SDU
TECHNOLOGY ENTREPRENEURSHIP AND INNOVATION

Variation in F0 shape within and between phonological categories of intonation

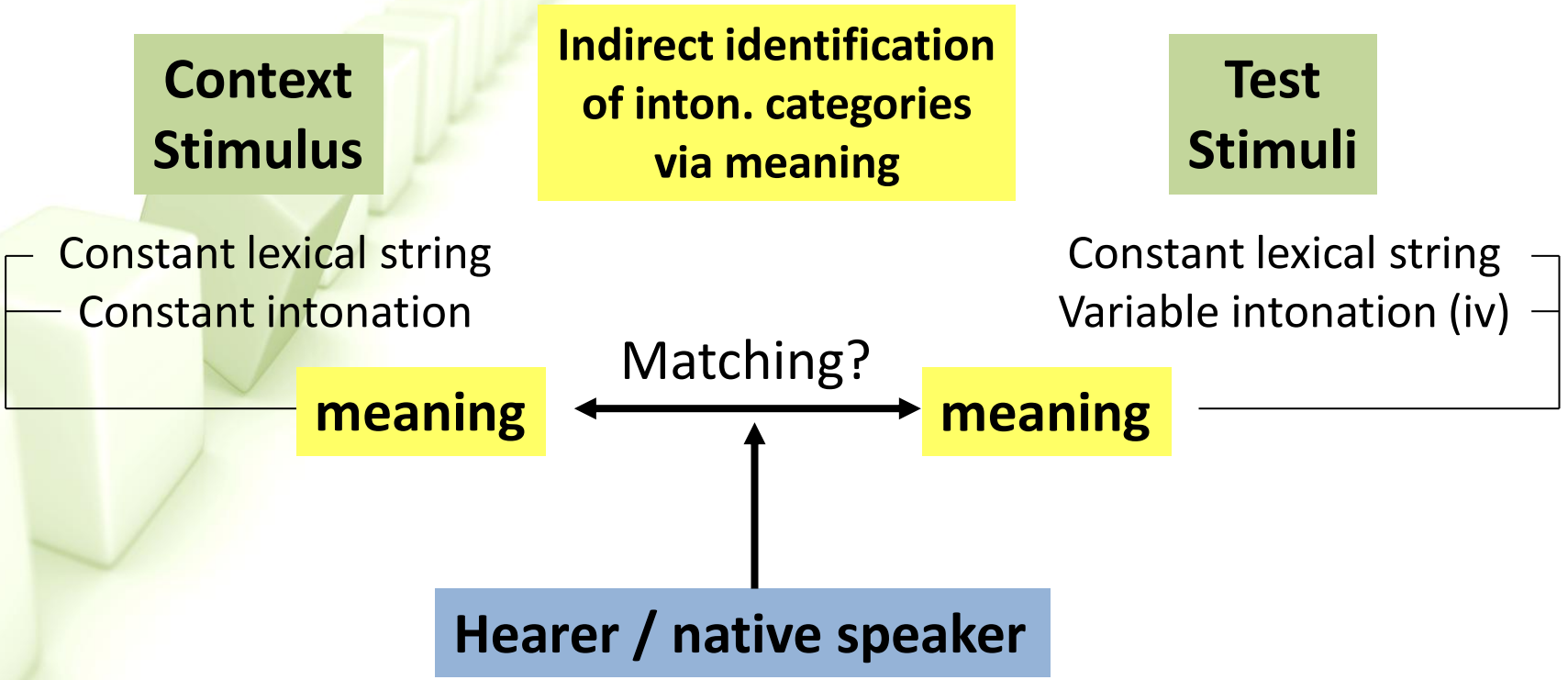
Oliver Niebuhr
Associate Professor of Communication and Innovation
Dept. of Technology Entrepreneurship and Innovation
Head of the Innovation Research Cluster Alsion
University of Southern Denmark (SDU)

Starting point: The 3 Pitch Accents of the Kiel Intonation Model

- Semantic-differential…
- and free-contextualization methods
- … were used to let naïve **listeners** outline the **meanings** of phonetically different pitch accent contours

- On this basis:
  - detailed phonetic analysis of accent **productions** (esp. alignment and scaling properties), elicited by giving the subjects appropriate semantic/ pragmatic contexts (e.g., situations, visual stimuli and/or previous utterances)
  - perception experiments with targeted manipulations of phonetic parameters → subjects identify the pitch accent categories (directly or indirectly) with reference to meanings → **indirect identification test**

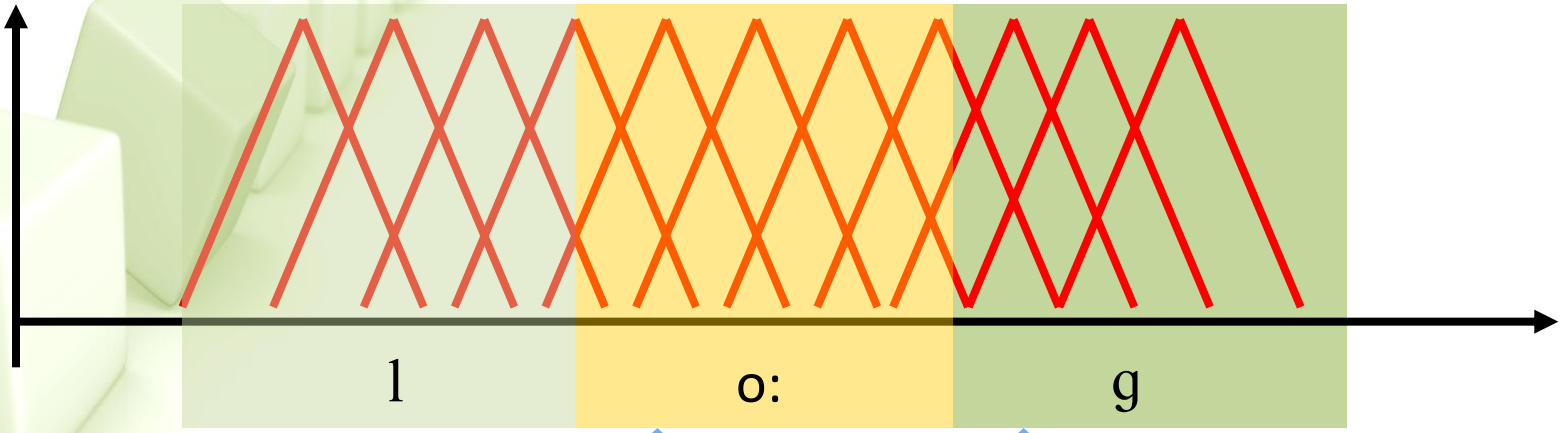
Starting point: The 3 Pitch Accents of the Kiel Intonation Model



- perception experiments with targeted manipulations of phonetic parameters → subjects identify the pitch accent categories (directly or indirectly) with reference to meanings → **indirect identification test**

Starting point: The 3 Pitch Accents of the Kiel Intonation Model

"Sie hat ja gelogen" (She's been lying), 11 stimuli



Abrupt change in pitch-accent meaning

Second change in pitch-accent meaning

Stimuli presented/judged in serial order and in randomized order with multiple repetitions

Starting point: The 3 Pitch Accents of the Kiel Intonation Model

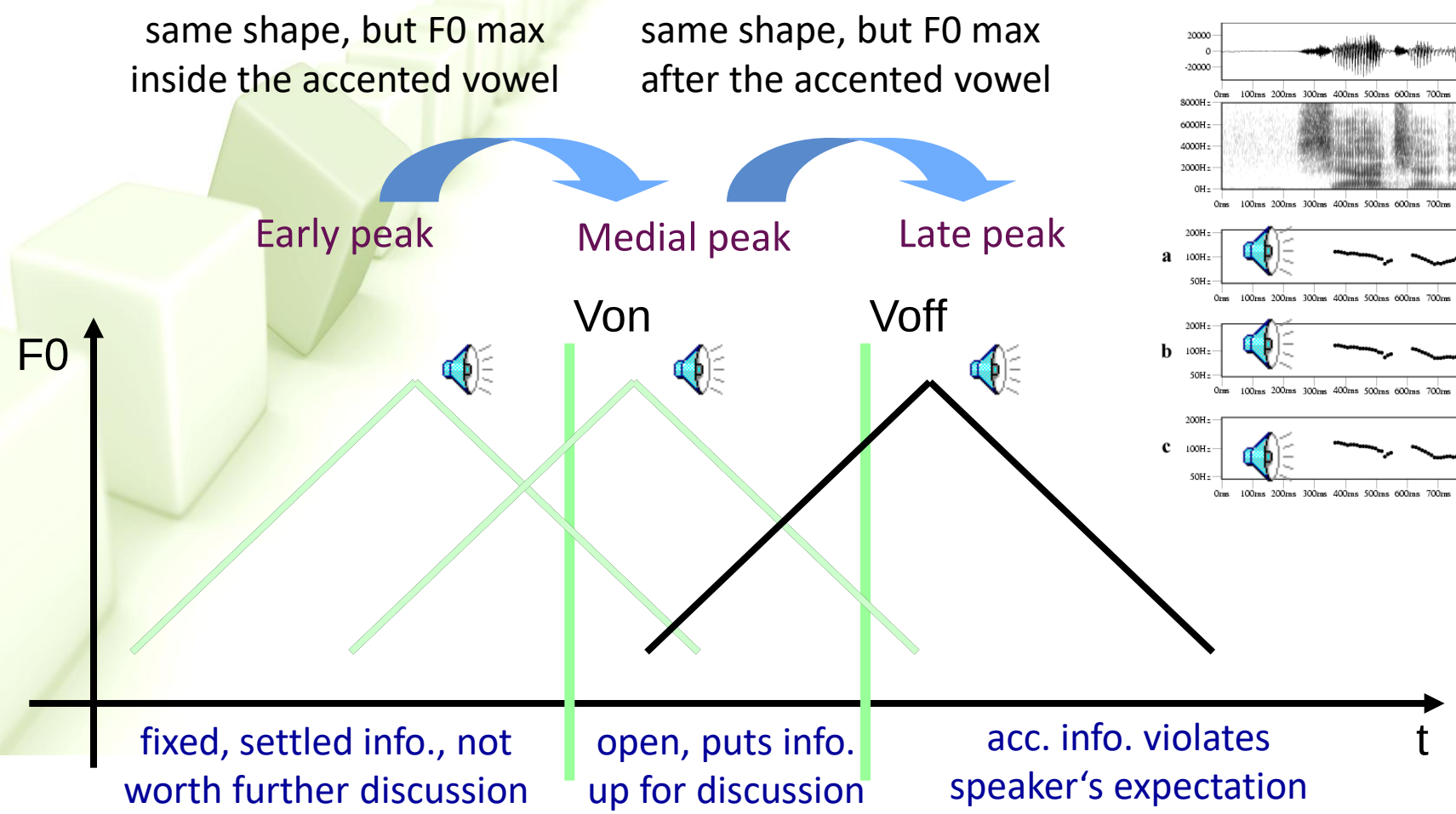
- The resulting phonological framework KIM
  - defines pitch-accent categories in terms of **contours**
  - the segmental string is regarded as a **separate, independent layer** of the speech signal

  - Pitch-accent **peaks** …
    - have a **pointed, rising-falling shape**
    - are phonologically distinguished by the synchronization of the F0-peak **maximum** relative to the boundaries of the **accented vowel**
    - → peak timing (alignment) is a direct phonological feature that refers to 2 **local** acoustic **landmarks**

Starting point: The 3 Pitch Accents of the Kiel Intonation Model

same shape, but F0 max inside the accented vowel

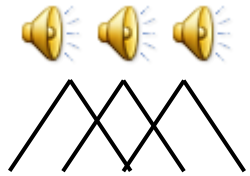
same shape, but F0 max after the accented vowel



Starting point: The 3 Pitch Accents of the Kiel Intonation Model

same shape, but F0 max inside the accented vowel

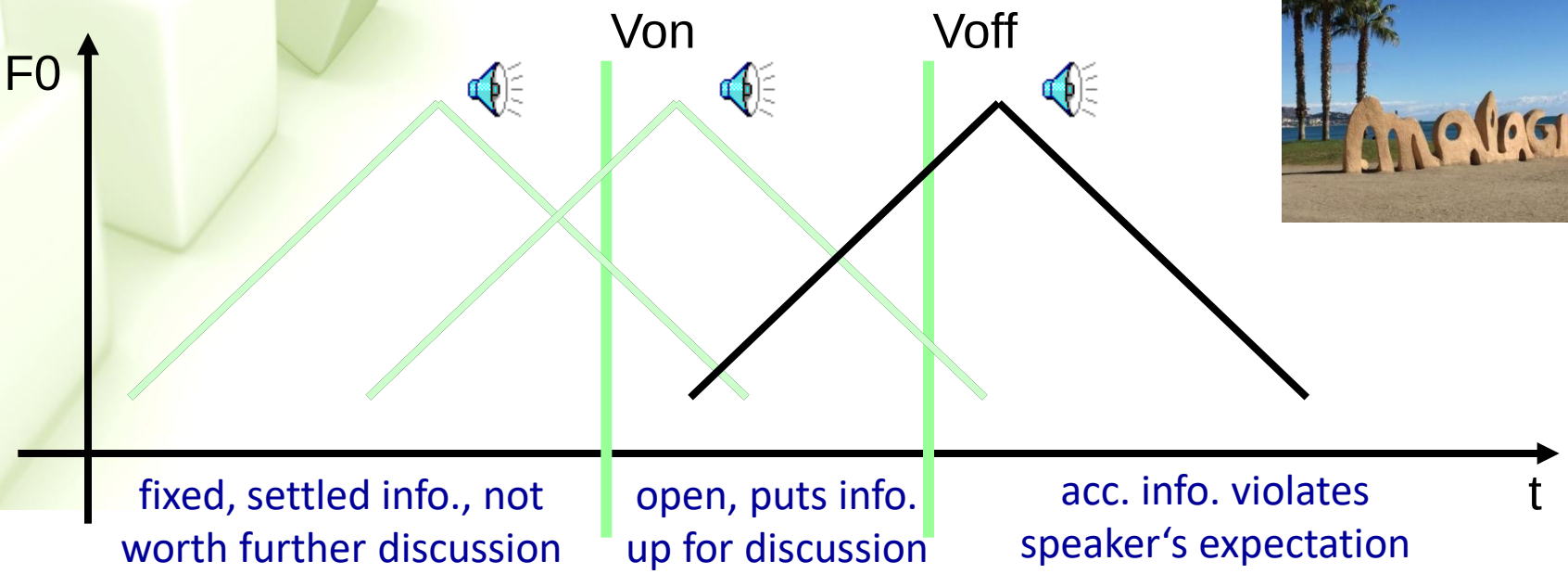
same shape, but F0 max after the accented vowel



Early peak

Medial peak

Late peak



fixed, settled info., not worth further discussion

open, puts info. up for discussion

acc. info. violates speaker's expectation

Starting point: The 3 Pitch Accents of the Kiel Intonation Model

- **We will mainly focus on these three pitch accent categories**
  - There is some consensus across phonological models of German intonation that these 3 categories should be distinguished (e.g., GToBI: H+L*, H*, H+!H*)
  - Very well investigated for almost 30 years now (since Kohler 1987)
  - They are used across languages
- Their attitudinal meanings are also well understood → Their specific interpretation can slightly vary with the semantic/pragmatic context

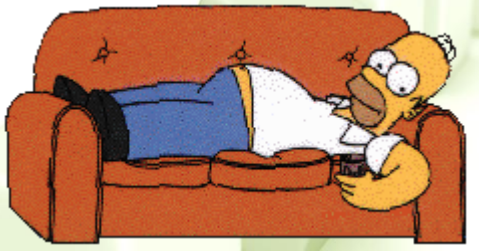
fixed, settled info., not worth further discussion

open, puts info. up for discussion

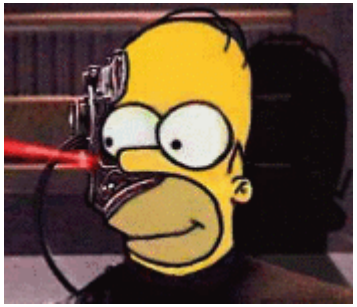
acc. info. violates speaker's expectation

Starting point: The 3 Pitch Accents of the Kiel Intonation Model

given



New



contrastive

- There is a correlation with information-structure, but no 1:1 relationship

fixed, settled info., not worth further discussion

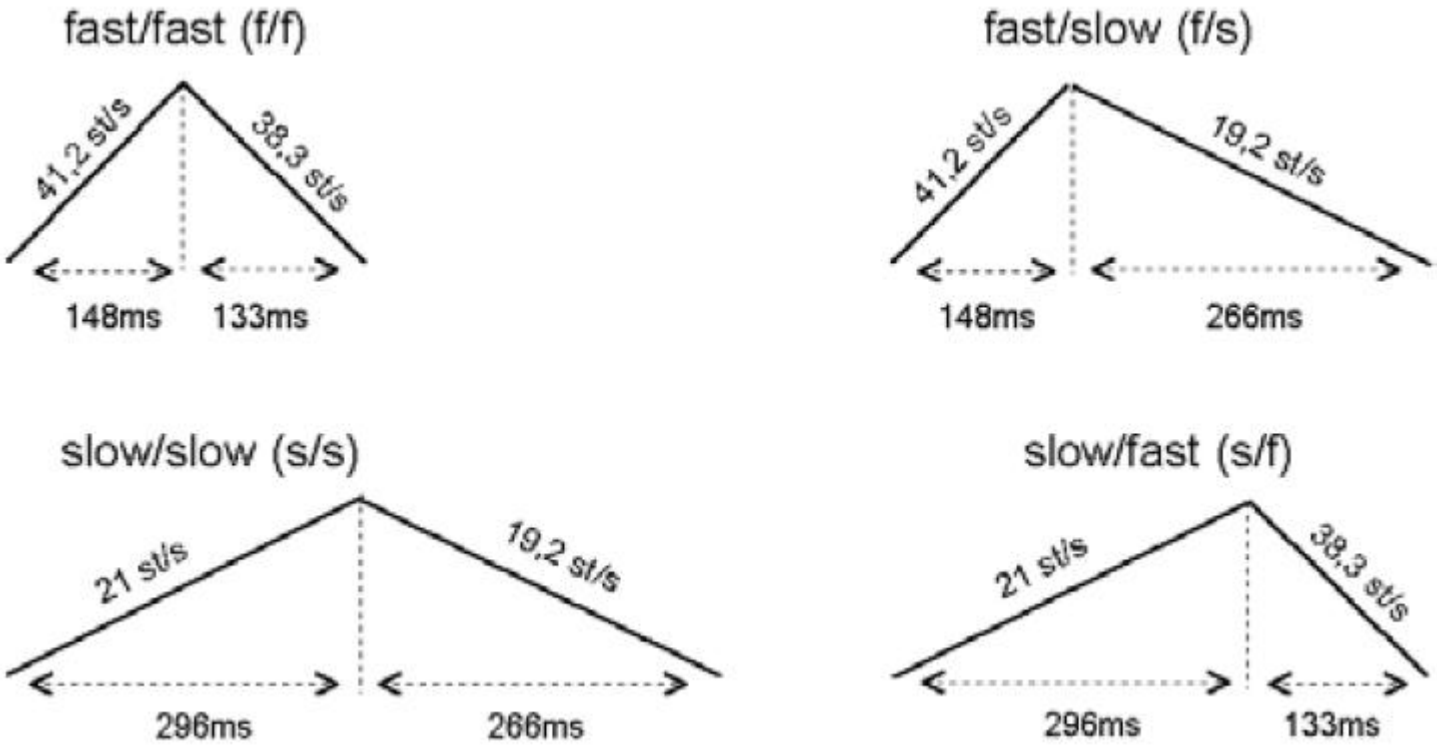
open, puts info. up for discussion

acc. info. violates speaker's expectation

Shaping Rises and Falls

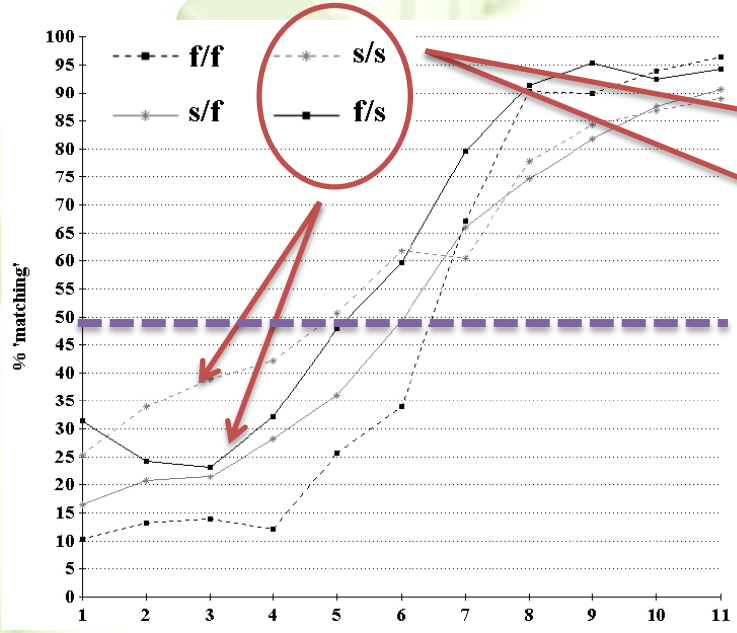
Shaping Rises and Falls

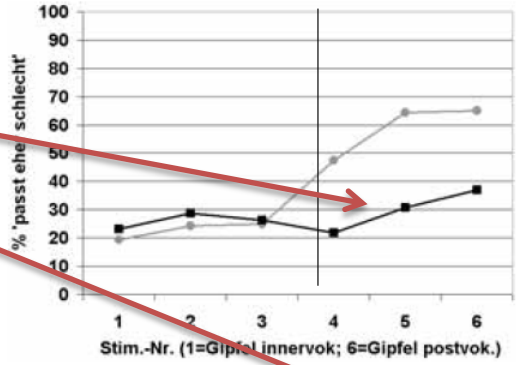
- Niebuhr (2003, 2007) replicated the F0-peak shift continua of Kohler (1987), but with 4 different peak shapes
- Target word "Malerin" (painter) → "Sie war mal Malerin"

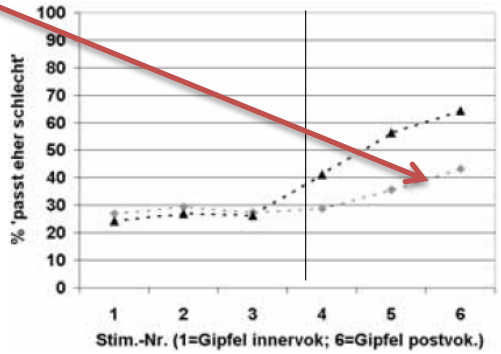


Shaping Rises and Falls

- Main finding: F0-peak shape has a strong effect on pitch-accent identification; the effect **crosses** the synchronization features



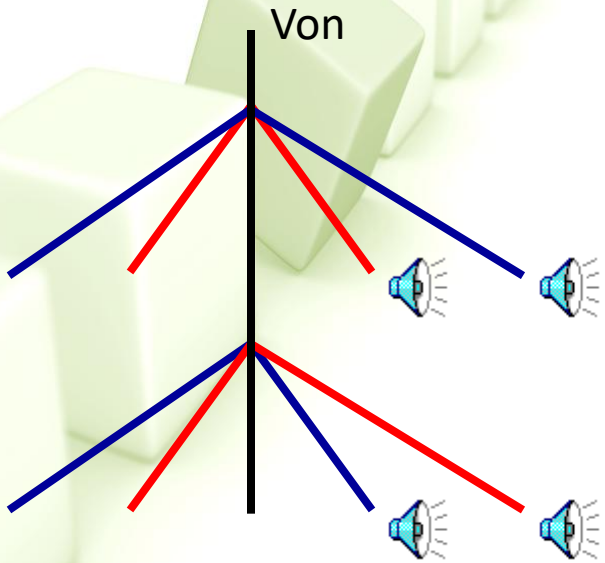




- Similar shape effects at the accented-vowel offset
- Here, slowly rising-falling F0 peaks are even **unable** to trigger a perceptual change from ‘medial’ to ‘late’
- Fast rise-fall → ‘late’ when F0 max 40 ms after Voff

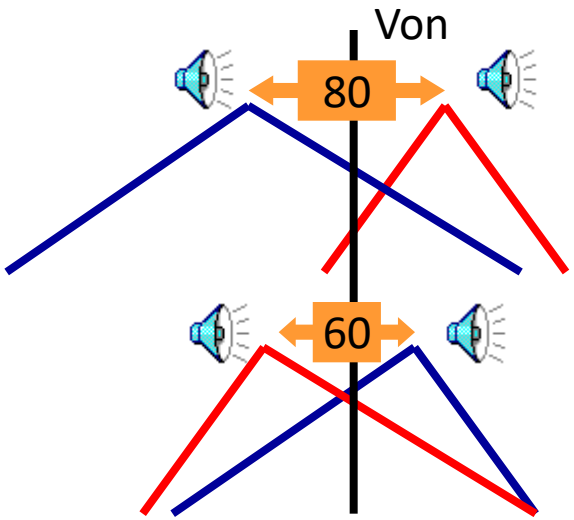
Shaping Rises and Falls

- Main finding: F0-peak shape has a strong effect on pitch-accent identification; the effect **crosses** the synchronization features



same synchronizations yield clearly different pitch-accent identifications

in order to yield a majority of 'medial' pitch-accent identifications …



- Similar shape effects at the accented-vowel offset
- Here, slowly rising-falling F0 peaks are even **unable** to trigger a perceptual change from 'medial' to 'late'
- Fast rise-fall → 'late' when F0 max 40 ms after Voff

Shaping Rises and Falls

- Main finding: F0-peak shape has a strong effect on pitch-accent identification; the effect **crosses** the synchronization features

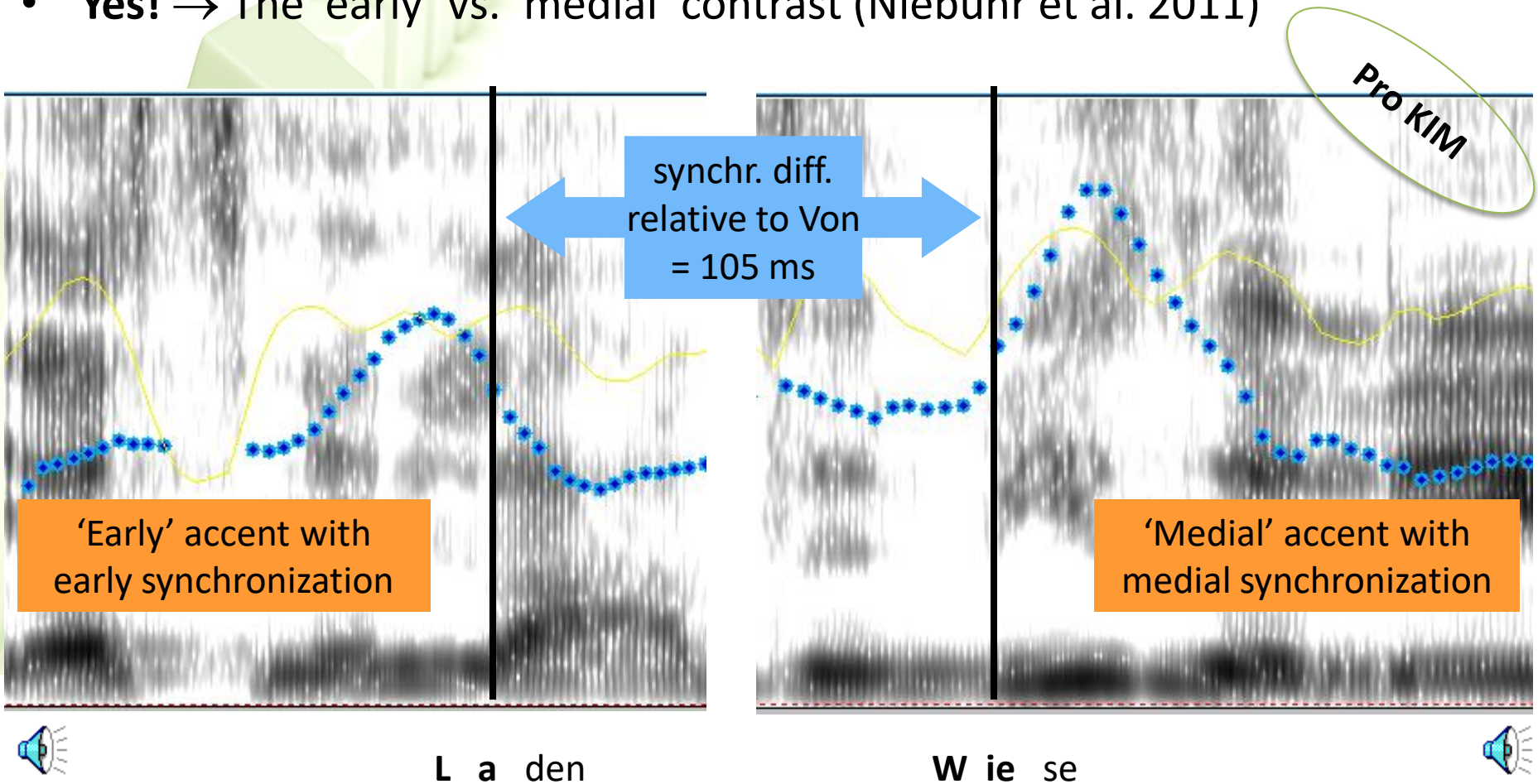
These clear perception result are not predicted by an intonaitonal model in which the only relevant feature is the peak maximum synchronization relative to the onset or offset of the accented vowel

But, maybe this is a mere perception phenomenon based on manupulations/variations that never occur in acutal speech production?

- Similar shape effects at the accented-vowel offset
- Here, slowly rising-falling F0 peaks are even **unable** to trigger a perceptual change from 'medial' to 'late'
- Fast rise-fall → 'late' when F0 max 40 ms after Voff

Shaping Rises and Falls

- Are these shape-differences used in actual speech production?
- **Yes!** → The 'early' vs. 'medial' contrast (Niebuhr et al. 2011)

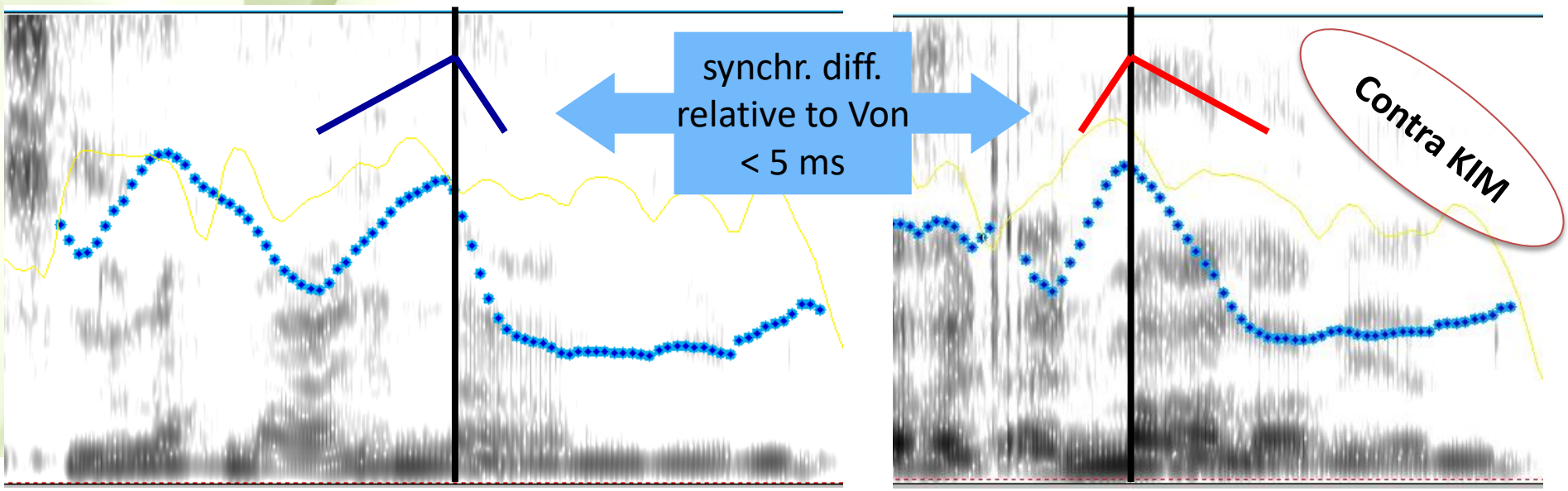


Shaping Rises and Falls

- Are these shape-differences used in actual speech production?
- **Yes!** → The 'early' vs. 'medial' contrast (Niebuhr et al. 2011)

'Early' accent with early synchronization

'Medial' accent with **EARLY** synchronization

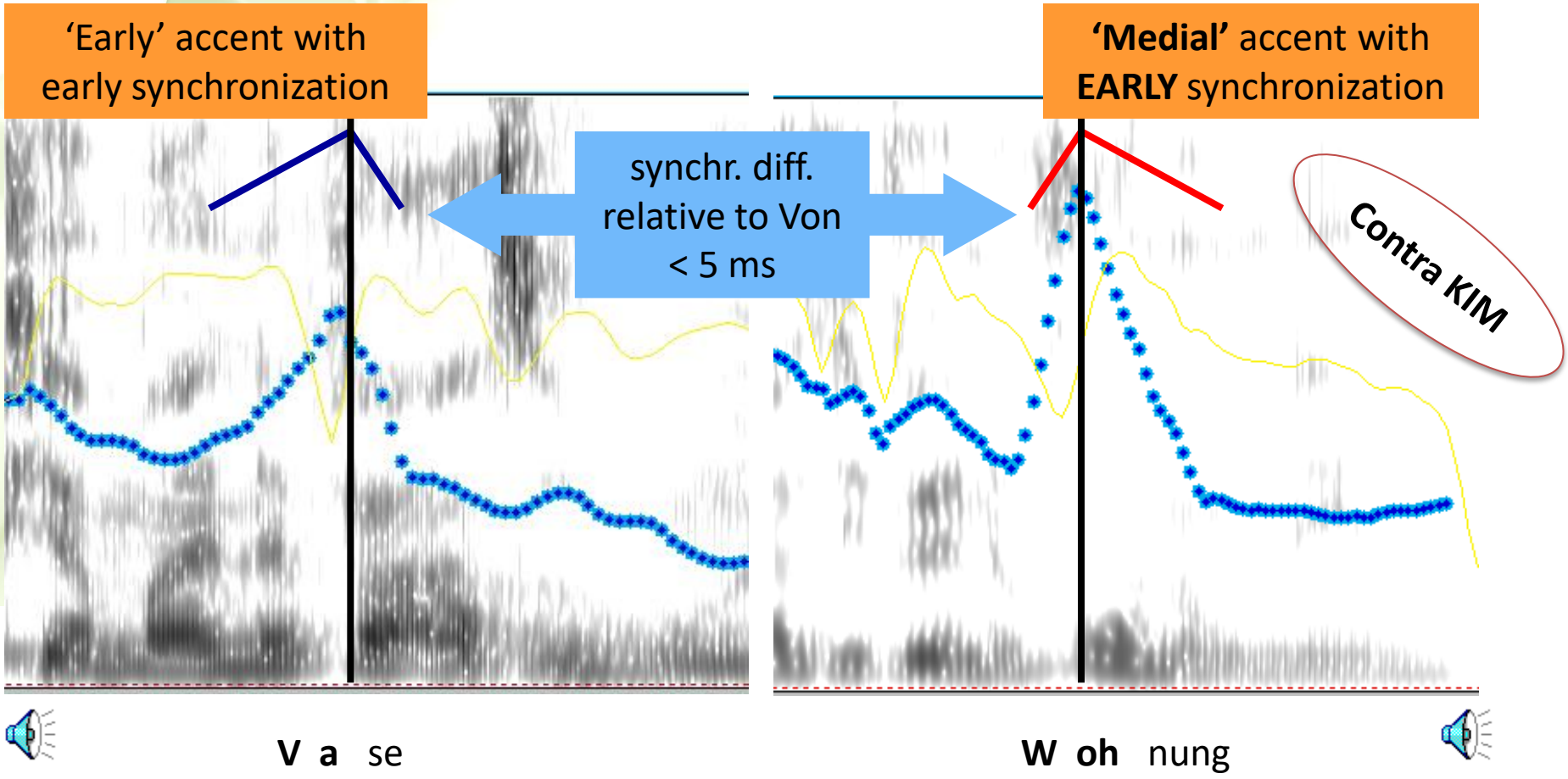


N a men

M a ler

Shaping Rises and Falls

- Are these shape-differences used in actual speech production?
- **Yes!** → The 'early' vs. 'medial' contrast (Niebuhr et al. 2011)

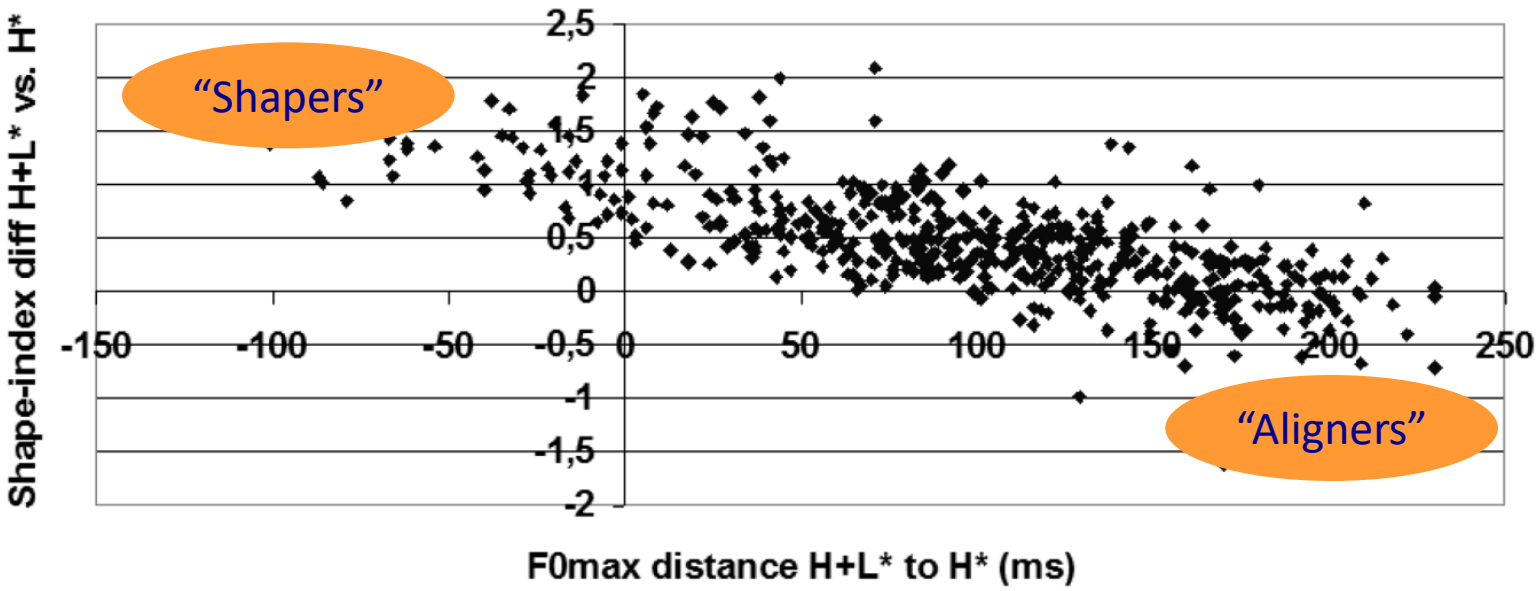


Shaping Rises and Falls

- Across 35 speakers of Northern Standard German:
- neg. correlation (r=-0.72; p<0.001) → **trade-off** between synchronization and shape in expressing the 'early'-'medial' contrast
- At the extremes: 2 strategies → **"Shapers" & "Aligners"**

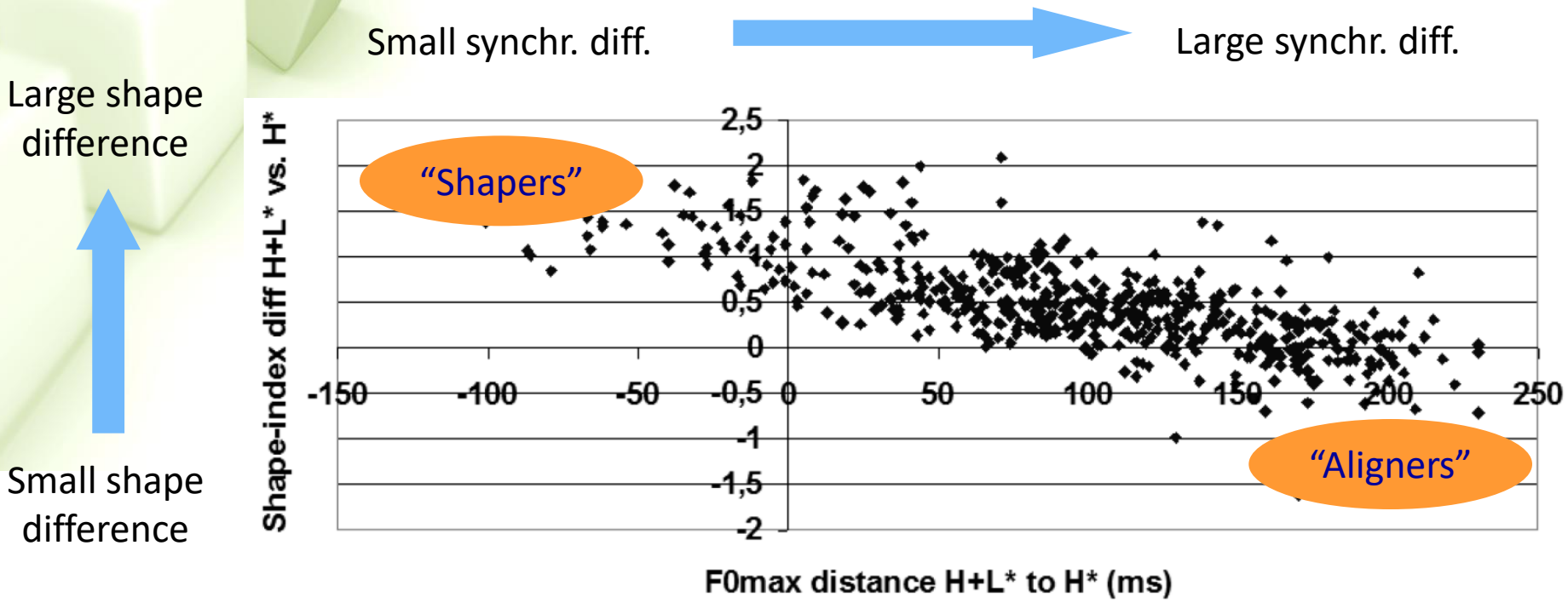
Small synchr. diff. → Large synchr. diff.

Large shape difference ↑ Small shape difference



Shaping Rises and Falls

- There are a lot more "Aligners" than "Shapers" in real life
- Only about **14 %** of the speakers make more use of peak shape than of peak alignment to distinguish pitch-accent categories
- **Yet, it is clear that the KIM's synchronization concept does not hold!**



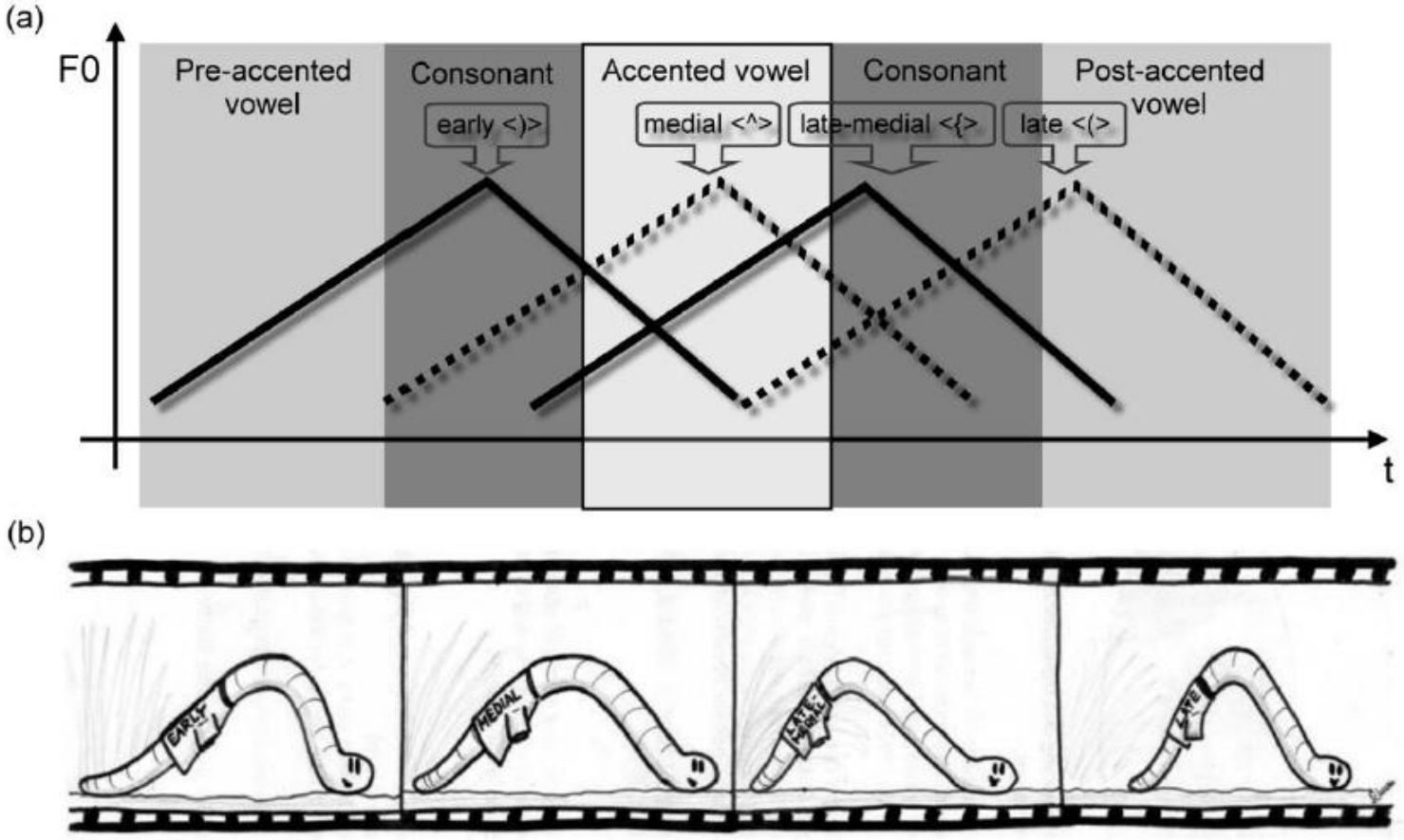
Shaping Rises and Falls

- There are a lot more "Aligners" than "Shapers" in real life
- Only about **14 %** of the speakers make more use of peak shape than of peak alignment to distinguish pitch-accent categories
- **Yet, it is clear that the KIM's synchronization concept does not hold!**

- And even those speakers who primarily use alignment show within-category shape variation, i.e. **pitch-accent specific peak shapes**, suitable to **support the perceptual identification** of the respective category!

- The "Worm Model" (Niebuhr 2017)

Shaping Rises and Falls



- The “Worm Model” (Niebuhr 2017)

Shaping Rises and Falls

- It seems that the 2 speaker-specific strategies are also present in previous production studies on alignment and segmental anchoring, but **masked by grand means** → Atterer & Ladd (2004)

Table 1
Mean alignment data for Experiment 1

Speaker	H (ref V1)	L (ref V0)	L (ref C0)
NF1	30.1	−33.6	57.0
NM2	8.2	−40.2	41.5
NM3	26.6	−69.1	10.4
NM4	3.4	−30.7	49.3
NF5	26.3	−27.6	48.2
NM6	13.7	−50.4	19.8
NF7	41.8	−24.3	41.0
Northern grand mean	21.4	−39.4	38.2
SF1	27.5	9.4	79.3
SM2	58.8	6.3	75.0
SM3	41.9	−6.4	64.2
SM4	24.3	−4.4	57.8
SM5	28.3	−26.3	57.6
SF6	8.4	−8.7	56.8
SF7	48.5	8.6	84.3
Southern grand mean	34.0	−3.1	67.9

The columns show the distance in ms between an *F0* label (L or H) and a segmental landmark (C0, V0, or V1). A negative value indicates that the *F0* label occurs before the segmental label.

=38.4ms

=50.4ms

46.6ms=

=35.7ms across segm.

Shaping Rises and Falls

- What factors determine where in the continuum from “Aligner” to “Shaper” a speaker is located?

- Study of Lindenroth (2013)
- 6 Speakers recited the “Lord’s Prayer” (= “Vater Unser”)

- Advantage:
  - Speaker know it by heart, no practicing and no reading necessary
  - Yields a sufficient number of early and medial peaks (n≥12 per categ.)

VP	Alter	♂/♀	Größe	Gewicht	Alkoholkonsum	Krankheiten	Fühlst Du Dich gesund?
1.	40	♂	1,82 m	81 kg	2-3x	nein	Ja
2.	36	♂	1,83 m	88 kg	5x wöchen.	nein	Ja
3.	43	♂	1,86 m	92 kg	1x wöchen.	nein	Ja
4.	34	♀	1,75 m	62 kg	2-3 wöchen.	nein	So lala
5.	26	♀	1,82 m	~90 kg	1x monatl.	nein	Ja
6.	34	♀	1,70 m	~81 kg	5x wöchen.	nein	Ja

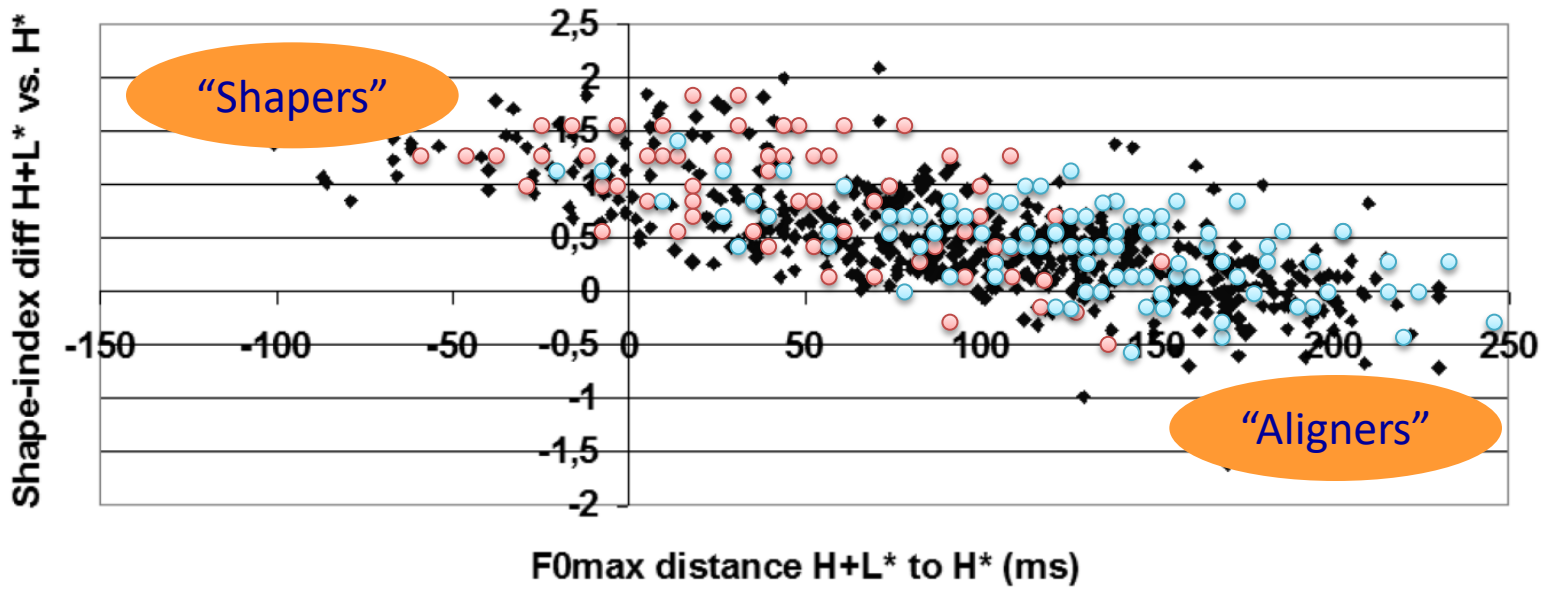
Shaping Rises and Falls

- What factors determine where in the continuum from "Aligner" to "Shaper" a speaker is located?

- Study of Lindenroth (2013)
- **Speaker gender seems to play a role! Females go more in the direction of "Shapers"**

Small synchr. diff. → Large synchr. diff.

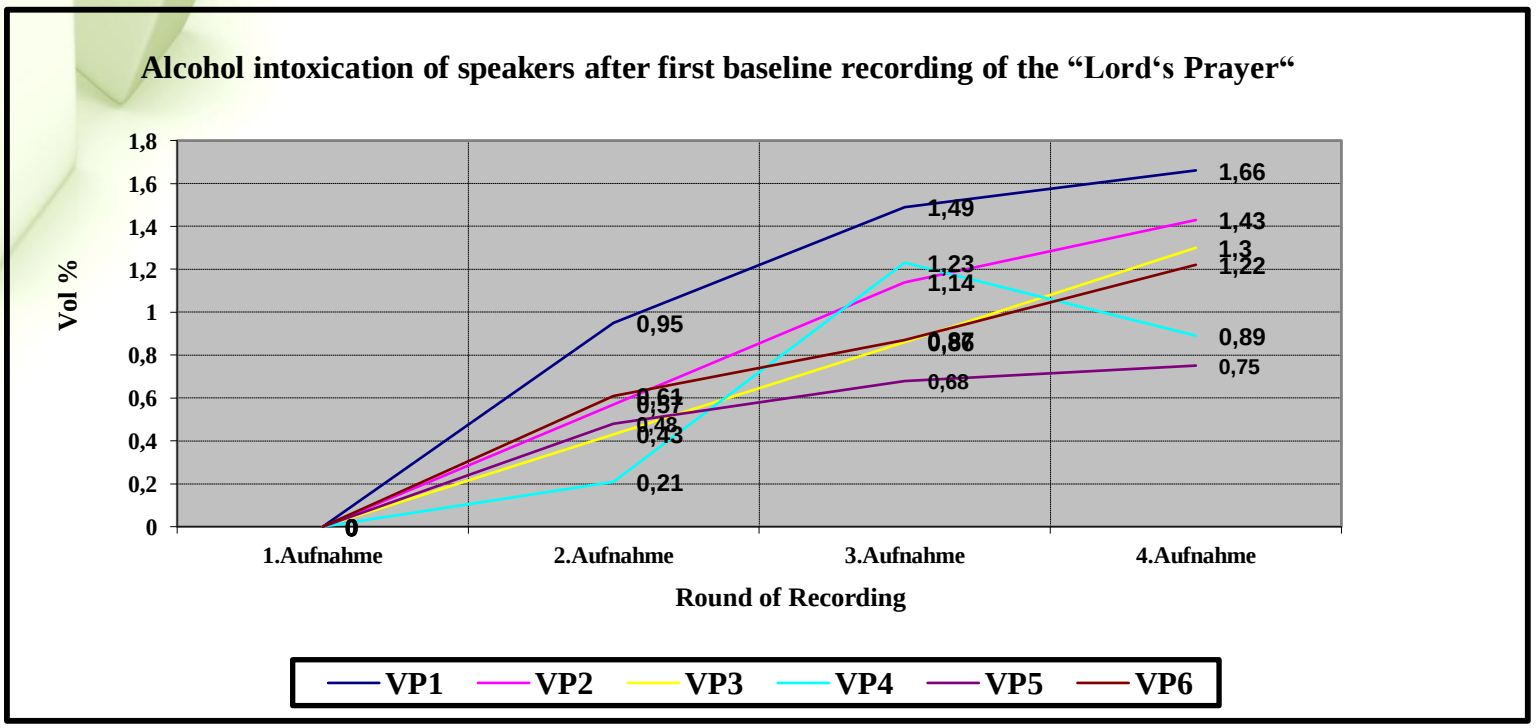
Large shape difference ↑ Small shape difference



Shaping Rises and Falls

- What factors determine where in the continuum from "Aligner" to "Shaper" a speaker is located?

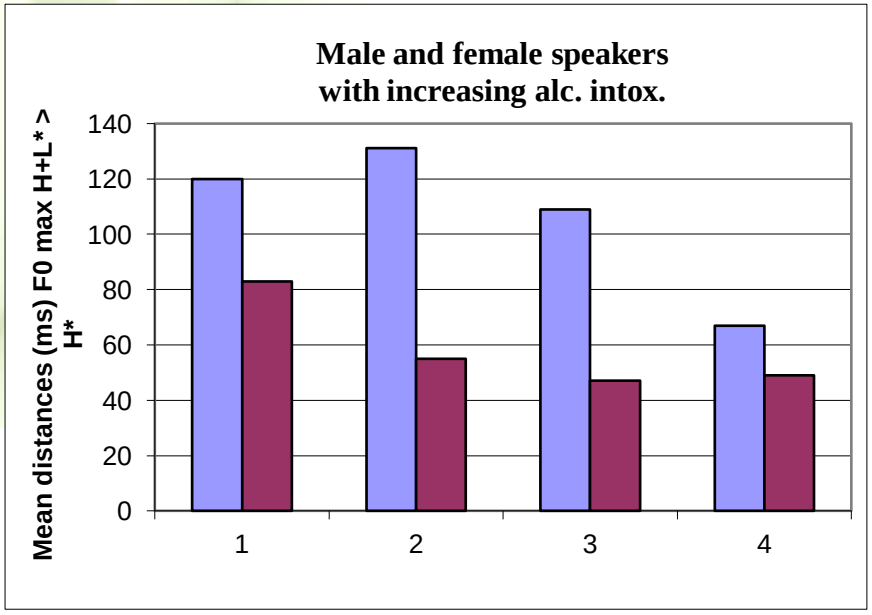
- Study of Lindenroth (2013)

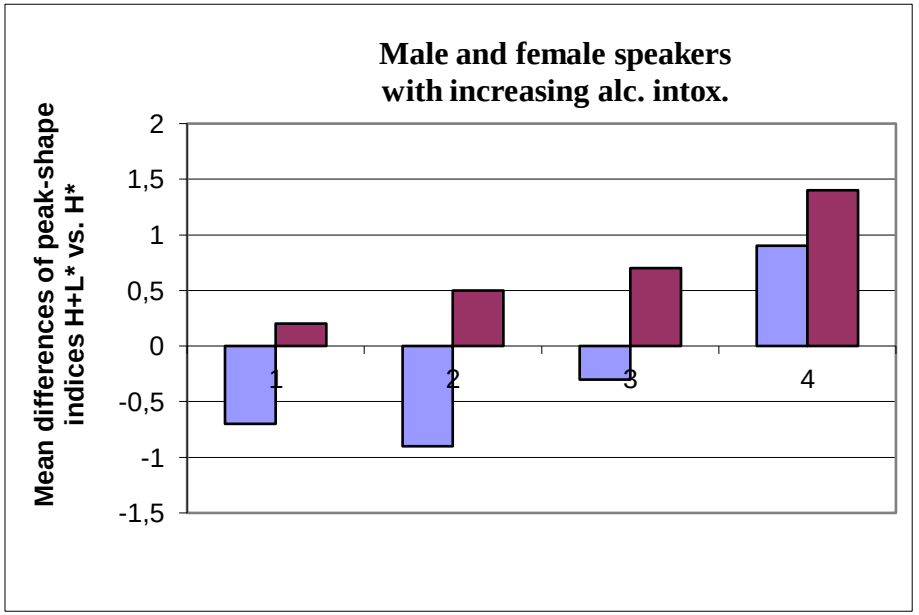


Shaping Rises and Falls

- What factors determine where in the continuum from "Aligner" to "Shaper" a speaker is located?

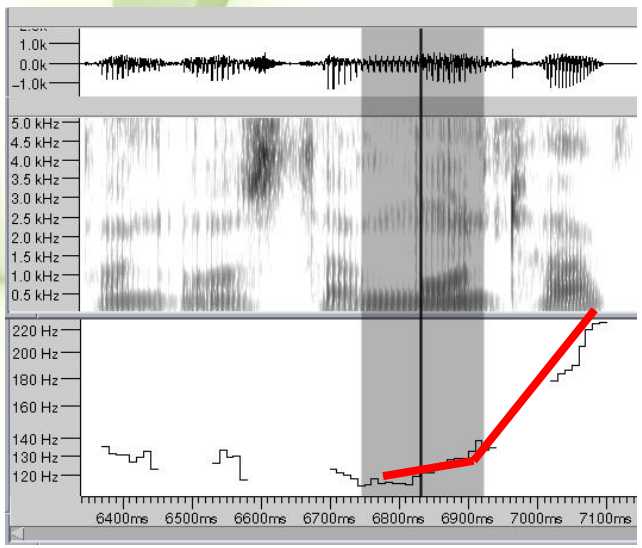
- Study of Lindenroth (2013)
- **Alcohol supports a shift in pitch-accent distinction from alignment to shaping** – less laryngeal control = less precise synchronization = shaping is easier (peaks also become generally flatter)

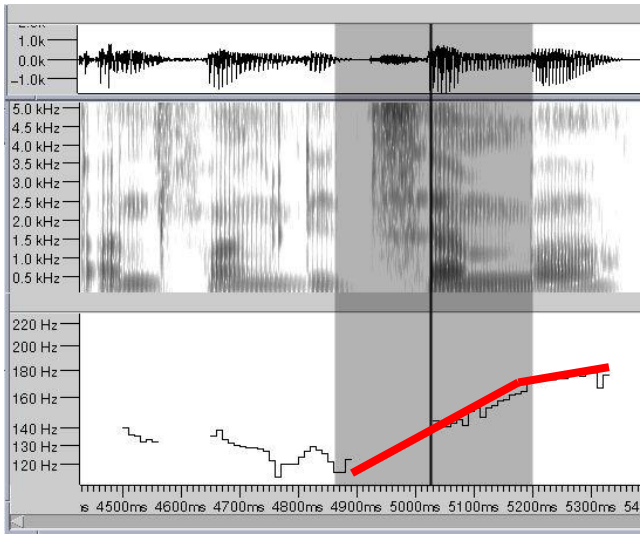




Shaping Rises and Falls

- Phrase-final rises: convex rises activate the dialg. partner (→ turn-yielding) ↔ concave rises restrain the dialogue partner (→ turn-holding)
- Shown in Dombrowski & Niebuhr (2005, 2010) by means of 177 phrase-final rises from the 'Kiel Corpus of Spontaneous Speech'
- English:

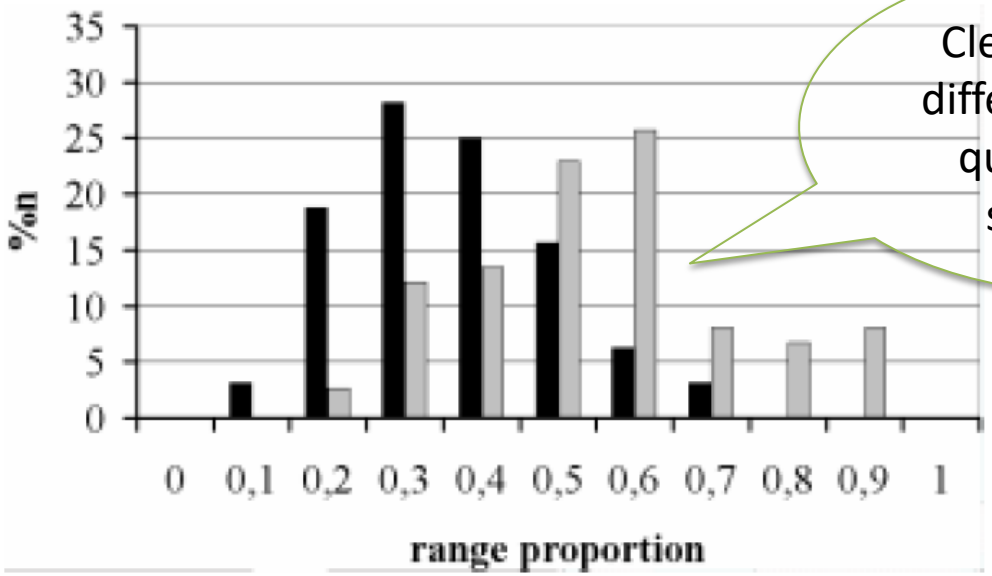




- A similar form-function relationship exists in combination with discourse markers in Estonian (Asu 2006) and with sentence mode in Swedish (Ambrazaitis et al. 2015)

Shaping Rises and Falls

- Phrase-final rises: convex rises activate the dialg. partner (→ turn-yielding) ↔ concave rises restrain the dialogue partner (→ turn-holding)
- Shown in Dombrowski & Niebuhr (2005, 2010) by means of 177 phrase-final rises from the 'Kiel Corpus of Spontaneous Speech'

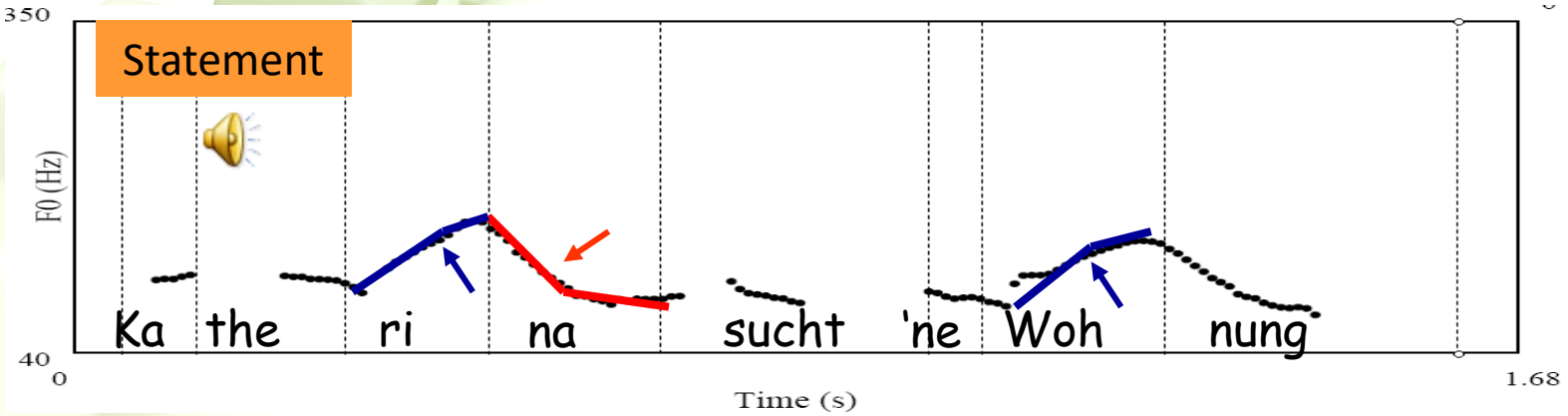


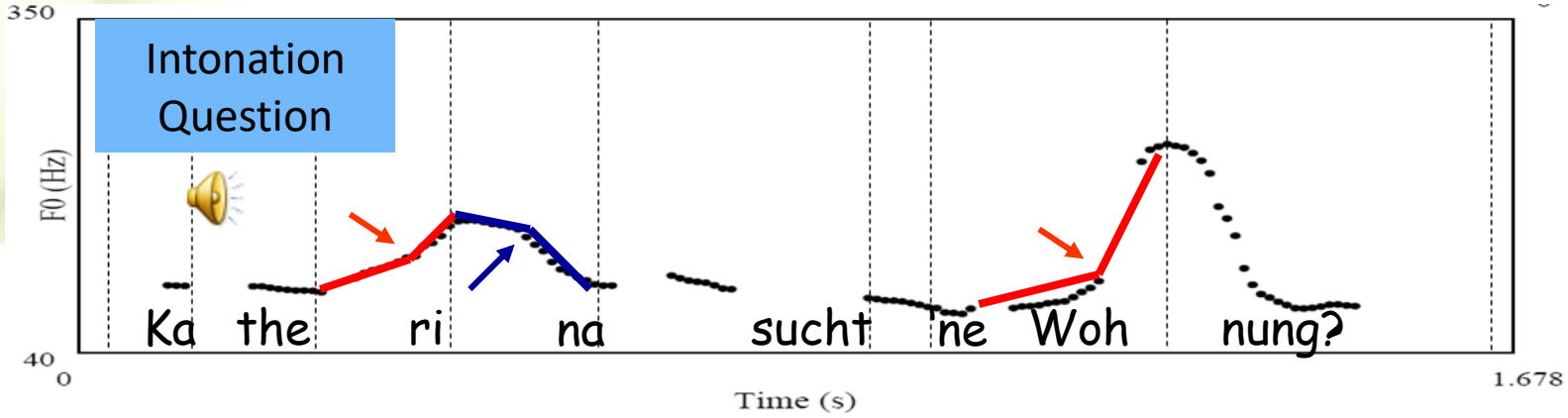
Clear functional difference in both questions and statements

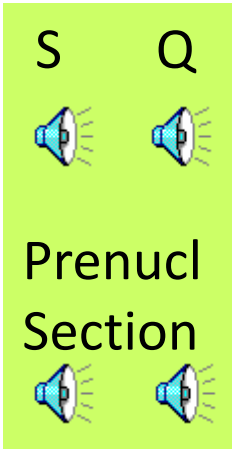
Figure 5a. Relative frequencies of the values for the *range proportion* in the turn-yielding (black) and the turn-holding (grey) conditions. The *range proportion* is scaled in tenths ranging from 0 to 1.

Shaping Rises and Falls

- In prenuclear and nuclear pitch accents: **rises are more concave and falls are more convex in questions** than in statements (Petrone & Niebuhr 2009, 2013).







Shaping Rises and Falls

- In prenuclear and nuclear pitch accents: **rises are more concave and falls are more convex in questions** than in statements (Petrone & Niebuhr 2009, 2013).

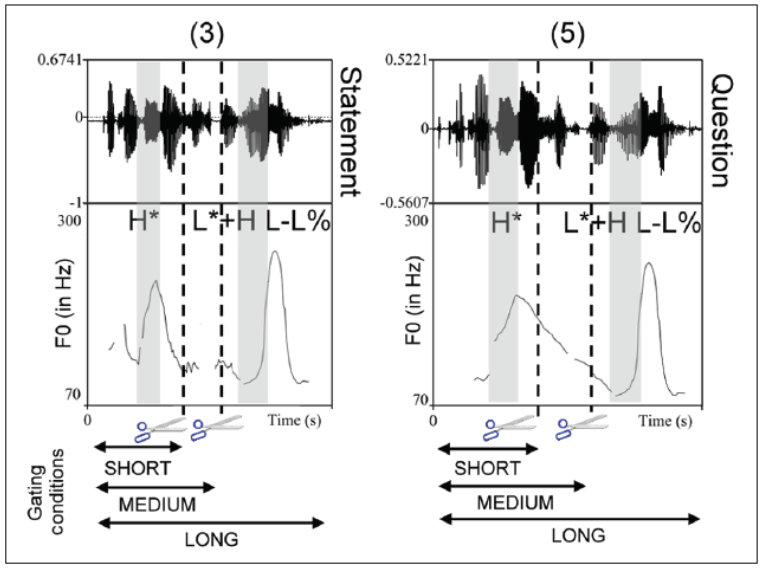
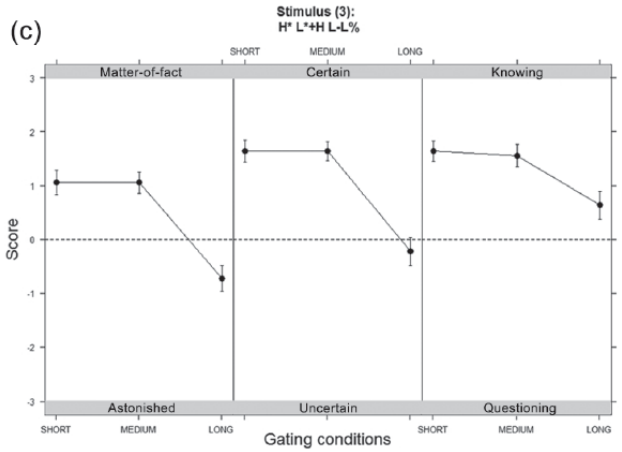
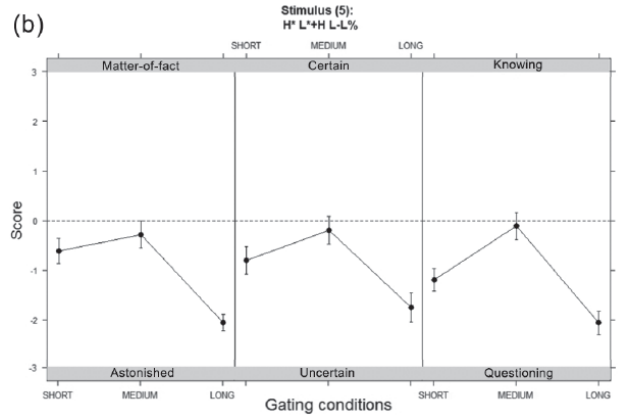


Figure 3. The three gating conditions SHORT, MEDIUM (cut after *Katherina* and after *Katherina sucht*) and LONG (uncut stimuli), exemplified by statement stimulus (3) and question stimulus (5).



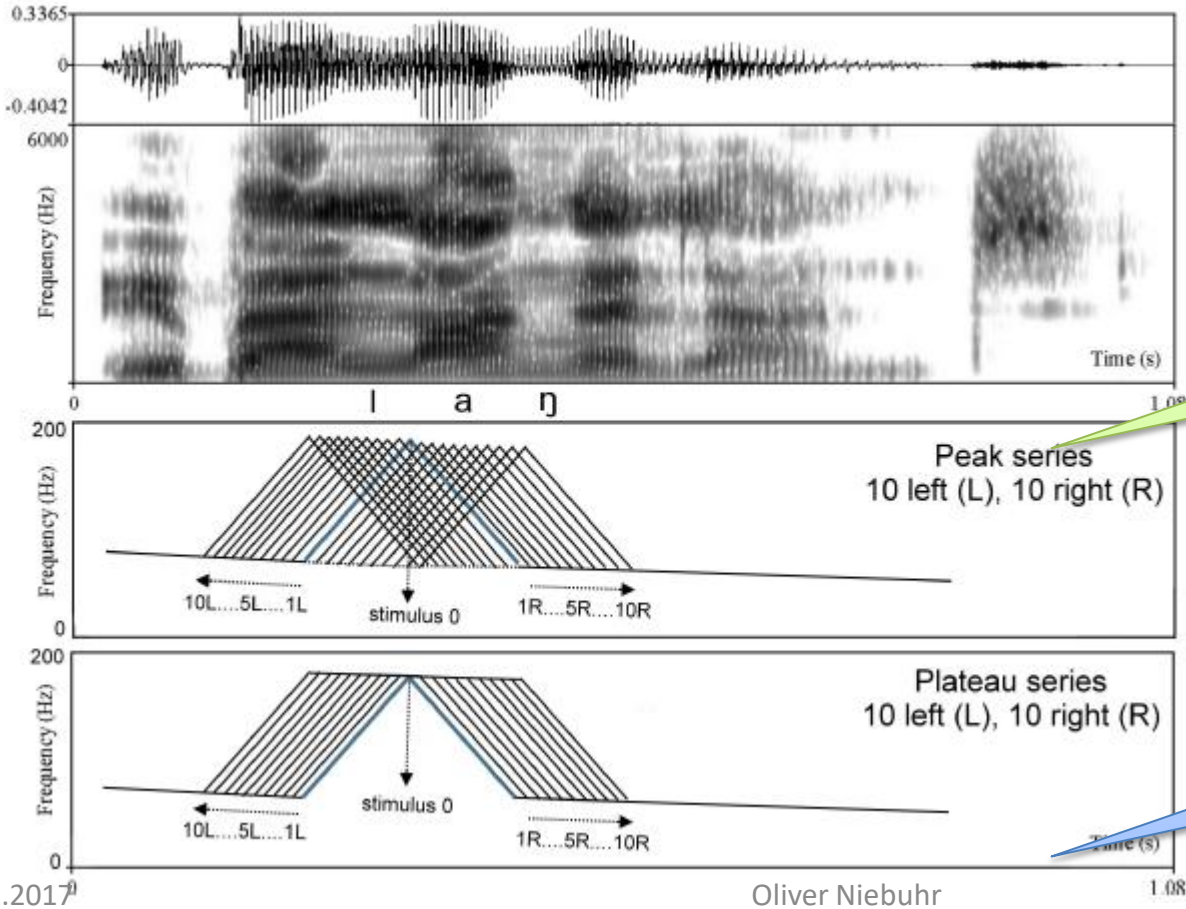


The 'short' stimulus condition is already sufficient to trigger rating differences related to sentence mode (astonished, uncertain, questioning)

Shaping Peak Maxima

Shaping Peak Maxima

- What about another type of shape difference? → **Peaks vs. plateaux?**
- Perception experiment by Niebuhr (2010)
- Test utterance "über Langeland" (via Langeland) → [ʔyːbɐ ˈlaŋəlant^{h}]

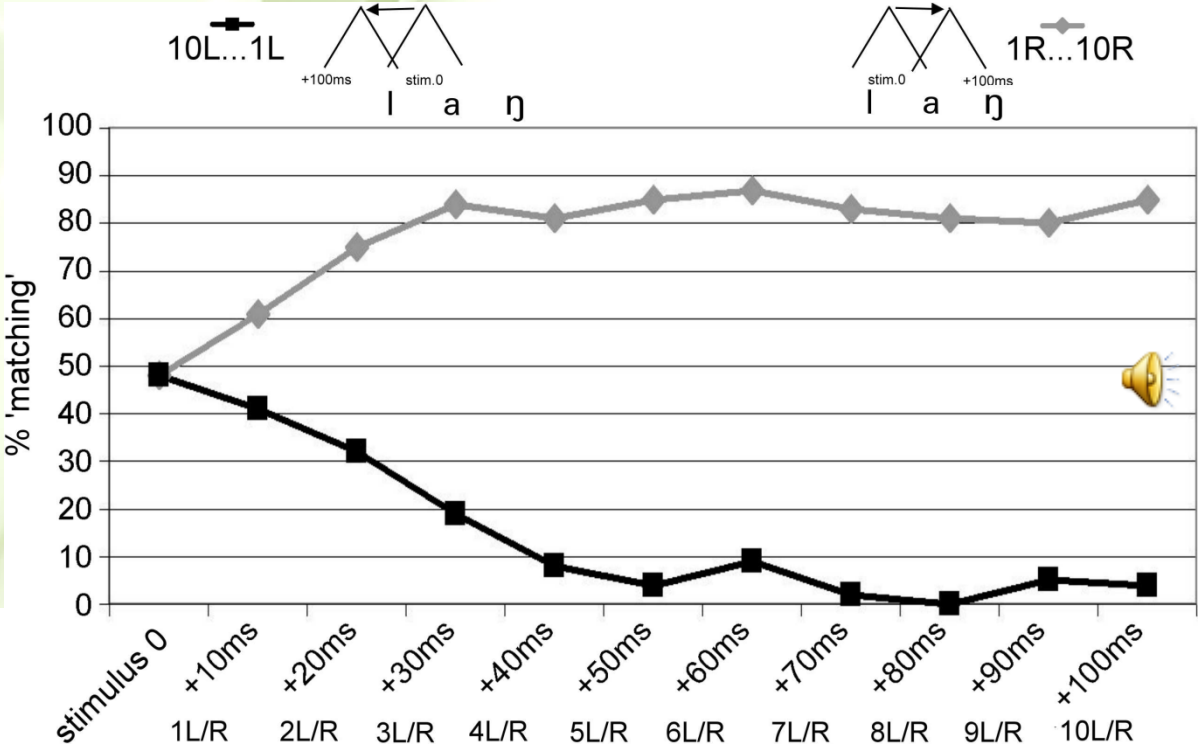


'Peak hypothesis', H1: The peak series will replicate the previous findings of Kohler (1987) and Niebuhr (2003) → F0 peaks whose rises are shifted more than 30 ms into the vowel will cause a (sudden) perceptual change from the early to the medial pitch accent

'Plateau hypothesis', H2: Plateaux strongly amplify the high-F0 content of the pitch accent → *both* the extension of the plateau into *and* away from the accented vowel in the plateau series will shift the pitch-accent identification towards the medial pitch accent

Shaping Peak Maxima

- What about another type of shape difference? → **Peaks vs. plateaux?**
- Perception experiment by Niebuhr (2010)
- Test utterance “über Langeland” (via Langeland) → [ʔyːbɐ ˈlaŋəlant^{h}]

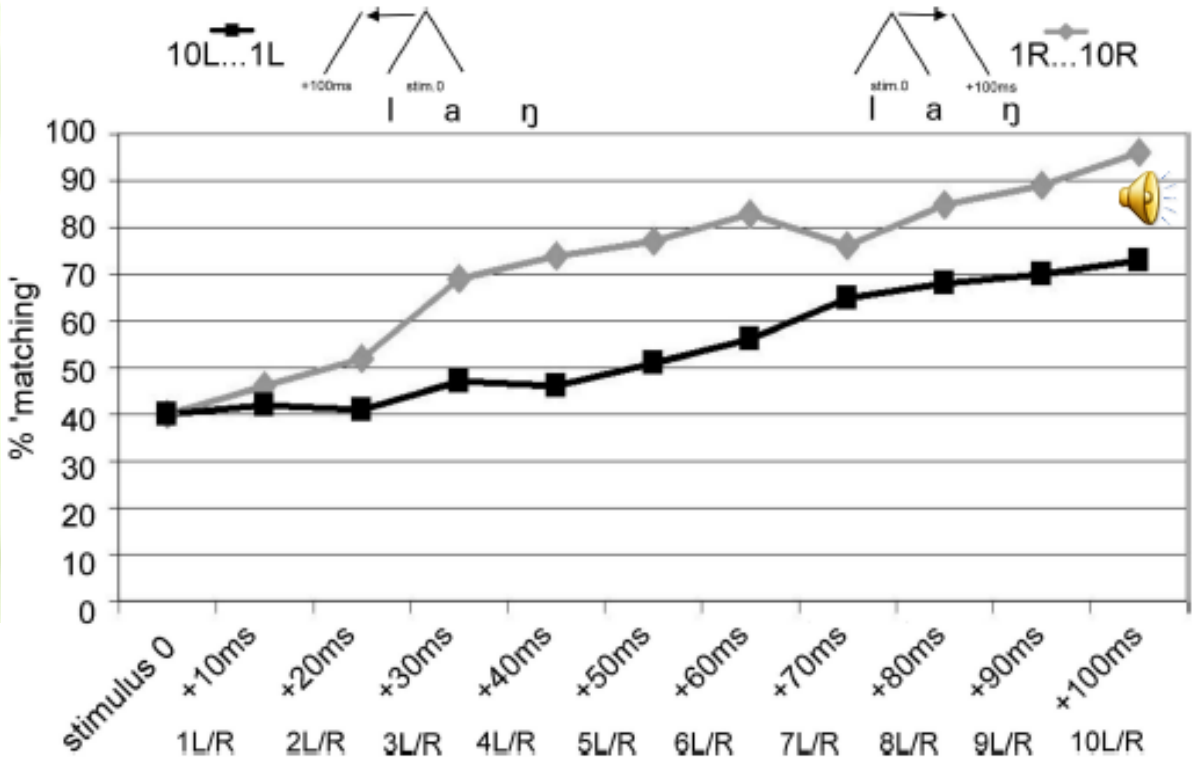


‘matching’ = medial peak identification

Results agree with H1
- Peak shift to the right = change from early to medial
- Peak shift to the left = early peak stays early

Shaping Peak Maxima

- What about another type of shape difference? → **Peaks vs. plateaux?**
- Perception experiment by Niebuhr (2010)
- Test utterance "über Langeland" (via Langeland) → [ʔyːbɐ ˈlaŋəlant^{h}]



'matching' = medial peak identification

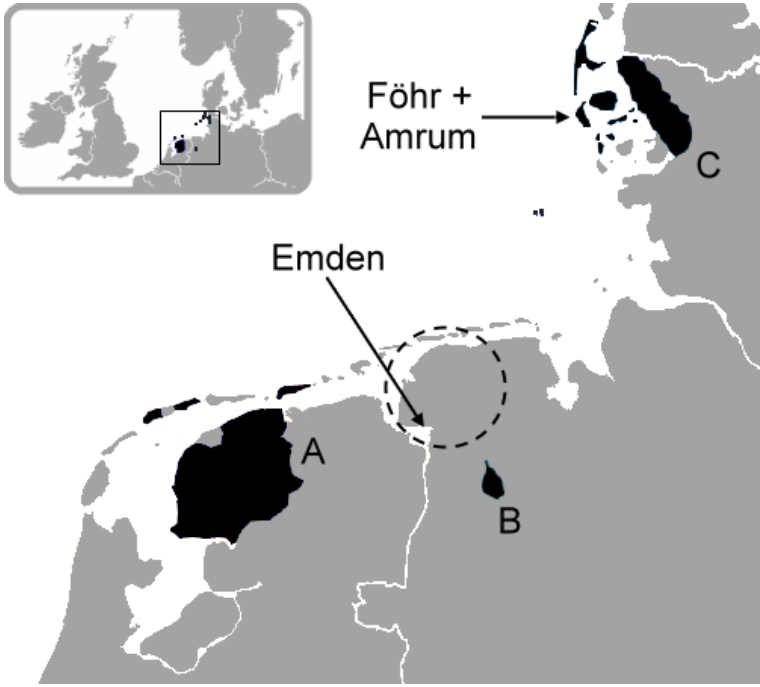
Results agree with H2
- Peak shift to the right = change from early to medial
- Peak shift to the left = change from early to medial, but less effective

Plateaux that lead away from the accented vowel can signal medial pitch accents!

Shaping Peak Maxima

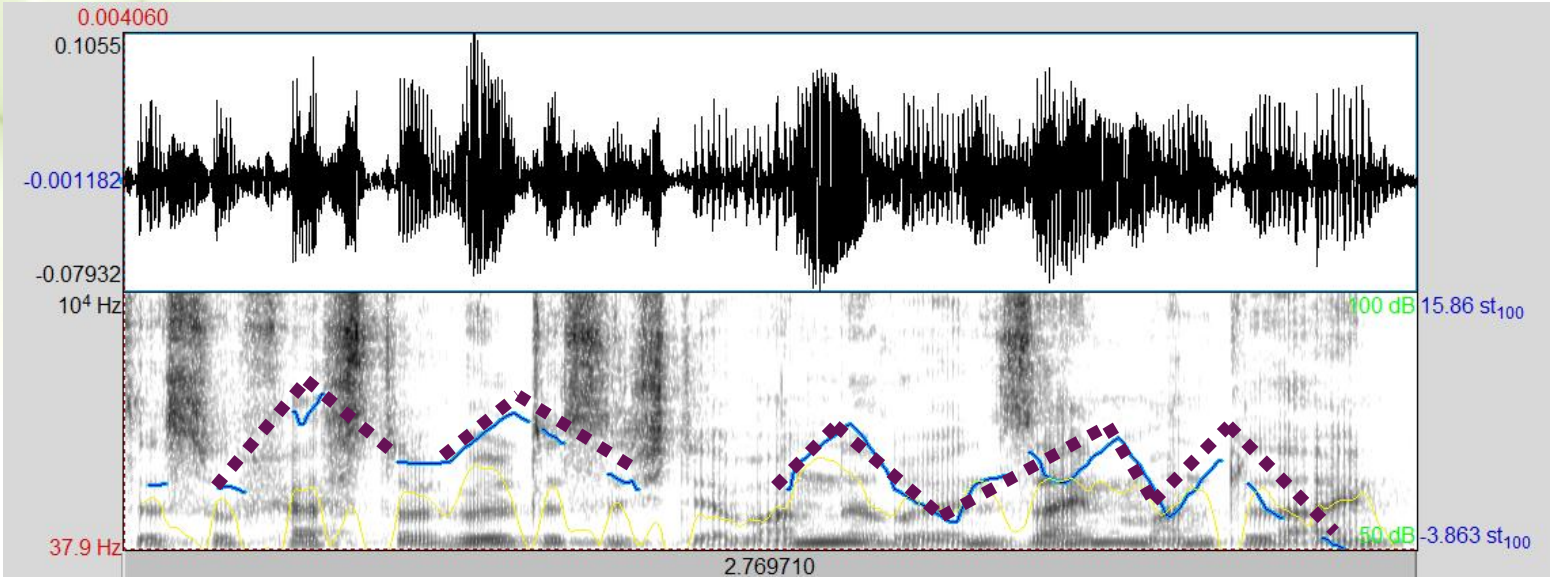


- The remaining speakers represent 3 main varieties:
  - West Frisian (A), about 400,000 speakers
  - East Frisian (B), about 2,000 speakers
  - North Frisian (C), about 8,000 speakers
- Phonetically and especially prosodically Frisian is an understudied language



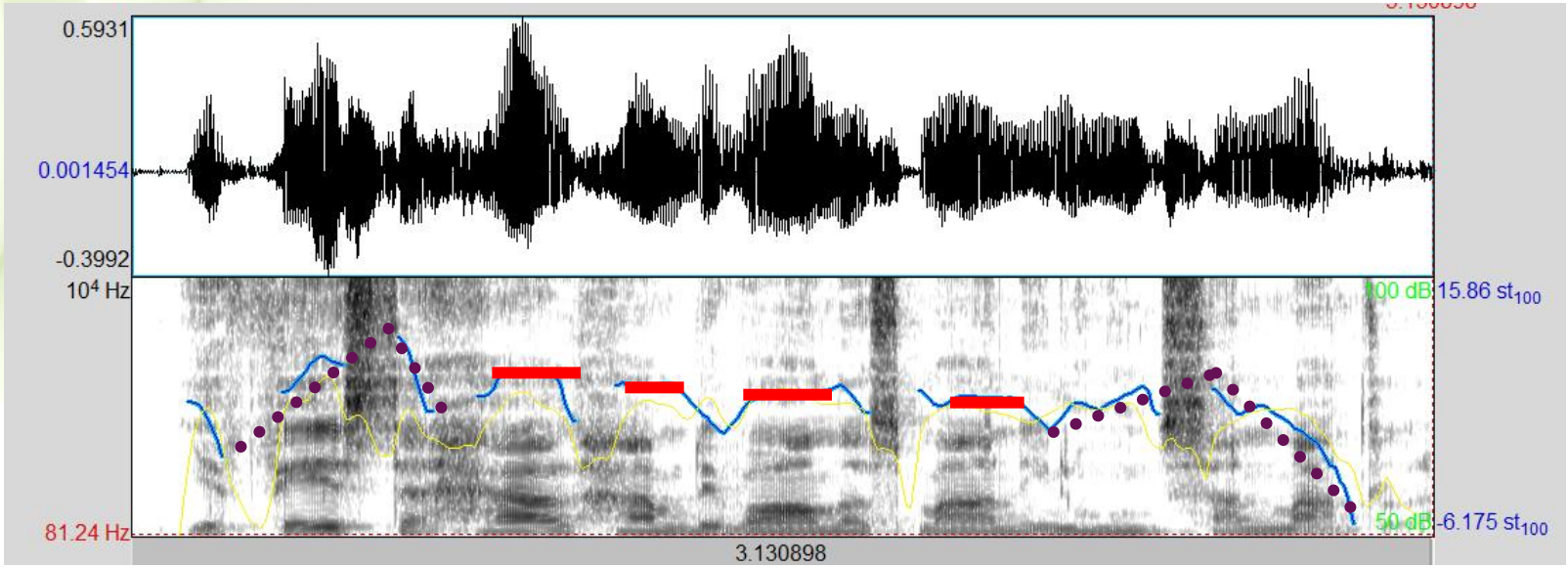
Shaping Peak Maxima

- Auditory observation: Some pitch accents sound particular in the ears of Northern Standard German listeners
- "[*but the more he blew] the more closely did the traveler fold his cloak around him*" (from 'The North Wind and the Sun')
  - Standard Northern German speaker
  - North Frisian speaker with Standard German Intonation (manip)
  - Original North Frisian intonation



Shaping Peak Maxima

- Plateau-shaped rather than pointed pitch accents → creating the impression of being “halted” or more “pounding” or “beating”
- Can we find more of these pitch accents ? (are they “just” idiosyncratic phenomenon ?) And if so, …
  - Where do they occur ?
  - How do they differ from the pointed pitch accents ?
  - Do they have a separate communicative function ?



Shaping Peak Maxima

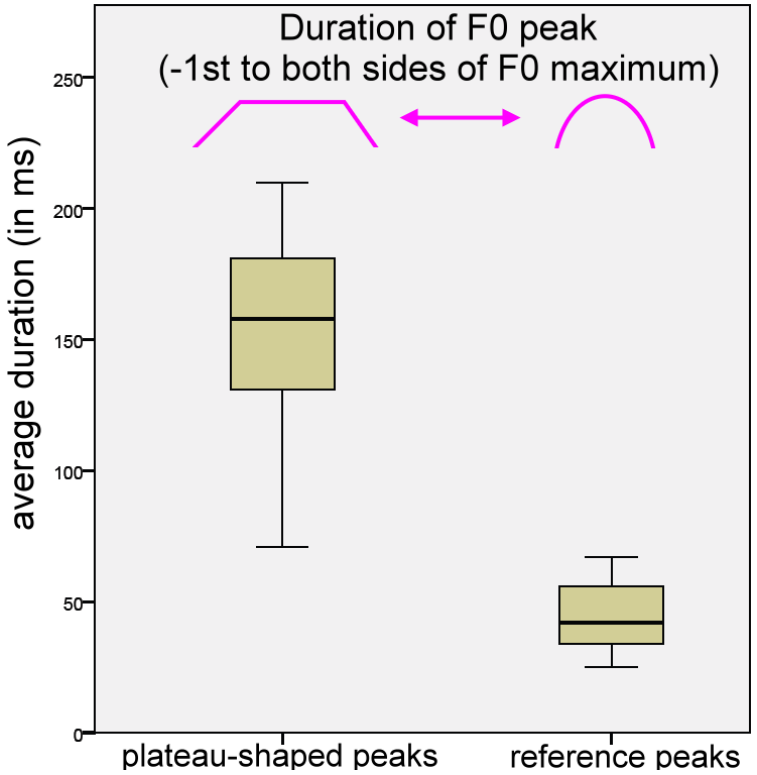
- Systematic **auditory search for plateau-shaped pitch accents** in existing corpora of read texts and spontaneous narratives, recorded during fieldwork by the second author
- Plus: complementary recording of short read dialogues, designed to elicit **broad, narrow, and contrastive focus intonations** on a set of segmentally controlled target words

Shaping Peak Maxima

- Auditory search yielded **numerous instances** across **all speakers**
- The sub-sample of **nuclear** pitch accents was acoustically analyzed
- …and related to a reference sample of "normally" sounding H* nuclear pitch accent counterparts, produced by the same speakers in the same recording session, on a similar syllable structure, and with a similar distance (in syllable numbers) to the preceding prenuclear accent and the following phrase boundary

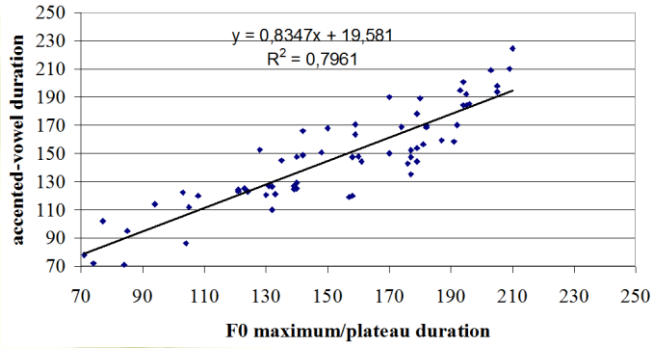
Shaping Peak Maxima

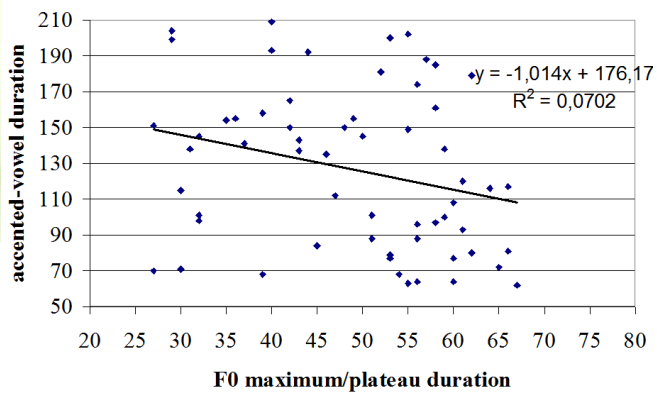
- There is a significant shape difference ($p<0.001$) between the two nuclear pitch-accent samples → **plateau-shaped peaks** (pitch maximum ≈ 150 ms) **vs. pointed peaks** (pitch maximum ≈ 50 ms)



Shaping Peak Maxima

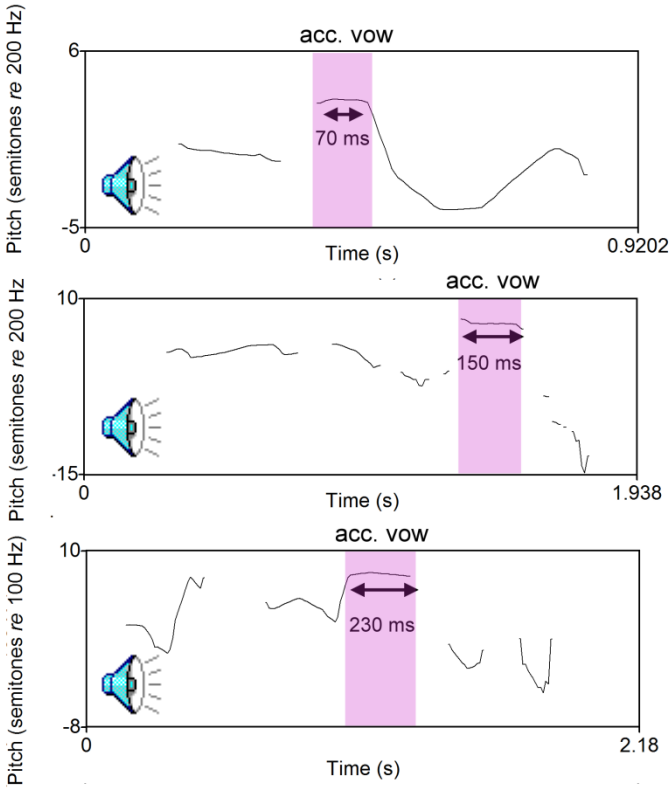
- The F0 peak maxima of pointed H* peaks were aligned towards the end of the accented vowel
- ↔ The **plateaux** of the plateau-shaped accents were **very precisely timed with the accented vowel**, independently of its quantity/duration





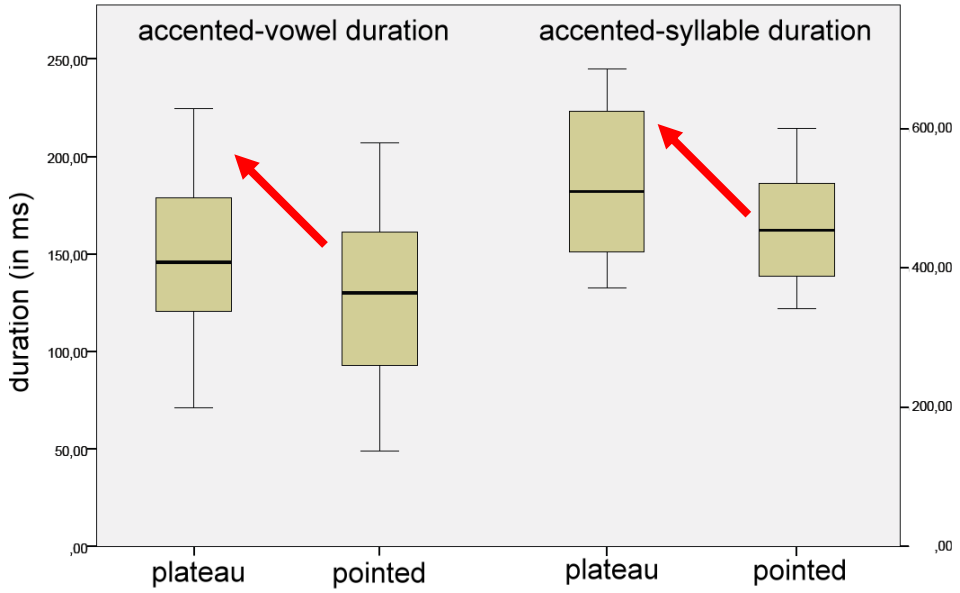
Sign. correlation plateau duration & acc.-vow. duration

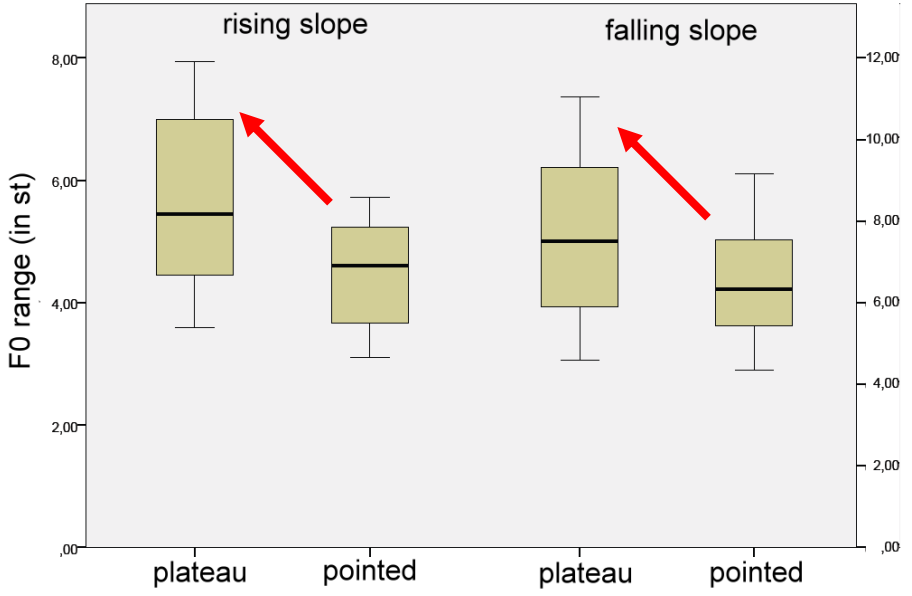
Absent for pointed peaks



Shaping Peak Maxima

- Plateau-shaped pitch accents co-occurred with **lengthening** of the accented vowel and syllable ($p<0.001$)
- The rising and falling **F0 movements** of plateau-shaped pitch accent peaks were slightly but significantly **longer** than those of pointed pitch accent peaks ($p=0.009$; $p<0.001$)





Shaping Peak Maxima

- With respect to communicative function,
- the plateau-shaped pitch accents seem to occur predominantly when speakers mark the accented piece of information as
  - Unexpected/astonishing,
  - Contrastive focus/topic,
  - Worth being remembered (…)

- "Claus Lembeck oober hed gud föörsurigt" (C.L. was well prepared, **though**')
- "Ik mei leewer ruadwin" (I myself prefer red wine **[= unlike you]**)
- "Hat feit jo goorniks liard" (She doesn't get **anything** straight)
- "Man aatj wiar domools uun Amerikoo" (My father was in **America** in those days **[= he was not at home]**)

Shaping Peak Maxima

- Plateau-shaped pitch accents
  - are used in contexts of surprise, contrast, lexical intensification
  - co-occur with vowel & syllable lengthening
- → Suggests that plateau-shaped accents represent an **emphatic type of pitch accent** in the North Frisian variety of Fering-Öömrang

- In Northern Standard German, emphasis is typically signalled by raising the F0 peak range → ≈ **5-6 st** (cf. Baumann et al. 2006; Niebuhr 2010)
- ↔ the plateau-shaped emphatic accents of Fering-Öömrang are only about ≈ **2 st** higher than their non-emphatic pointed counterparts

- → Suggests that speakers of Fering-Öömrang **use peak-maximum extension as a substitute** of peak-maximum raising!
  - Perceptually plausible strategy, as *"plateau-shaped accents sound higher than peaks"* (Knight 2008:223), cf. also Niebuhr (2011)

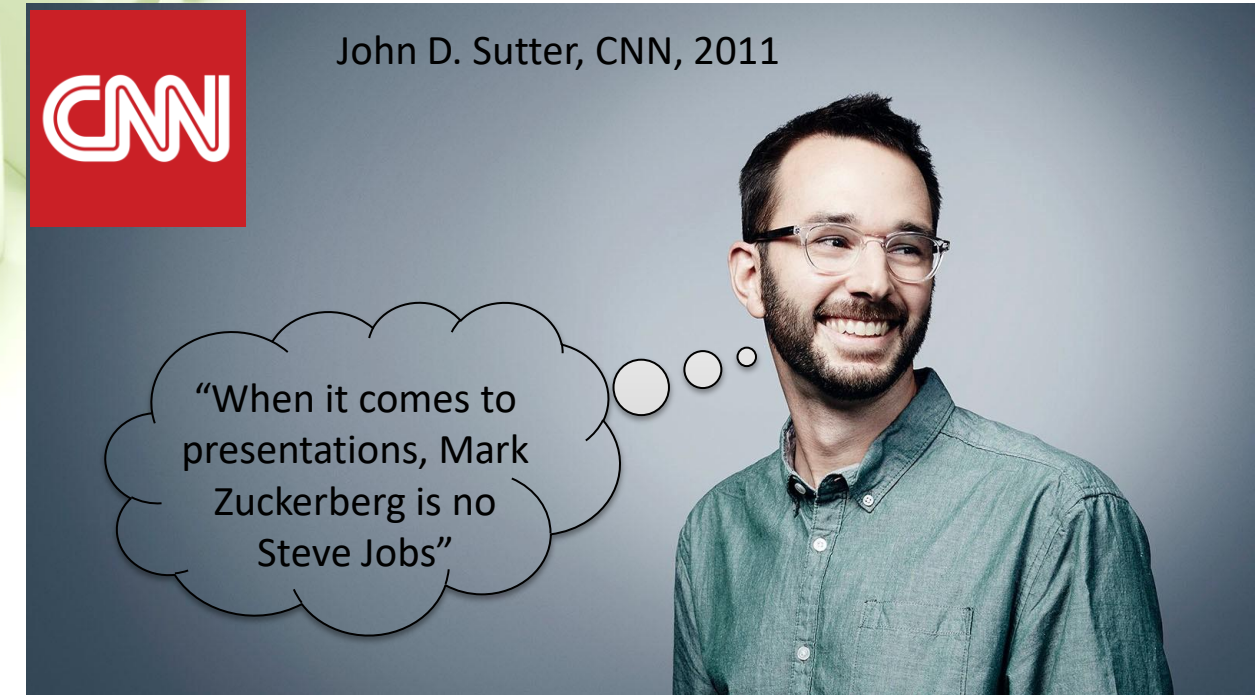
Shaping Peak Maxima

- Compared with speakers of other West Germanic languages like English and German, Fering-Öömrang speakers seem to be **“F0 flatteners”** rather than **“F0 lifters”**

Similar to “Shapers” and “Aligners”
→ Another continuum between
“Lifters” and “Flatteners”
w.r.t. changing peak height?

Shaping Peak Maxima

- "Lifters" and "Flatteners" could also be a within-language and **between-speaker variable…**
- … that is related to creating an **expressive/charismatic impression**.

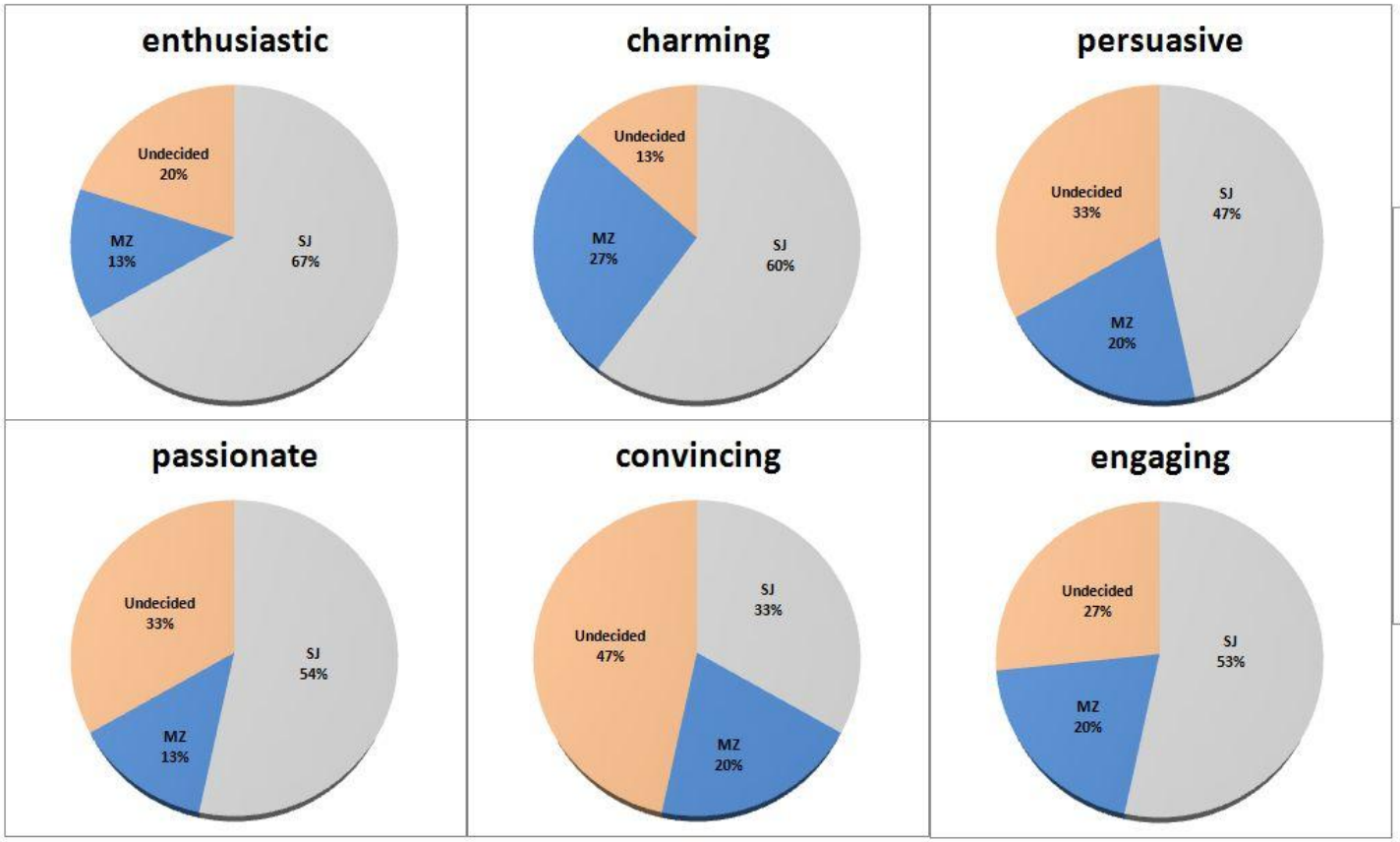


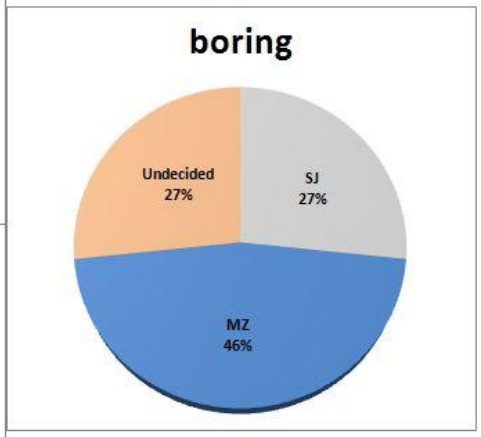
http://edition.cnn.com/2011/TECH/innovation/07/07/zuckerberg.facebook.presentation/

Shaping Peak Maxima

- Sutter's statement **confirmed by our own perception experiments**
- 18 sbs. judged the speech performance of two identical robots (Keepons)

  → **What are the acoustic reasons for these judgment differences?**





= Keepon with Steve Jobs' Voice sounds...

= Keepon with Zuckerberg's Voice sounds...

Shaping Peak Maxima

- **Investor-oriented speech:** Outlining the company's place in the market, focusing on its success, supported by figures and tables.

- **Customer-oriented speech:** Describing the product's new features and advantages over previous and competitor products.

Speech Material

- Keynote speeches at product launching events of Apple's two most important innovations of the last decades.
- iPhone 4 (2010), iPad 2 (2011).
- Data: 22 min. of speech from Hi-Fi YouTube videos; ≈ 12,000 sound segments; 5,000 syllables.
- 12 min. investor-oriented and 10 min. customer-oriented speech.

- Keynote speeches at Facebook's "F8" developers conferences held in 2014 and 2015.
- Data: 21 min. of speech extracted from Hi-Fi YouTube videos; ≈ 13,700 sound segments; 5,700 syllables.
- 10 min. investor-oriented and 11 min. customer-oriented speech.

F8 FACEBOOK DEVELOPER CONFERENCE

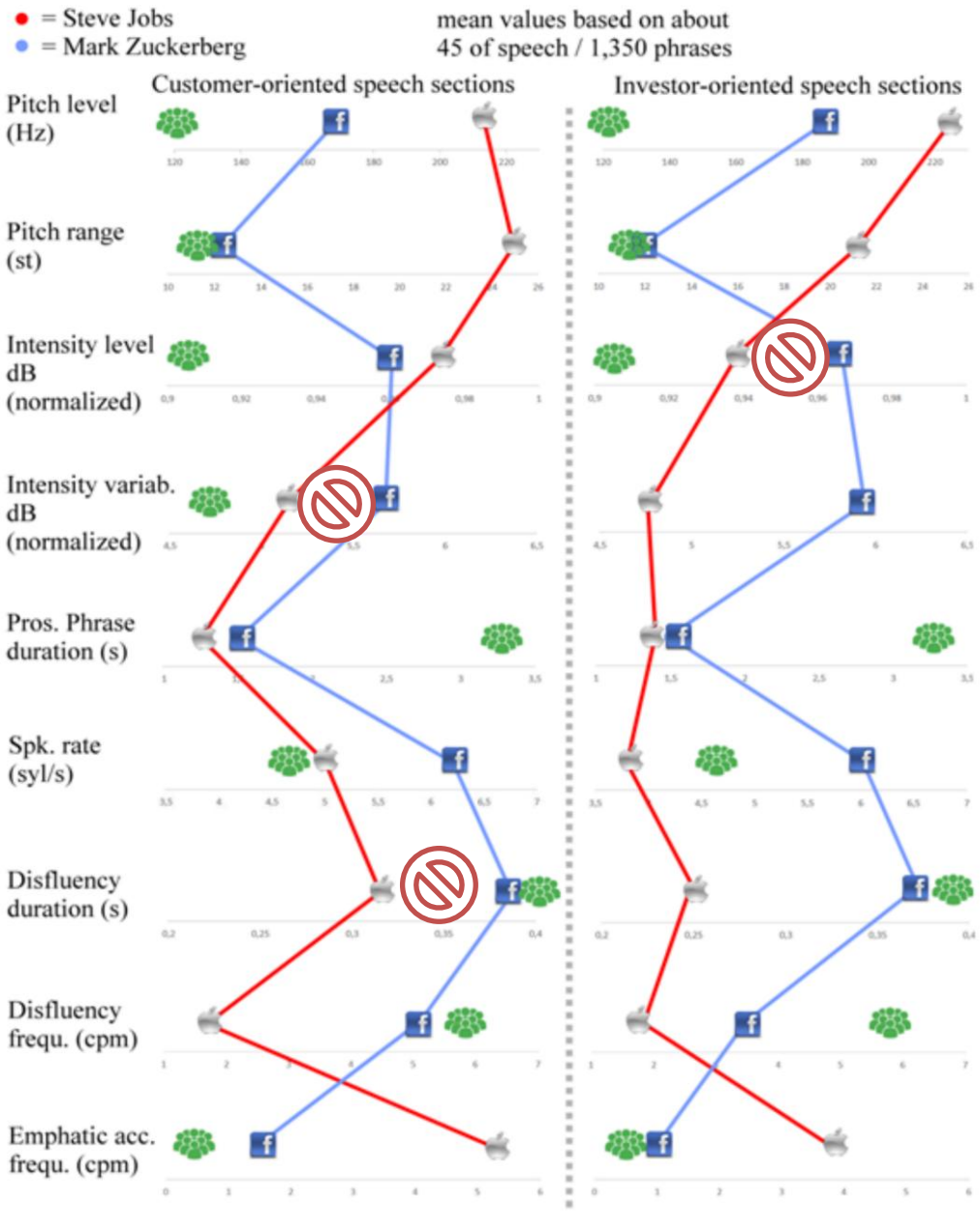
Shaping Peak Maxima

- Jobs and Zuckerberg differ in most acoustic-prosodic parameters

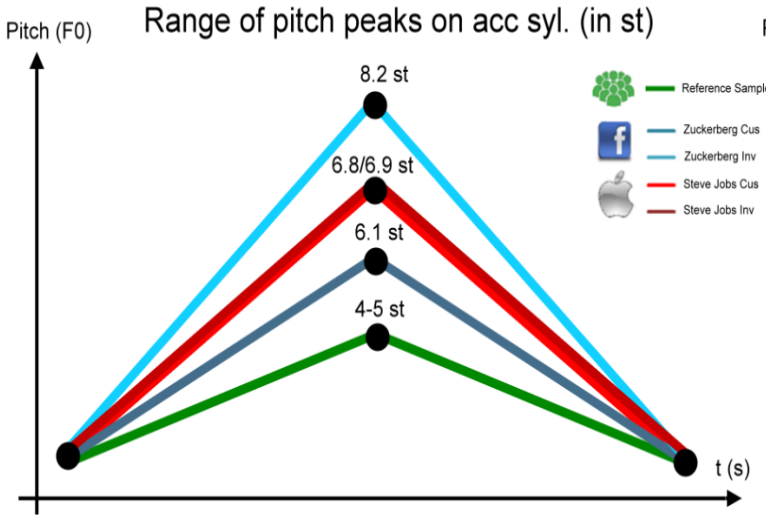
- Both speakers stand out against the ref. sample, but Steve Jobs more often

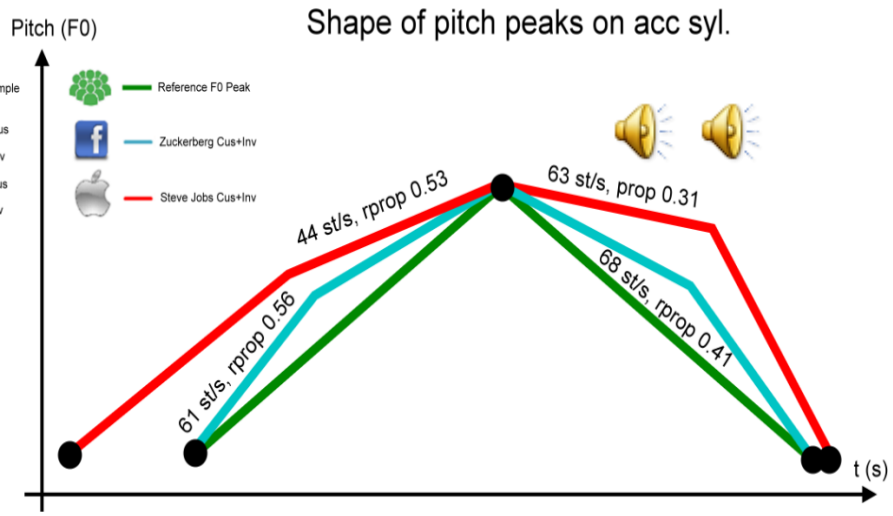
- **Compared to Zuckerberg**

- **Jobs is characterized by…**
- Higher pitch level, larger range
- Higher intensity level
- Less disfluciencies
- Shorter phrases
- More (200-300%) emphatic accents



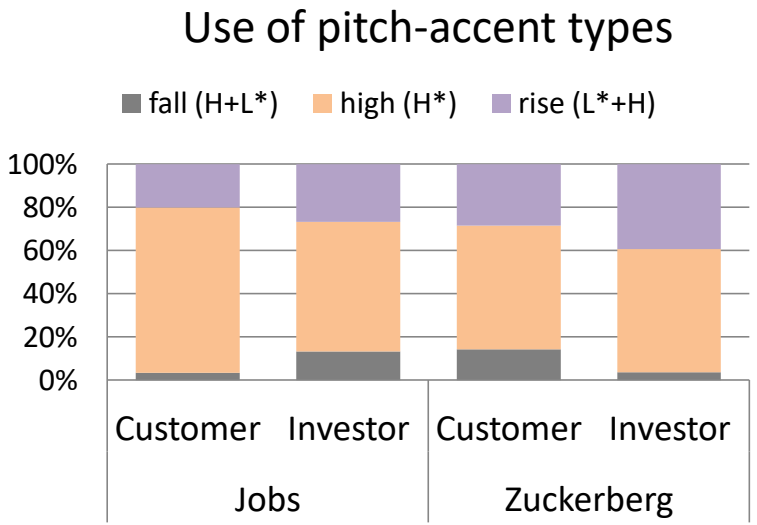
Shaping Peak Maxima



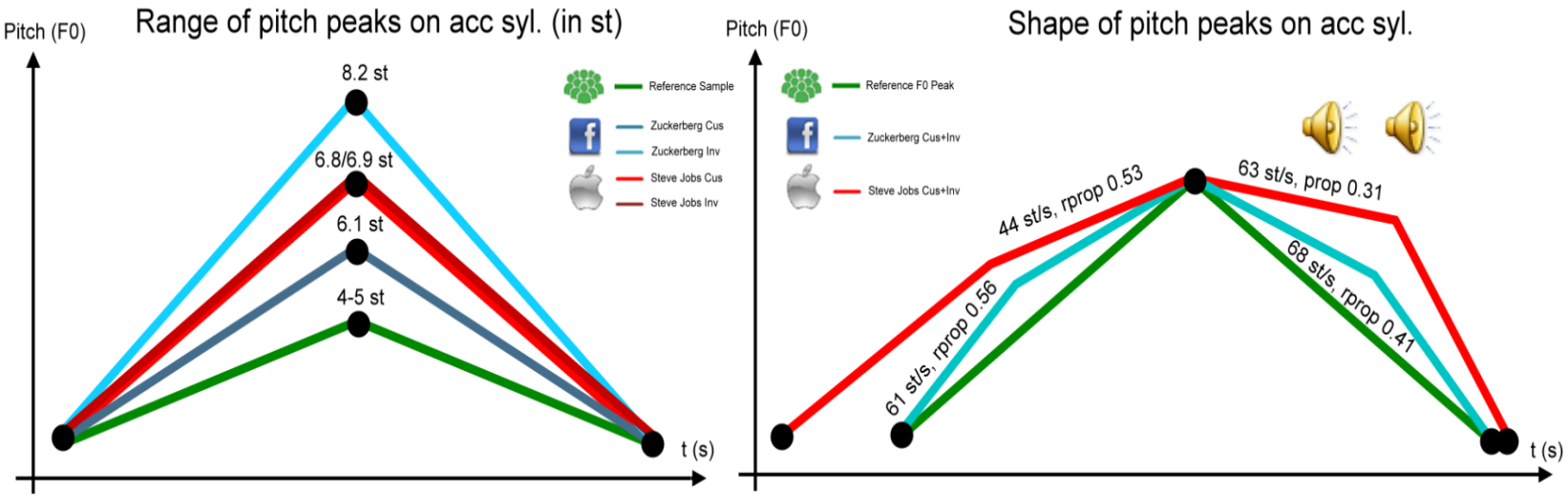


- **Compared to Zuckerberg**

- **Jobs is characterized by…**
- Higher pitch level, larger range
- Higher intensity level
- Less disfluciencies
- Shorter phrases
- More (200-300%) emphatic accents



Shaping Peak Maxima



- Is peak shape (= wider, shallower, more plateau-like maxima) a separate factor in sounding chrismatic?
- Or is this shape difference a concomitant effect of differences in pitch-accent type and emphatic-accent (type) frequency?

Shaping the underlying sound structure

Shaping the underlying sound structure

- Interactions of F0, duration and intensity in the production and perception of pitch accents are almost unknown

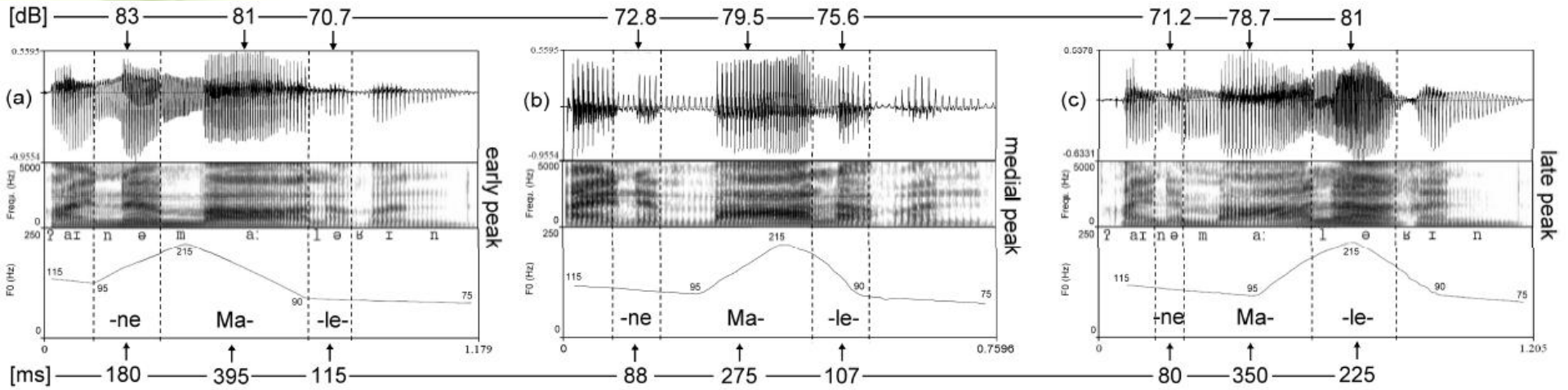
  - Analyzing production data, Kohler (1991:144) notes that there is a *"natural parallelism"* between the F0 and intensity curves of German pitch accents
  - With reference to his fieldwork experience, Himmelmann (2006:165) states: *"while it is true that tonal and intonational categories are primarily marked by changes in pitch, other auditory parameters such as length, loudness, and voice quality often also play a role in the marking of these categories"*
  - Peng et al. (2009) found for listeners with cochlear implants: *"intonation recognition was enhanced by cooperating F0 contour and intensity cues, but was adversely affected by these cues being conflicting"*

Shaping the underlying sound structure

„Sie‘s [mal Male]rin gewesen“

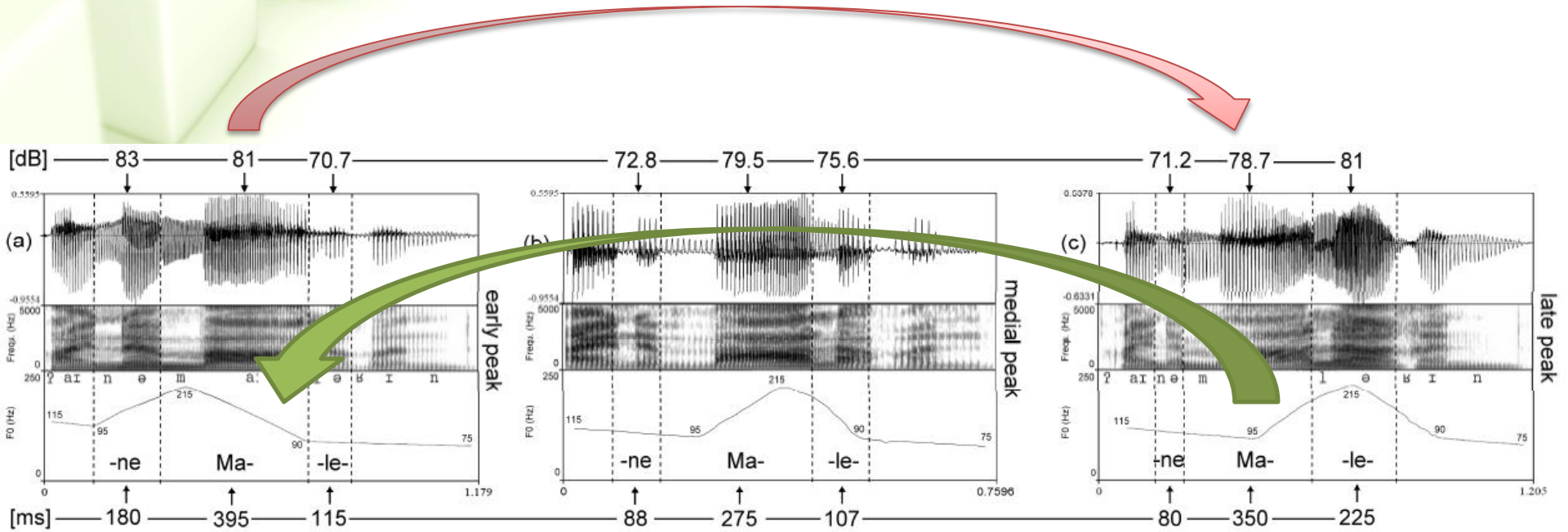
‘early’ peak:
- PreAcc syllable “long & loud”
- PostAcc syllable “short & soft”

‘late’ peak:
- PreAcc syllable “short & soft”
- PostAcc syllable “long & loud”



Shaping the underlying sound structure

- Perception experiment: “Eine Malerin” stimuli with **original and swapped** F0 and D&I patterns presented with multiple repetitions in overall randomized order and judged on semantic scales by 21 native German listeners

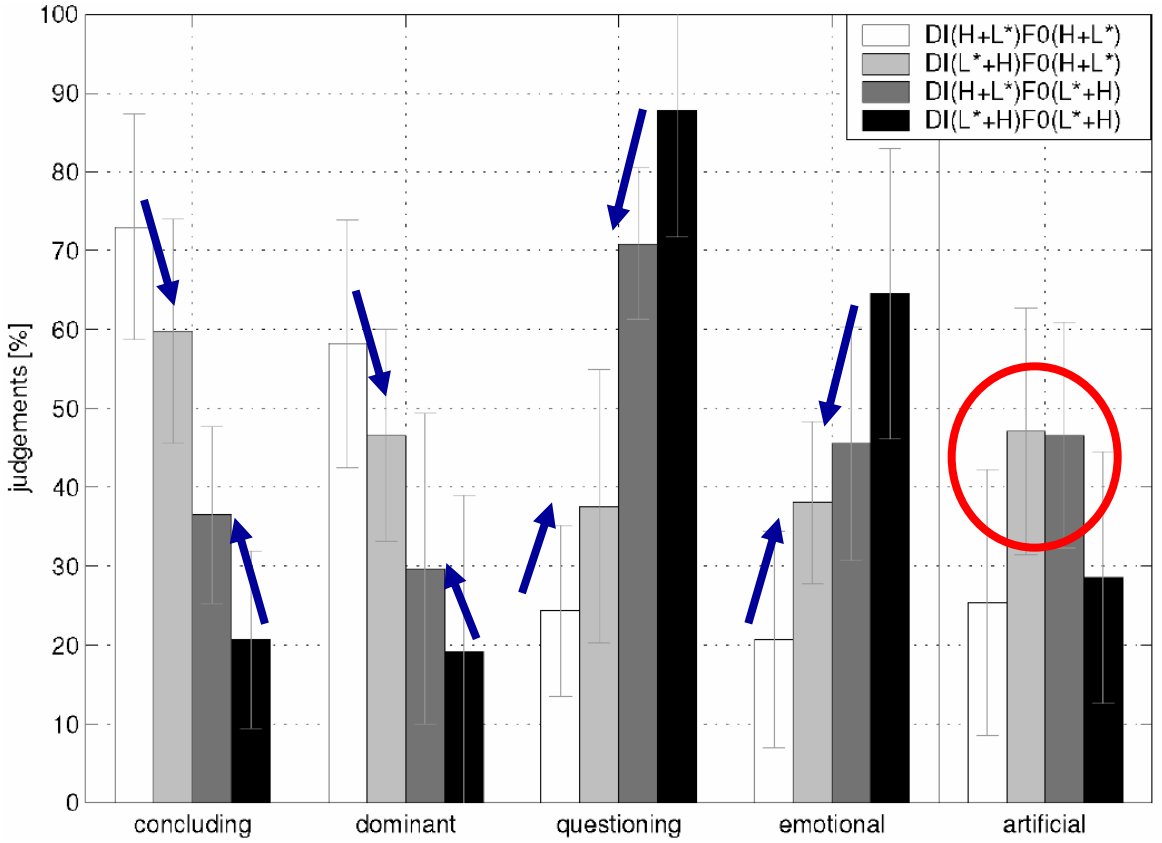


Shaping the underlying sound structure

- Perception experiment: "Eine Malerin" stimuli with **original and swapped** F0 and D&I patterns presented with multiple repetitions in overall randomized order and judged on semantic scales by 21 native German listeners

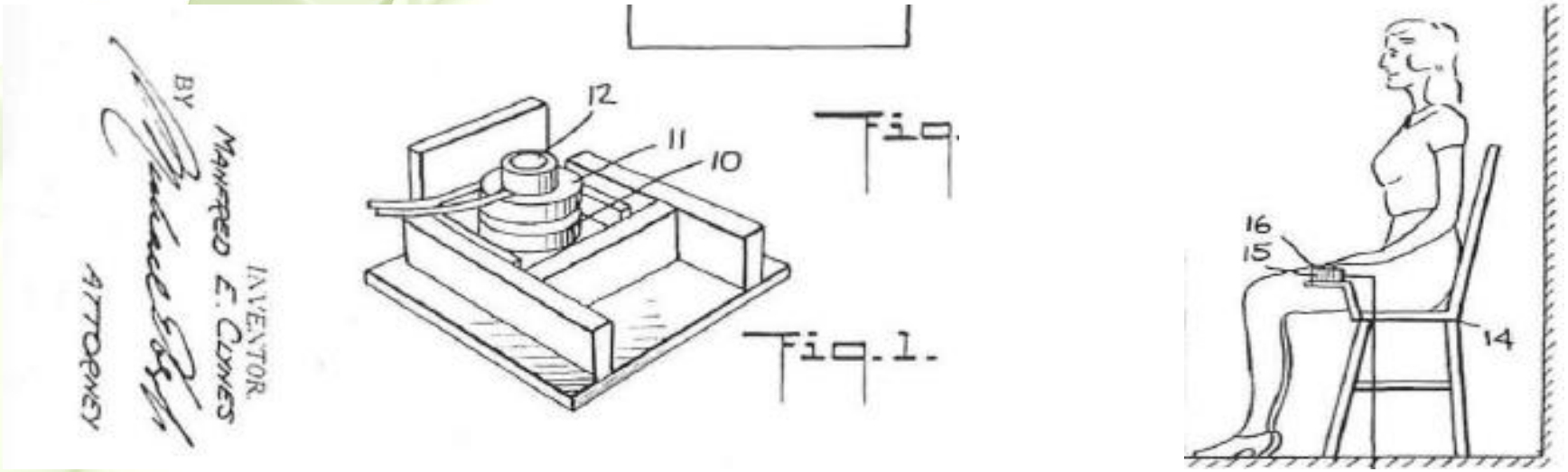
- ○ stimuli with swapped patterns sound more artificial
- → same F0 peak with different underlying D&I = changes on all sem. scales
- meaning components change towards the other pitch-accent category

(Niebuhr & Pfitzinger 2010)



Shaping the underlying sound structure

- Finger-tapping experiment of Boysen (2012) provided further supporting evidence for something like **"pitch-accent specific micro-rhythms"** across the triplet of pre-accented, accented, and post-accented syllable

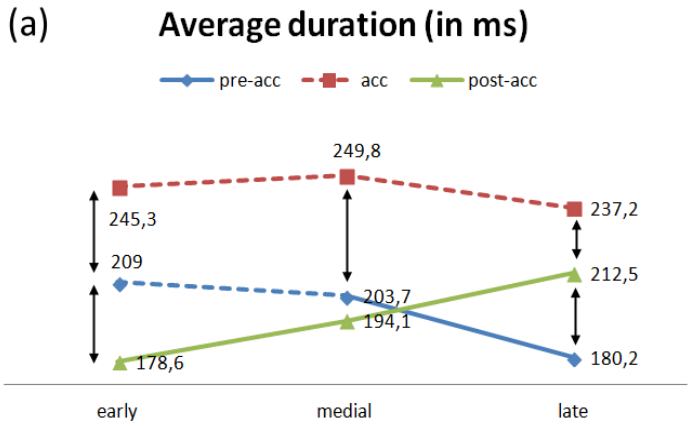


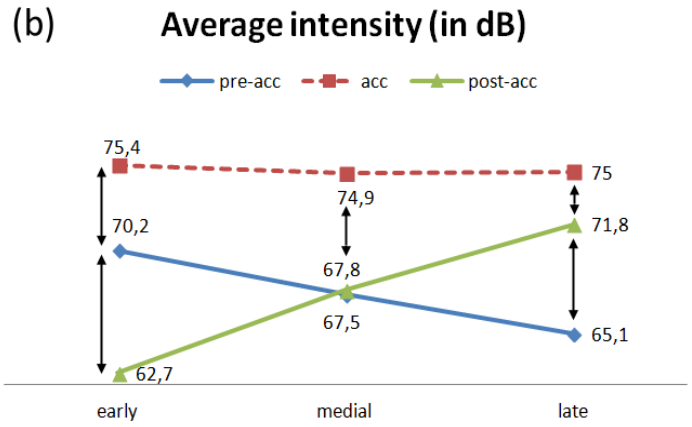
- 16 participants produced 20 sentences in a syllable-by-syllable yet fluent fashion with early, medial, or late accents on nuclear-accent target words
- They Pressed a button for each syllable. The device ("Sentograph", Clynes 1925) recorded pressure duration and pressure strength

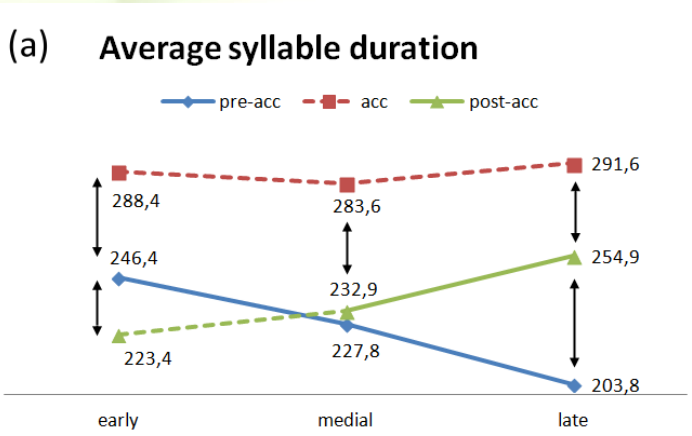
Shaping the underlying sound structure

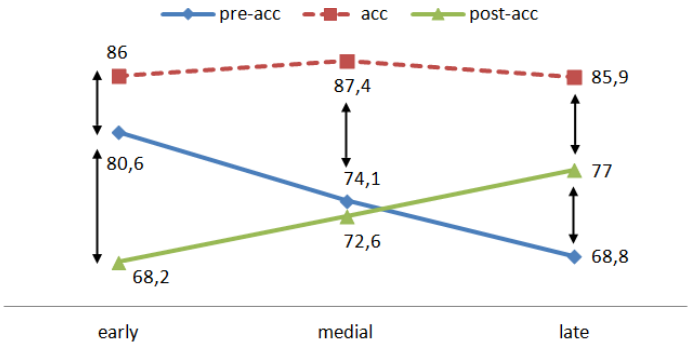
- Finger-tapping experiment of Boysen (2012) provided further supporting evidence for something like **"pitch-accent specific micro-rhythms"** across the triplet of pre-accented, accented, and post-accented syllable

Finger-tapping results



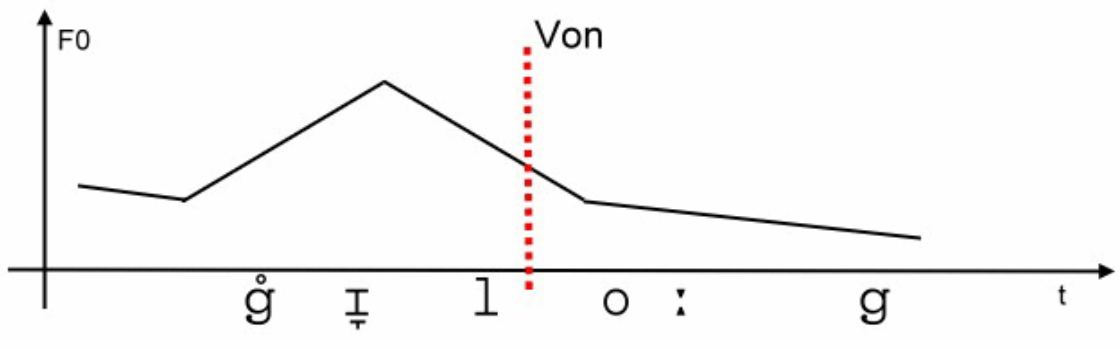


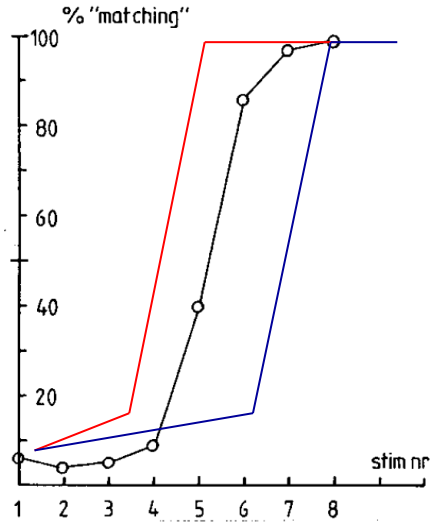




Results of acoustic analysis

Shaping the underlying sound structure





↔ Less well-known follow-up experiments showed that the location of the category boundary varies for different stimulus utterances (Kohler 1991)

- "Sie hat ja ge**lo**gen" = lateral + vowel
- "Sie ist ja ge**ritt**en" = fricative + vowel ⇒ later boundary
- "Sie hat ja ge**jo**delt" = approximant + vowel ⇒ earlier boundary

- Why? → Niebuhr (2006, 2007): maybe, it is not the segment boundary between C+V in terms of a **spectral change** (formant transitions) that matters, but the **intensity change into and out of the accented vowel**.

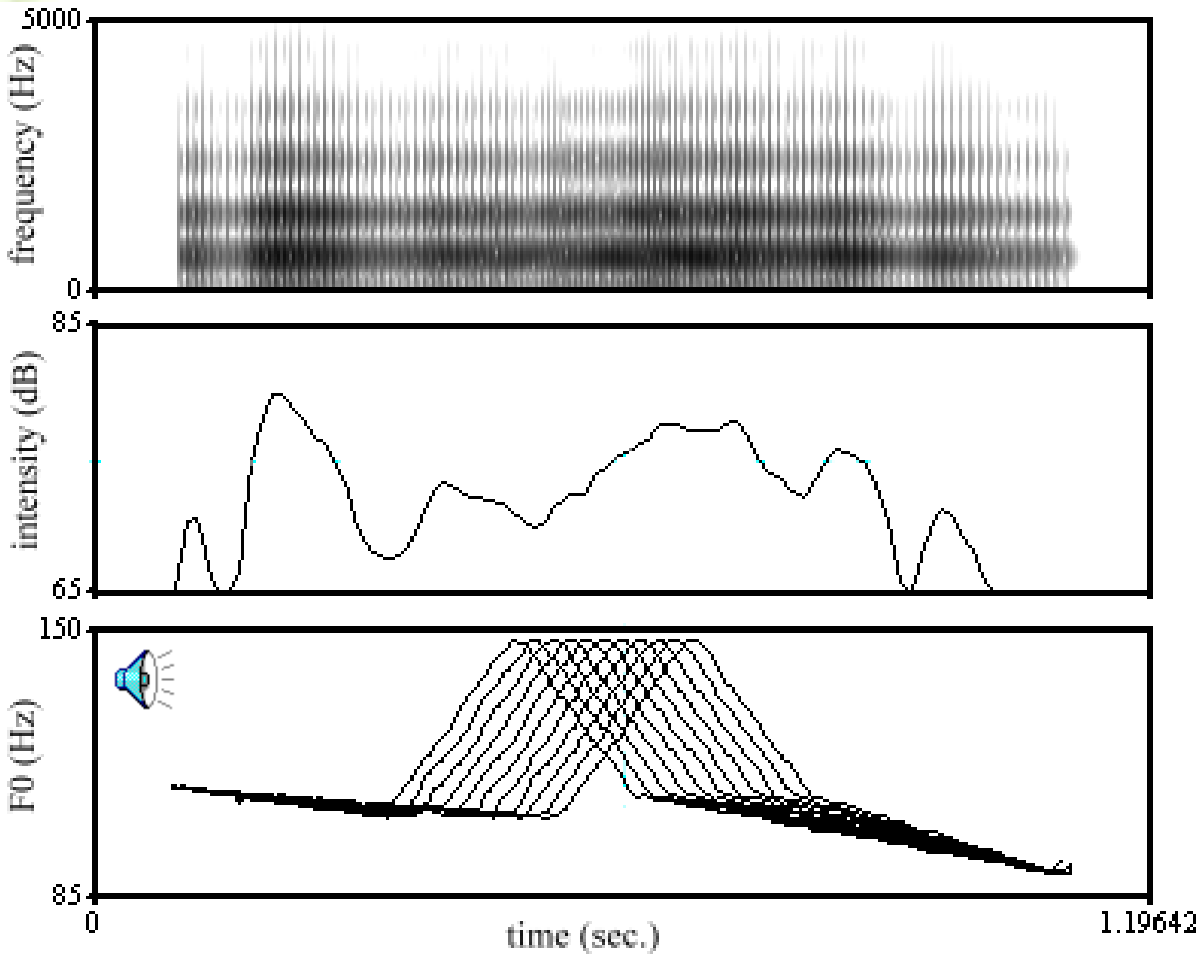
Shaping the underlying sound structure

- Two F0-peak shift series were resynthesized
  - one using the stimulus utterance “Sie war mal Malerin”
  - the other series kept exactly the F0 and intensity patterns of the “Malerin” series, but on a constant Schwa-like vowel quality (=“hum” in PRAAT)
  - → the two stimulus series (“Malerin” and “HUM”) differ only with regard to the presence / absence of the segmental string.

- Parallel perception experiments with separate groups of subjects
  - indirect identification task for the “Malerin” stimuli
  - AXB task for “hum” stimuli (A=first, B=last stim. in the series)

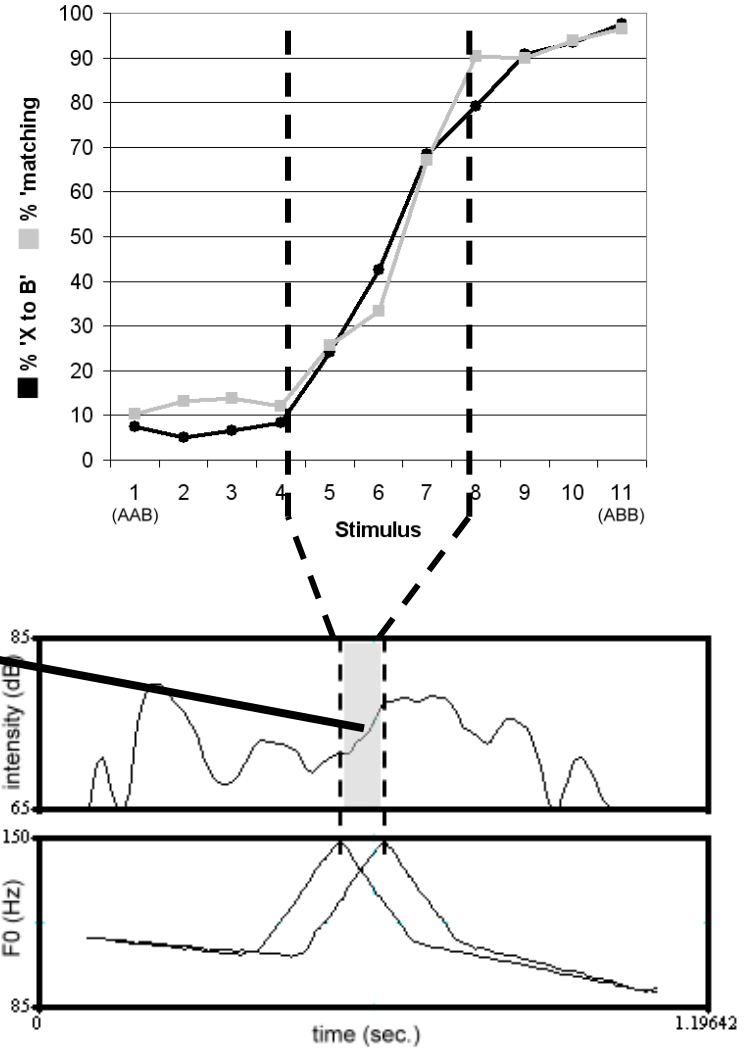
Shaping the underlying sound structure

"Malerin" series



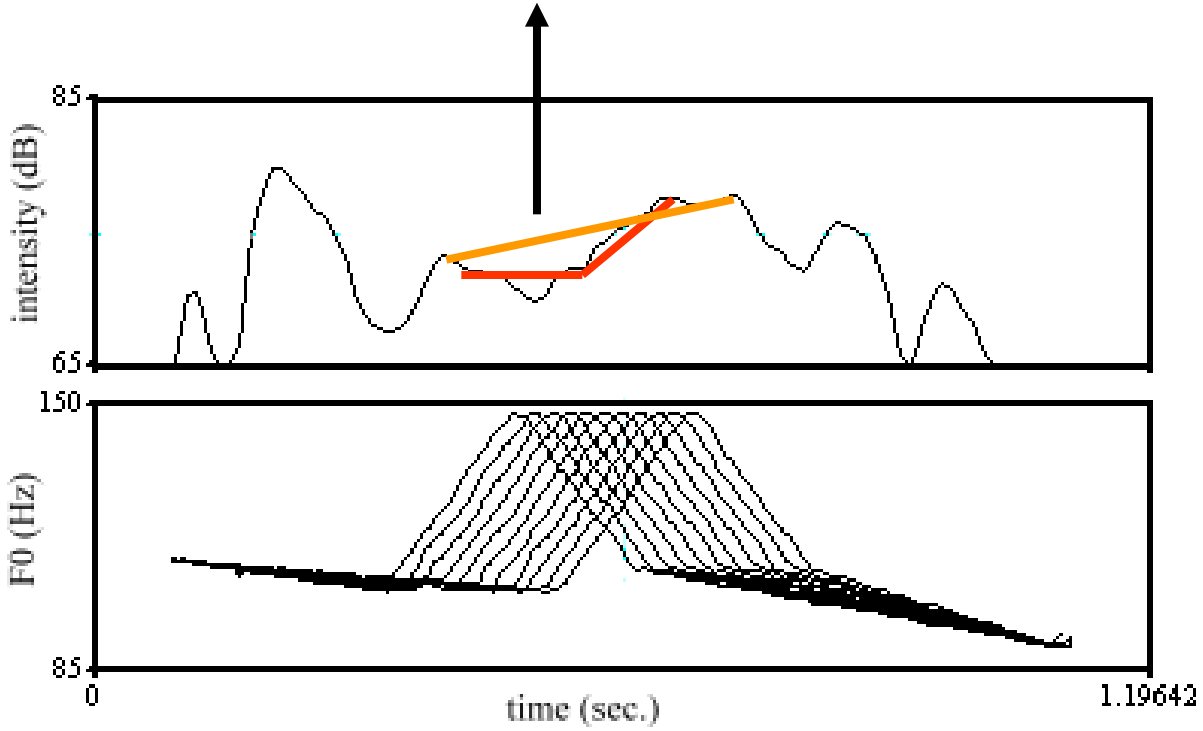
Shaping the underlying sound structure

What happens, if we manipulate the steepness of the intensity increase, which reflects the CV transition?

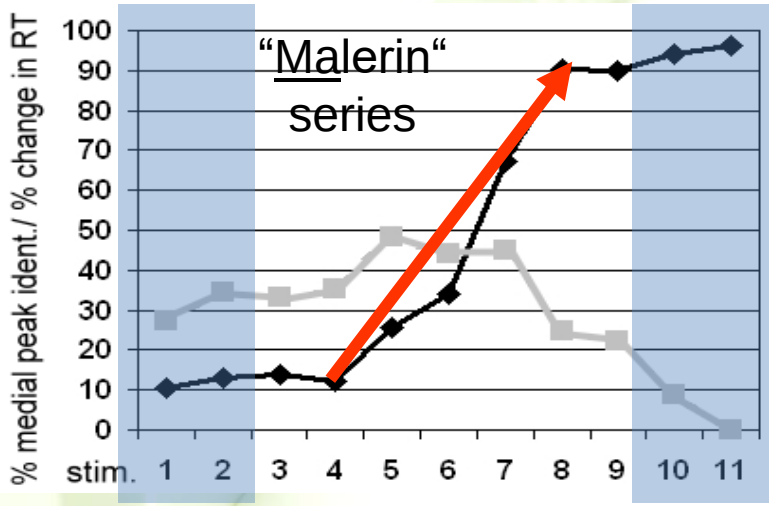


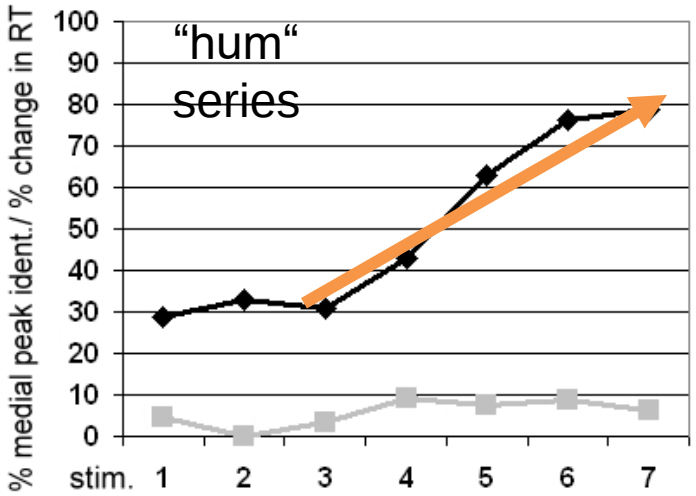
Shaping the underlying sound structure

Again two perception experiments, based on (a) “Malerin“ and (b) “hum“ stimulus series.



Shaping the underlying sound structure





- "Malerin" and "hum" series yield again similar (stat. identical) results
- The dynamics of the change from 'early' to 'medial' **decreases** with **decreasing dynamics** of the intensity change. The 50% categ. **boundary** is shifted as well.

- A comparable effect of the intensity dynamics on pitch-accent perception was found for a peak-shift series from **'medial' to 'late'**, based on a manipulation of the decreasing intensity at the **VC boundary**

Shaping the underlying sound structure

- Whether or not a pitch-accent contrast meets the criteria of **categorical perception**, depends – in NSG – on the underlying intensity dynamics! → The "categorical pitch perception" of Kohler (1987) can be turned into a gradual one.

- Syllable durations and intensity patterns **across and within** syllables contribute to the perception of pitch accents

- V**owel boundaries** (rather than the syllable boundaries) are crucial for distinguishing different pitch accent categories

- The across-syllable patterns can be thought of as **pitch accent specific micro-rhythms**.

Conclusion/Outlook

- Pitch accents can have accent-specific shapes in terms of the slopes of the rising and falling movements = within-category variation, can serve as an acoustic cue to accent-type identification
- Rising slopes – of pitch accents and boundary tones – can take convex or concave shapes = adds another meaning difference = between-category variation
- Peak shapes – pointed or plateau maxima – can represent both within-category variation (H* implementation differences and emphatic-accent characteristics) and between-category variation (meaningful scaling differences are created plateaux, charisma creation through plateau)
- Pitch accents create an imprint in the segmental string in terms of pitch-accent specific micro-rhythms = within-category variation in syllable duration and intensity

- **Intonation models have to find a way to take shape into account!**

Conclusion/Outlook

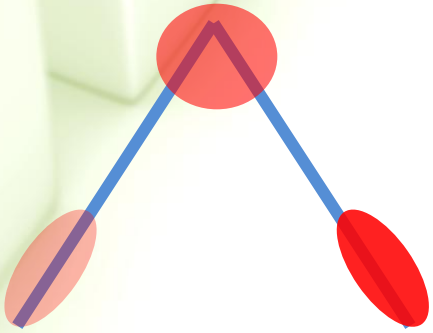
- **The Contrast Theory** (Niebuhr 2007, 2013)
- detailed theory about the perception and cognitive processing of intonation
- based on cross-modal analogies from the **psychology of visual perception**

- **Intonation models have to find a way to take shape into account!**

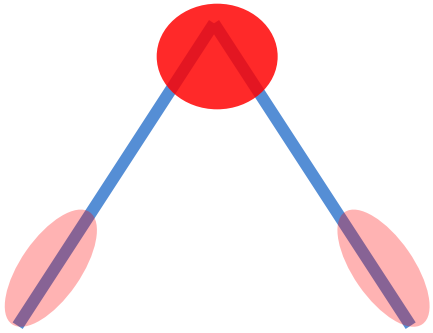
Conclusion/Outlook

- Basic assumption: An intonational unit like an accent contour is a **Gestalt of perceived pitch** that can consist of two basic building blocks
  - steady tones
  - tonal movements

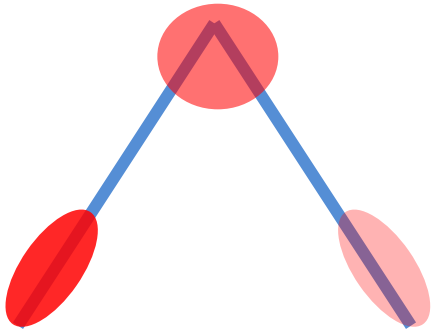
  } created by „optimal tonal perception“ (House 1990) → Pitch Gestalt
- Second Gestalt: **perceived prominence pattern**, created by sequence of steady tones and elbows of tonal movements



‘Early’



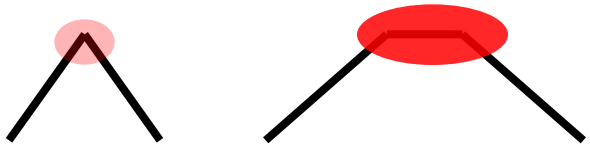
‘Medial’



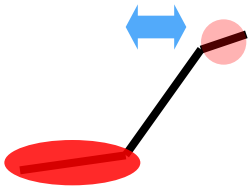
‘Late’

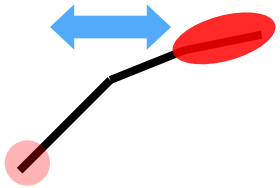
Conclusion/Outlook

- Pitch Gestalt based on F0 and aperiodic pitch elements
- Prominence Gestalt reflected in accent-specific micro-rhythm
- How can we change the Prominence Gestalt ?
  - change the **duration** of pitch elements



  - change the **distance** between pitch elements (smaller distances emphasize differences in the stimulus → enhances prominence differences , cf. visual perception)
  - changes in **F0 shape** – e.g., from convex to concave – are an effective means to vary the duration and proximity and hence of the prominence levels of pitch elements (proximity affects perceptual grouping and level of contrast enhancement)





Conclusion/Outlook

- Pitch Gestalt based on F0 and aperiodic pitch elements
- Prominence Gestalt reflected in accent-specific micro-rhythm
- How can we change the Prominence Gestalt ?
  - change the **intensity levels** that underlie the pitch elements

→ **alignment/synchronization** is the **most efficient way** to combine a Pitch Gestalt with the required Prominence Gestalt
→ it exploits intensity differences that are there anyway (in the form of sequences of sound segments)

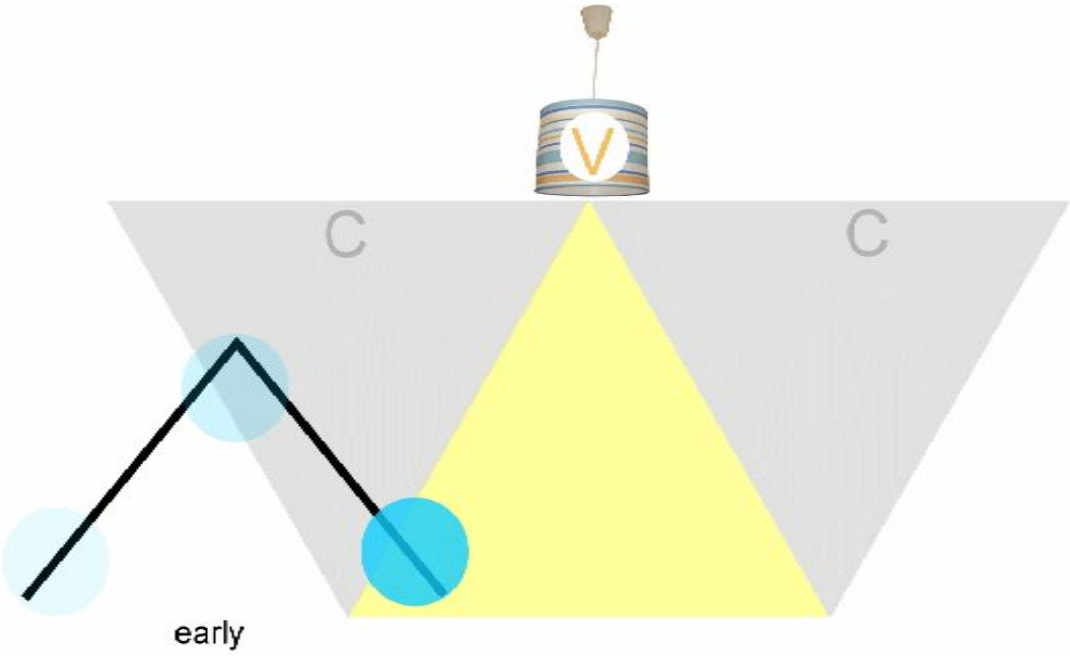
→ unlike in the Kiel Intonation Model, **alignment/synchronization in the Contrast Theory** has not the status of a direct phonological feature; but the vowel boundaries remain crucial (intensity) landmarks

→ Alignment changes due to changes in syllable structure can be accounted for in the Contrast Theory

Conclusion/Outlook

- Pitch Gestalt based on F0 and aperiodic pitch elements
- Prominence Gestalt reflected in accent-specific micro-rhythm
- How can we change the Prominence Gestalt ?
  - change the **intensity levels** that underlie the pitch elements

Different types of vowels and consonants affect the intensities and transitions between the „lighter“ and „darker“ areas of the syll.
→ requires **adjustments in alignment** (or shape).



Thank you for your attention