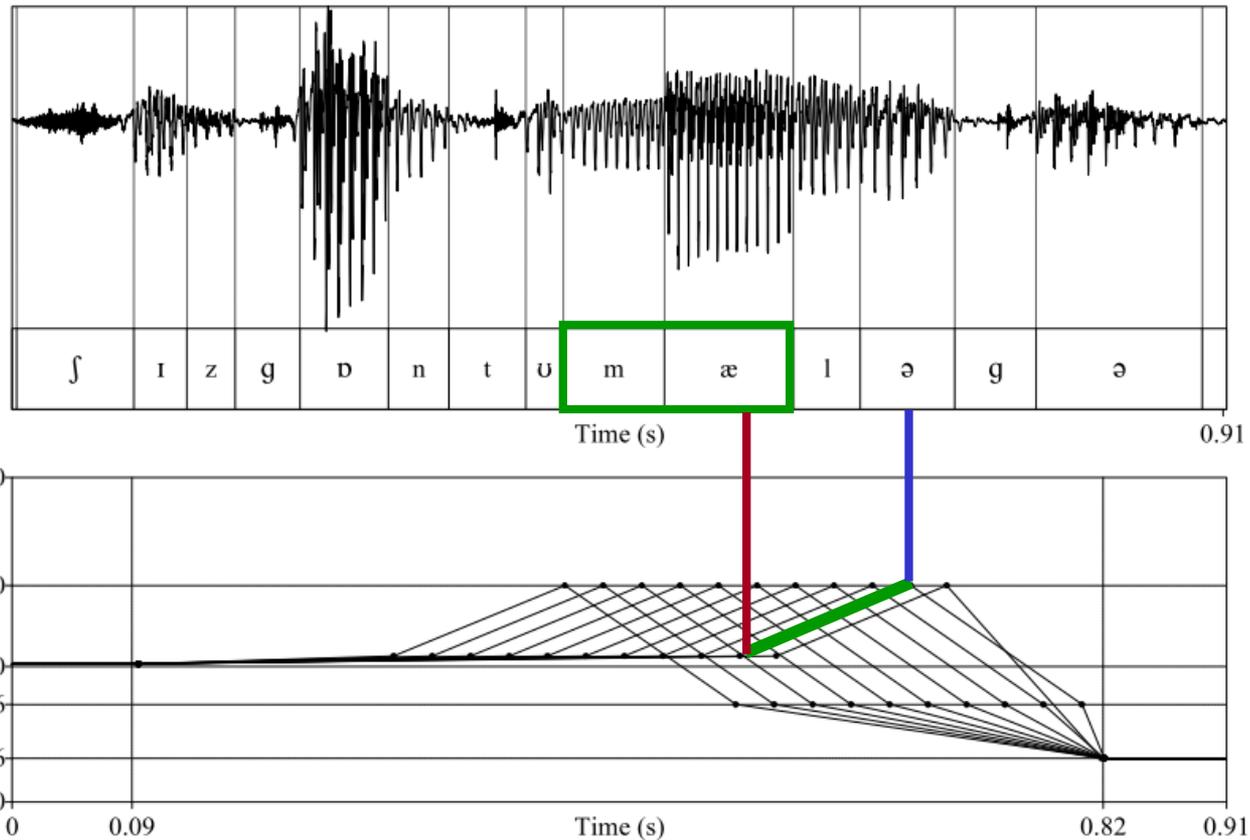


~ “medial peak” :  $H^* \quad (H^* + L / L + H^*)$

# Exploiting time domain

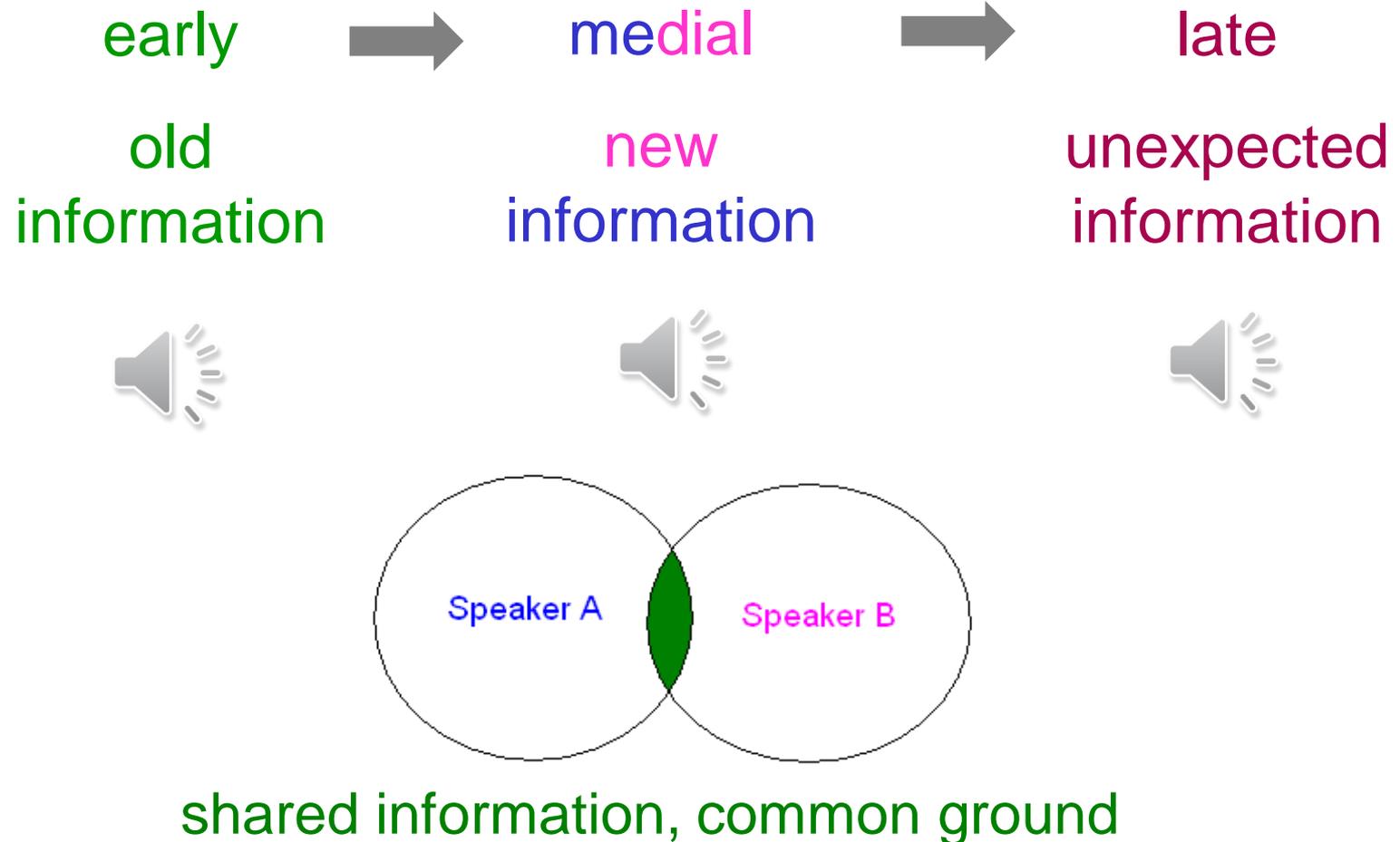
H

L



~ “late peak” : L\*+H

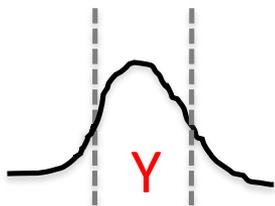
# Meaning of the temporal peak shift



# Medial vs. late peaks

A: This company has never employed anyone from Germany.

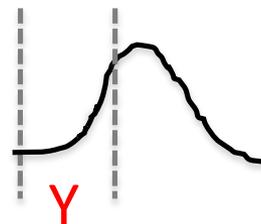
B: Jürgen's from Germany.



Of course it has! It employed Jürgen.

$(L+)H^* L-L\%$

B: Jürgen's from Germany.



We both know Jürgen, I suggest him as a possible choice but am also open to discussions...

$L^*+H L-L\%$

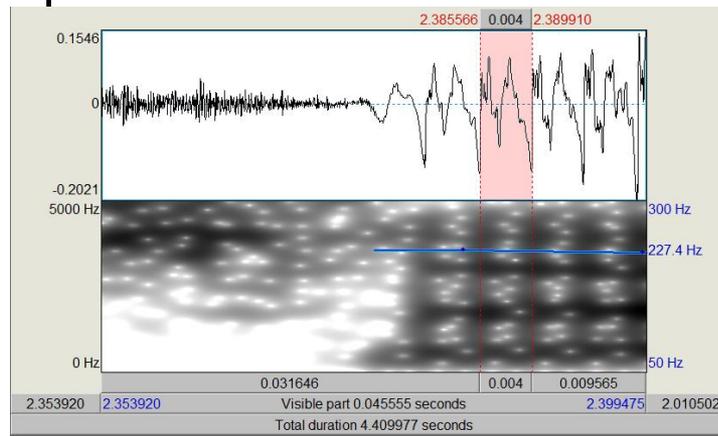
# Global changes of pitch

These are the types of pitch changes we are interested in when doing *intonational phonology*

- **Global:** take place in the accented and the preceding/following syllables
- **Functional:** indicate not only the placement of phrasal prominences but also other pragmatic meanings (e.g. old/new information)
  - Speaker-specific characteristics like age, gender, mood/emotion are also coded by pitch but are irrelevant for **intonation (= linguistic use of pitch)**
- **Reproducible:** are not idiosyncratic but apply to the whole community, can be produced/ imitated/ perceived/ interpreted by all speakers of the variety

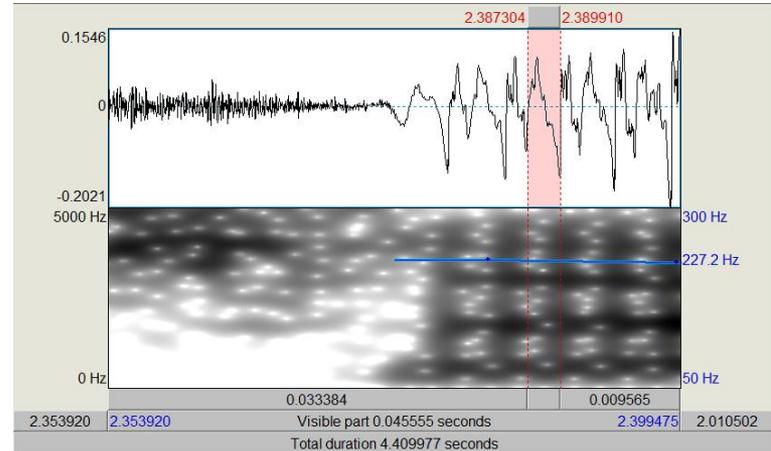
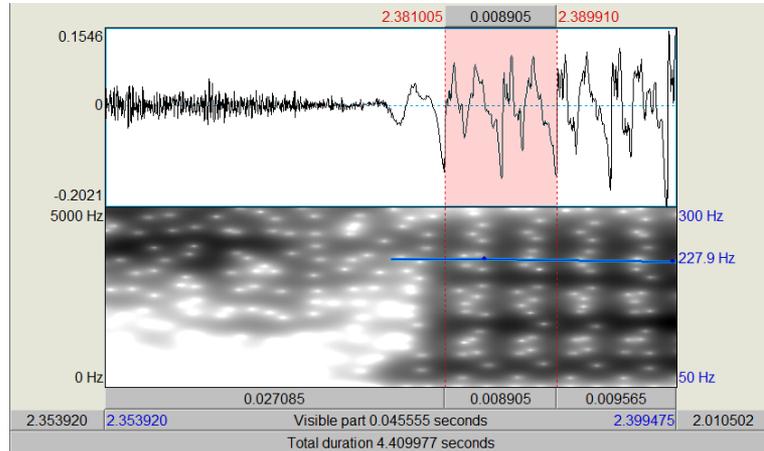
# Acoustic work with pitch

- F0 (the fundamental frequency of speech) is the acoustic correlate of perceived pitch
- “Pitch track algorithms” (e.g. in Praat) estimated F0 over time as a function of **vocal fold vibration**
- How? Autocorrelation approach
  - A periodic waveform is correlated with itself since one period looks much like another
  - Find the period by finding the ‘lag’ (offset) between two windows on the signal for which the correlation of the windows is highest
  - Lag duration (T) is 1 period of waveform
  - Inverse is F0 (1/T)

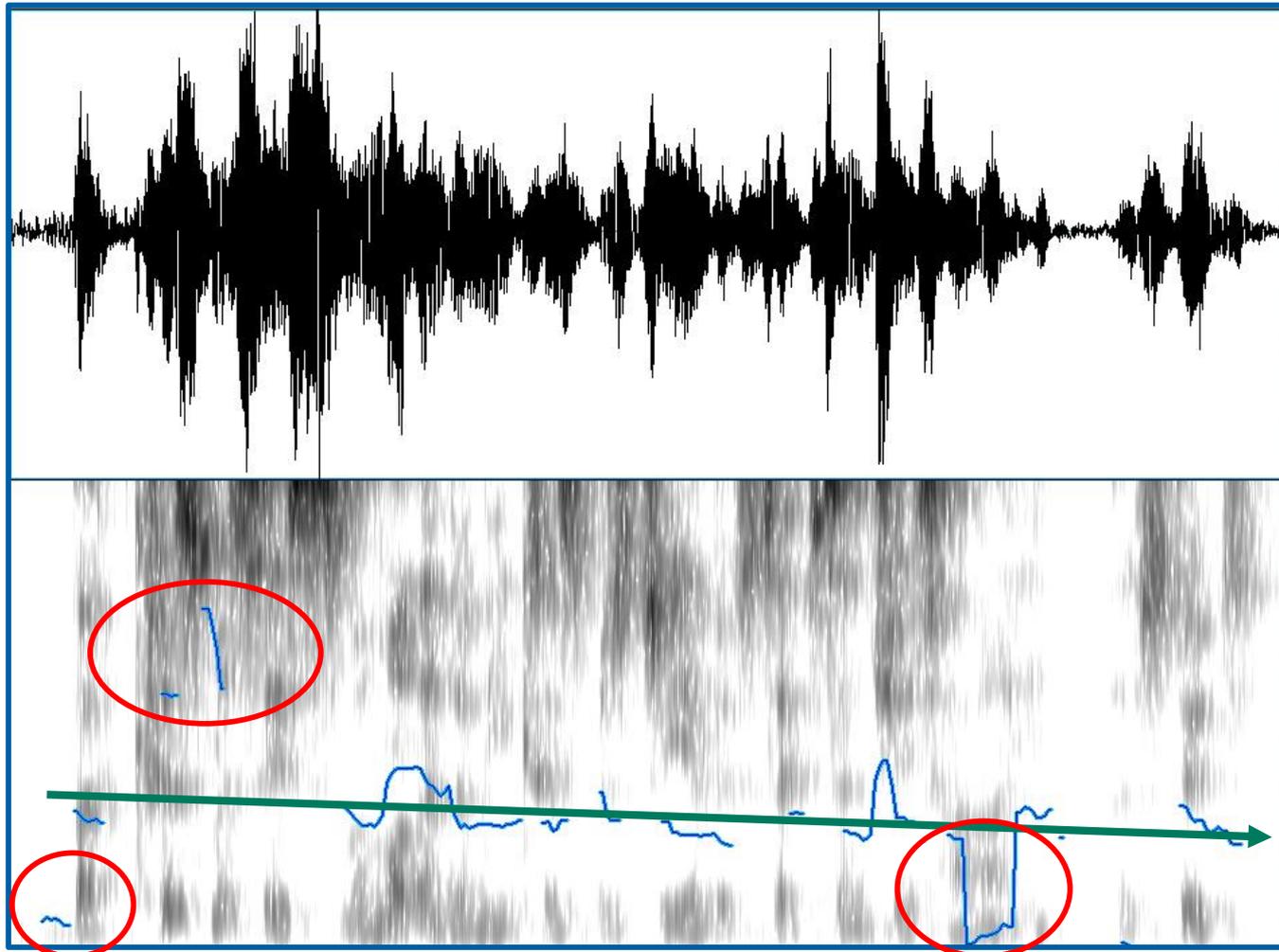


# Autocorrelation is prone to errors

1. **Halving**: shortest lag calculated is too long → estimated cycle too long, too *few* cycles per sec → **underestimated** F0
2. **Doubling**: shortest lag too short and second half of cycle similar to first → cycle too short, too *many* cycles per sec → **overestimated** F0

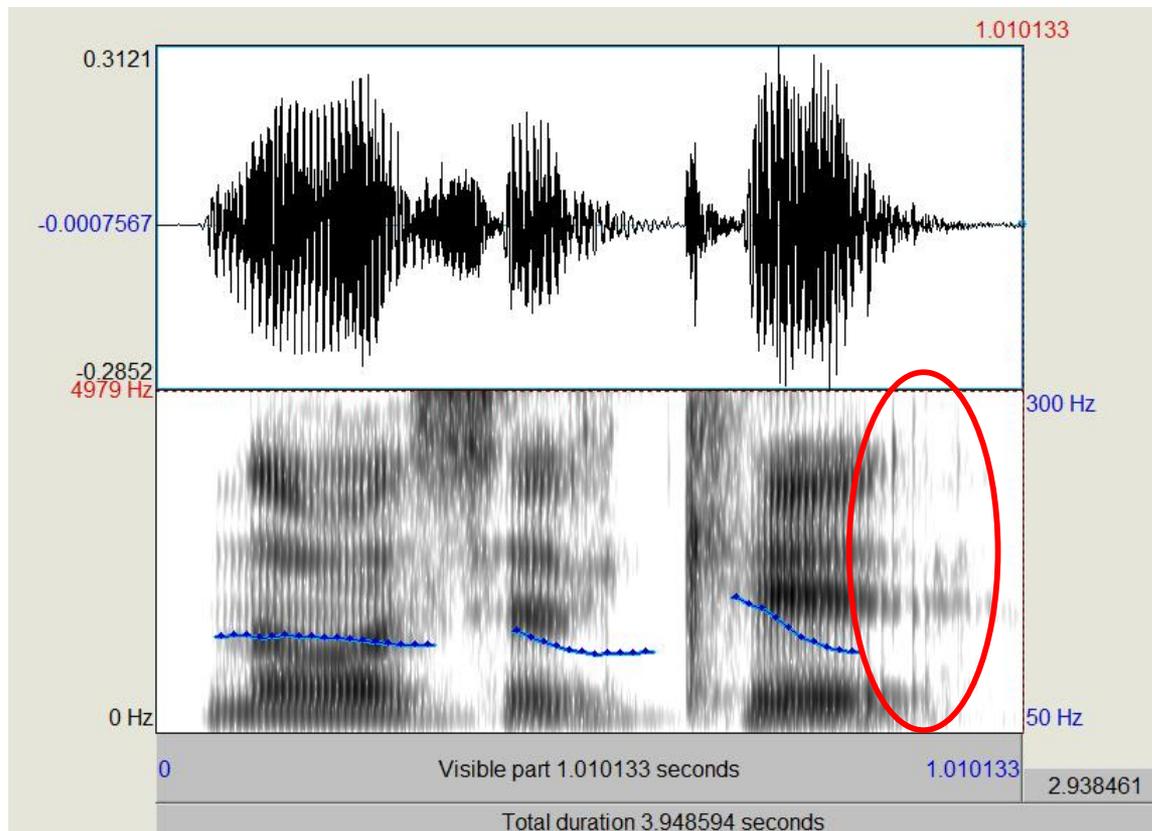


# An example from Glaswegian



# No F0-tracking

- Voiceless segments (no vocal fold vibration)
- Creaky voice/glottalling (irregular vocal fold vibration)



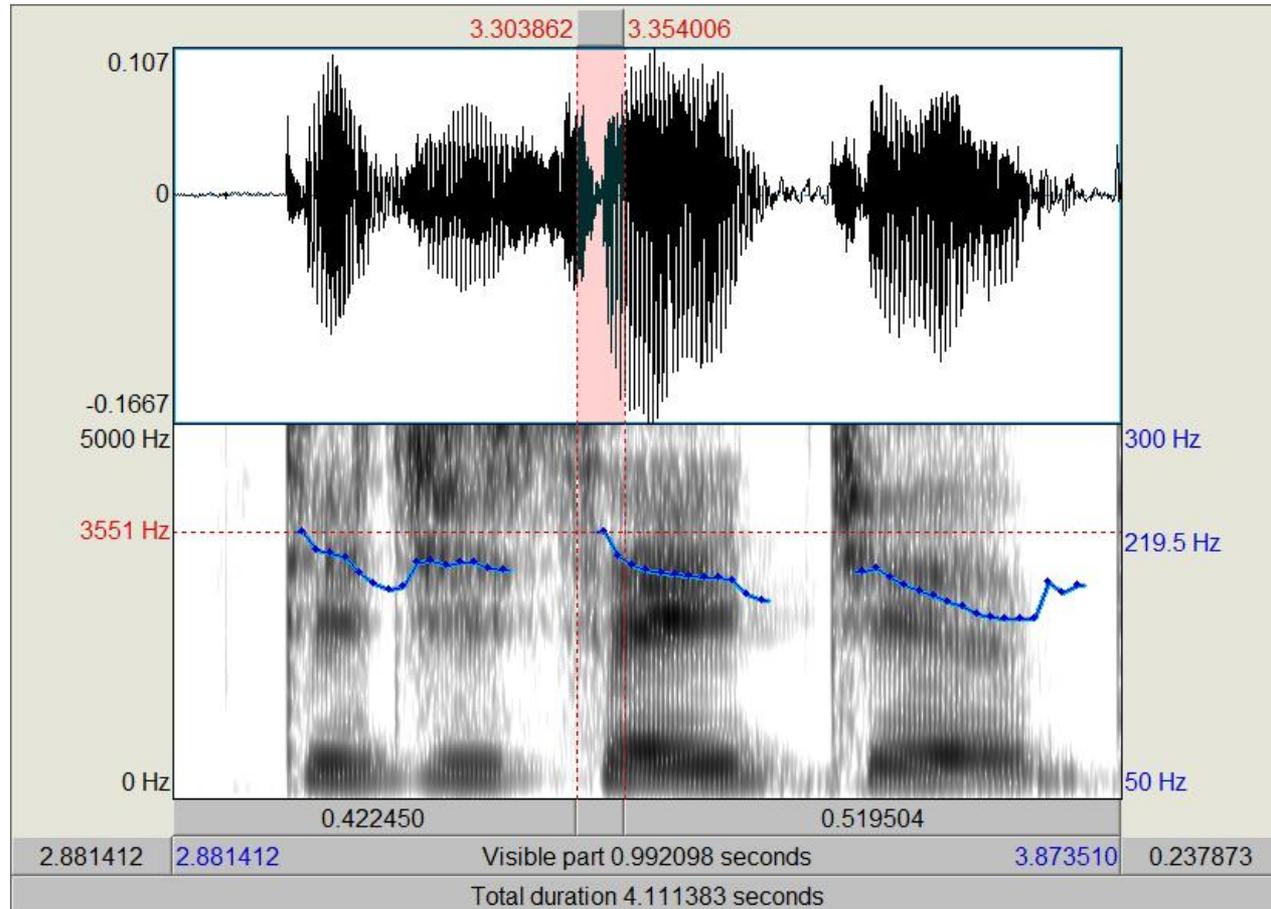
# Genuine but misleading f<sub>0</sub>-fluctuations (1)

*You may know my niece. vs. You can see his track.*

So-called “microprosodic” effects due to the presence of consonants (particularly obstruents)

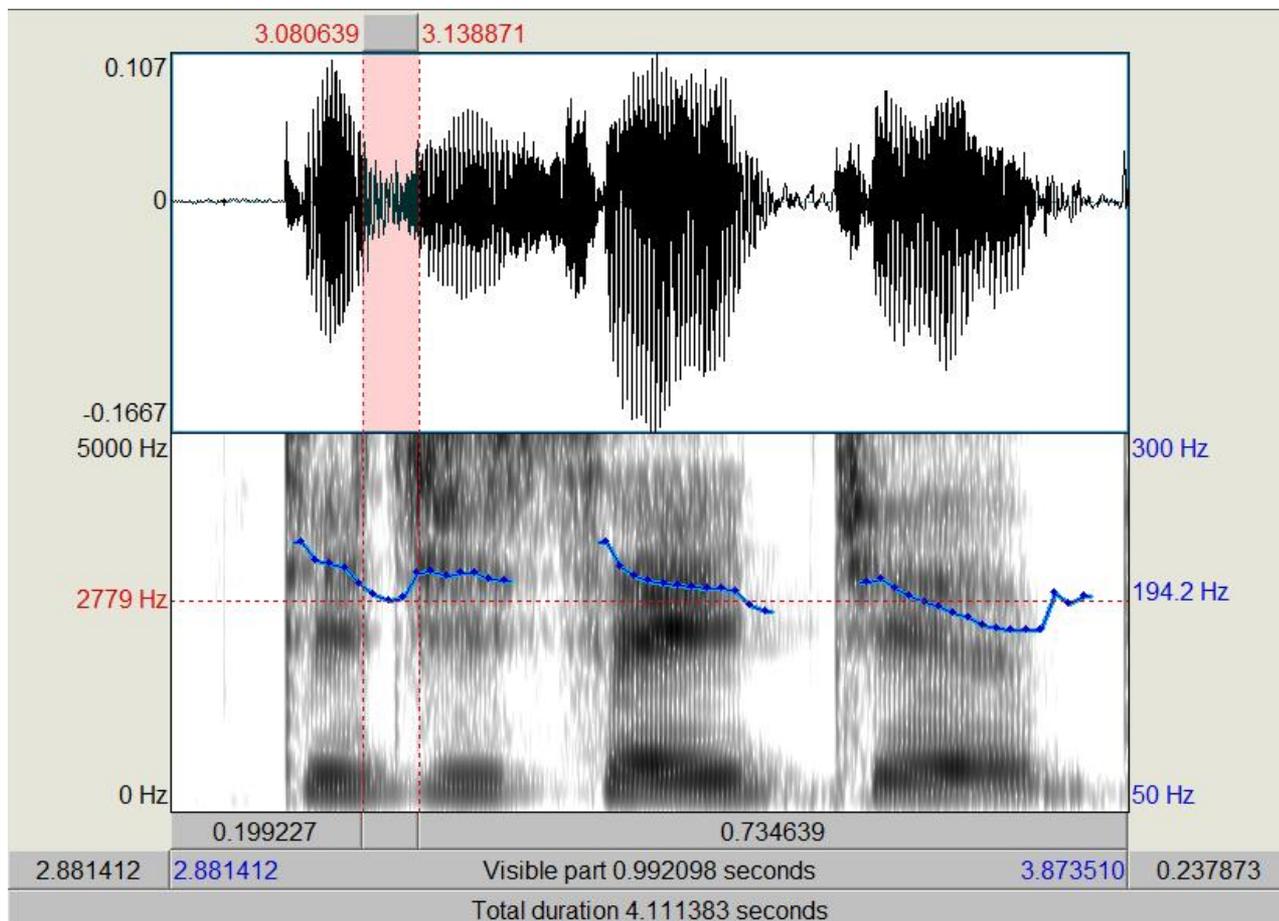
# Genuine but misleading f0-fluctuations (2)

Canonical example:  
F0-raising due to a (preceding) **voiceless** obstruent



# Genuine but misleading f0-fluctuations (3)

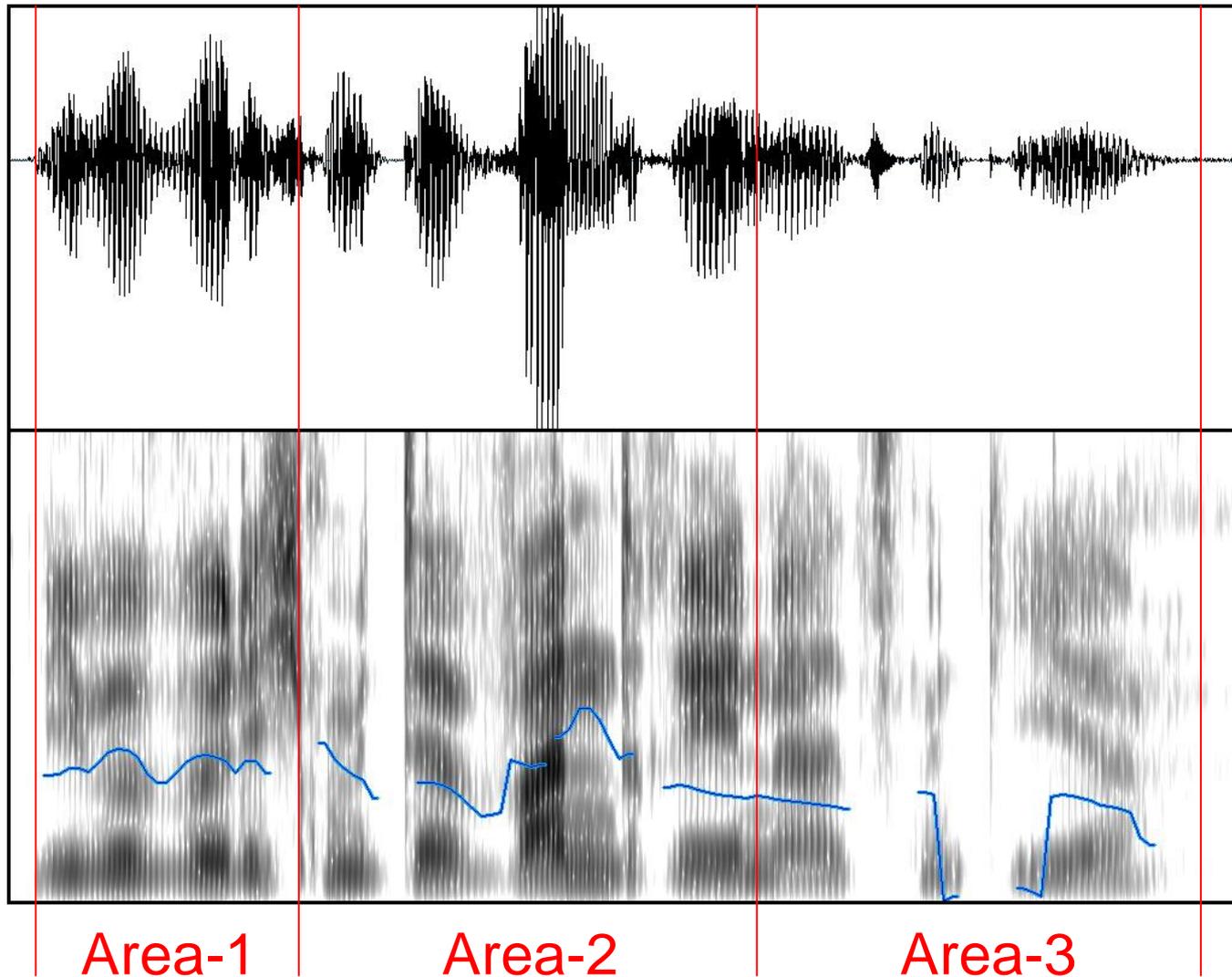
Canonical example:  
F0-lowering due to a **voiced** obstruent



## Genuine but misleading f<sub>0</sub>-fluctuations (4)

- Relevant to the perception of the segmental categories but not intonational categories
- Systematic (and can give rise to tonogenesis)
- Strictly localized and often steep change (a few ms of duration, magnitude can vary ~5-30 Hz)

## Socrative task #2: Inspection of F0-tracks



## Socratic task #2: Inspection of F0-tracks

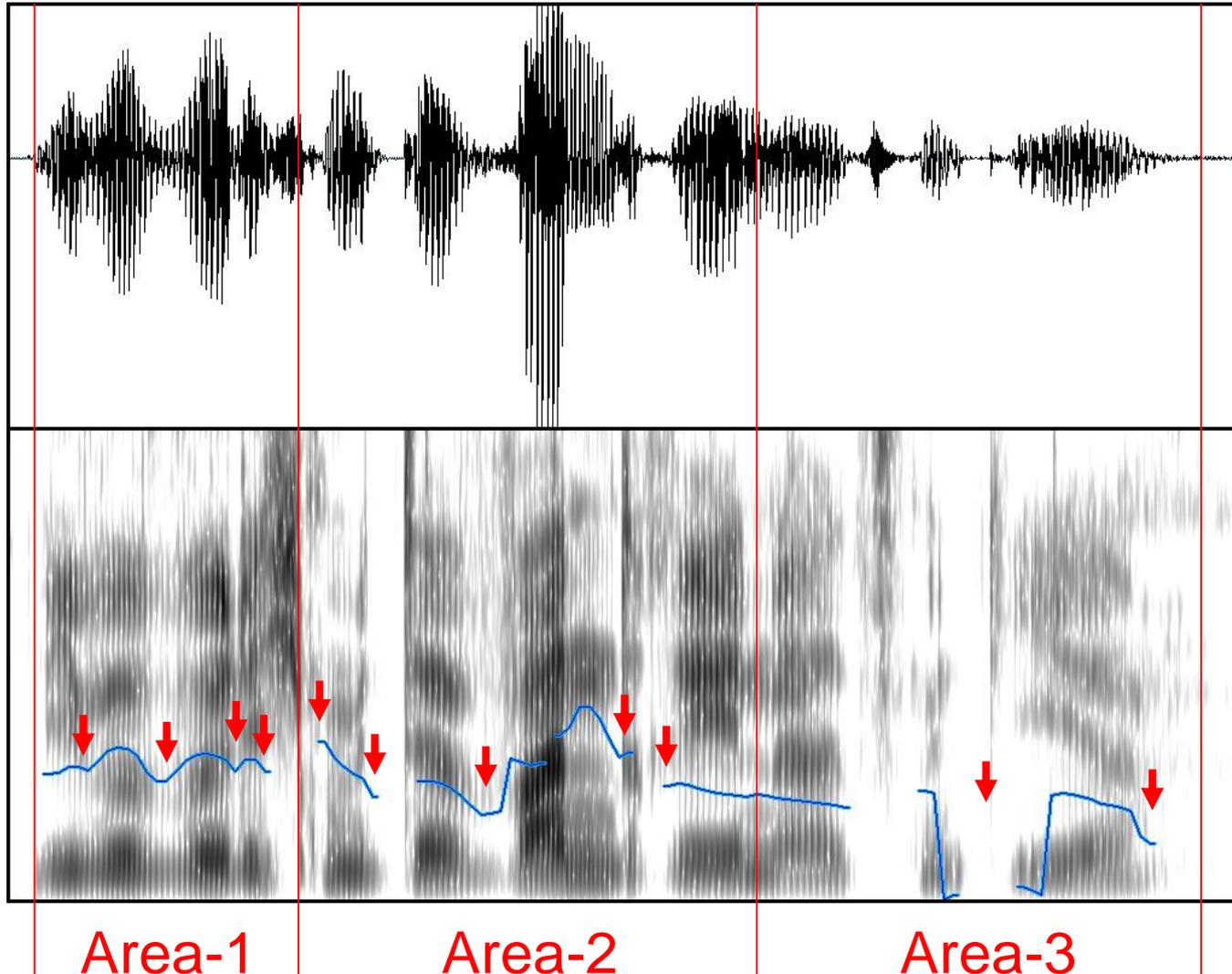
- Working in pairs (or groups of 3)
- Discuss the F0-track of each of the highlighted areas (1-3).
- Any problematic F0-perturbations in this example?

Socratic input:

1. *How many?*
2. *What type?*
  - halving
  - doubling
  - F0-raising
  - F0-lowering

# SOCRATIVE #2

# Socratic task #2: Inspection of F0-tracks



1. 4, lowering
2. 5, raising & lowering
3. 2, halving & lowering

# F0-variation we see in the signal can be

## Phonological

- Not predictable from the segmental environment  
*Yes! Yes? Yes...*
- Used to express linguistic meanings (pragmatics, sentence mode etc.)
- Cannot be summarised in a set of rules
- Association with prominence \* or boundary (%) - abstract

## Phonetic

- Predictable from the environment  
bro**th**er vs. brass**e**rie
- No contribution to linguistic contrasts (compensated for in perception)
- Is often summarised in a set of rules
- Changes surface but not the abstract underlying representations

# Phonetic factors

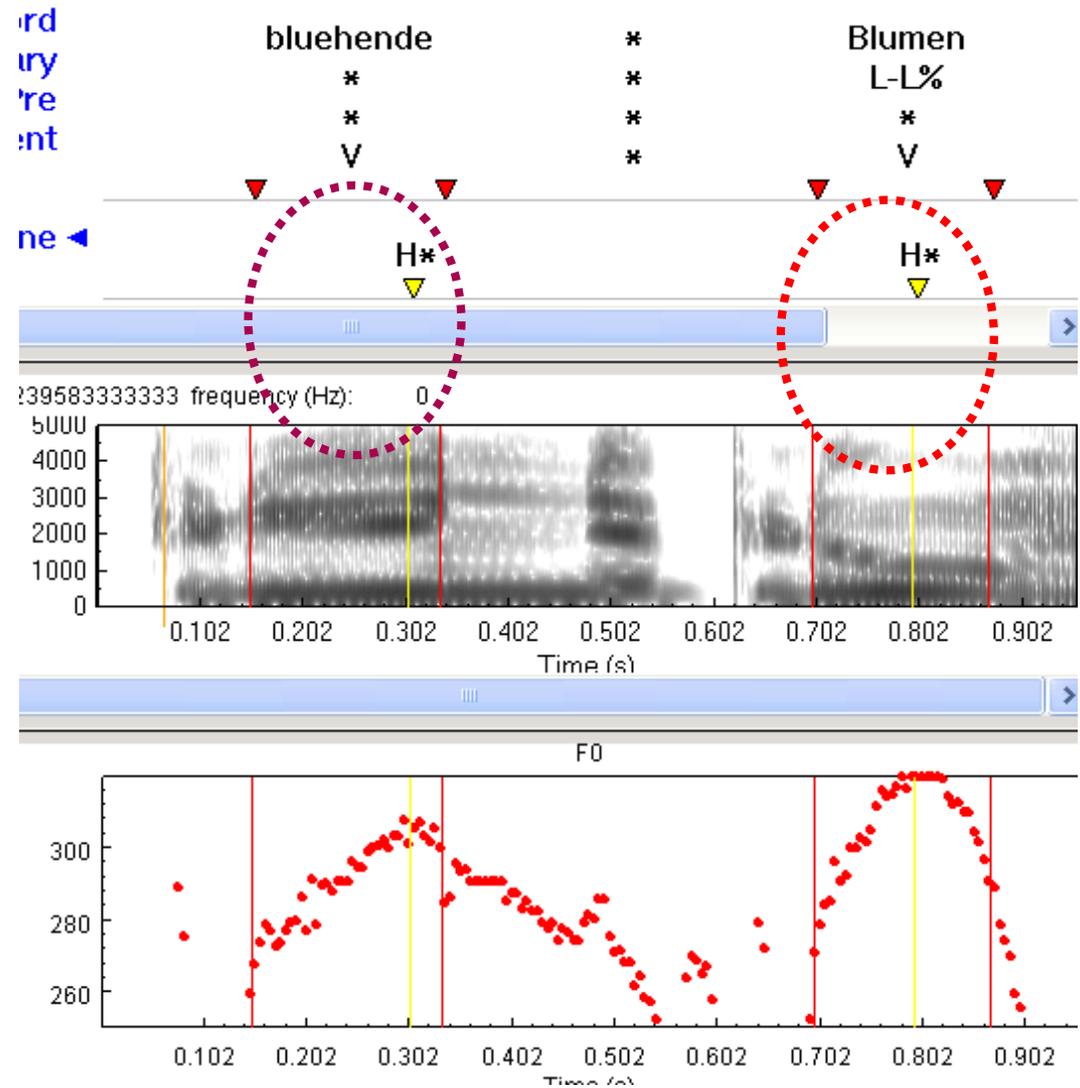
- Phonetic factors modifying the surface representations can be *language-specific*
- However, there is a set of factors that systematically applies to intonational categories in *most intonational languages*

# Position within the prosodic phrase

The same phonological category (here: H\* in German) shows different alignments:

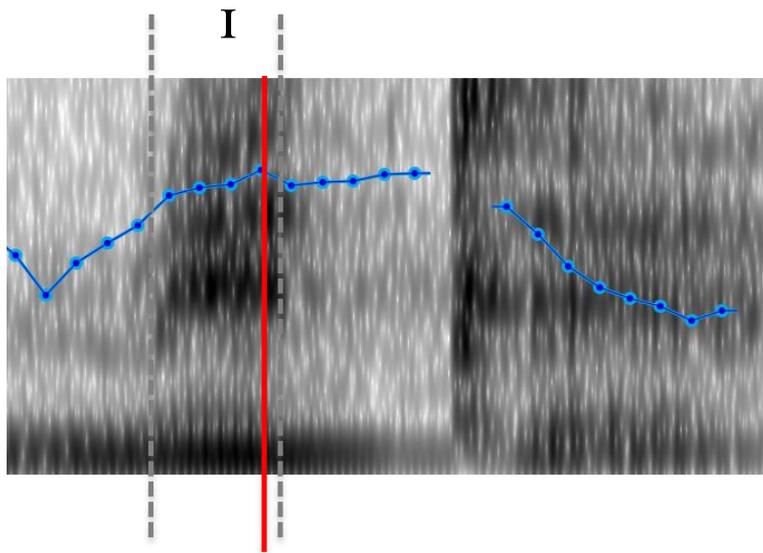
(1) Peak of H\* is later in pre-nuclear „bluehende“

(2) Peak of H\* is earlier in nuclear „Blumen“

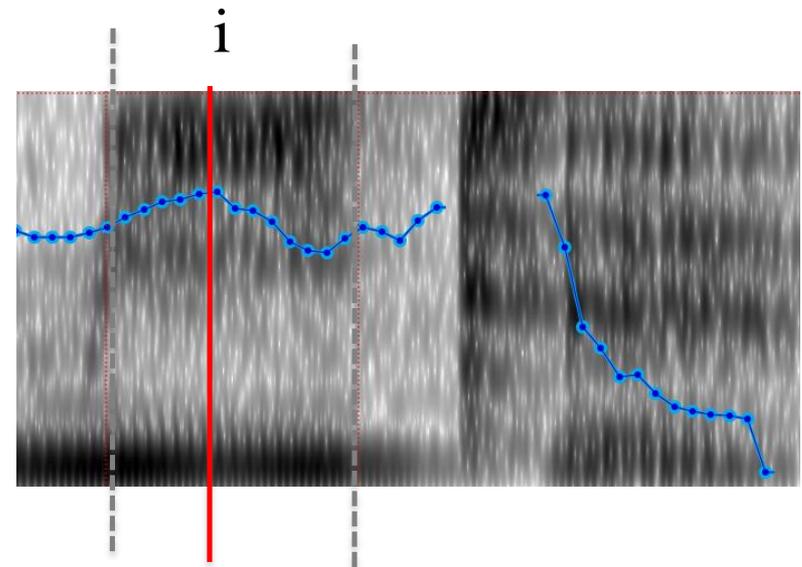


# Timing of the accented vowel

Alignment of high targets (as in H\*) is often proportionally later in short than in long vowels



Mitte (German „middle“)



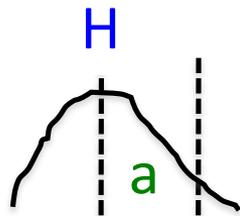
Miete (German „rent“)

Ladd, D.R., Mennen, I., Schepman, A. (2000)  
*Examples courtesy Jonathan Harrington*

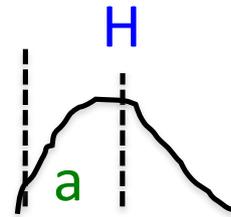
# Proximity of another pitch accent

H (in H\*) can be shifted to the left (i.e. aligned earlier) if closely followed by another pitch accent („tonal crowding“)

Ma Lemm



Mamalie Lemonick

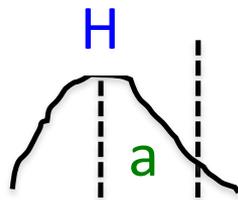


*Silverman & Pierrehumbert (1990)*

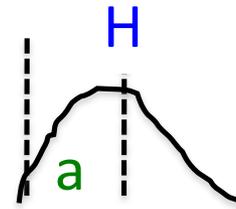
# Proximity of a word boundary

H (in H\*) can be shifted to the left (and aligned earlier) if there is a word boundary

Ma Le Mann



Mama Lemm

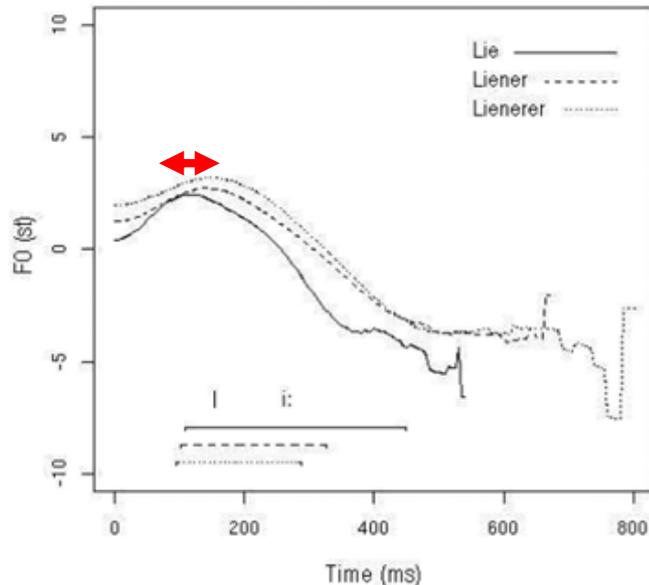


*Silverman & Pierrehumbert (1990)*

# Proximity to a phrasal boundary

If a pitch accent is close to the upcoming phrasal boundary, its peak will be shifted to the left (aligned further away from the boundary):

- Put it into the **bowl**.
- Put it into the **bowl** over there.



German sentences:

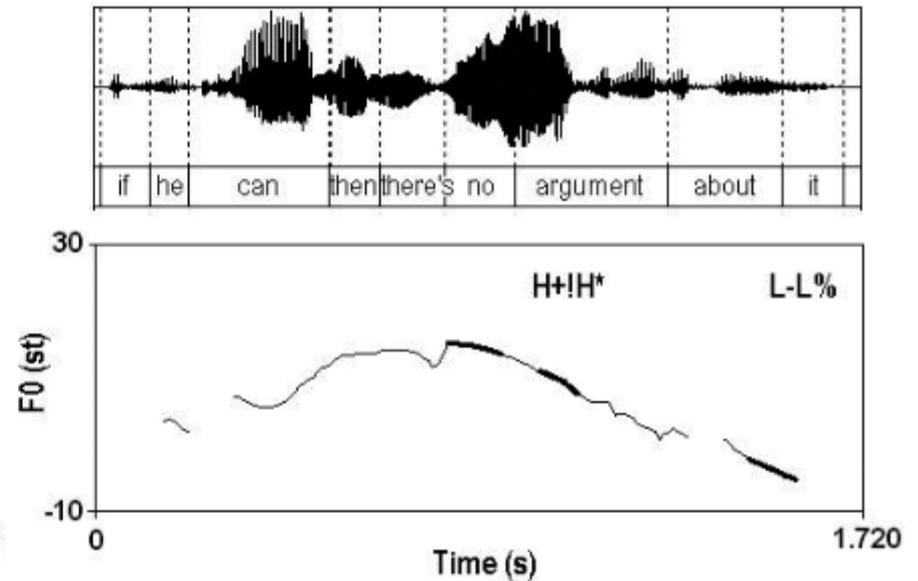
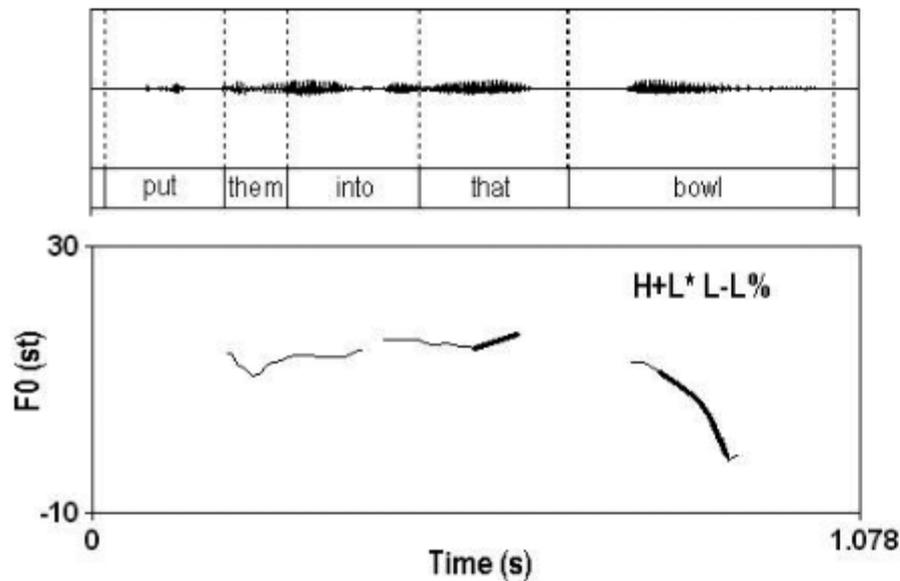
- *It was Mr **Lienerer**.*
- *It was Mr **Liener**.*
- *It was Mr **Lie**.*

Rathcke & Harrington (2010)

# Proximity to a phrasal boundary

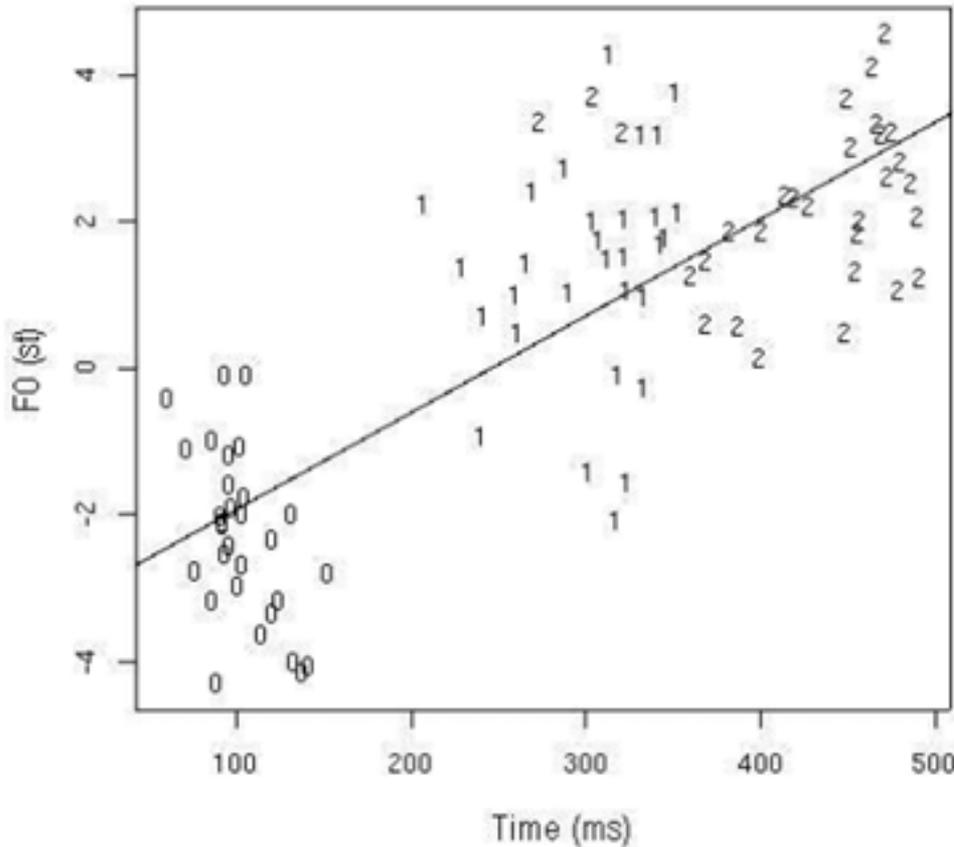
Not only changes to alignment, also scaling!

GToBI/ToBI: H+!H\* and H+L\* distinction in German and English



*From Rathcke & Harrington (2010)*

# Pitch lowering effect of L%



- *It was Mr **Lienerer**. L%*
- *It was Mr **Liener**. L%*
- *It was Mr **Lie**. L%*

Time  $\longrightarrow$  L%

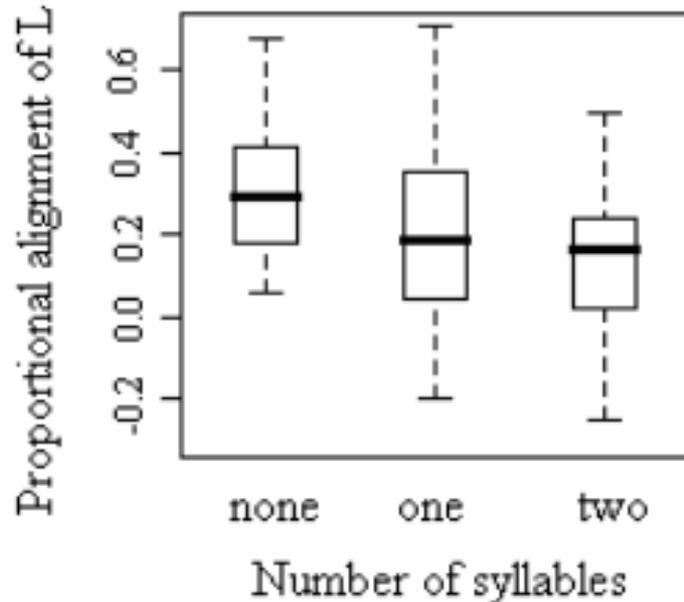
- positive correlation of scaling and the distance between the syllable and the phrasal edge
- no H+!H\* vs. H+L\* contrast

*From Rathcke & Harrington (2010)*

# Does this only happen to nuclear accents?

Myths:

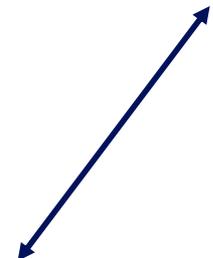
- L\* are more stably anchored with the accented syllable than H\*
- Prenuclear rises are more stable than nuclear ones



([ **bl**ablablablaba ])

([ bla**bl**ablablaba ])

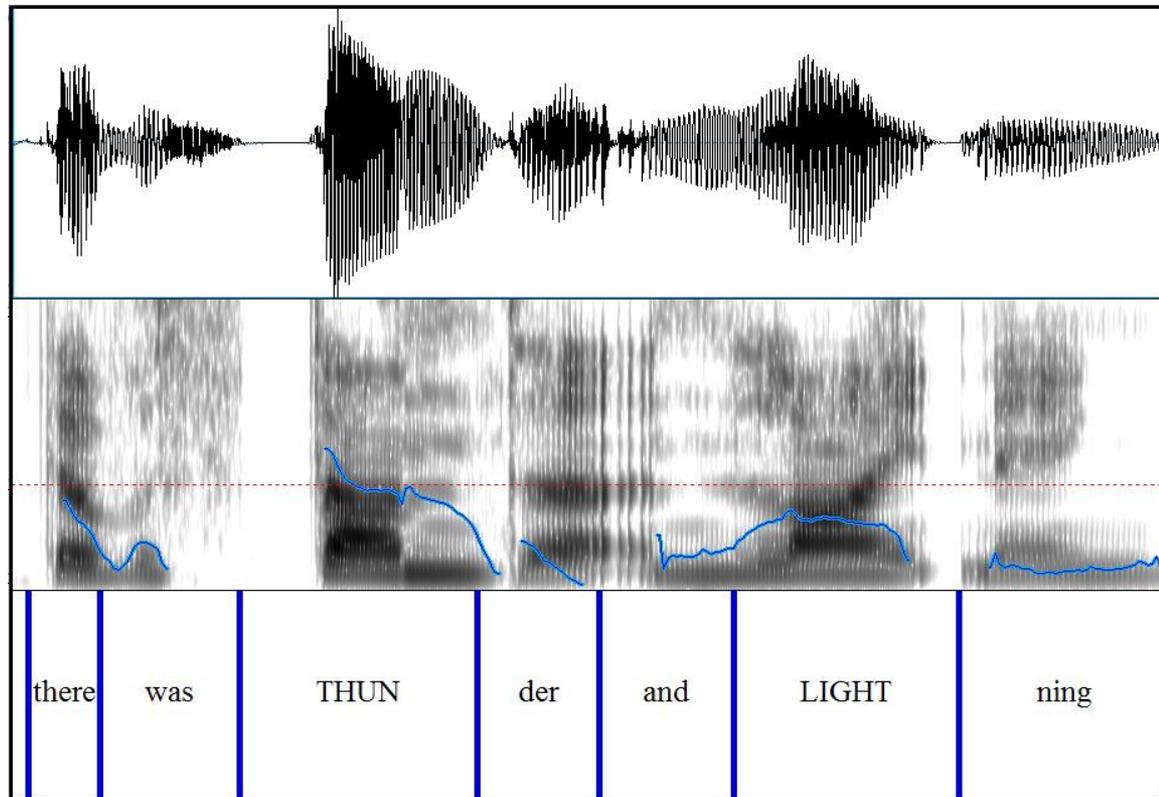
([ bla**bl**ablaba ])



*From Kleber & Rathcke (2008)*

# Socrative task #3: identifying pitch accents

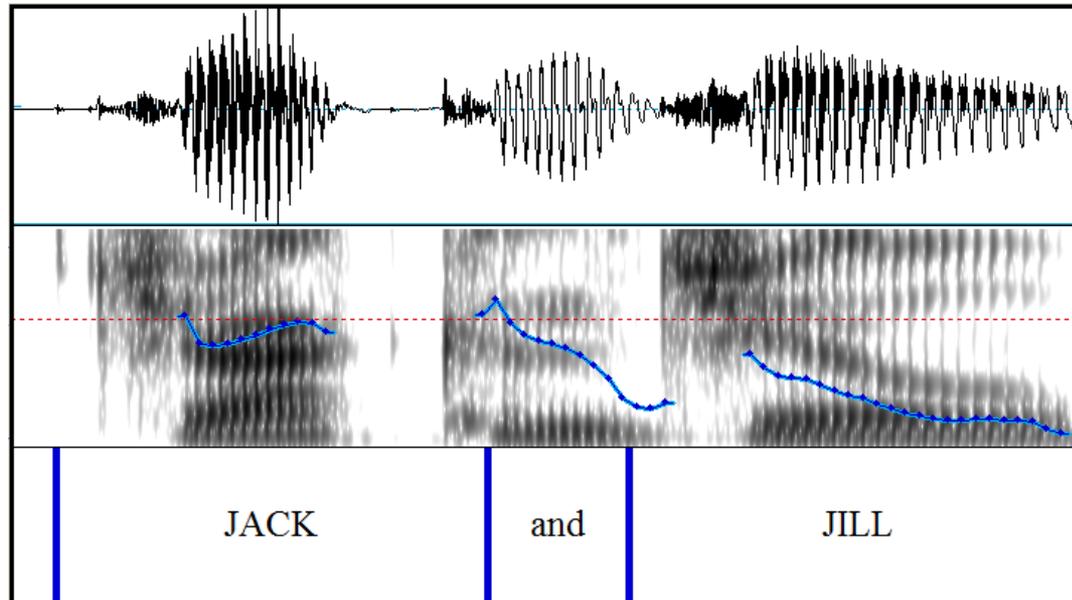
(3a) Same or different? Can you guess the type of pitch accent?



*Data from Rathcke & Smith (2015)*

## Socratic task #3

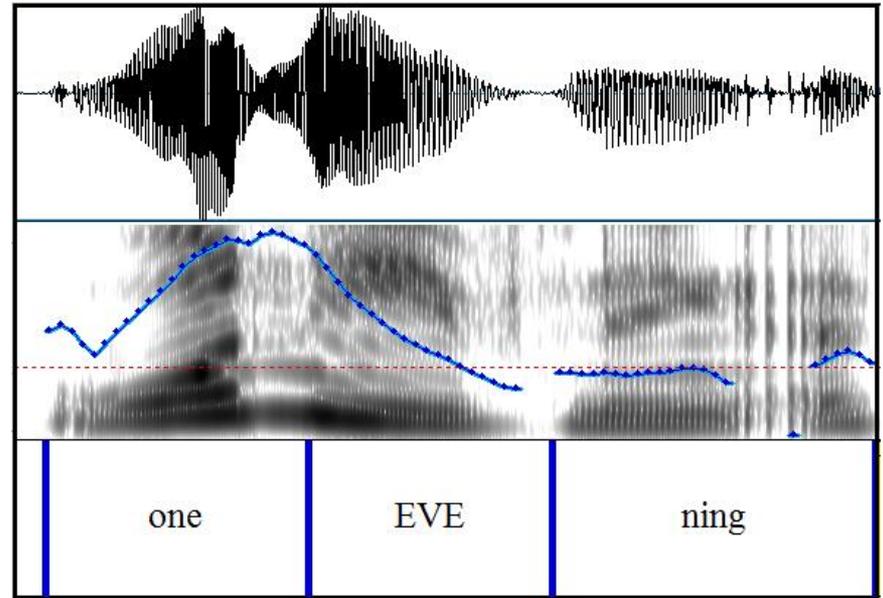
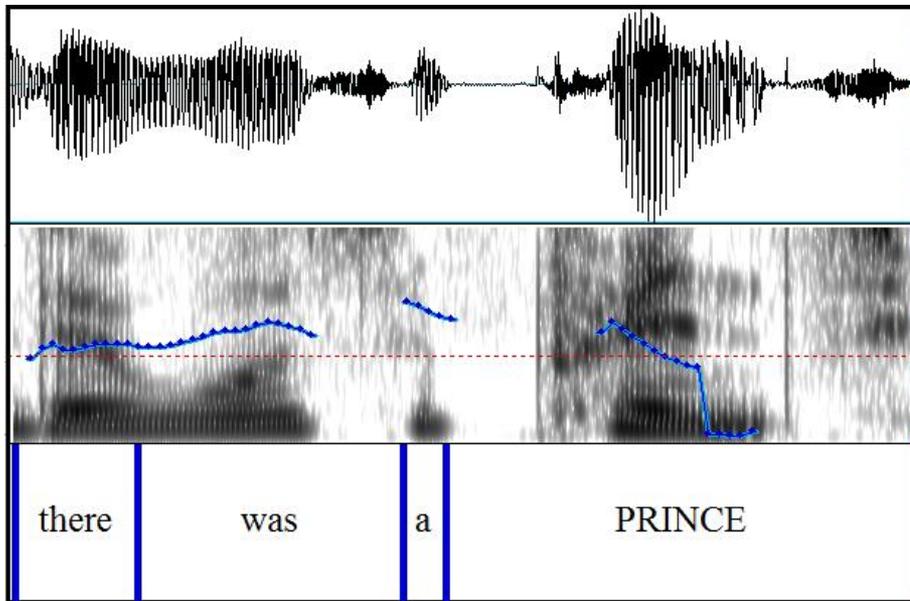
(3b) Same or different? Can you guess the type of pitch accent?



*Data from Rathcke & Smith (2015)*

## Socratic task #3

(3c) Same or different? Can you guess the type of pitch accent?



*Data from Rathcke & Smith (2015)*

# **SOCRATIVE TASK #3**

# Time pressure

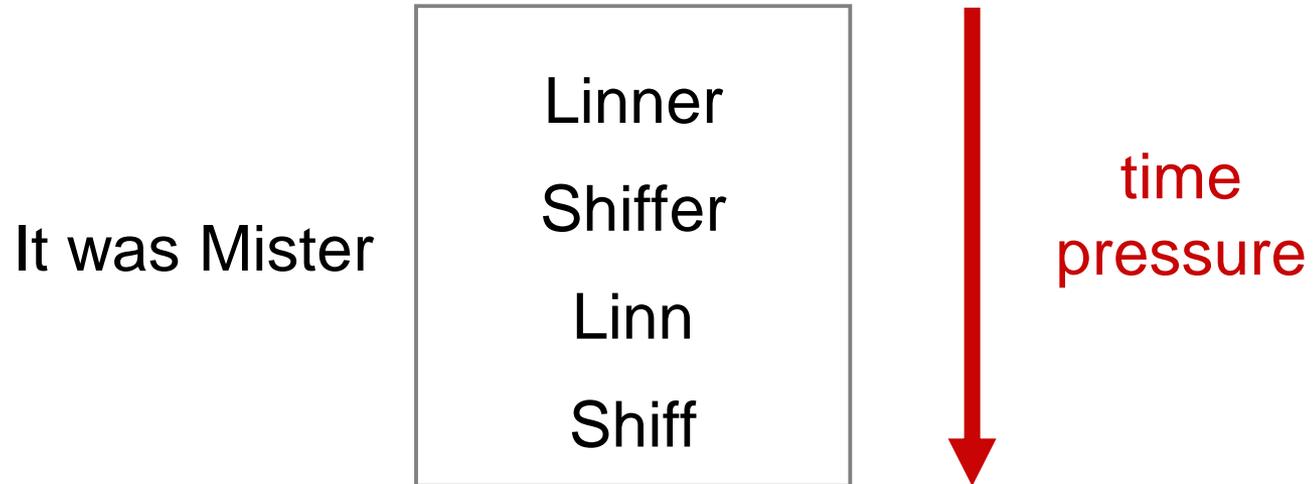
These phonetic modifications arise from the issue of *time pressure*:

- F0-information needs sufficient **time** and **sonority** to evolve

Availability of several syllables that have vowels or sonorous (= F0-bearing) segments allows a complete (or „prototypical“) realisation of pitch patterns

- Lack of **sonority** leads to an increase in time pressure  
→ the target pitch patterns show deviations, F0-modifications arise

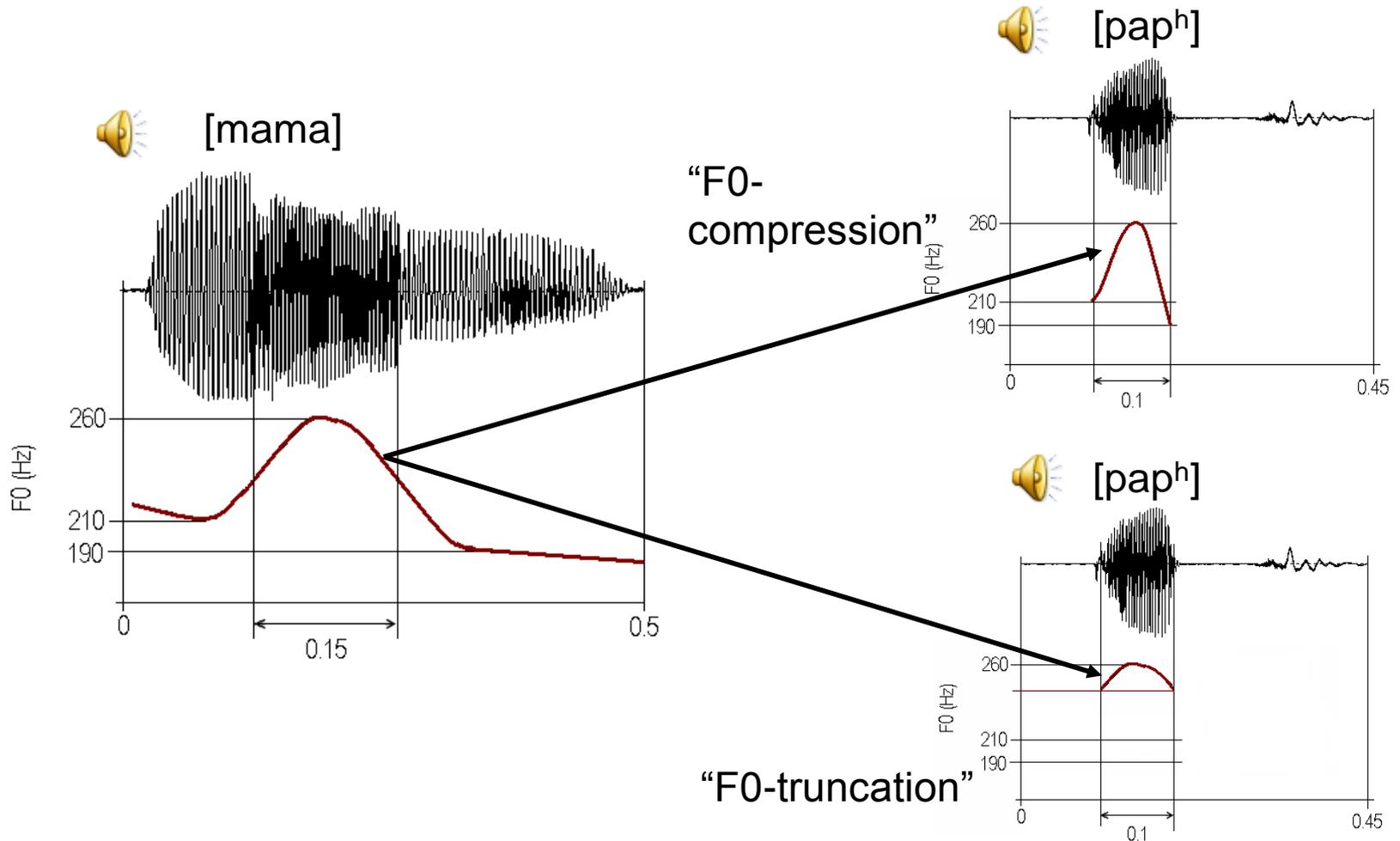
# Time pressure at phrase boundaries



- Microprosodic influences in *Shiffer* and *Shiff*
- Lack of a postnuclear syllable: “Schiff” in phrase-final position: a short/lax vowel, 1/10 of the duration of sonorous *Linner*

*Grabe (1998), Rathcke (forthcoming)*

# Consequences of insufficient time/sonority

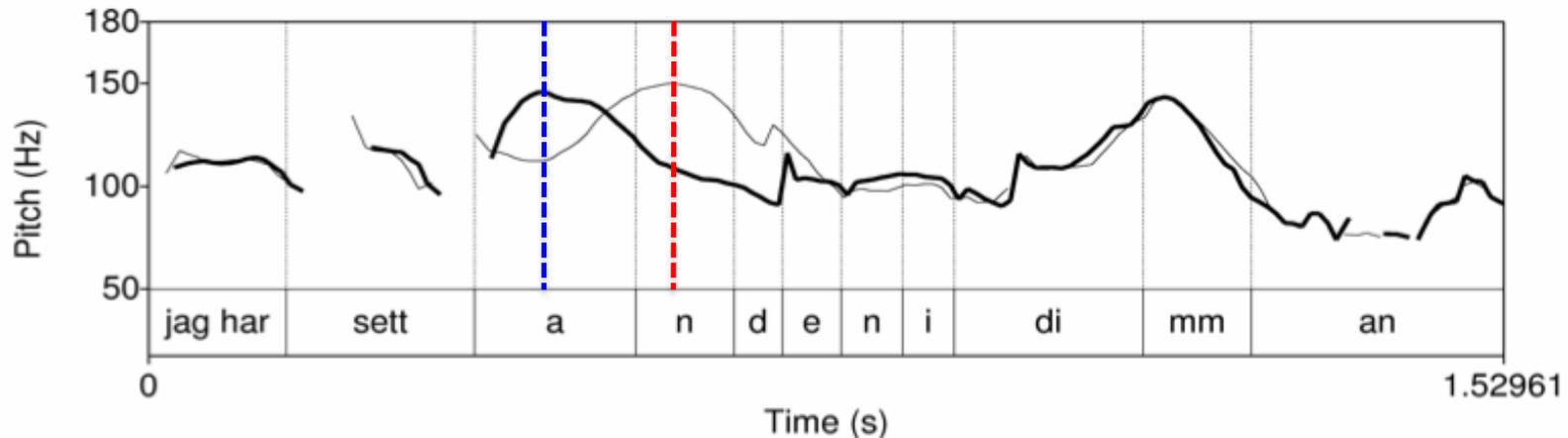


Erickson & Alstermark 1972; Bannert & Bredvad 1975; Grønnum 1991

# Accent-1 and Accent-2 in Swedish



I saw a duck/a ghost in the fog.

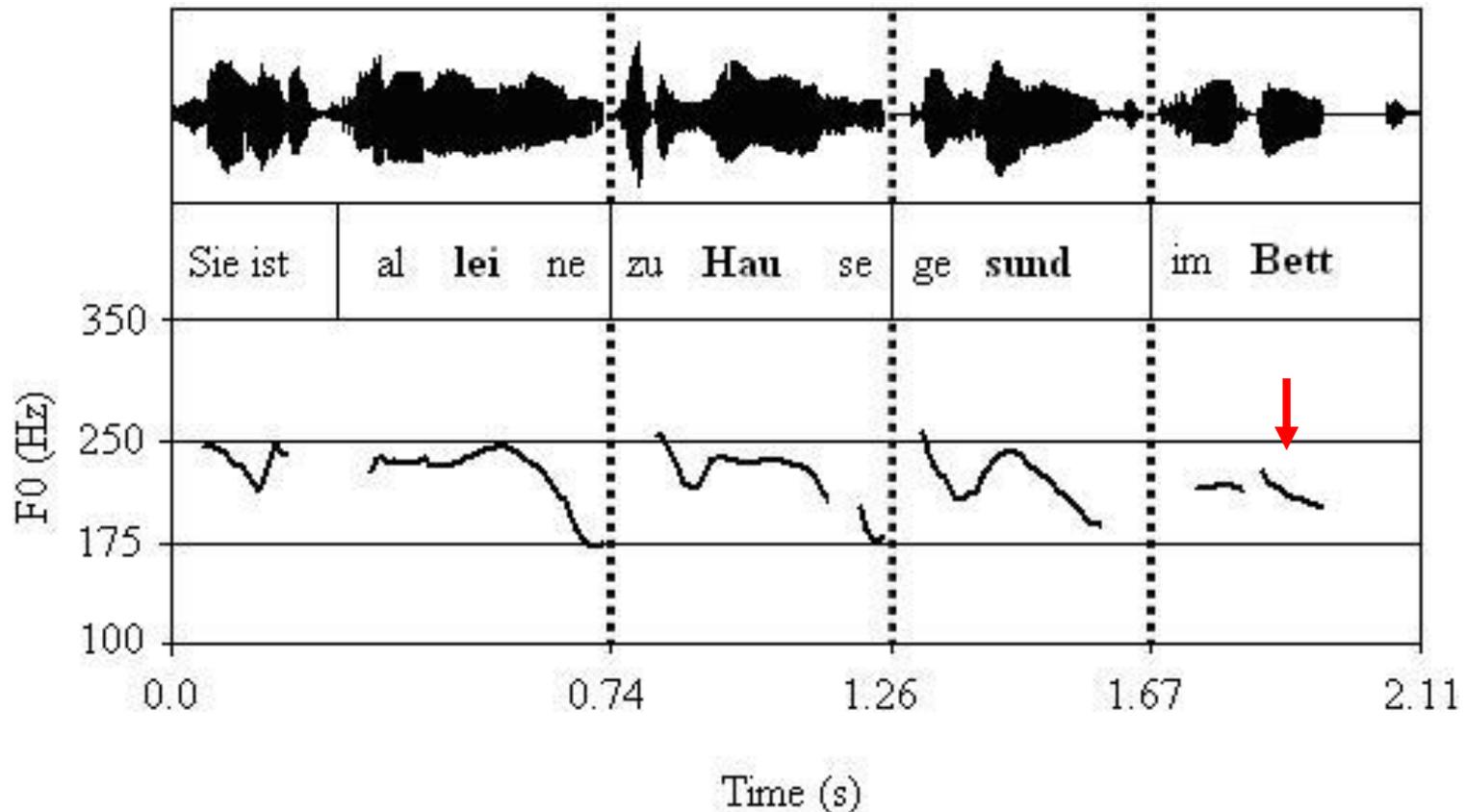


- Monosyllabic words vs. polysyllabic (and longer) words
- Dialects of Swedish differ in what they do with F0: dialects can either compress or curtail.

Erickson & Alstermark 1972; Bannert & Bredvad 1975; Grønnum 1991

# Example: H\* L-% in German

Sie ist **alleine**... zu **Hause**... gesund... im **Bett**



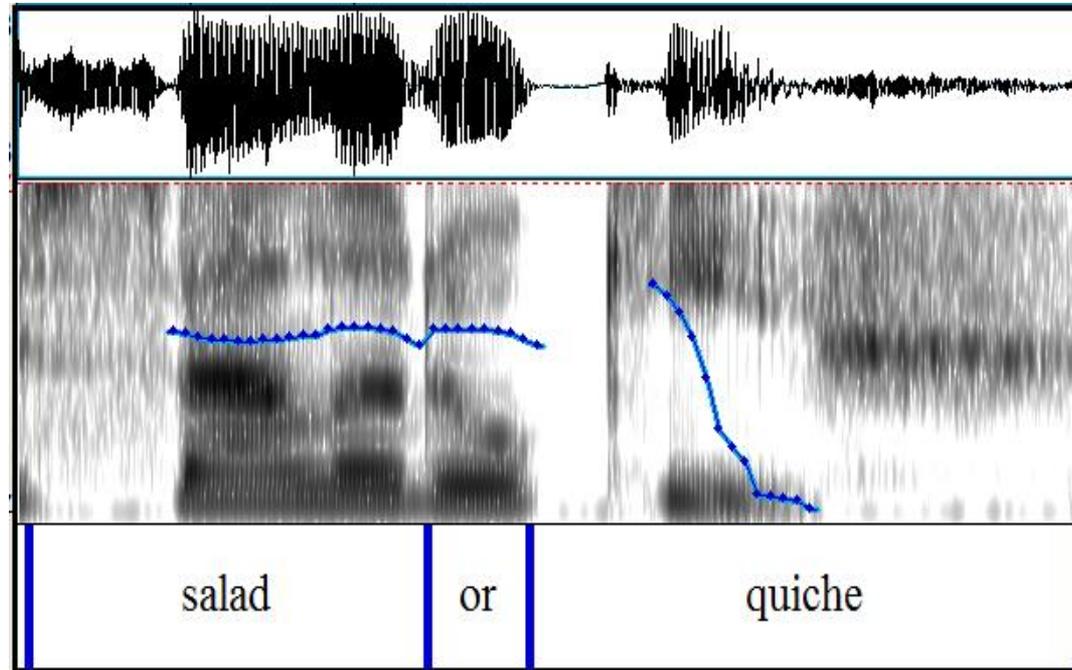
From Rathcke (2013), after Grabe (1998)

# A typological stance

Compressing/truncating languages (Ladd 1996)

- English (meaning SSBE): “*a compressing language par excellence*” (Ladd 1996:133; Grabe 1998)
  - regardless of pitch pattern, compression is found
- Truncation far more common: e.g. Hungarian, Palermo Italian (Grice 1995), German (Grabe 1998), Spanish (Ortega-Llebaria 2009)
  - some pitch patterns are truncated, others may be compressed

# English: “compression language par excellence”



SSBE (“BBC accent”) but many other varieties truncate as well

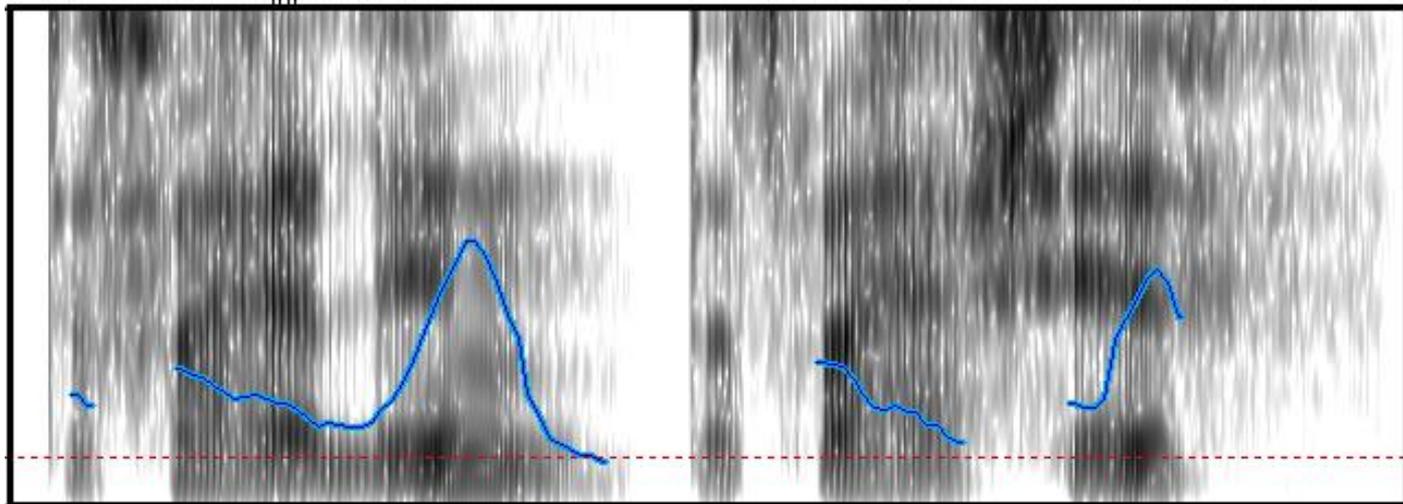
# German “truncates falls but compresses rises”

- Grabe (1998):  $H^*+L$  0% (GToBI  $H^* L$ -%) as used in statements and  $L^*+H$  H% (GToBI  $L^* H$ -%) as used in yes-no questions
- More falls and rises in German:
  - GToBI (Grice et al. 2002, 2005) as well as KIM (Kohler 1991, Niebuhr 2007):  **$H+L^*$  / early peak** and  **$L^*+H$  / late peak**
- Accentual rises -  $L^*+H$  are compressed ( $L\%$  is truncated) and accentual fall in  $H+L^*$  show compression (Rathcke 2009, forthcoming)

# Different faces of truncation

Truncation is not a unified phenomenon!

- German: gradual truncation – H\* L%    L\*+H L%  
Depending on the duration of sonority, stronger or weaker truncation (target undershoot of different magnitudes)



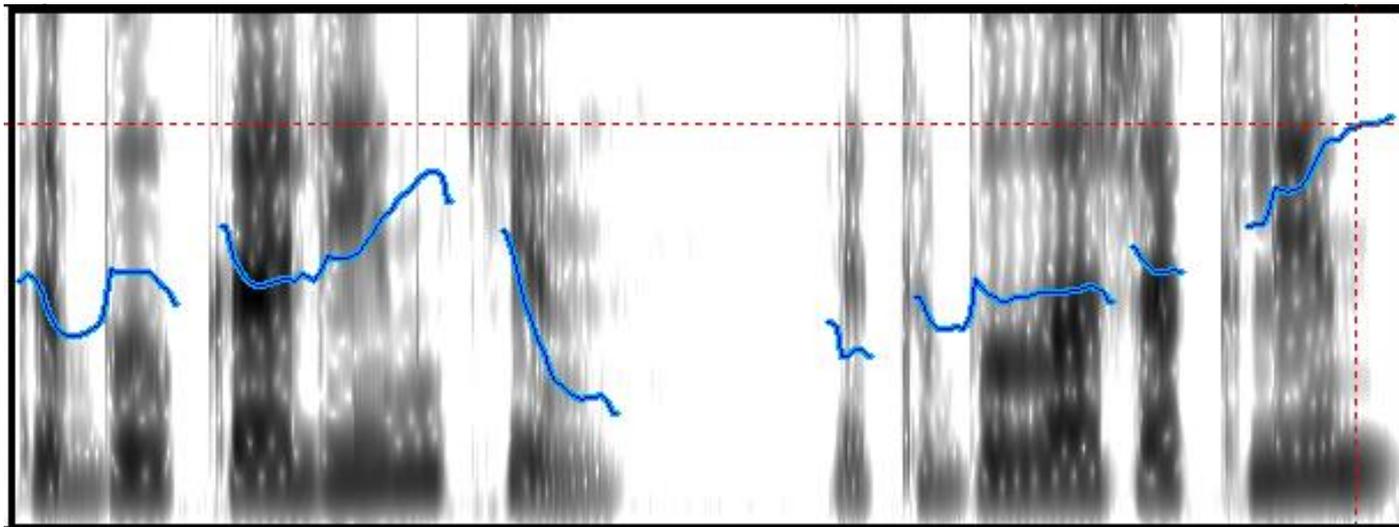
*Example from Rathcke (forthcoming)*

# Different faces of truncation

Truncation is not a unified phenomenon!

- Russian: gradual truncation ( $H^* L\%$ )  
categorical truncation ( $L^*+H L\%$ )

If there is no postaccentual syllable, there is no f<sub>0</sub>-fall



*Data from Rathcke (forthcoming)*

# Different faces of truncation

Truncation is not a unified phenomenon!

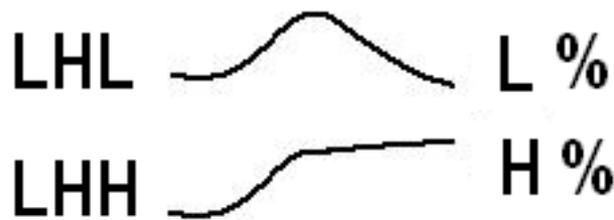
- German: gradual truncation
- Russian: gradual & categorical truncation
- Spanish: flexible truncation  
truncation is a speaker-dependent choice (vowel lengthening or insertion, cf. Prieto and Ortega-Llebaria 2009)

**The dichotomy truncating/compressing languages does not seem to be very helpful.**

# Potential consequences of truncation

*How do languages manage to maintain a functioning system of intonational categories in various contexts?*

LHL LHH



*Rathcke (2013)*

# Cross-linguistic perception of truncation

Despite truncation, f0 patterns are correctly identified as underlying H% or L% by native listeners, guided by *very subtle, language-specific* cues (Rathcke 2013)

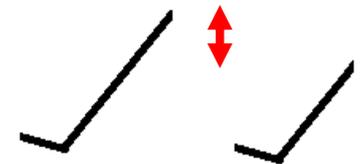
## German:

- German listeners respond to slight F0-fall
- ~1.5 st sufficient to perceive a L%



## Russian:

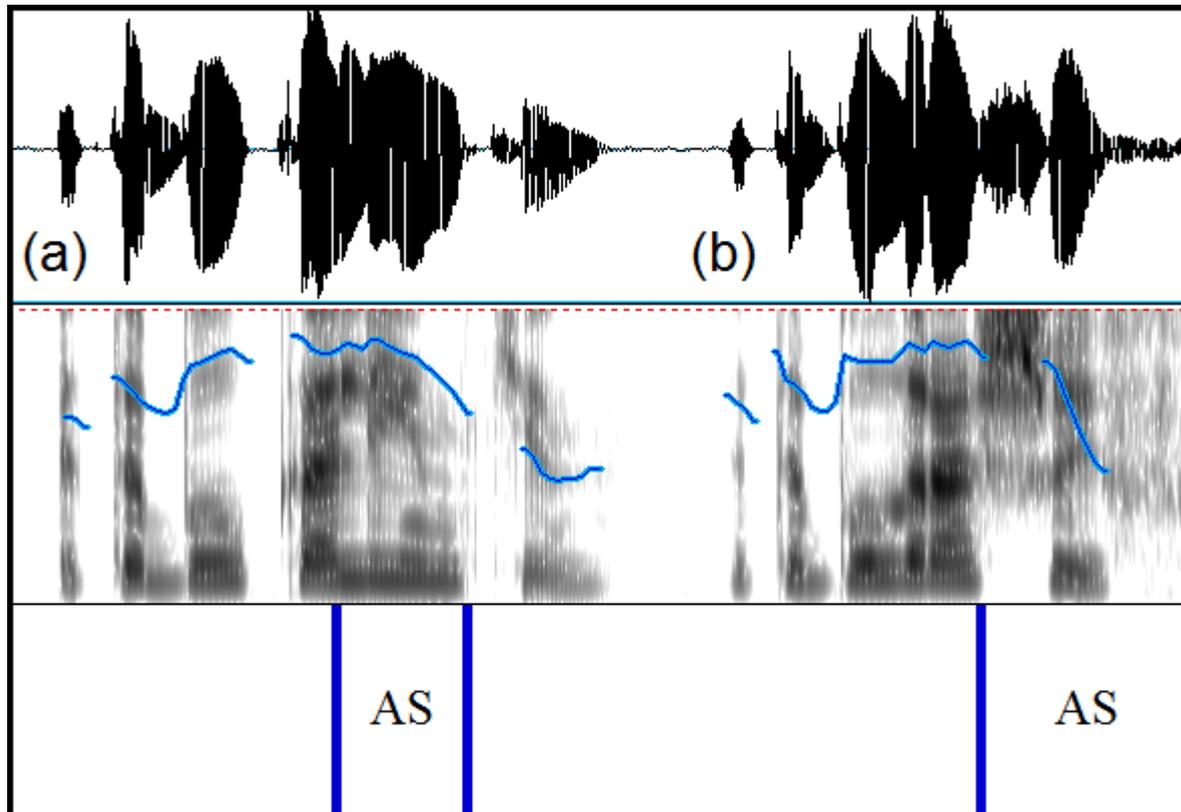
- Russian listeners respond to slight scaling differences
- Surprisingly, higher pitch cues L% and lower pitch cues H%



# Socrative #4

- Truncation? Compression? Realignment?

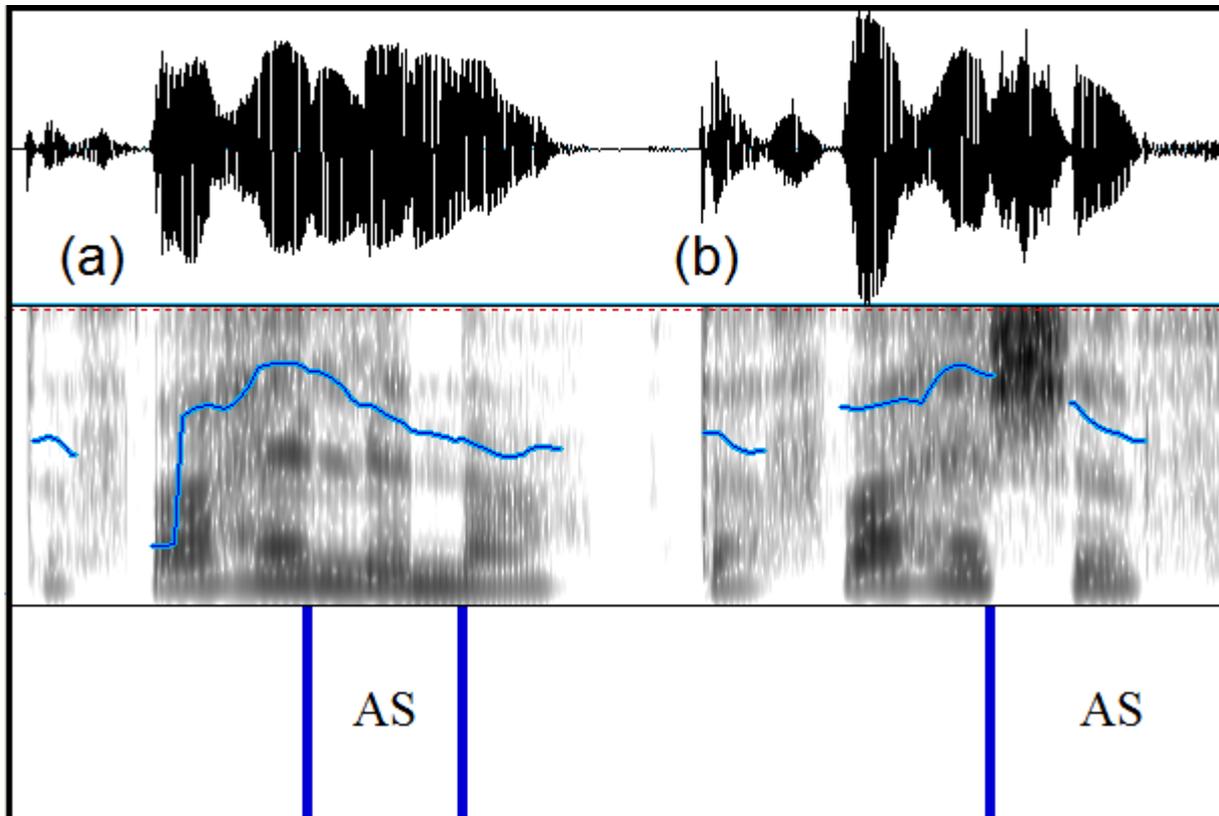
#1



# Socrative #4

- Truncation? Compression? Realignment?

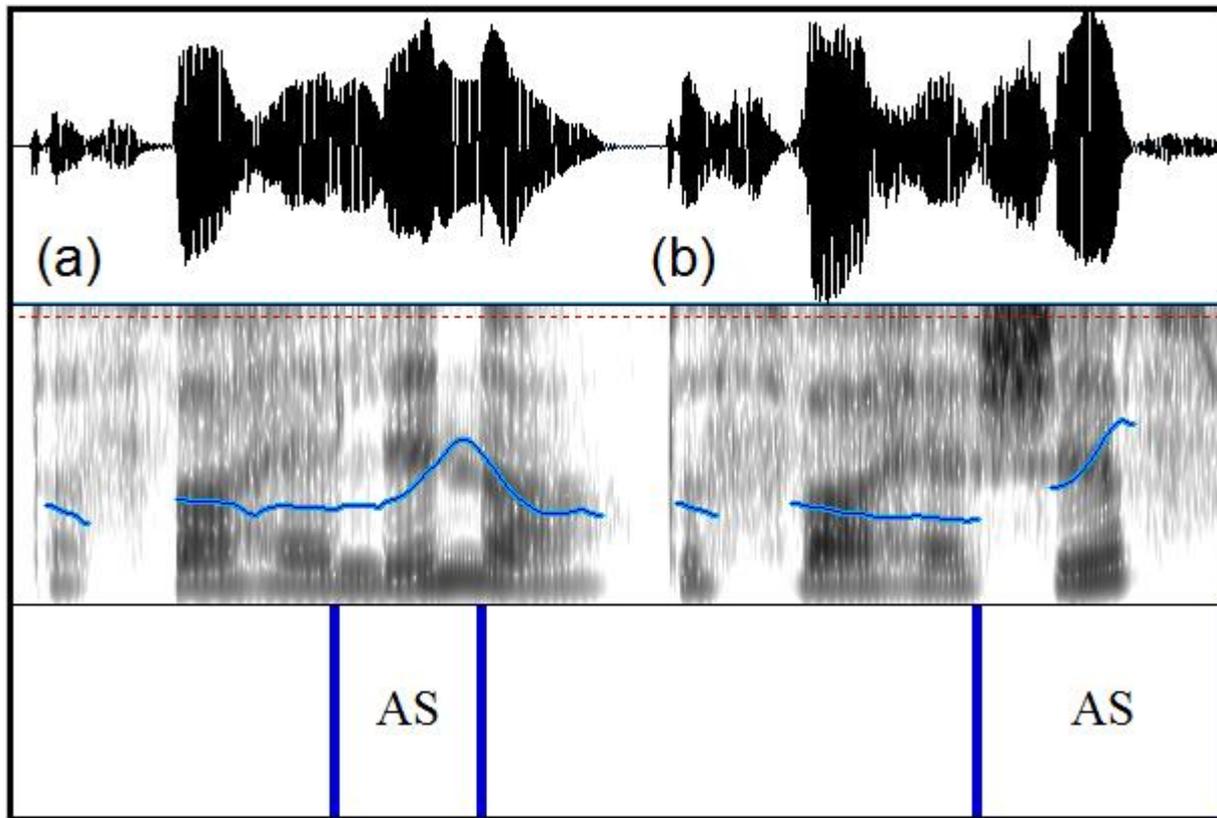
#2



# Socrative #4

- Truncation? Compression? Realignment?

#3

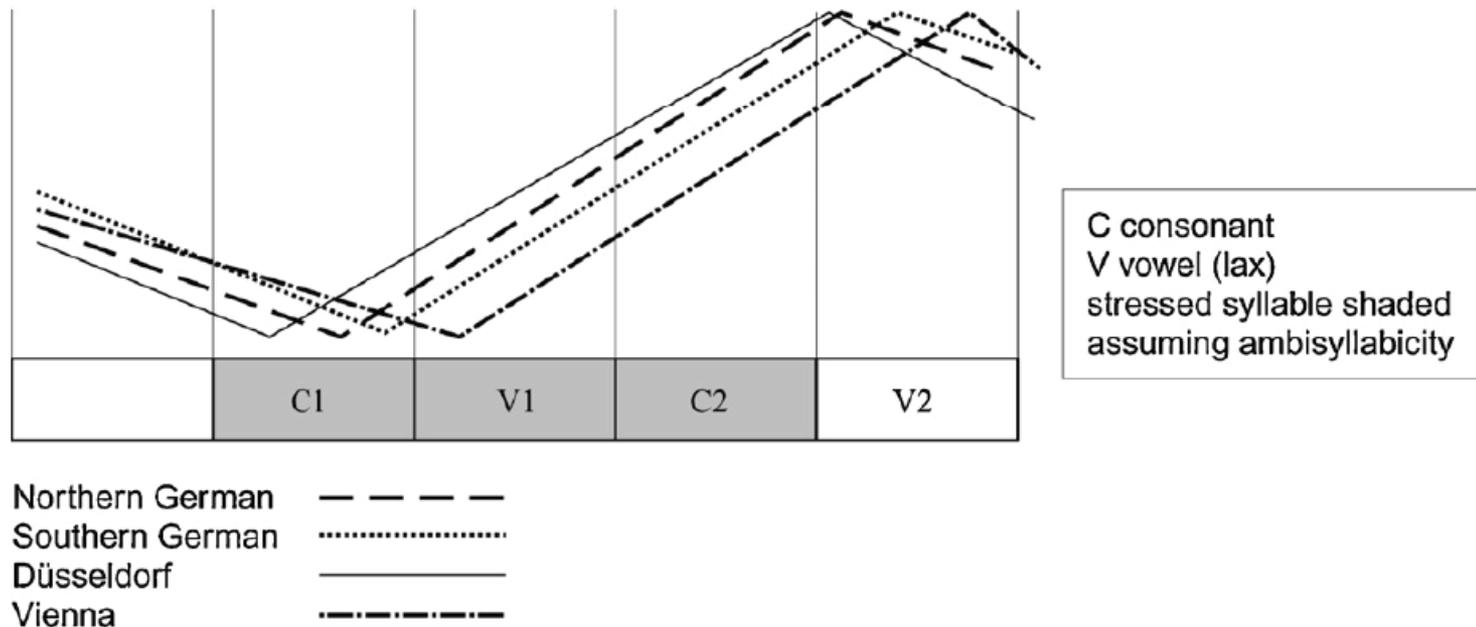


# SOCRATIVE #4

## More on F0-variation

*Dialect* can also have an influence on the phonetic realisation of pitch categories

- e.g. prenuclear rise in German

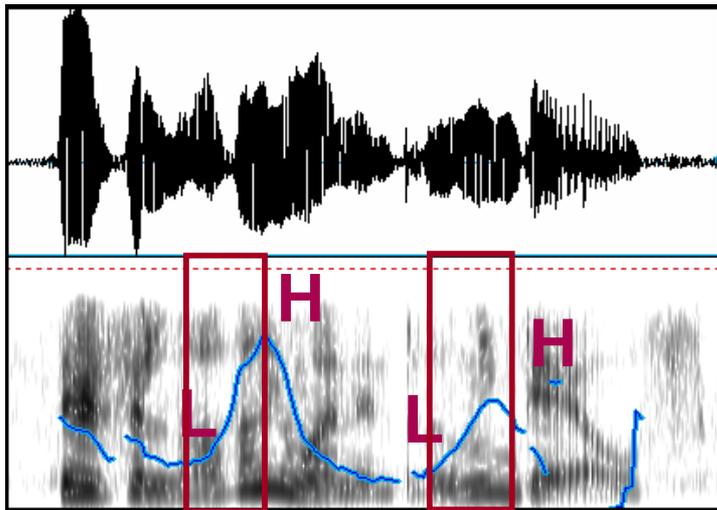


*Graph from Baumann & Rathcke (2013)*

*Data from Atterer & Ladd (2004) and Mücke et al. (2008)*

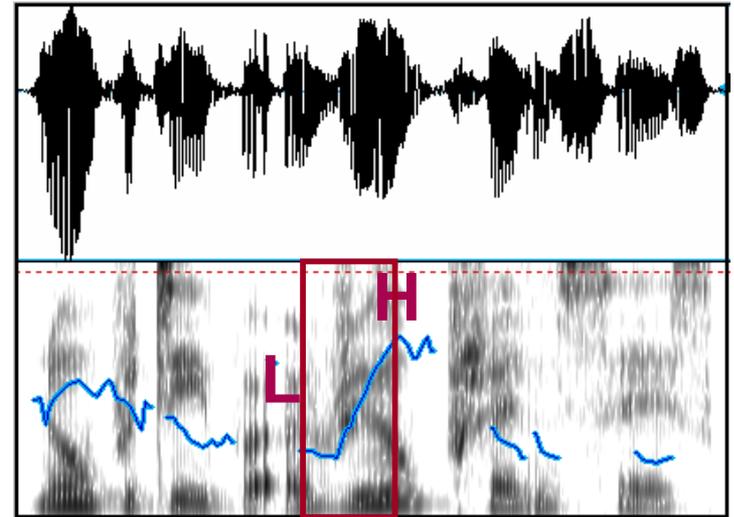
# A cross-dialectal example: L\*+H

Main pattern: low target on the accented syllable is immediately followed by relatively sharp rise to a peak, difference in “scoopiness”



Mainstream American:  
“lovely” & “Bloomington”

*From ToBI materials*



Glasgow: “real”

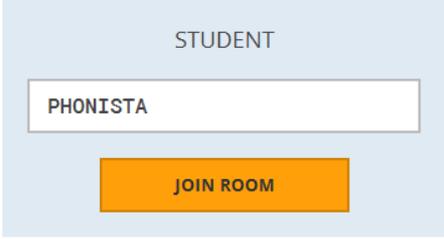
*From Rathcke & Smith (2015)*

# Core take-home messages

1. Be wary of F0 tracks in any signal processing programme
  - Tracking errors, octave jumps
2. Try to work with sonorous sentences
  - Number of events (segments, syllables) controls time available for f0-realisation
  - Voiceless/voiced segments define the amount of sonority
3. Don't trust what you see!
  - Check for microprosody and avoid measuring at the very start/end of a vowel (if surrounded by obstruents)
  - Check for positional effects
  - Listen! Compare prototypical and problematic realisations
  - Try to imitate what you hear (analysis-by-synthesis)
  - Be prepared to revise your decision (particularly if developing a new system, cf. H+L\* and H+!H\*)

# Space Race

## *A Big Truncation Contest*



STUDENT

PHONISTA

JOIN ROOM

- German has gradual truncation (H\* L-L%), responds to duration of the sonority available
- Which words will have more truncation than others?  
*Sie ist ...*

*{gesund; zu Hause; im Bett; allein; lieb, alt; doof; hübsch; satt; sauer}*

- Using these **10** words, create **9** pairs of words, starting with the words that would have the least amount of truncation, followed by the word with incrementally higher amount of truncation
  - Pair #1: nanny > nan (2 syllables vs. 1 syllable)
  - Pair #2: nan > Brad (sonorants vs. stops)
  - Pair #3: Brad > bred (intrinsic vowel duration)

# SOCRATIVE #5

# Fantastic work everyone!



Congratulations to the Winners of the Big Truncation Contest!

University of  
**Kent**