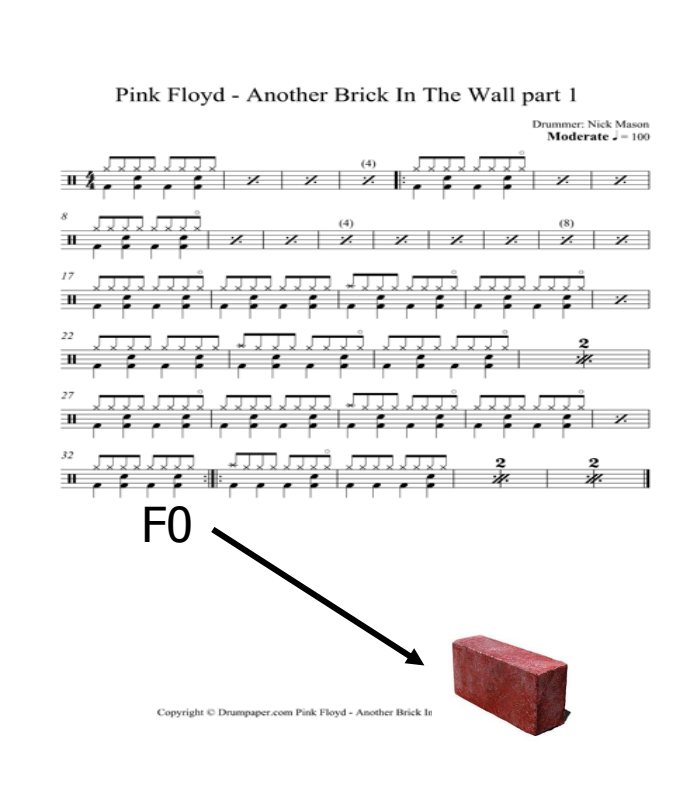
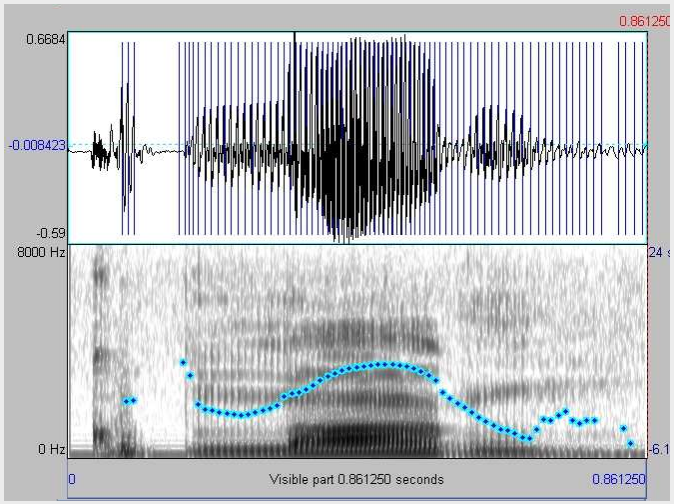
Another brick in the fall – On the contribution of sound segments to sentence intonation



Lingvistik
Språk- och litteraturcentrum
Lunds Universitet
12.05.2011

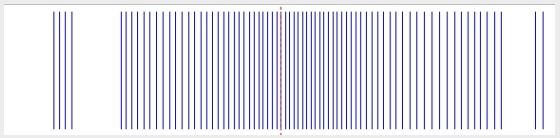
Segments and intonations

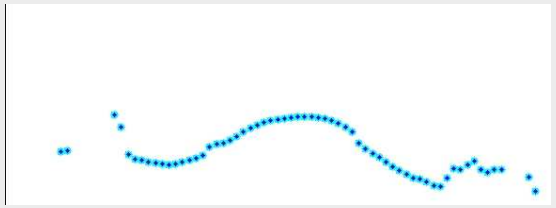


Intonation → **Words, Phones, Features**

- Well known fact: The F0 course contributes to the coding of speech sounds at the segmental level. For example,
  - F0 movements at vowel onsets or offsets can function as cue to the phonological fortis/lenis distinction (e.g., Kohler 1979)
  - F0 relative to vowel formants determines vowel quality (e.g., Traunmüller 1985)
  - Position of F0 elbows supports detection of word boundaries (e.g., D'Imperio 2000)
  - Local F0 dips can be phonetic representatives (i.e. „allophones") of /t/
  - (…)

Segments and intonations

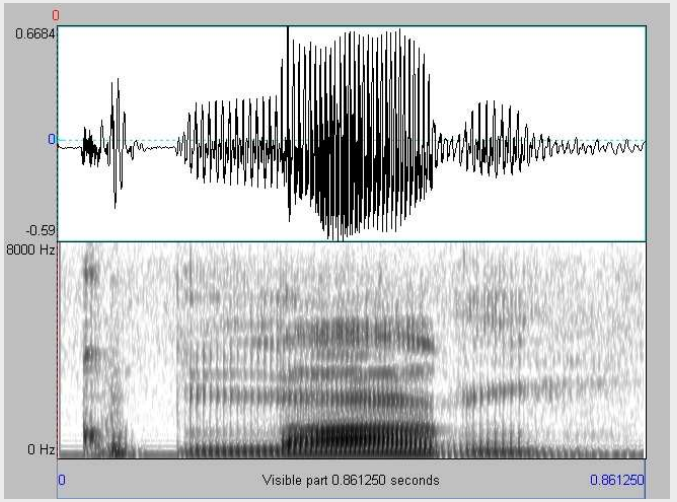




Intonation

!

The speech sounds at the segmental level also make important contributions to conveying intonational patterns and meanings



Words, Phones, Features

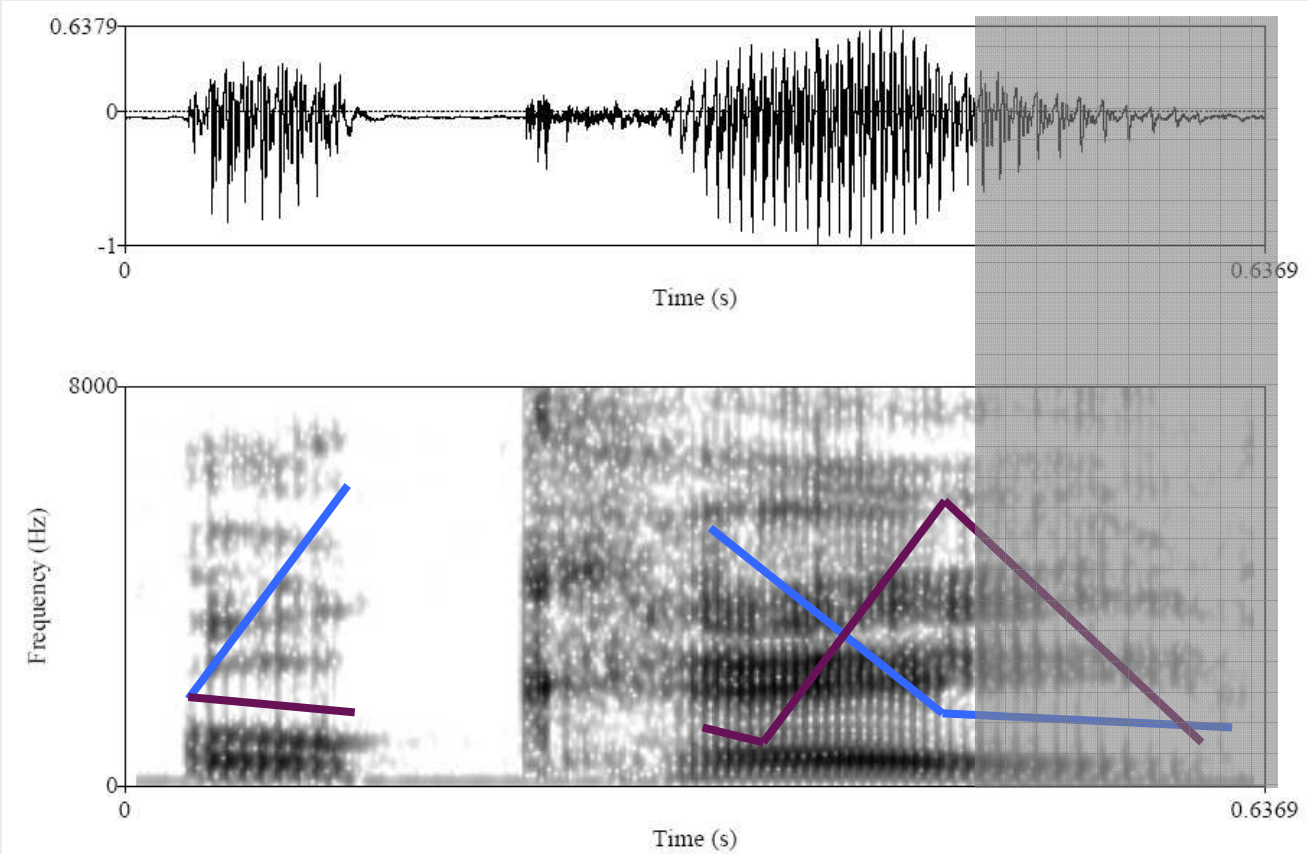
- However, from the perspective of intonation, speech sounds at the segmental level are typically regarded as „troublemakers“ for the elements of intonation.
  - The F0 contour shows micro-perturbations, i.e. the articulation causes variation in the F0 course
  - F0 relative to vowel formants determines vowel quality (e.g. Traunmüller 1985)
  - Nuclear F0 dips can be phonetic representatives (i.e. allophones) of /t/ (House 1990)
  - (...)

- Well known fact: The F0 course contributes to the coding of speech sounds at the segmental level. For example,
  - F0 movements at vowel onsets or offsets can function as cue to the phonological fortis/lenis distinction (e.g. Kohler 1979)
  - Position of F0 elbows supports detection of word boundaries (e.g. D'Imperio 2000)
  - Formant patterns influence pitch perception (intrinsic vowel pitch; Stoll 1984; Niebuhr 2004), formant transitions hinder decoding tonal movements (House 1990)
  - Voiceless segments interrupt and/or truncate the F0 course (Grabe 1998; Barnes et al. 2011 (...)

Segments and intonations

„Okay!"

Early peak / H+L*
= „I see/agree and I will do what you want, end of conversation"



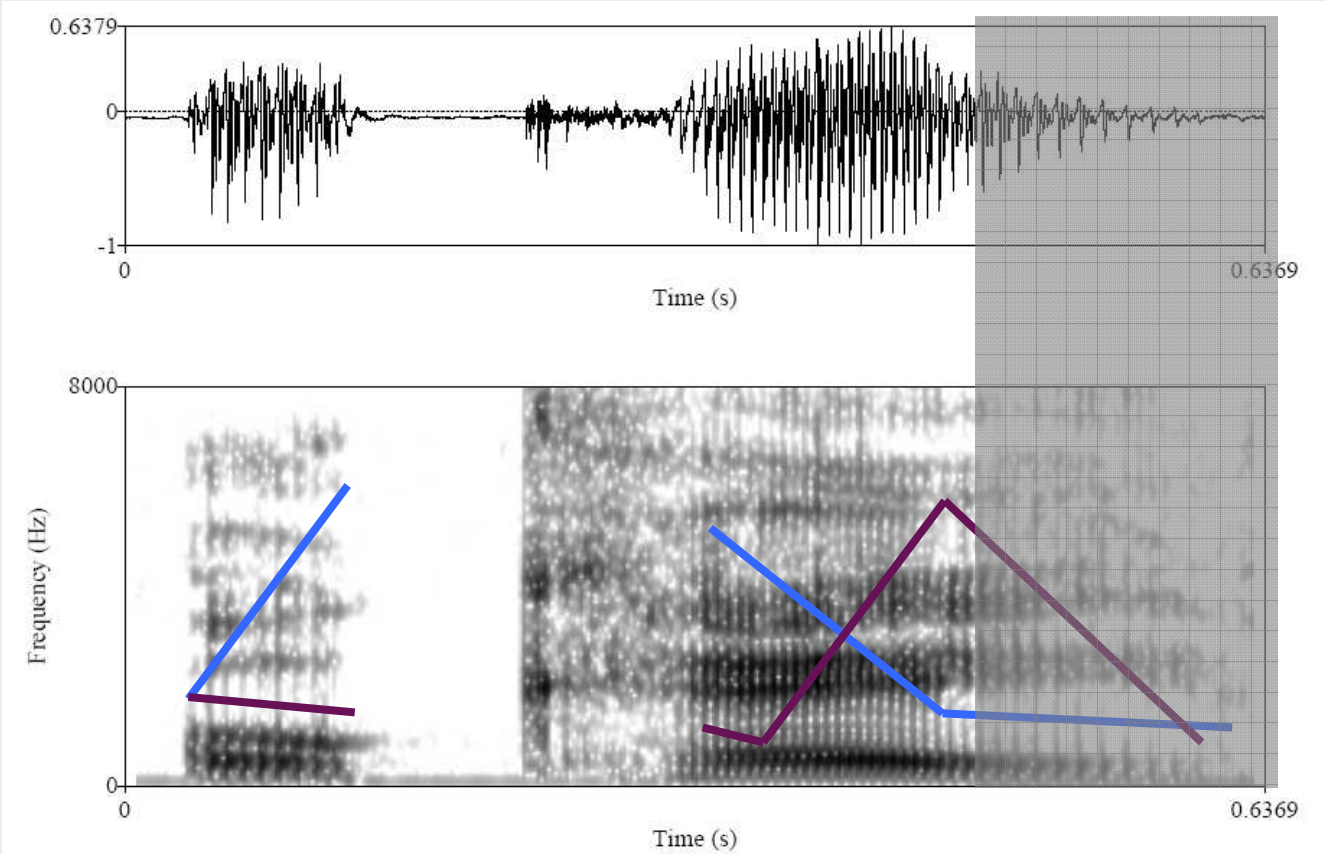
Late peak / H+L*
= „I do not see/agree but I will do what you want" (surprised or indignated)

z.B. **„Sie schickt"**
(She sends [it])

Segments and intonations

„Okay!"

- According to Grabe, falling F0 movements are truncated by voiceless consonants in phrase-final pos. in German (cf. „compression" vs. „truncation")

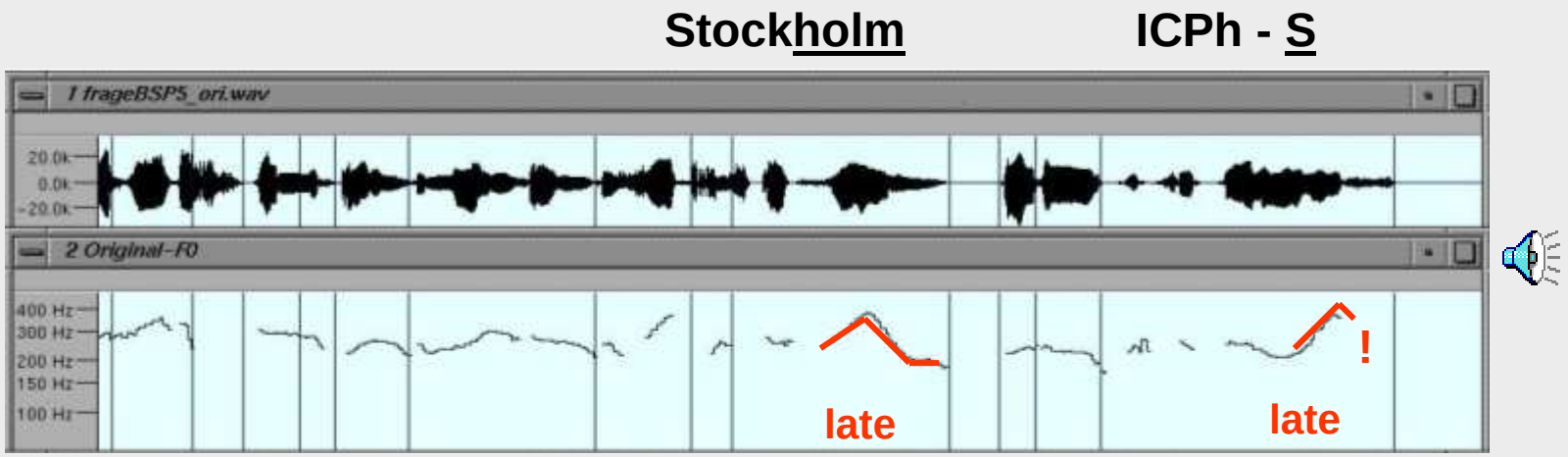


- So, does this mean that the perceive a truncated terminal fall as a rise or high-plateau intonation and hence as a question?
- → **„Okay??"**

z.B. **„Sie schickt"**
(She sends [it])

Segments and intonations

- No! .... For (at least) three reasons:
  - → In most cases a small residual of the F0 fall remains there.
  - → German questions and statements do not only differ in the direction of the utterance-final F0 movement, but also in speech rhythm, speaking rate, voice quality, F0 level and shape, etc.
  - → **The listener can add the truncated F0 piece, the fall is acoustically truncated, but perceptually complete!!!**



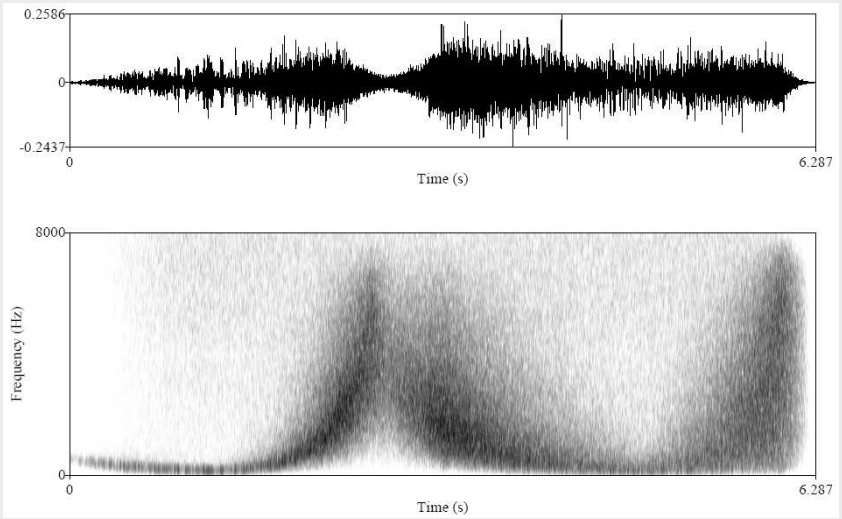
Segments and intonations

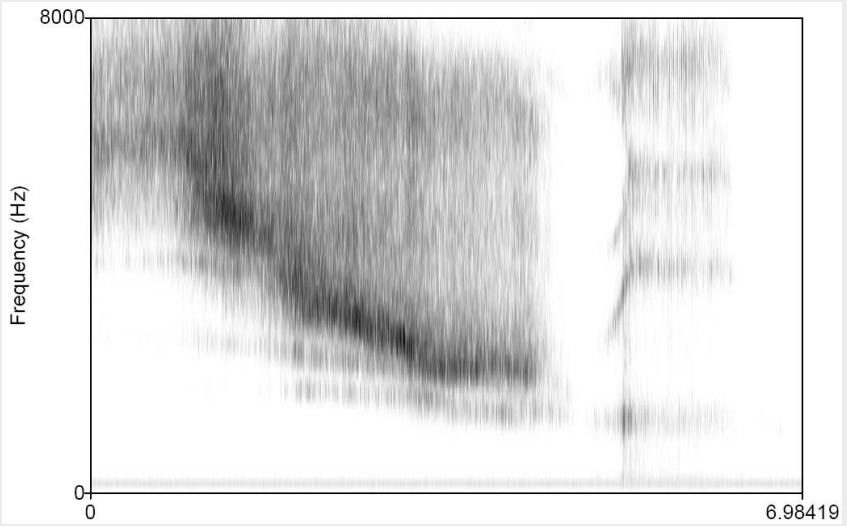
- Perception always means construction. We create a conscious percept on the basis of the incoming information and signal-external top-down information.
- Apart from the fact that we can use the context to restore the truncated fall, is there something in the speech signal that supports the completion of F0 falls?

- Pitch is not just a matter of F0. Noise, like in voiceless fricatives and aspirations, can also create a range of pitch impressions, depending on the spectral energy distribution
- → **„sibilant pitch", „spectral pitch"**

Segments and intonations

- Perception always means construction. We create a conscious percept on the basis of the incoming information and signal-external top-down information.
- Apart from the fact that we can use the context to restore the truncated fall, is there something in the speech signal that supports the completion of F0 falls?





„Segmental intonation“ ?!

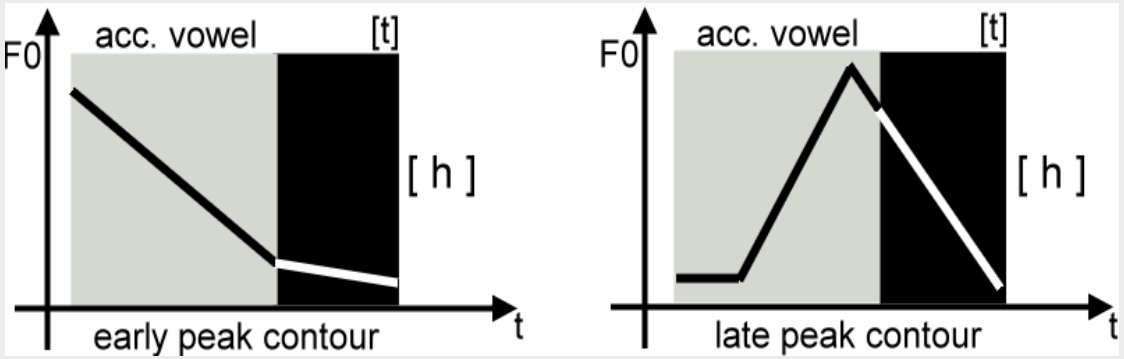
- Against the backdrop of such noise pitches:
- **Acoustic analysis** of a read-speech corpus, which was recorded in the early 80s by K.J. Kohler & R. Gartenberg.
- (Its original purpose: Investigation of microprosodic and other segmental effects on the phonetic realization of pitch accents and phrase-final intonation movements)
- My research subject: **utterance-final /t/ aspiration noises**

- Target words: Two-word utterances, ending in "___*ickt"* ([ɪkˈtʰ]) and starting with the personal pronoun "*Sie".* The word "___*ickt"* is prosodically realized with an early or a late pitch accent.
  - For example:
  - ***"Sie schickt", "Sie strickt", "Sie schrickt", "Sie tickt",* etc.**

„Segmental intonation“ ?!

- Against the backdrop of such noise pitches:
- Altogether analyzed 300 instances = 30 occurrences of the 5 target words (a produced – in randomized order – with both early and late accents by 10 different male speakers, including KJK.

(a) (b)



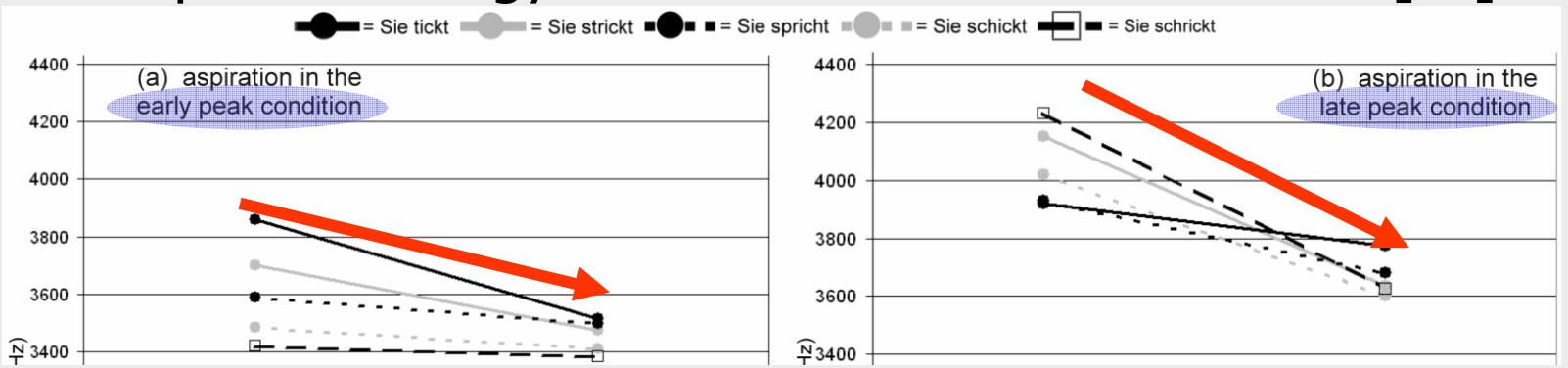
(Niebuhr, 2008)

- Tokens were analyzed with regard to duration and energy of aspiration noise, lower spectral energy boundary of the noise, spectral energy maximum and final F0 value before plosive

„Segmental intonation“ ?!

- Results part 1:
- There are differences between the “phonetic details” of the final /t/ aspiration noises that occur at the end of the "__*ickt”* words in early-peak and late-peak contexts.

Spectral energy maxima at onset and offset of [t^{sh}]



Spectral energy changes are consistent with perceptual impression that the pitch falls across the /t/ aspiration, particularly in the highly truncated late-peak contexts.

„Segmental intonation“ ?!

- Results part 2a:
- There are differences between the “phonetic details” of the final /t/ aspiration noises that occur at the end of the "__*ickt”* words in early-peak and late-peak contexts.

		(a) final F0 (Hz)	(b) dur asp (ms)	(c) Imax asp (dB)	(d) SEB (Hz)	(e) Emax spec (Hz)
early	Mean	74.6	96.2	58.5	1801 (1757)	3611 (3457)
	sd	8.1	9.8	2.2	53.8 (47.6)	284.7 (108.6)
late	Mean	127.5	106.2	54.8	1803 (1773)	4050 (3663)
	sd	6.3	19.7	2.5	48.6 (43.9)	413.6 (350.1)
F test	*F*	1.6767	0.2713	0.7864	1.2242 (1.1788)	0.4737 (0.0962)
	p	0.1064	0.0011	0.2803	0.3121 (0.3451)	0.0366 (<0.0001)
t test	*t*	−25.875	−2.321	5.5950	−0.1380 (−1.2360)	−4.3713 (−2.8100)
	dof	48	36	48	48 (48)	43 (29)
	p	<0.0001***	0.0261*	<0.0001***	0.8908 (0.2225)	<0.0001*** (0.0088**)

- In addition to stronger pitch decrease
- The /t/ aspiration noise after late pitch accents is significantly…
  - …longer
  - …higher (= higher spectral energy maximum in the noise)
  - …and softer (smaller overall energy of the noise)
- …compared with the aspiration noises following early pitch accents

„Segmental intonation“ ?!

- Why this additional difference?

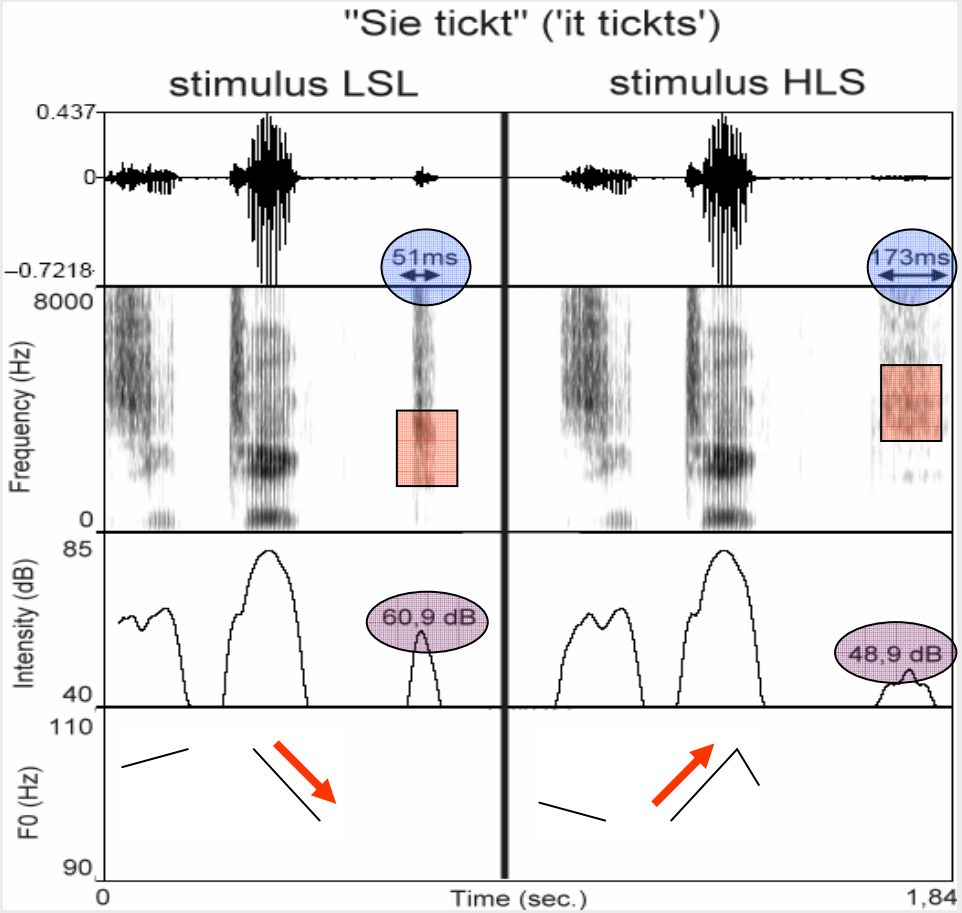
*“Sie **schickt**“*
‘früh‘ ,‘spät‘

Gartenberg und Panzlaff-Reuter (1991):
Late pitch accents increase the duration of the accented vowel

Dombrowski et al. (2006):
Late pitch accents go along with a softer, breathier voice quality in the accented vowel

Niebuhr (2007):
Late = stronger prominence of high pitch than Early (although low pitch dominates for both)

→ Do the additional differences contribute to the coding of (the meanings of) early and late pitch accents?



„Segmental intonation“ ?!

- Impetus for a perception experiment with judgments based on the semantic differential paradigm
- Starting from the utterance ***"Sie tickt"*** **(It ticks, [e.g., bomb or clock])** → most likely to occur as a separate utterance
- Systematic manipulation of the utterance-final aspiration noise
  - 2 Basic conditions:
  - “Long” (p) und “Soft” (ce)
  - “Short” (p) und “Loud” (ce)
  - Each of them combined with “High” and “Low” (o)
  - → **4 experimental conditions** (+4 with inverted duration and energy patterns, = HS, HL, LS, LL each with H and L)
- A constant, slightly declining (but perceptually virtually flat) F0 contour in all stimuli

„Segmental intonation“ ?!

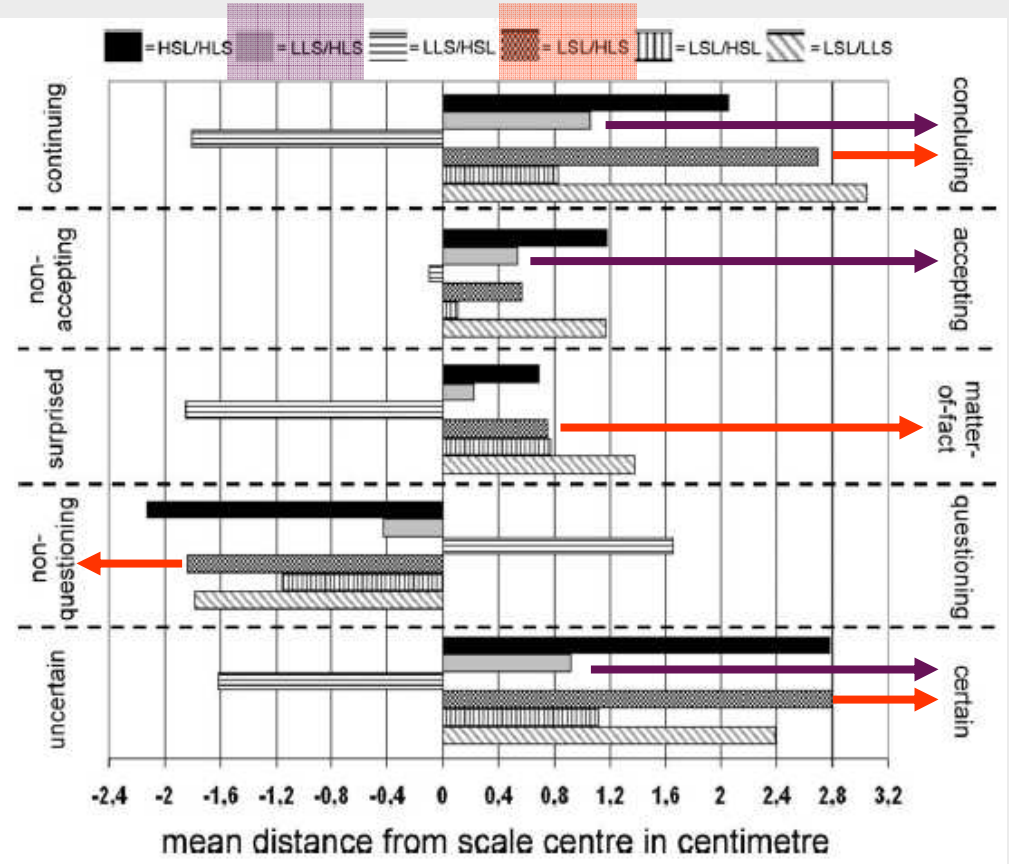
- What are the communicative meanings of the early and late pitch accents, i.e. in which respects do the two accent categories differ?
- Dombrowski (2003) and Kohler (2005) did perception experiments with the semantic differential (on the full set of Standard German pitch accents, including the “early and late peak”)
- Their results show the communicative meanings of the early and late accents differ significantly along 6 semantic scales

früh spät	Concluding Continuing	Accepting Nonaccepting	Matter-of-fact Surprised	Nonquestioning Questioning	Certain Uncertain

- The "*Sie* tickt” stimuli of there present experiment were judged on the basis of these 6 scales
- A total of 25 native speakers of (Northern Standard) German participated in the experiment and judged the 240 stimuli (8 stimuli x 5 repetitions x 6 scales) in an overall randomized order.

„Segmental intonation“ ?!

- Results part 2b:
- The meaning profiles created by the utterance-final aspiration noises (F0 was constantly flat) match well with those that were found on an F0 basis for the early and late pitch accents
- **Conclusion 1:** The pitch accents are not only represented in the speech signal by F0, …

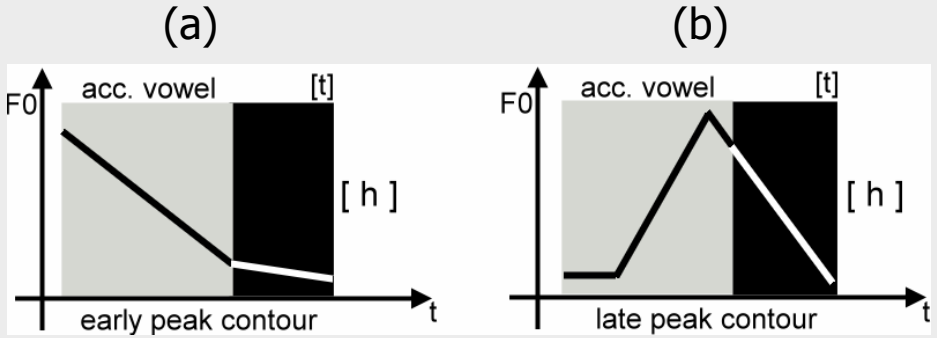


…but are also coded by prototypical sound features of the traditional segmental string.

„Segmental intonation“ ?!

- **Conclusion 2:** In addition to expressing pitch-accent meanings, the decrease of the spectral energy maximum across the aspiration noise “continues” the truncated utterance-final F0 fall and may trigger a perceptual completion of the pitch movement in the listeners.

“*Sie **schickt***“
(a) , (b)



(Niebuhr, 2008)

- What happens if we replace the two naturally produced aspiration noises by a third noise that is phonetically intermediate in terms of D+E and in which the spectral energy maximum does not successively decrease?

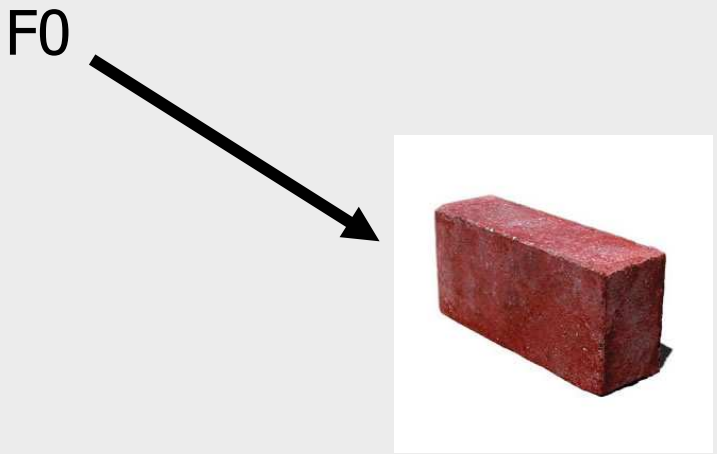
“*Sie **schickt***“, (a)… and (b)?

- Recently confirmed experimentally by Klaus Kohler by means of the initial „ICPhS“ example.

„Segmental intonation“ ?!

- This means that “truncation vs. compression of final F0 falls” seems to be only a sensible concept in the acoustic domain…
- ↔ In contrast, in perception the utterance-final speech sounds (here: /t/ aspiration) constitute…

„…another brick in the fall“



„Segmental intonation“ ?!

- In view of the results from utterance-final /t/ aspiration, it was assumed that also other sound classes show systematic fine phonetic differences depending on the intonational context.
- → The focus of a follow-up study was on phrase-final intonation movements (“boundary tones”) in questions (rising, H%) and statements (falling, L%)
- Which sounds are particularly promising in terms of intonation-induced phonetic variation?
  - Like /t/ aspiration /ʃ/ can create “sibilant pitch”. Moreover: /ʃ/ is realized in German as rounded [ʃʷ], and this rounding is typically already present on the preceding vowel → e.g., “Tisch” (table) = [tʰʏʃʷ].
    - So, does /ʃ/ have a lighter quality before utterance-final H% rises than before utterance-final L-% falls? If rounding is involved, can this sound-quality differences already be found in the vowel preceding /ʃ/?

Even more… „Segmental intonation“ ?!

- In view of the results from utterance-final /t/ aspiration, it was assumed that also other sound classes show systematic fine phonetic differences depending on the intonational context.
- → The focus of a follow-up study was on phrase-final intonation movements (“boundary tones”) in questions (rising, H%) and statements (falling, L%)
- Which sounds are particularly promising in terms of intonation-induced phonetic variation?
  - The voiceless velar/uvular fricative /x/ is also able to convey a range of different noise pitches. Moreover, as /x/ occurs after back rounded vowels like [u] in German, it typically also rounded, i.e. [x^{w}].
    - So, does /x/ have a lighter quality before utterance-final H% rises than before utterance-final L-% falls? If rounding is involved, can this sound-quality differences already be found in the vowel [u] preceding /x/?

Even more… „Segmental intonation“ ?!

- In view of the results from utterance-final /t/ aspiration, it was assumed that also other sound classes show systematic fine phonetic differences depending on the intonational context.
- → The focus of a follow-up study was on phrase-final intonation movements (“boundary tones”) in questions (rising, H%) and statements (falling, L%)
- Which sounds are particularly promising in terms of intonation-induced phonetic variation?
  - The suffix <-er> is typically realized in Standard German as an open vocoid, e.g. [ɐ]. However, it is well known that the exact sound quality varies greatly (from [ɑ̟] to [ɛ]) depending on the surrounding consonants and the speaker’s dialect. So, is intonation another source of variation? That is, does the <–er> vocoid show a lighter quality (more open/fronted) before H% rises than before L% falls?
  - Like [ɐ], schwa (/ə/) is also a phonetically very variable sound in German. Is the phonetic quality of schwa also adapted to the intonational context and hence lighter before H% rises than before L% falls?

Even more… „Segmental intonation“ ?!

- → Study on the acoustic representation of pitch contexts (or “targets”) in fricative and vowel sounds in German = “intonation segments” (=H%, L%) and “segmental intonations” (pitch impression created by speech sounds)
- **Based on pairs of target words:**
  - *"Tisch", "Fisch"* (table, fish) → /ʃ/ preceded by /ɪ/
  - *"Buch", "Tuch"* (book, towel) → /x/ preceded by /u/
  - *"lecker", "Bäcker"* (delicious, backer) → vocalized <-er>
  - *"Tage", "Schramme"* (days, scratch) → /ə/
  - In the vocoid sounds LPC measurements of **F2** were done at three points in th segment: Onset +20 ms, middle, offset -20 ms
  - In the fricative sounds the **Centre of Gravity (CoG)** was calculated every 7 ms. Then, based on this (variable) number of values the mean CoG and the CoG range were determined for each sound.
  - The **durations** of all sounds were measured.

Even more… „Segmental intonation“ ?!

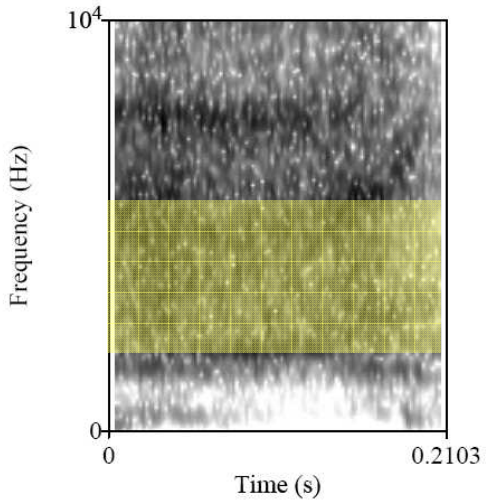
- A corpus of spontaneous-sounding speech was recorded, using an improved version of the method of Kohler und Niebuhr (2007)
- This means:
- 4 Dialogue texts with an informal speaking were created on everyday topics
- The target words were integrated into the texts (without marking them in any way). The utterance-final rises and falls as well as the accentuations of the targets were evoked solely by means of the semantic-pragmatic contexts
- The dialogues were produced after 30 min practice by pairs of speakers who knew each other very well for a long time (i.e. good friends)
- Except for the target words, the speakers were allowed to adjust the texts to their own vocabulary by replacing, adding or omitting words.
- One of the speakers was the experimenter himself (me). He tried with his own informal way of speaking (a) to create an informal atmosphere in the lab and (b) to elicit a similar speaking style in the subject’s productions.
- Every dialogue was produced 4 times; only the last 2 repetitions were used for acoustic analyses.
- So far, 10 speakers were recorded (→ n=40), recordings of 10 more speakers are planned.
- The method results in a compromise between real-life communication behaviour and the conextual control of read speech

Even more… „Segmental intonation“ ?!

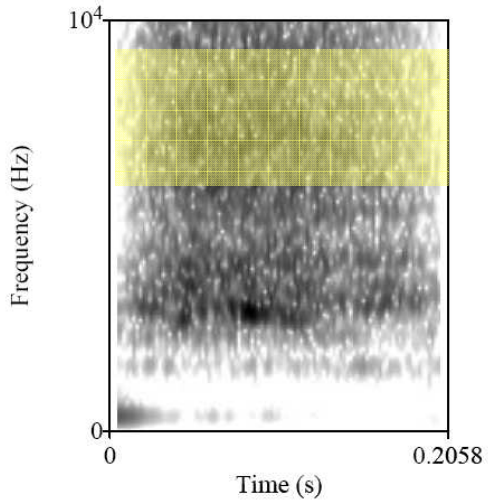
- Results for ***"Tisch"*** **and** ***"Fisch"***
- The sibilant /ʃ/ was produced considerably lighter when preceded by H% than when preceded by L-%. This perceptual impression is reflected in significantly different mean CoGs. That is, the mean CoGs are higher in H% than in L% contexts. The fricative durations do not differ significantly.

		/ɪ/	/ʃ/
Mean CoG (Hz)	L-%		4.500
	H-%		5.300
F2 (middle, Hz)	L-%	1.600	
	H-%	1.900	

L-%



H-%

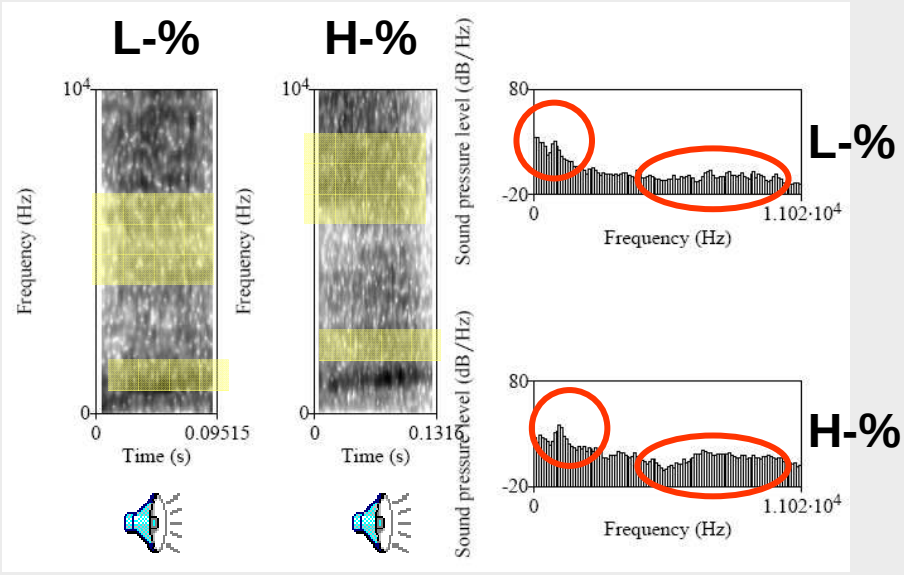


- The realizations of /ɪ/ also differ clearly depending on H% and L%. When followed by H% F2 (middle) is significantly higher. Perceptual analyses suggest that this difference involved the parameter ‘lip rounding’.

Even more… „Segmental intonation“ ?!

- Results for ***"Buch"*** **and** ***"Tuch"***
- The frikative /x/ shows a lighter noise quality in H% than in L% contexts. Correspondingly, the acoustic mean CoGs were significantly higher for H% rises than for L% falls. The fricative durations do not differ significantly.

		/u/	/x/
mean CoG (Hz)	L-%		2.400
	H-%		3.500
F2 (middle, Hz)	L-%	500	
	H-%	700	



- The /u/ productions are also slightly, but significantly different depending on H% and L%. When followed by H% F2 (middle) tended to be higher (=p<0.1). According to perceptual impressions this phonetic difference involves lip rounding.

Even more… „Segmental intonation“ ?!

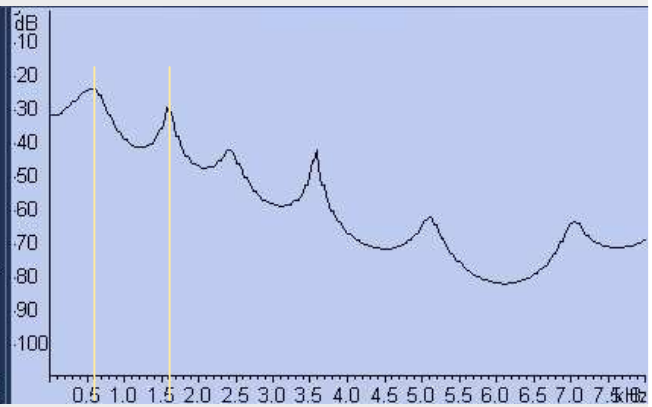
- Results for ***"lecker"*** **and** ***"Bäcker"***
- The sound quality of <-er> shows significant differences between H% and L%, but not across the whole voicoid. Only towards the offset of the sound segment (-20 ms) F2 tends to be higher for H% than for L% (=p<0.1).
- → In consequence, the vocoids spanned by the H% intonation rises are realized with slight diphthongation ([ɐɛ]) and are in tendency longer (=p<0.1) than the vocoids spanned by the L% intonation falls.

		<-er>
F2 (offset, Hz)	L-%	1.250
	H-%	1.600
Dauer (ms)	L-%	105
	H-%	130

L-%

700 1300

H-%



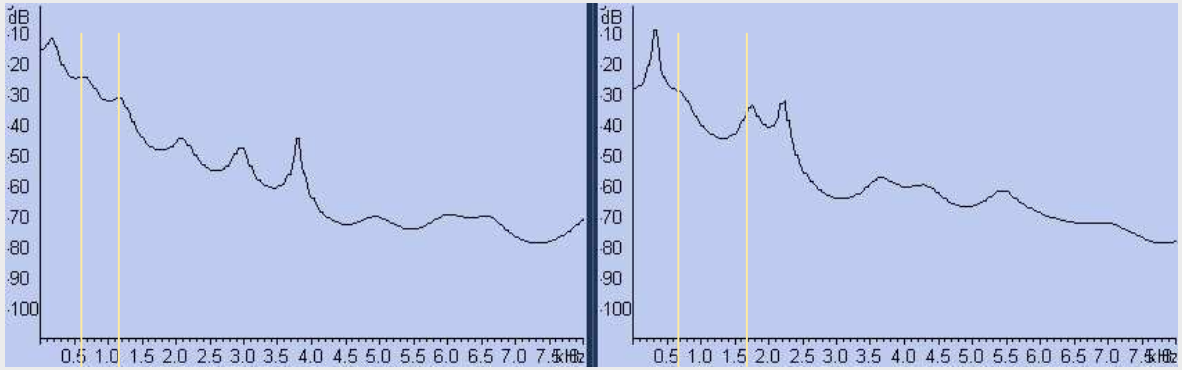
650 1650

Even more… „Segmental intonation“ ?!

- Results for ***"Tage"*** **and** ***"Schramme"***
- The /ə/ realizations are produced lighter in connection with H% than in connection with L%. This is true in tendency (=p<0.1) for the intial measurement after vowel onset and holds in terms of a clear significant difference for the measurements in the middle of the vowel and towards the vowel offset. The schwa durations are not significantly different in connection with coinciding H% and L% intonation movements.

		/ə/
F2 (onset, Hz)	L-%	1.400
	H-%	1.500
F2 (middle, Hz)	L-%	1.600
	H-%	1.900
F2 (offset, Hz)	L-%	1.500
	H-%	1.900

L-% **H-%**



600 1200 **650 1750**

Assimilation + „Segmental intonation“ ?!

- So far, we have only dealt with sound variations in utterance-final intonation contexts.
- Are such “segmental intonations” are peculiarity of this structural position? Or can we also find similar sound variations earlier in the intonation phrase?
- This is the topic of a recent follow-up study (to be presented at the ICPhS 2011). However, the corresponding acoustic analyses primarily aimed at investigating regressive sibilant assimilation patterns in /sʃ/ or [z̬ʃ] sequences of German, compared with /ss/ or [z̬s] sequences.
- We know that – like, for example, in English and French – regressive assimilation of place of articulation (alveolar-to-postalveolar) is a gradual process.
- It can also be assumed that stronger /s/-to-[ʃ] assimilation results in perceptually “darker” sibilant noise (up to [ʃʷʃʷ])

Assimilation+„Segmental intonation“ ?!

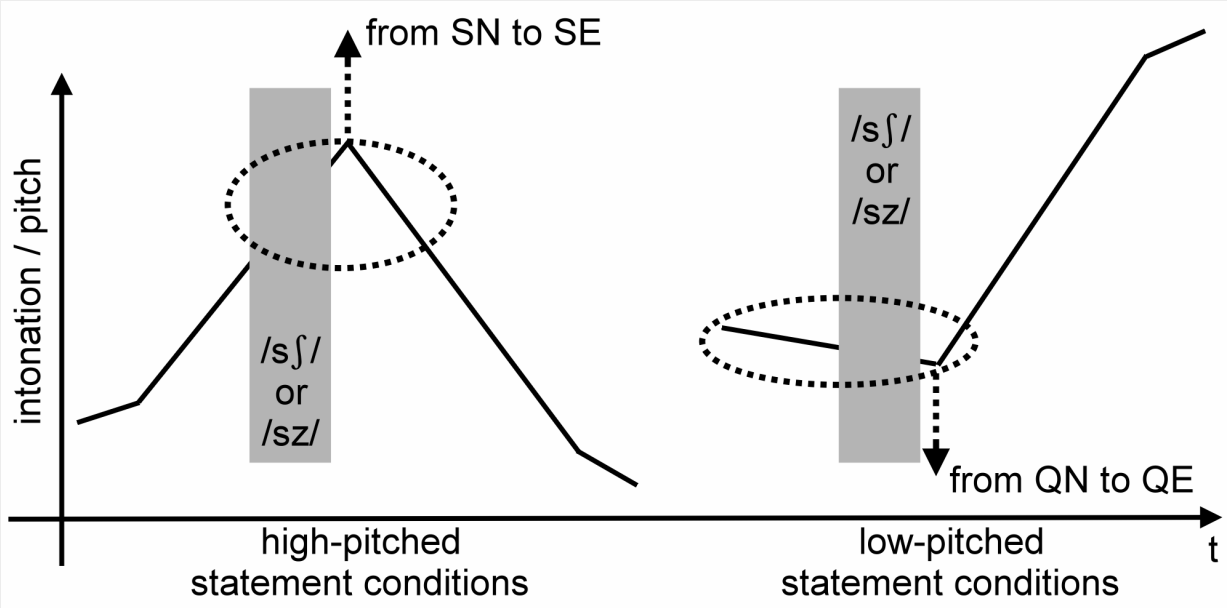
- We used a sub-section of an existing Standard German speech corpus (part of **“KIESEL”** = Kieler Sammlung emphatischer Lesesprache, currently about 4h of speech)
- Specifically, data of 8 female speakers…
- …producing target pairs of function words and nouns like
  - *"aus Schweden”* ([aʊs ʃʷv̥eːdⁿn], from Sweden) → /sʃ/
  - *"als Sänger”* ([alts z̥ɛŋɐ], as a singer) → /sz̥/
  - *"bis Sachsen”* ([b̥Is z̥aksʸnʸ], to Saxony) → /sz̥/
  - *"als Spender”* ([alts ʃʷpɛndɐ], as a donor) → /sʃ/
- In both a neutral and an emphatic speaking style
- The nouns showed the nuclear pitch accent.

Assimilation+„Segmental intonation“ ?!

- We used a sub-section of an existing Standard German speech corpus (part of **“KIESEL”** = Kieler Sammlung emphatischer Lesesprache, currently about 4h of speech)
- Specifically, data of 8 female speakers…
- …producing target pairs of function words and nouns like
  - *"aus Schweden”* ([aʊs ʃʷv̥eːdⁿn], from Sweden) → /sʃ/
  - *"als Sänger”* ([alts z̥ɛŋɐ], as a singer) → /sz̥/
  - *"bis Sachsen”* ([b̥Is z̥aksʏnʏ], to Saxony) → /sz̥/
  - *"als Spender”* ([alts ʃʷpɛndɐ], as a donor) → /sʃ/
- …in both a neutral and an emphatic speaking style (N vs. E)
- …and as both declarative statements and questions (S vs. Q).
- The nouns showed the nuclear pitch accent.
- = A total of almost 400 tokens, 96 per sibilant condition.

Assimilation+„Segmental intonation“ ?!

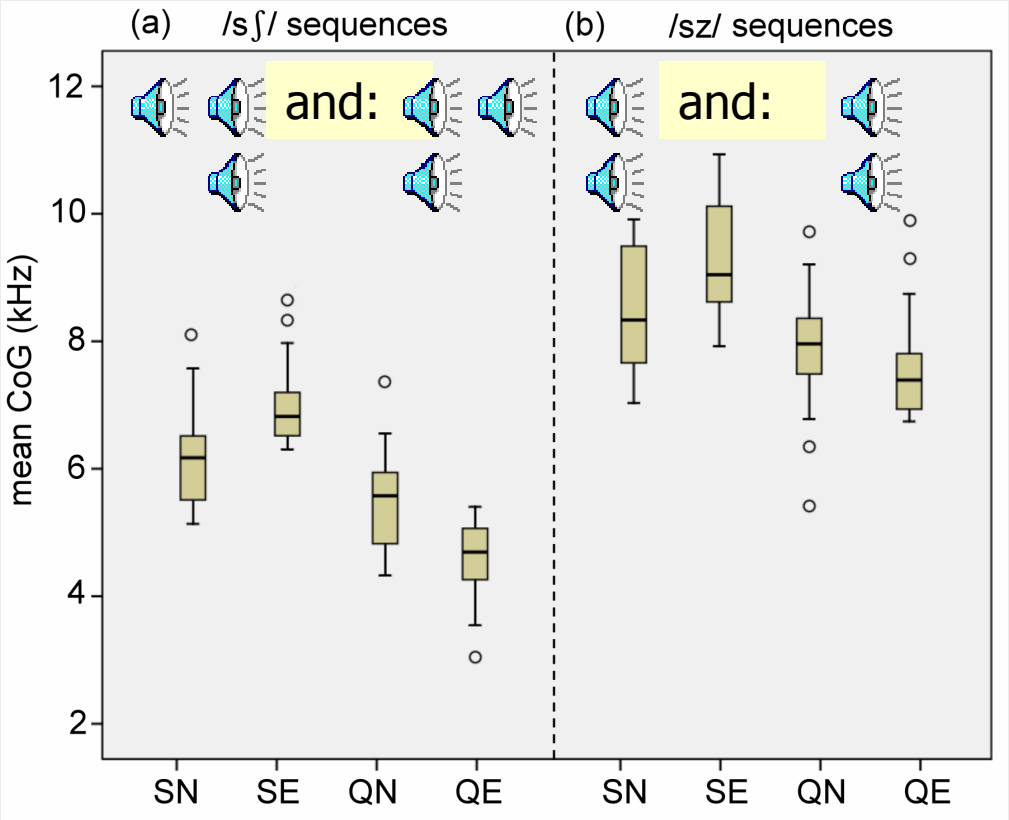
- Nouns in questions were produced with low pitch accent (i.e. L*) ↔ Nouns in statements were produced with high pitch accent (i.e. H*)
- Emphasis lowered L* and raised H*.



- Like in the previous study, CoG means, CoG ranges and sibiliant-sequence durations were determined.

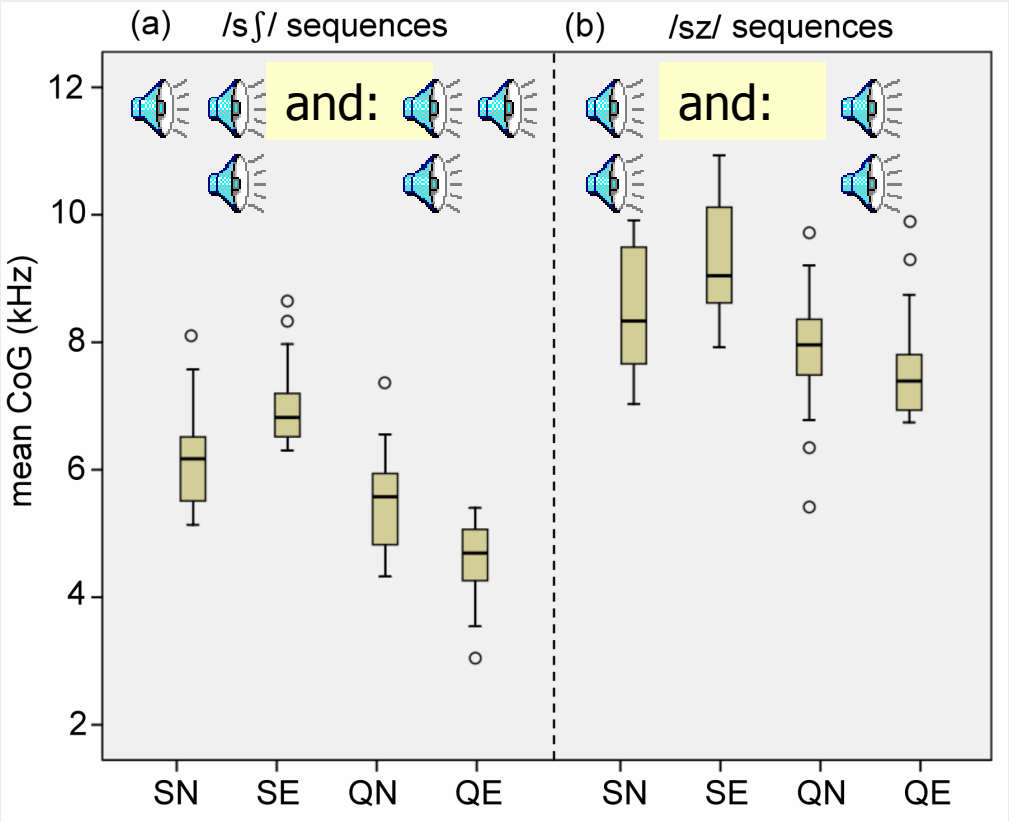
Assimilation+„Segmental intonation“ ?!

- Results:
- Pitch-accent type and emphasis both strongly (and significantly) affected the mean CoGs.
- Higher mean CoGs for before H* than before L* accents
- This pitch-accent difference increases under emphasis, parallel to the F0 difference between H* and L*



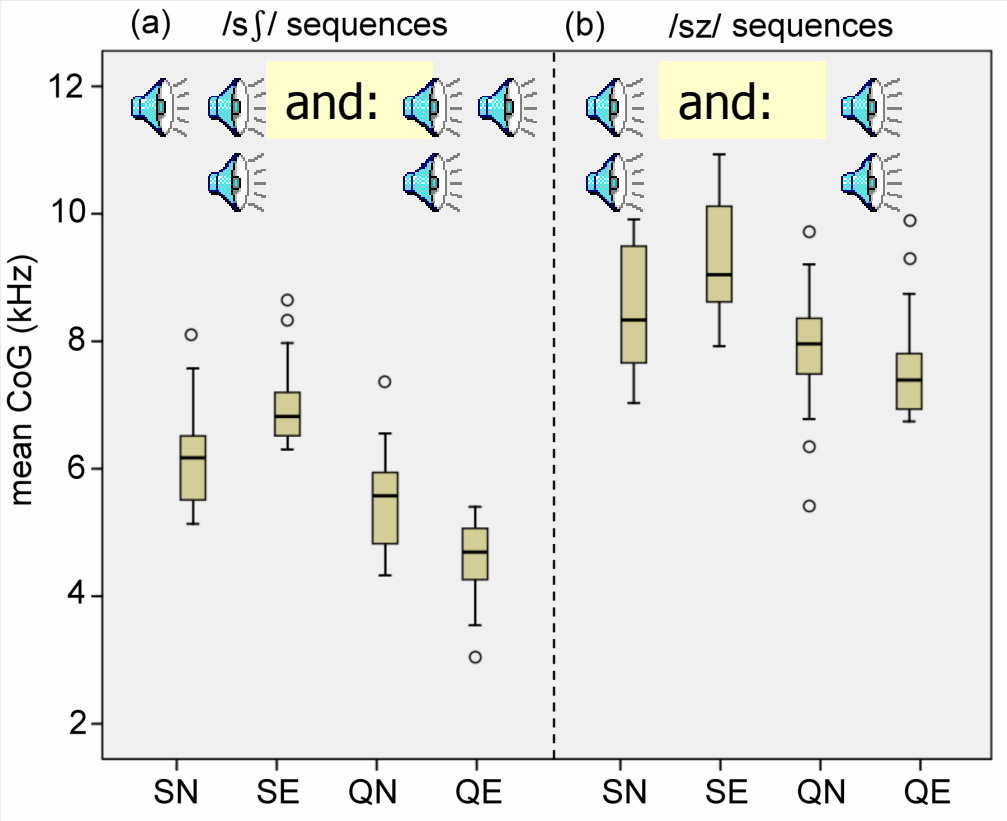
Assimilation+„Segmental intonation“ ?!

- Results:
- Pitch-accent type and emphasis both strongly (and significantly) affected the mean CoGs.
- The co-variation of nuclear pitch-accent scaling and mean CoG was stronger in the assimilating /sʃ/ than in the [sʒ] sequences.
- Lower mean CoGs in /sʃ/ went along with smaller CoG ranges (stronger spectral assimilation)



Assimilation+„Segmental intonation“ ?!

- Results:
- Pitch-accent type and emphasis both strongly (and significantly) affected the mean CoGs.
- The sibilant-sequence durations differed due to emphasis and S/Q, but not between /sʃ/ and [sz̥] → no elision



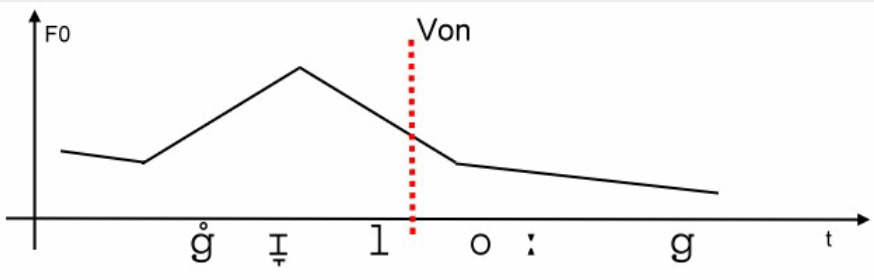
Assimilation+„Segmental intonation“ ?!

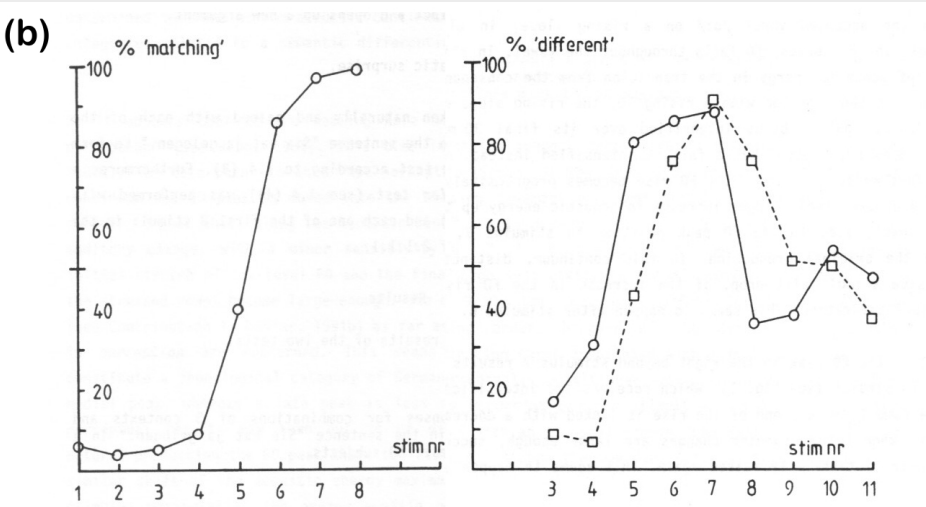
- Conclusion:
- The phenomenon of “segmental intonation” is not restricted to utterance-final speech sounds, but occurs also earlier in the prosodic phrase…
- …at least for sibilant sequence.
- In the case of /sʃ/ sequences the degree of spectral (and also temporal) assimilation seems to be used as a means to adjust the sibilant pitch of the sequence to the intonational contexts.

- The use of an essentially segmental process for intonational purposes strongly crosses the traditional dichotomy between the segmental and the prosodic layers of the speech signal.

Assimilation+„Segmental intonation“ and vowel intensity

- A final note on the seminal intonational study of Kohler (1987): On the one hand, it is well known that his F0-peak shift triggered an abrupt „categorial“ perceptual change from early → non-early at the accented-vowel onset





- But on the other hand: it is less well known that later perception experiments (Kohler 1991) yielded also less abrupt perceptual changes at later positions in the F0-peak shift continuum!
  - "Sie hat ja ge**lo**gen" = lateral + vowel ⇒ Earlier Boundary
  - "Sie ist ja ge**ritt**en" = frikative + vowel
  - "Sie hat ja ge**jo**delt" = approximant + vowel ⇒ Later Boundary

„Segmental intonation“ ?! …and vowel intensity

- Obvious question:
- Why these different outcomes in different utterances for an identical F0-peak shift continuum?

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

- The intensity change is more abrupt for sequences of nasal+vowel or lateral+vowel than for a sequence of approximant+vowel.
- And the abruptness depends on the phonetic quality of the vowel itself.

„Segmental intonation“ ?! …and vowel intensity

- Starting from this idea,
  - Two F0-peak shift series were resynthesized (11 steps à 20 ms)
  - One using the stimulus utterance *"Sie war mal **Ma**lerin"* (She was once a painter)
  - The other utterance keeps exactly the F0 and intensity patterns of the “**Ma**lerin” series, but on a constant Schwa-like vowel quality (=“HUM” in ‘praat’)
  - So, basically, the two stimulus series (“**Ma**lerin” and “HUM”) differ solely in terms of the presence / absence of the segmental string.
  - Two parallel perception experiments with two separate groups of subjects (native speakers of Standard Northern German)…
  - …using an Indirect Identification test for the “**Ma**lerin” series
  - …and an AXB test for the “HUM” series

„Segmental intonation“ ?! …and vowel intensity

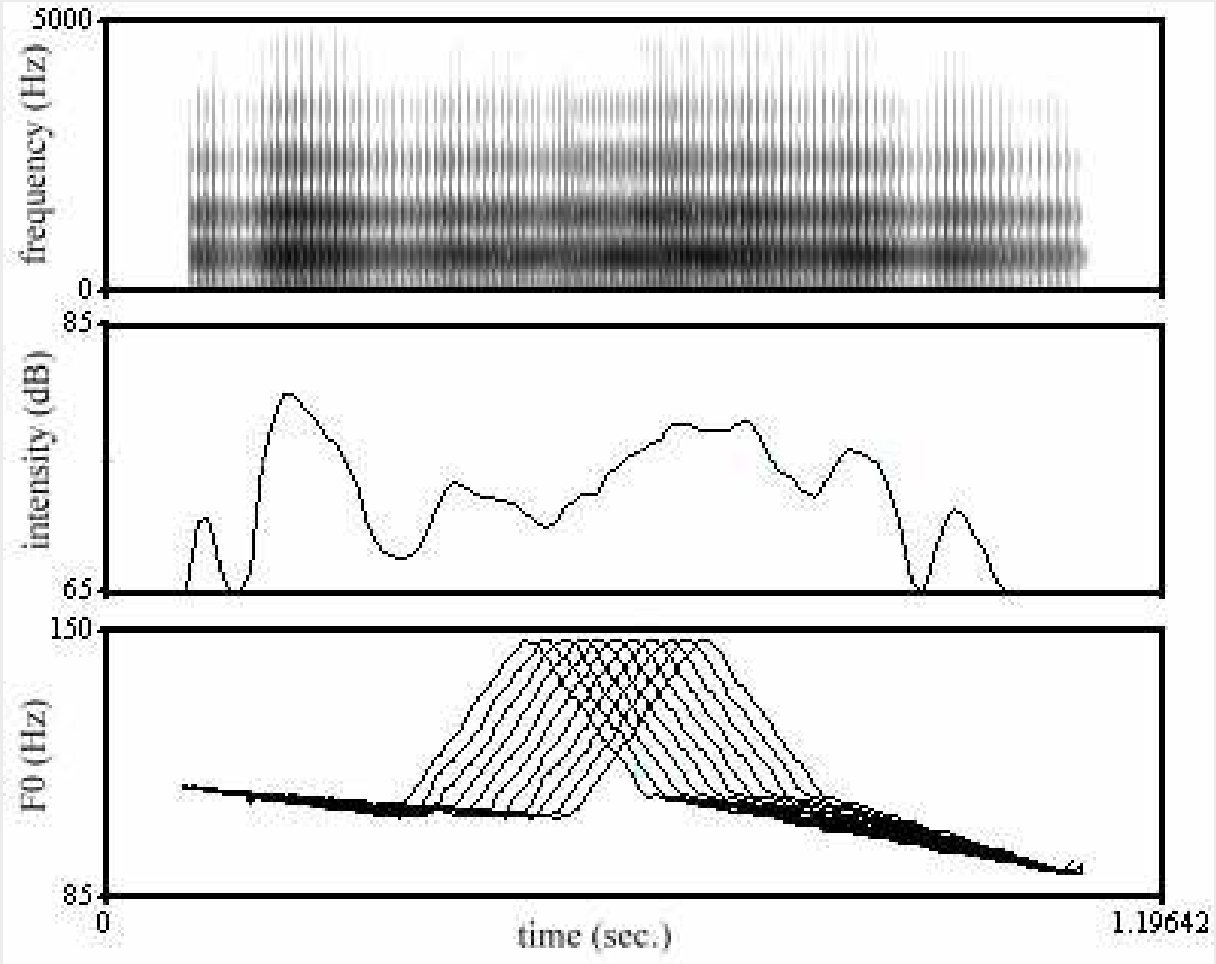
“**Ma**lerin“ series

11 stimuli

→ perceptual change at…?

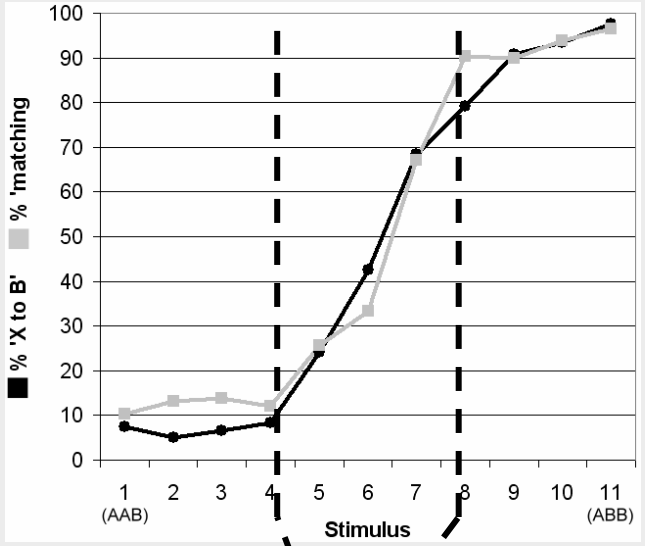
The 11

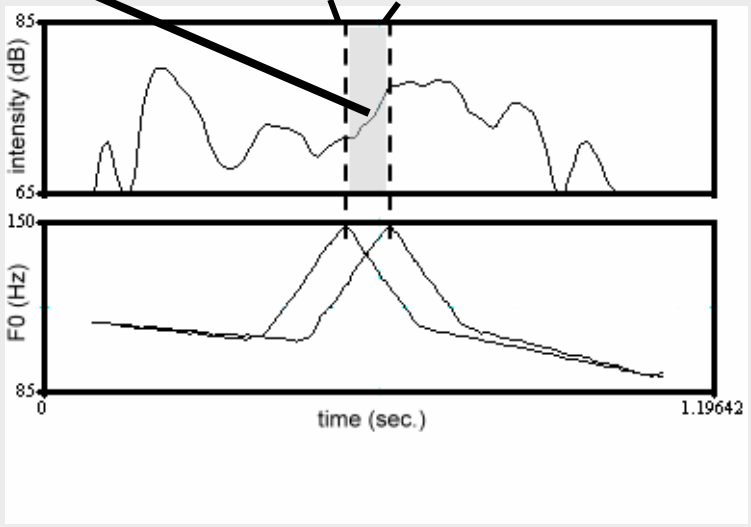
corresponding HUM stimuli



„Segmental intonation“ ?! …and vowel intensity

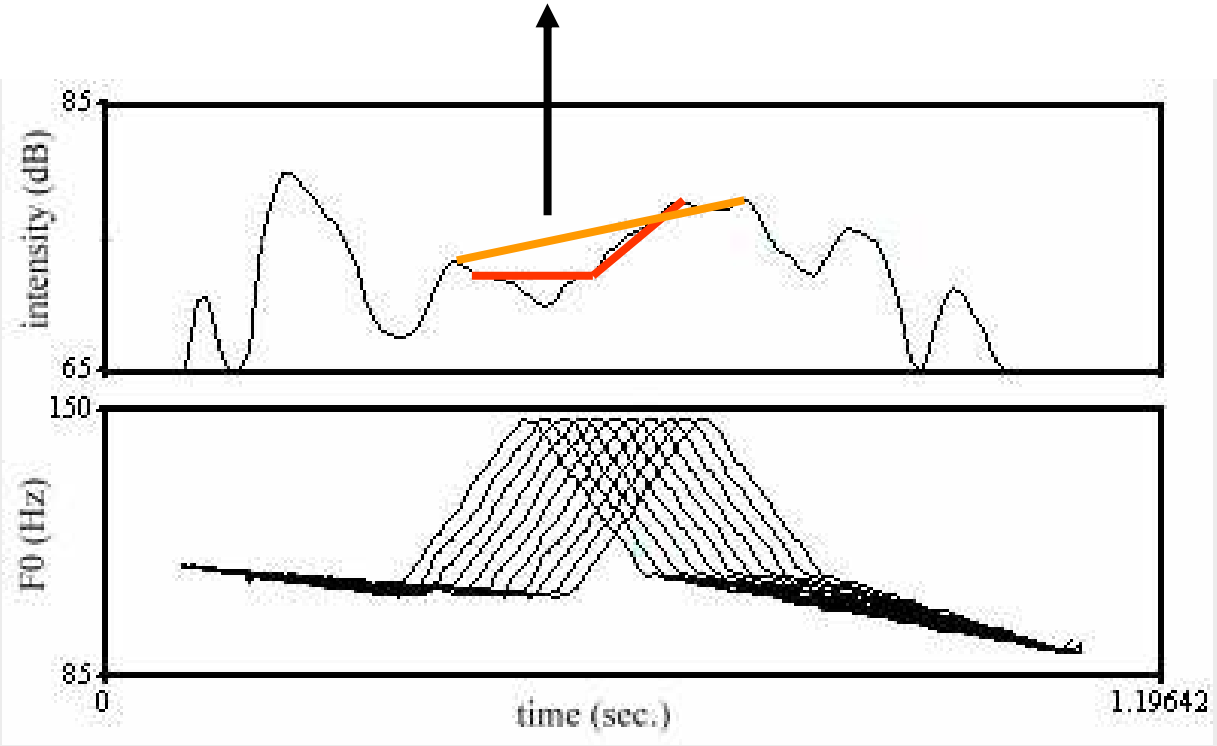
So, what happens, if we manipulate the rising slope of the intensity curve that reflects the CV transition?



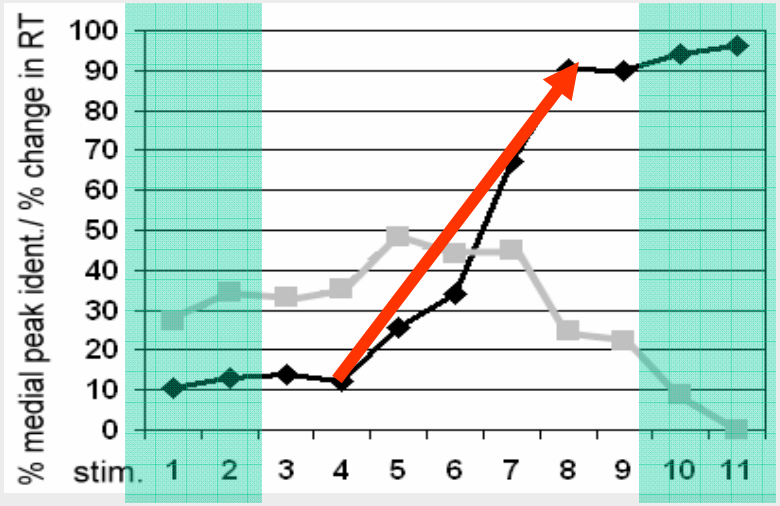


„Segmental intonation“ ?! …and vowel intensity

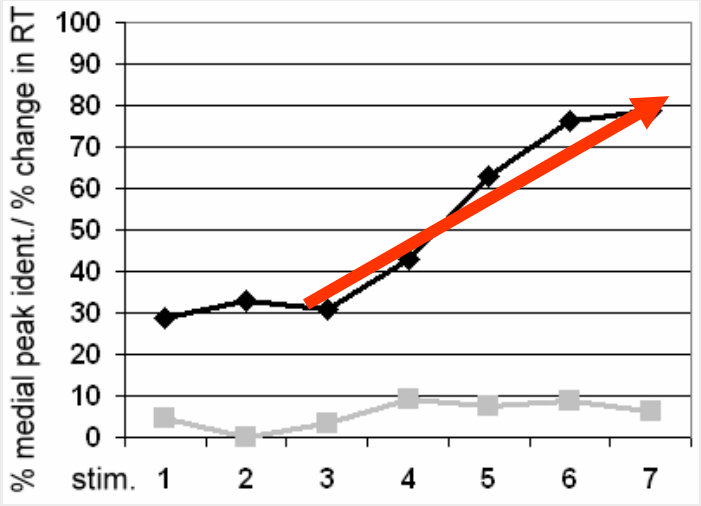
Again two perception experiments, based on (a) “**Ma**lerin“ and (b) “HUM“ stimulus series.



„Segmental intonation“ ?! …and vowel intensity



“**Ma**lerin“ series



- The dynamics of the perceptual change from early to non-early accent decreases with decreasing dynamics of the underlying intensity change.
- The same effect shows up, if a less pointed F0 peak is shifted.
- A completely similar effect of the dynamics of the F0 and intensity courses on the pitch-accent perception can be found for peak-shift series from **medial to late** pitch accent in German, based on a manipulation of the decreasing intensity at the **VC boundary**

„Segmental intonation“ ?! …and vowel intensity

- Conclusions of Niebuhr (2006, 2007): The picture sketched by Kohler (1987, 1991) must be refined
  - The abruptness of the perceptual change between two “adjacent” pitch accents is **not** determined by the accent types themselves.
  - The change from early to non-early can become gradual
  - ⇔ The change from non-early to late can be made “categorical”
  - The prosodic signalling of German pitch accents is based on an **interplay between F0 and intensity changes (or levels)**.
  - ⇒ But the findings also support the central claim of Kohler that the synchronization of the pitch accent peak relative to the vowel boundaries is decisive for the pitch-accent identification

mmm

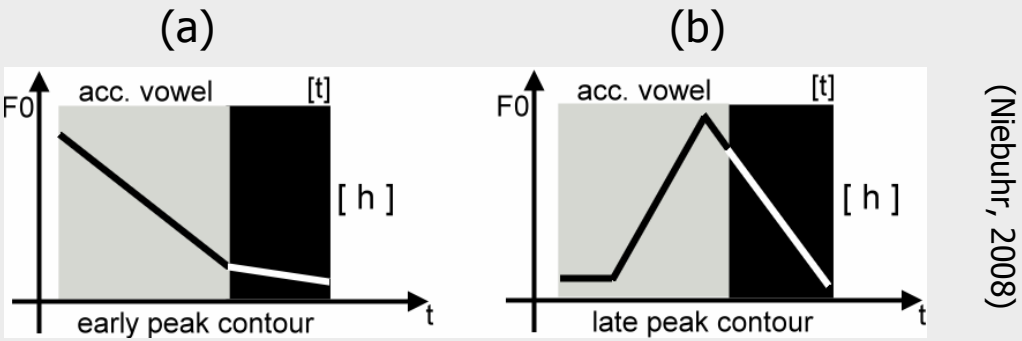
„Phonetic detail“ in Intonation

- General summary
- The (so-called) “segmental layer” of the speech signal is not (merely) a “troublemaker” for the coding of pitch accent and phrase-final intonation movements. Rather, speech sounds show fine phonetic detail that make a separate contribution to the signalling of intonational patterns and their meanings.
- This contribution can be of two different kinds:
  - Direct: (a) via changes/F0-related adaptations in the spectral energy distribution of both frivatives/aspirations and vocoid sounds (creating what is known as *vowel/intrinsic pitch, sibilant pitch*, etc.)
  - Indirect: (a) in terms of “articulatory metaphers” of intonational meanings (e.g., meanings of early is reflected in short, “harsh” sounds, meaning of late is reflected in longer, more cautious sounds); (b) in terms of a manifestation of tonal structures (e.g., diphthong dynamics and sequences of pitch accent and phrase-final intonation, etc.)

„Phonetic detail“ in Intonation

- General summary
- It is likely that the “segmental intonations” play a major role for our cross-linguistic perceptual impression that intonation is *„certainly subjectively continuous"* (Daniel Jones 1909:275), despite about 30% of the speech sounds are voiceless.

“*Sie **schickt***“
(a) , (b)



(Niebuhr, 2008)

“*Sie **schickt***“, (a)… und (b)?

*„Er arbeitet al**s** **S**chneider*"

und:

„Phonetic detail“ in Intonation

- General summary
- Even though it is the first time that “segmental intonations” are systematically investigated in production and perception studies…
- …the notion of “segmental intonation” is pretty old!

- In his book “the phoneme and its use”, Jones notes that *"voice pitch"* (1950:9) must be included in the list of *"phonetic contexts"* that cause allophonic variation in both consonants and vowels.