Region of Southern Denmark

SDU MADS CLAUSEN INSTITUTE

How to ask without a question and to reply without an answer – Prosodic meaning dimensions in German dialogues

Oliver Niebuhr
SDU Electrical Engineering, Mads Clausen Institute
University of Southern Denmark (SDU), Sønderborg, Denmark
olni@sdu.dk

Region of Southern Denmark

The “myth” of question intonation

SDU
MADS CLAUSEN INSTITUTE

Flensburg city hall
19 June 2009

“When can we finally go home” (?)

“If you want me to answer then you have to ask a **real** question”

So, what is
a real question then?!

The “myth” of question intonation

- **3 assumptions**
  - (1) **Intonation** is the only relevant prosodic parameter in Q/S signaling
  - (2) The Q/S signals of intonation are located **at the end** of sentences/phrases
  - (3) The Q/S signals are represented by the difference between **rise and fall**

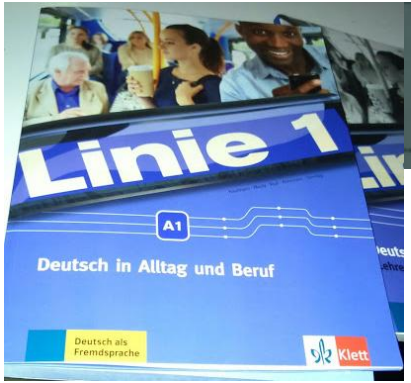
- These 3 assumptions were very widespread in the early linguistic/phonetic research on sentence mode in German – and many other West Germanic languages, see Trubetzkoy (1938), Siebs (1969), Bolinger (1978), Crystal (1970), for (critical) discussions on that topic

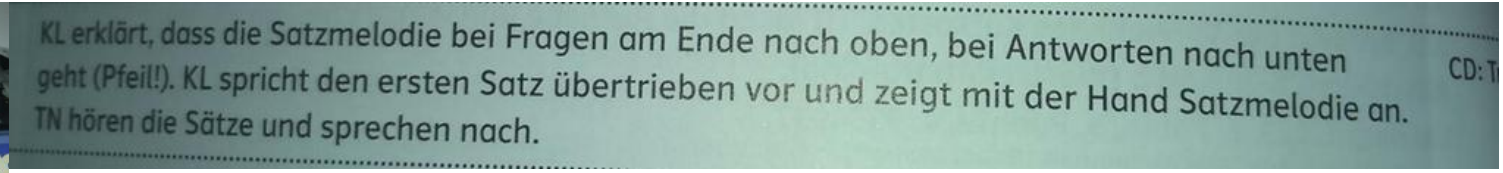
- They are still common today, especially in German-as-a-second-language courses and textbooks

The “myth” of question intonation

- Heinz Griesbach (2000:231): *„Fragen erkennt man bei mündlichen Äußerungen an der Frageintonation, im schriftlichen Ausdruck am Fragezeichen am Ende des Satzes“* (aus Bauplan Deutsch: Übungsgrammatik und Satzbauhelfer. Frankfurt: Libri)

- Gerhard Helbig (2001:430): *„Eine besondere Schwierigkeit für Finnen ist die steigende Frageintonation“* (aus Deutsch als Fremdsprache: Ein internationales Handbuch, Band 1. Berlin: de Gruyter)



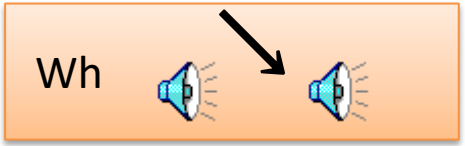

KL erklärt, dass die Satzmelodie bei Fragen am Ende nach oben, bei Antworten nach unten geht (Pfeil!). KL spricht den ersten Satz übertrieben vor und zeigt mit der Hand Satzmelodie an. TN hören die Sätze und sprechen nach.

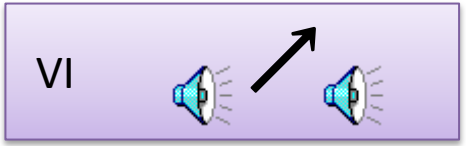
- Klett (2015)

The “myth” of question intonation

- With the upcoming empirical research in the speech sciences, the claim of a “question intonation” was further refined and replaced by an interaction between morphosyntactic structure and final intonation

  - Wh-questions fall at the end
  - Questions with question-syntax (VI) rise at the end



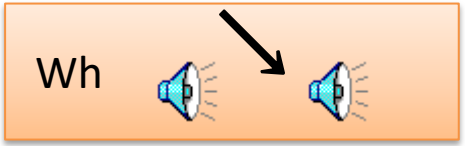


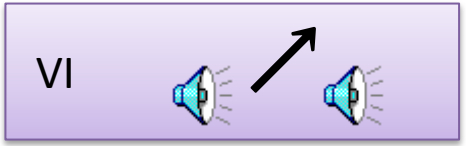
- German: von Essen (1964), Isacenko und Schädlich (1970)
- English: Halliday (1967)
- Dutch: `t Hart et al. (1990)
- (…)

The “myth” of question intonation

- However, the strictness of this claim has always been subject of critical discussion

- Peters (2005:100): *“bestimmte syntaktische Konstruktionen [werden]* ***eher*** *mit der einen als mit der anderen [satzfinalen Intonation] kombiniert“*
- Stock und Zacharias (1973:138): there is “***no obligatory connection***” between lexically and syntactically marked questions and the direction of their final intonation movement



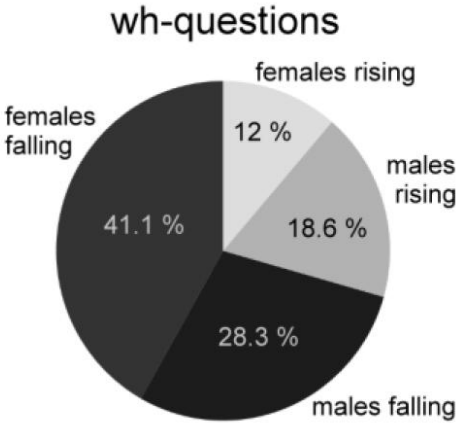


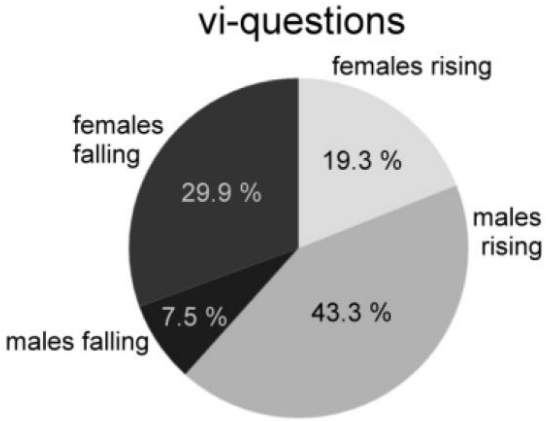
- What *“rather”* and “*not obligatory*” exactly means was quantified in several acoustic-prosodic analyses of larger speech corpora
- Selting (1995), Kügler (2003), Kohler (2004), Peters (2005), Michalsky (2016)

The “myth” of question intonation

- Kohler (2004): based on the ‘Kiel Corpus of Spontaneous Speech‘ (Peters et al. 2005)
  - = 42 speaker; 82 dialogues; 1,387 turns; 293 relevant questions
  - 119 instances of Wh-questions; 174 instances of VI questions

n=293	↘	↗
Wh	57%	43%
VI	21%	79%





- The corpus analyses of Kügler (2003) and Selting (1995) yielded similar percentages

The “myth” of question intonation

- So, obviously the final rising/falling intonation in (German) questions is not simply a redundant question signal! But, what is it then?

  - Kügler (2003), Selting (1995), Peters (2005), Michalsky (2016): **The final intonation movement has a separate communicative function within (German) questions**
  - It is about discourse control and turn taking
  - Rise/Fall = **Speaker allows dialogue partner to give a longer/shorter answer**

- *“Are you from Denmark?”* → Please only say ‘yes’ or ‘no’ (fall) vs opening a conversation (rise) = please tell me a bit more about you (does not even require an answer)

- The correlation of ↗ and ↘ with Wh and VI questions results from the fact that answers to VI questions are inherently longer than answers to Wh questions. The ‘yes’ or ‘no’ on a VI question must be commented.

The “myth” of question intonation

- Further functional interpretation:
- Rise = listener-oriented, interlocutor can take control, no particular answer expected
- Fall = speaker-oriented, speaker keeps control, “give me the answer that I like to hear”
  - *“Würde Ihnen das passen?”* With ↘ = “yes” is expected.
  - *“Was würden Sie denn davon halten?”* With ↘ = “okay” is expected
  - *“An welchen Tagen hätten Sie Zeit?” With* ↗ = “please suggest any date you like”

The “myth” of question intonation

- The findings of Fries (1964) fit in with this latter explanation:
- Speech data from 39 AE radio quit shows (similar to *“Guess Who?!”*)
- Only 38.3% of all questions showed a final rise, i.e. almost 2/3 were falling. Why?
- The game required the contestants to elicit a “yes” from the candidate → Contestants designed their questions so as to get the expected answer

- To conclude, assumption (3) can be rejected
  The direction of the final intonation movement (rise vs fall) is not about sentence mode itself.
  It is about the speaker-hearer relationship (in questions) (Michalsky 2016)

cf. the Flensburg city hall example

The “myth” of question intonation

- **However, can this conclusion also be valid for questions that are not morpho-syntactically marked and in which the final intonation is not redundant?**

- Isacenko & Schädlich (1970:32): *„if an utterance contains no syntactic or lexical cue to identify it as a question […] then only the last […] rising tone-switch provides the necessary information to allow the hearer to identify it as a question“*.
- Stock & Zacharias (1973:137): *„Ein Satz, der strukturell als Nichtfrage erscheint, muss mit Steigkadenz gesprochen werden, wenn er eine Fragefunktion übernehmen soll“*
- Lieberman (1967:136): *„Declarative questions must show a final rise“*.

- Would be a strange functional asymmetry —— **and it is not true!**
  “Du meinst Deine Tante aus Schweden?”
  “Auf dem Bau?”
  “Es ist also die aus Salzburg?”
  “Du meinst das ‘Sieben‘ ?”

Prosodic characteristics of questions

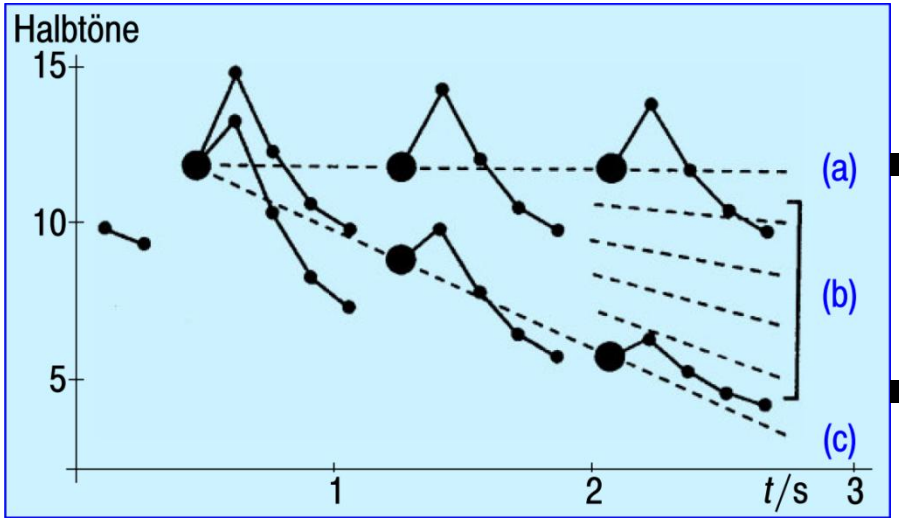
- So, if rise vs fall has the same separate communicative function in (morphosynt.) unmarked as (morphosynt.) marked questions, how are questions and statements then differentiated in unmarked declarative questions?

- Starting point of our own research on prosodic question signals

- First question: what do other languages do?
  - **Faster speaking rate** in Neap. Ital., Malay, Cantonese etc. (D'Imperio 2000; Petrone 2008; van Heuven & van Santen 2005; Ma et al. 2010)
  - **Breathier voice quality** in Bari Ital. (Grice und Savino 1997**)**
  - **Fewer prenuclear accents** and hence **flatter intonation** – also due to the fact that unmarked questions are functionally often checks/confirmation questions… = focus on a single specific information (Rost-Roth 2004)

Prosodic characteristics of questions

- **Less steep F0 declination**, general tendency to **increase high/suppress low F0** in Danish (Thorsen 1980, 1983), Hausa (Lindau 1986), Mandarin (Shen 1990), etc.



Question

Statement

Prosodic characteristics of questions

- **Does German make use of these phrase-level prosodic characteristics as well in order to distinguish statements from (morphosynt. unmarked) questions?**
  - Enacted everyday conversations with target sentences included
  - All sentences are 5-7 syllables long
  - All have a declarative SVO structure

    "Du <u>warst</u> in <u>Mai</u>land/<u>Ne</u>pal ./?"

    "Er <u>ar</u>beitet als <u>Sänger</u>/<u>Schnei</u>der ./?"

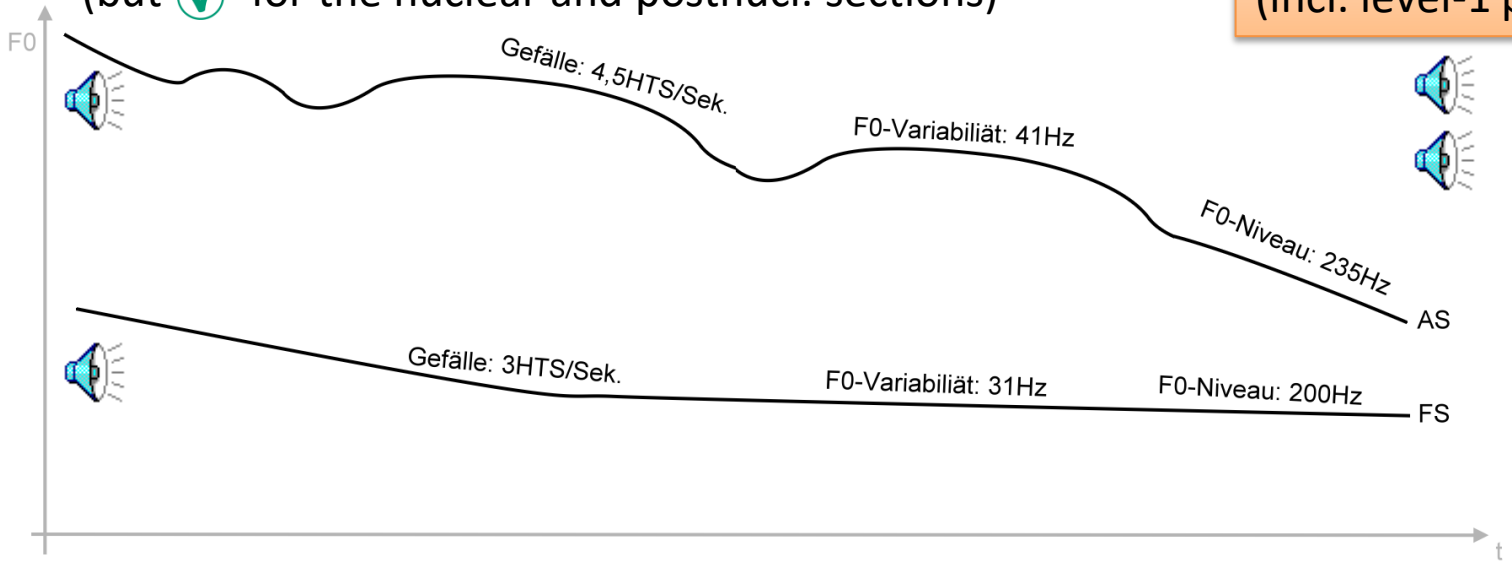
- Across all enacted conversations, each target sentence occurred once as a statement and once as a question
- 8 pairs of (female) German native speakers, head-mounted microphones

- 96 statements and 96 questions were available for acoustic-prosodic analysis
- Measurements were made manually with PRAAT or Wavesurfer (H1-H2 only in /a/ vowels)

Prosodic characteristics of questions

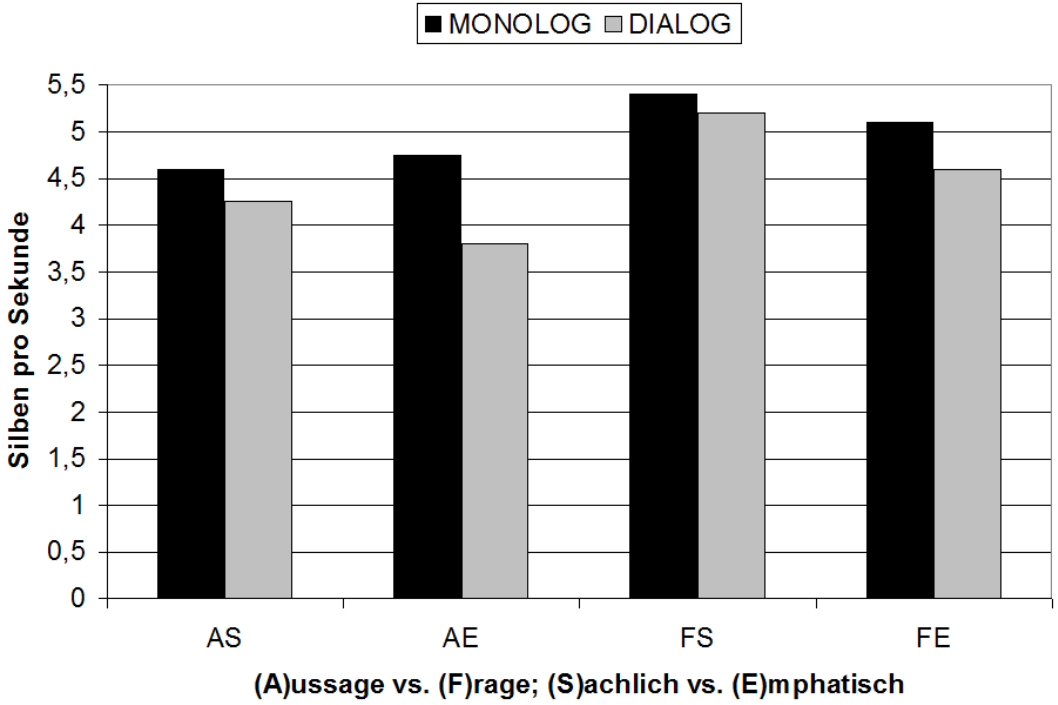
- 100% of statements had a final fall ↔ 39% of questions had a fall and 61% had a final rise
- F0 measurements:
  - In the prenuclear region: Q = flatter F0 than in S? ✓
  - In the prenuclear region: Q = lower F0 than in S? ✗
    (but ✓ for the nuclear and postnucl. sections)

87 % fewer prenuclear accents in Q than in S (incl. level-1 promin.)



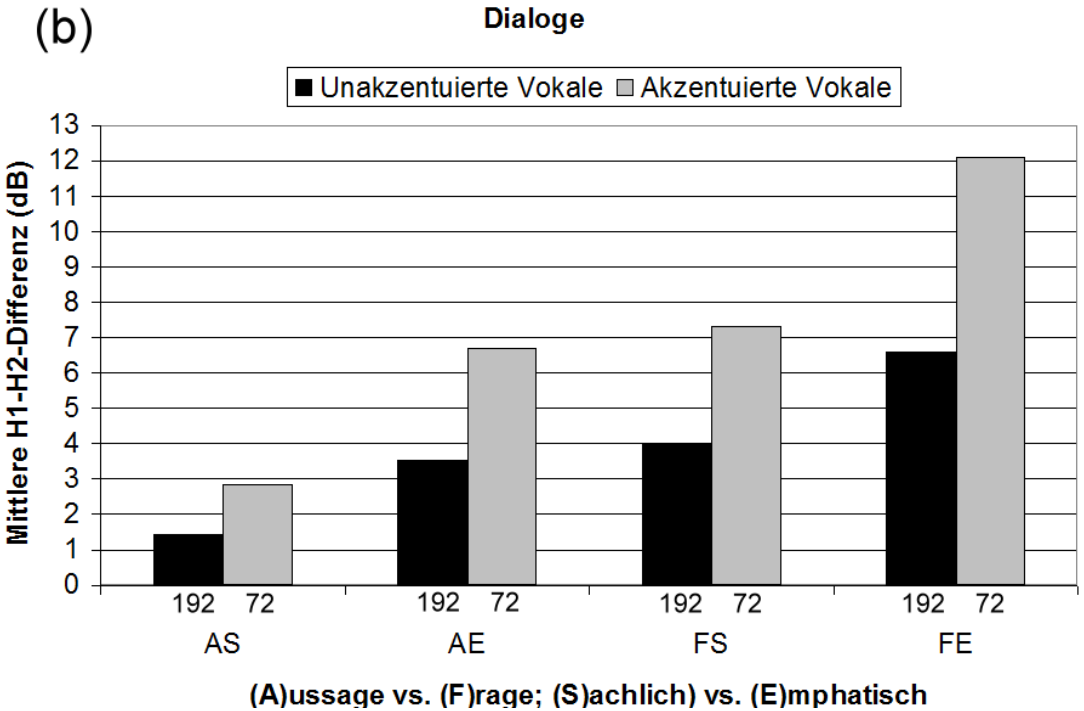
Prosodic characteristics of questions

- 100% of statements had a final fall ↔ 39% of questions had a fall and 61% had a final rise
- Speaking rate (syl/s):
  - Q faster than S ? ✓



Prosodic characteristics of questions

- 100% of statements had a final fall ↔ 39% of questions had a fall and 61% had a final rise
- Voice Quality (H1-H2):
  - Q breathier than S ? ✓



Prosodic characteristics of questions

- **What about the perceptual relevance of these findings?**
- Cross-check in a perception experiment with 18 native German listeners
- Original stimuli of the production study
  - *[Du wohnst in] [Mailand]*
  - *[Du wohnst in] [Nepal]*
  - *[Du arbeitest als] [Sänger]*
  - *[Du arbeitest als] [Schneider]*

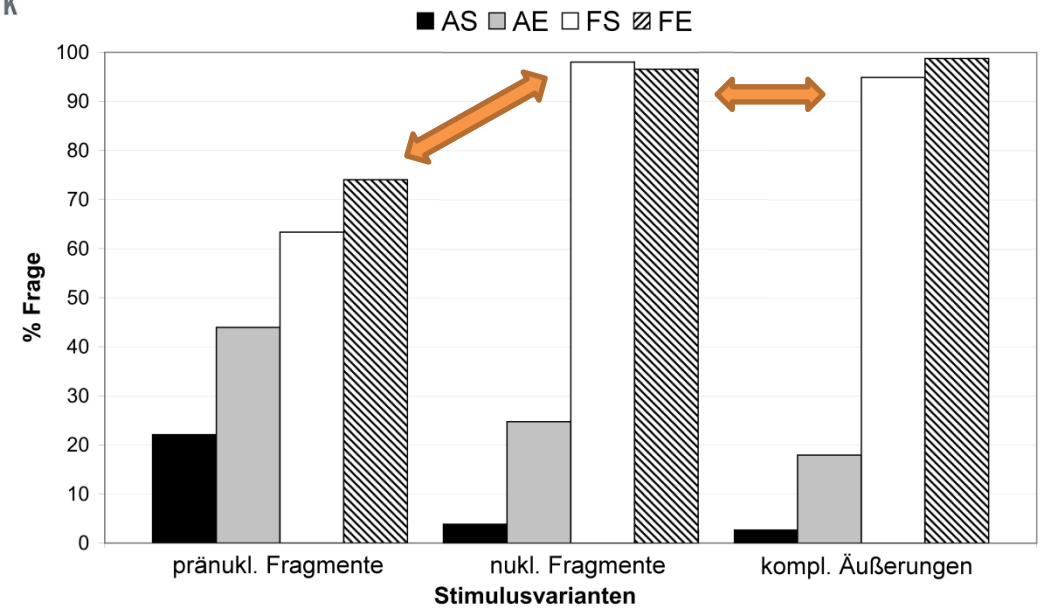
- **3 Stimulus conditions: (1) full sentence, (2) pre-nuclear only, (3) nuclear only**
- The 3 conditions were judged in separate experimental session (1 week pause in between)
- Each session = 96 stimuli
- The participants were split up into 6 sub-groups of 3 people → The 6 possible orders of the 3 conditions were distributed across the 6 subgroups
- Stimulus presentation orders within each condition were individually randomized

Prosodic characteristics of questions

- **What about the perceptual relevance of these findings?**
- Cross-check in a perception experiment with 18 native German listeners
- Original stimuli of the production study
  - *[Du wohnst in] [Mailand]*
  - *[Du wohnst in] [Nepal]*
  - *[Du arbeitest als] [Sänger]*
  - *[Du arbeitest als] [Schneider]*

- Hypothesis (1): the full-sentence stimuli are correctly identified as either Q or S
- Hypothesis (2): Q-S identification is worse in the pre-nuclear-only and nuclear-only conditions
- Hypothesis (3): Q-S identification is above chance level in the pre-nuclear-only and nuclear-only conditions

Prosodic characteristics of questions



- Hypothesis (1): the full-sentence stimuli are correctly identified as either Q or S
- Hypothesis (2): Q-S identification is worse in the pre-nuclear-only and nuclear-only conditions
- Hypothesis (3): Q-S identification is above chance level in the pre-nuclear-only and nuclear-only conditions

Region of Southern Denmark

Prosodic characteristics of questions

SDU MADS CLAUSEN INSTITUTE

- **3 assumptions**
  - (1) Intonation is the only relevant parameter that is involved in Q/S signaling ✗

    The signaling of Q vs S (in German and many other languages) involves a whole **bundle of prosodic features**

  - (2) The Q/S signals of intonation are located at the end of sentences/phrases ✗

    Prosodic cues to Q vs S are already present **early in the prenuclear section**

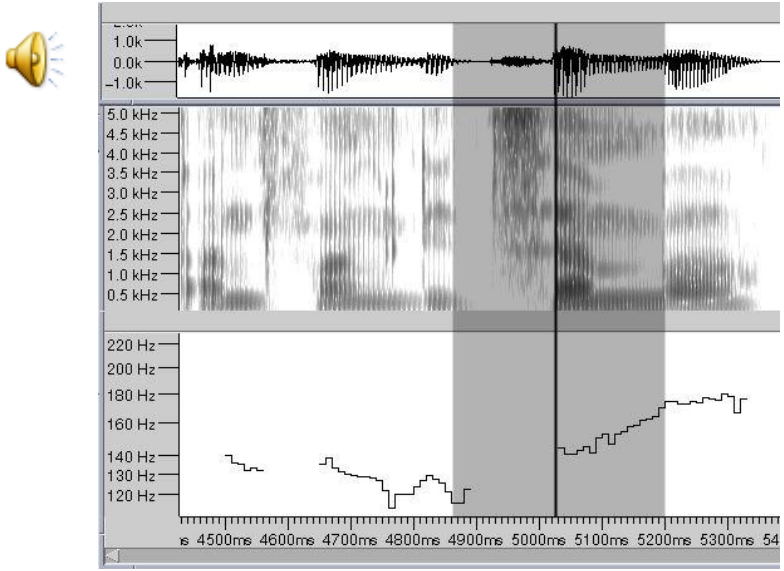
  - (3) The Q/S signals are represented by the difference between rise and fall ✗

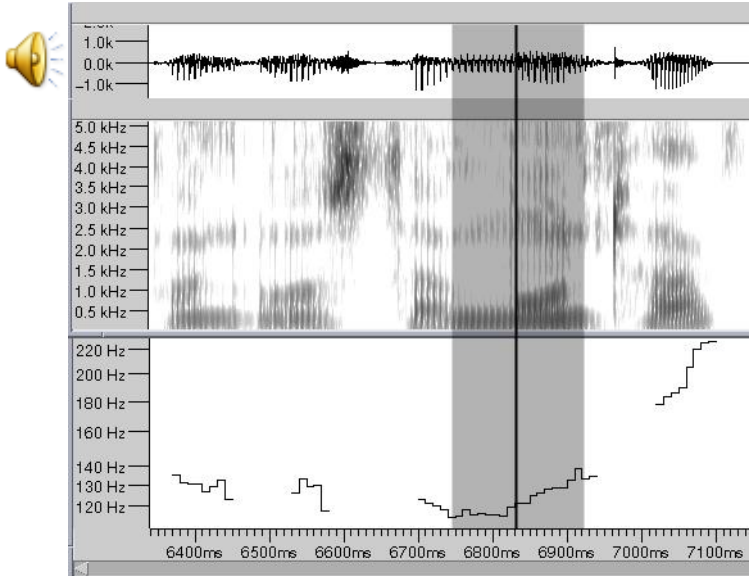
    As Q vs S signaling depends on other prosodic cues than the phrase-final intonation. The final rise vs fall difference has a **separate communicative function** that seems to apply across all types of questions, incl. morphosyntactically unmarked ones

! And wouldn't it make sense to let the interlocutor know as early as possible whether an utterance is a S or a Q? → Wh and VI are initial Q signals, so why should prosody be the only final Q signal?

On the role of movement curvature (1) in phrase-final position

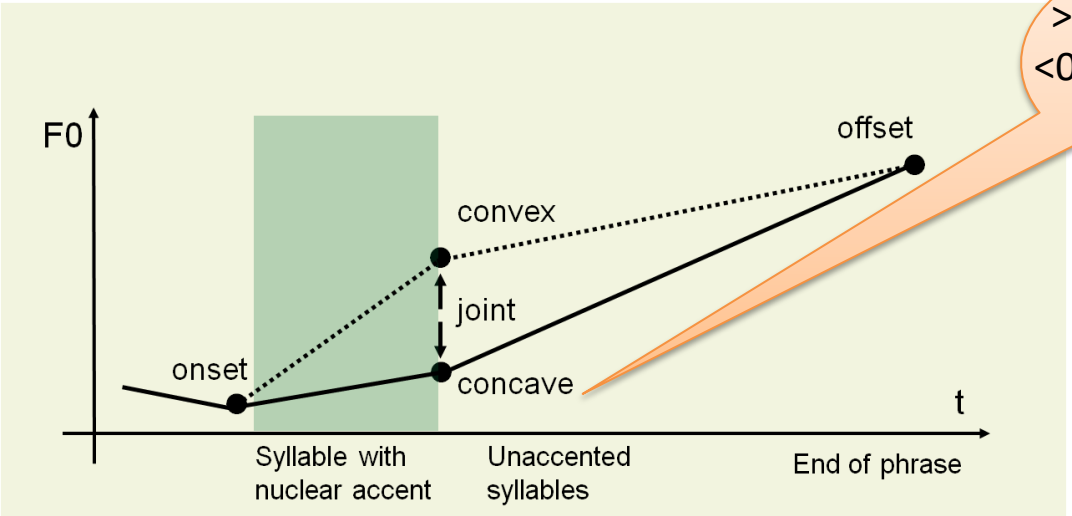
- Coincidental observation in the Kiel Corpus of Spontaneous Speech: Phrase-final rises in turn-yielding and turn-holding contexts can have the same F0 range and F0 valley-alignment…
- … but differ systematically in the **shape of the rise**!
  - Seen from above: convex = turn-holding ↔ concave = turn-yielding
  - Later also found for Estonian discourse markers by Asu (2006)



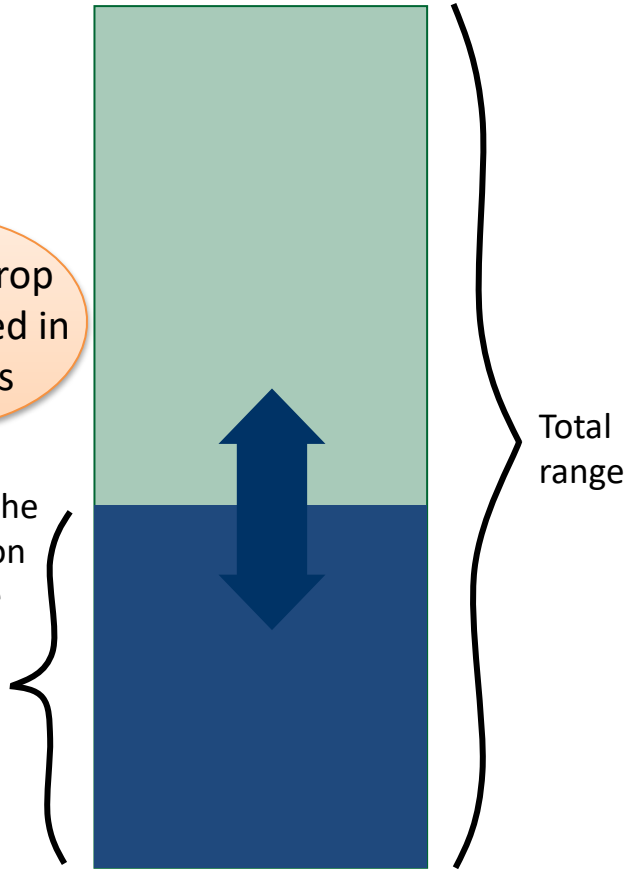


On the role of movement curvature (1) in phrase-final position

$$\frac{\text{Range in first section}}{\text{Total Range}} = \text{Range Proportion}$$
(rprop)



>50% of rprop <0.5 occurred in questions



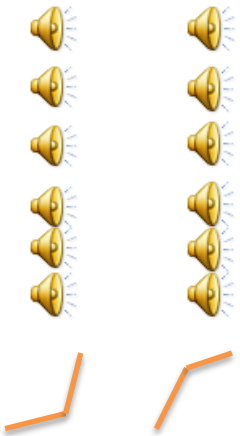
On the role of movement curvature (1) in phrase-final position

- So, can a rise with a concave shape under otherwise constant conditions be a question indicator? First step: Does a **concave rise sound more questioning** in VI-Qs?

- Perception experiment based on six syntactically marked questions:
- (1a) *Bist du im Urlaub?* (‘Are you on holiday?’)
- (1b) *Sind Sie der Eigentümer?* (‘Are you the owner?’)
- (2a) *Liegt das bei Lübeck?* (‘Is that near to Lübeck?’)
- (2b) *Liegt das in Niedersachsen?* (‘Is that in Lower Saxony?’)
- (3a) *Sind Sie Angela?* (‘Are you Angela?’)
- (3b) *Sind Sie Angelika?* (‘Are you Angelika?’)

rprop

0.2 vs 0.8



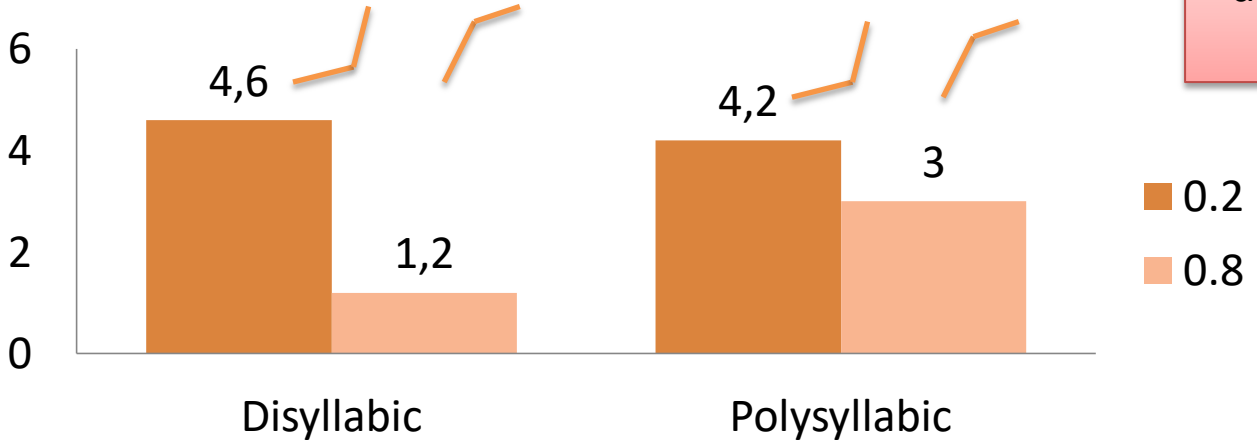
- Combination of two variables with 2 and 3 steps:
  - disyllabic vs. polysyllabic condition
  - lexical semantics of the questions (fact, place, name)

On the role of movement curvature (1) in phrase-final position

- So, can a rise with a concave shape under otherwise constant conditions be a question indicator? First step: Does a **concave rise sound more questioning** in VI-Qs?

- 31 native German listeners were asked to judge – on a 7-point scale – the question character of the stimuli (7 = most, 0 = least)
- 144 stimuli presented in an overall randomized order

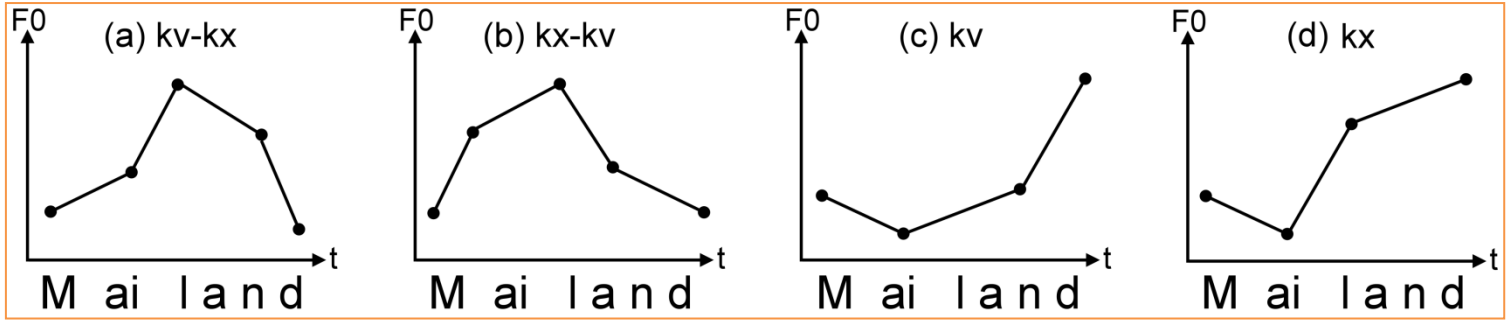
Yes, concave rises do sound more questioning



On the role of movement curvature (1) in phrase-final position

- So, can a rise with a concave shape under otherwise constant conditions be a question indicator? Second step: Can a **concave rise be a Q indicator** in declarative sentences?

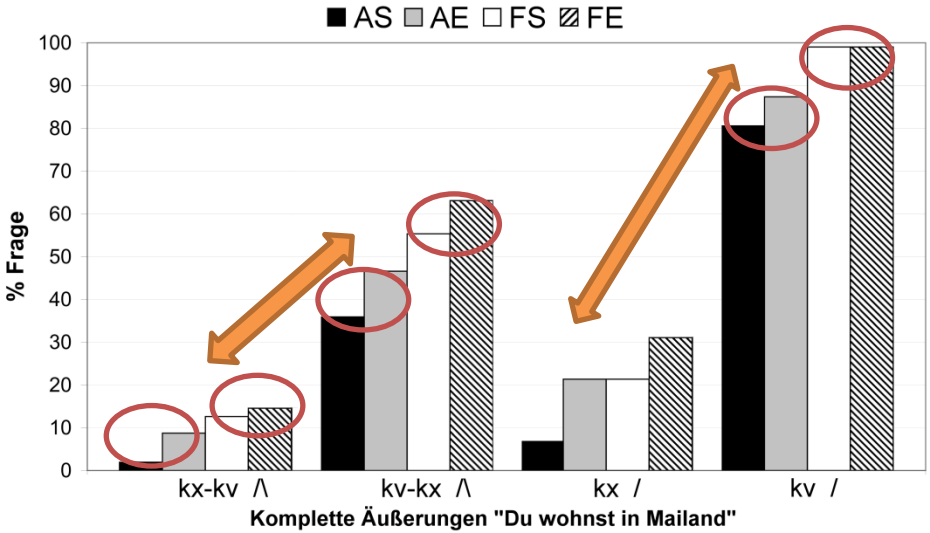
- Stimuli based on the sentence *"Du wohnst in Mailand"* from the initial production study
- 4 curvature conditions, resynthesized for original S and original Q productions

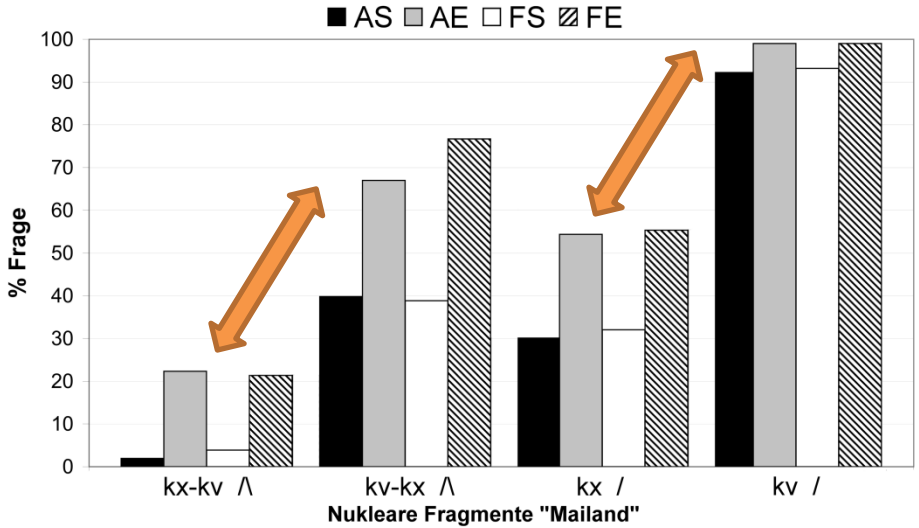


- 2 presentation conditions: Full sentences and nuclear fragments *"Mainland"*
- Again in separate sessions (separated by about 1 week)
- 17 native German listeners were asked to identify the stimuli as either Q or S

On the role of movement curvature (1) in phrase-final position

- So, can a rise with a concave shape under otherwise constant conditions be a question indicator? Second step: Can a **concave rise be a Q indicator** in declarative sentences?

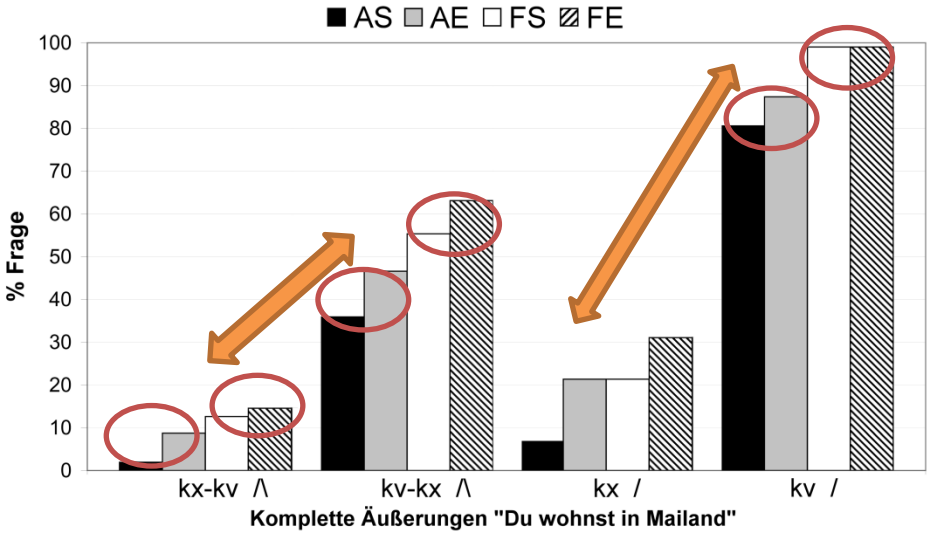


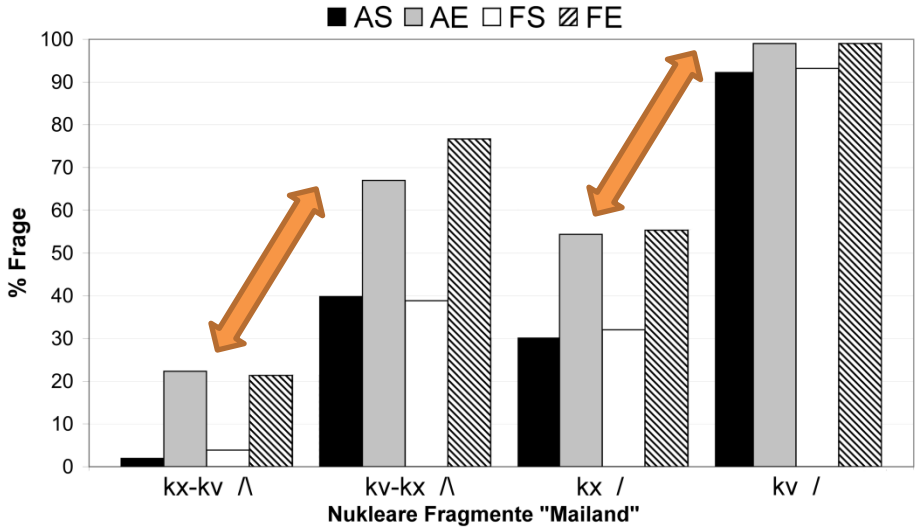


- (Elliptic) nuclear-only stimuli yielded overall more Q identifications than full decl. stimuli
- Yes, unlike convex final rises, **concave final rises are a Q indicator** in elliptic/decl. sentences
- There must have been further prosodic cues → lower Q identification for original S stimuli

On the role of movement curvature (1) in phrase-final position

- So, can a rise with a concave shape under otherwise constant conditions be a question indicator? Second step: Can a **concave rise be a Q indicator** in declarative sentences?

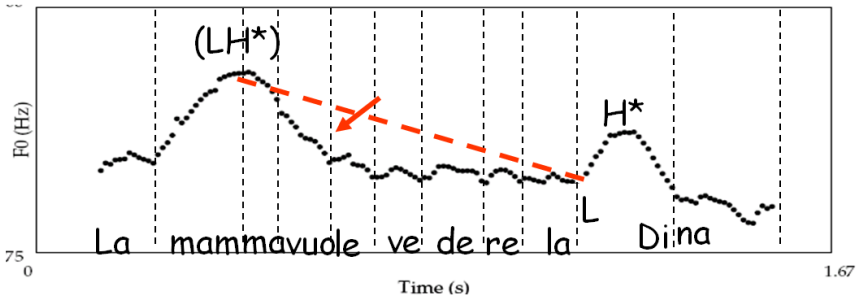


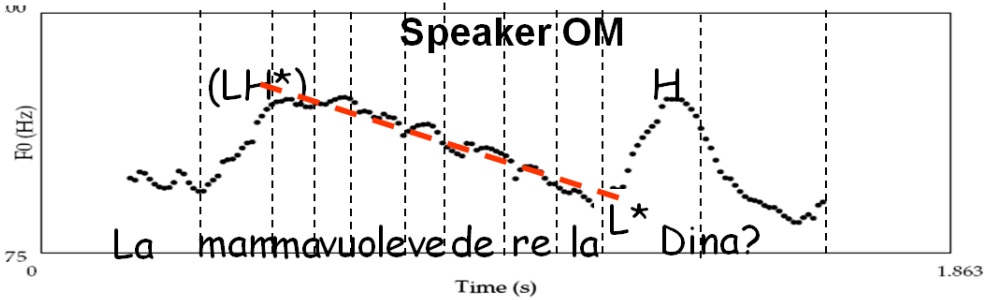


- **Concave rises are also Q signals in rising-falling nuclear peak contours**… in combination with a **convex** fall!

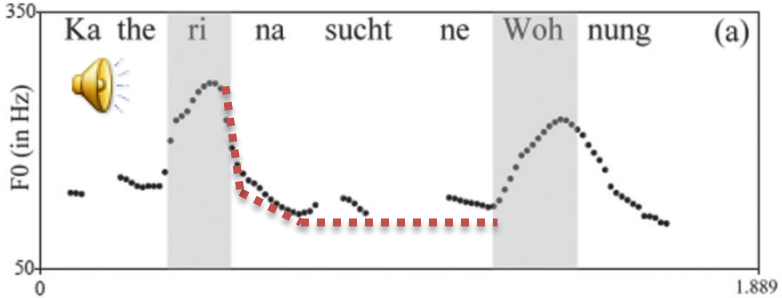
On the role of movement curvature (2) in prenuclear position

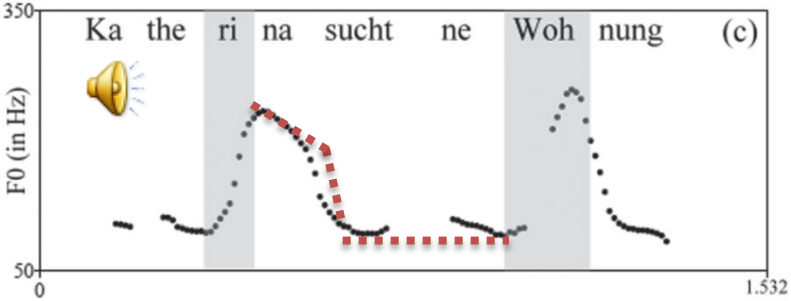
- **Convex falls can also be Q signals in/after prenuclear accents**
- Neapolitan Italian, Petrone (2008)

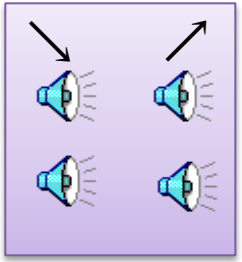




- And German? → Test case: sentences with declarative syntax



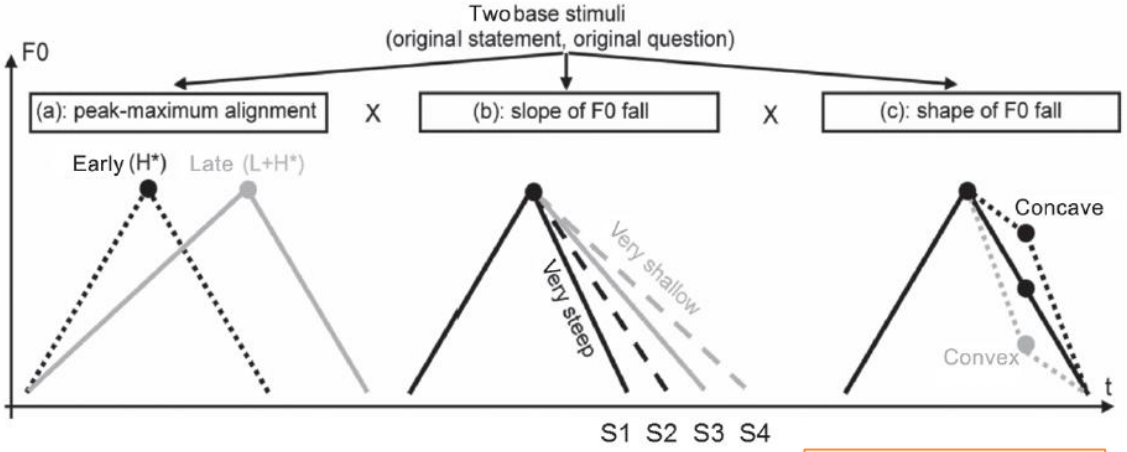




On the role of movement curvature (2) in prenuclear position

- Perception experiment with F0 manipulations only in the prenuclear section of *"Katharina sucht ne Wohnung"*

- The nuclear section was made constant for both original Q & S (intermediate contour)
- 33 native German listeners
- Indirect-identification (matching) experiment: context+stimulus



C1: *Tatsächlich? Das ist ein gewagter Schritt!* (Really? That's a risky step!);
C2: *Fragst du mich das wirklich? Das weiß doch inzwischen jeder.* (Do you really ask me this? Everybody knows it by now.).

Q = matching

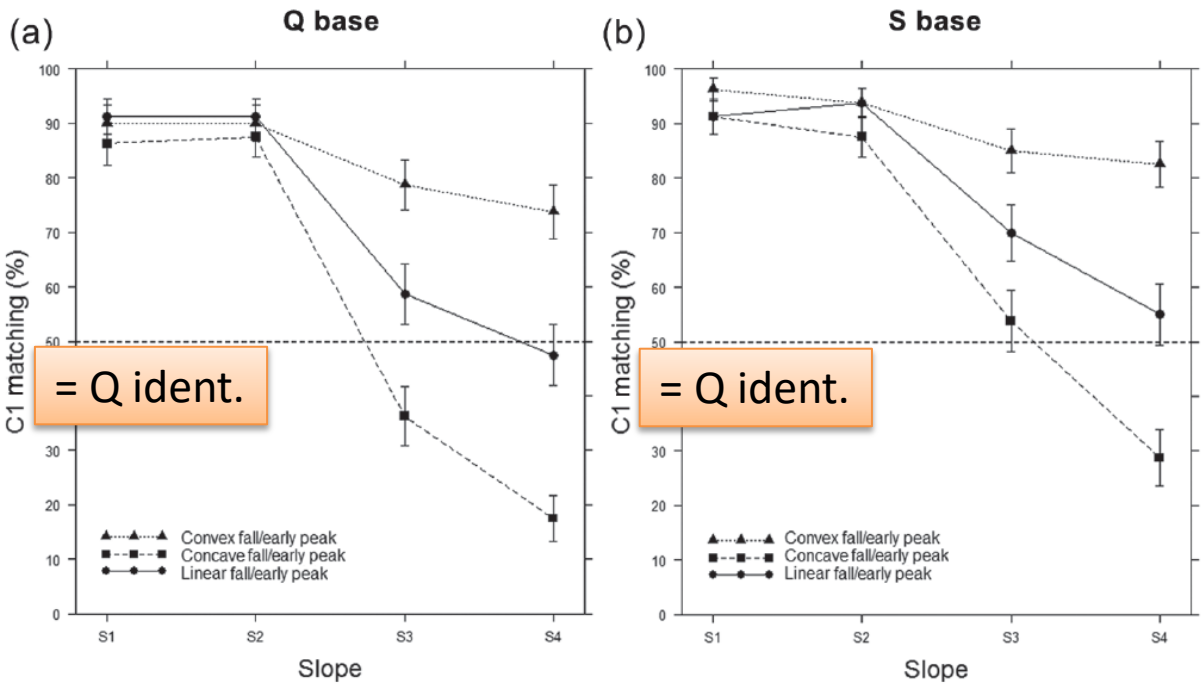
S = matching

- Between-subjects design: 16 sbs judged stimuli in C1 ctxt ↔ 17 sbs judged stimuli in C2 ctxt
- 24 stimuli per group, presented 10 times (=240 stimuli) in an overall randomized order

On the role of movement curvature (2) in prenuclear position

- Perception experiment with F0 manipulations only in the prenuclear section of *"Katharina sucht ne Wohnung"*

- Strong effect of curvature on Q identification in German (like in Neap. Italian)
- **convex fall = Q indicator**
- The effect is present for both orig Q and orig S stimuli
- **The shallower the falling slopes the stronger is the curvature effect**



Pauses after requests (VI questions)

- **Roberts and Francis (2013)**
  - A: “Can you give me a ride to the school? I need to pick up the flyers there.”

$\leq$ 600 ms

B: “Sure.”

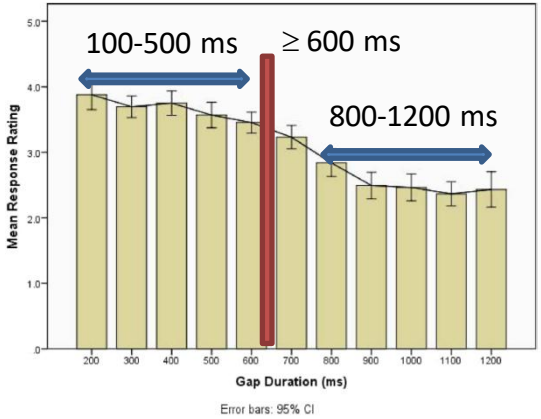
High perceived willingness of B
The affirmative answer is sincere

Pauses after requests (VI questions)

- **Roberts and Francis (2013)**
  - A: "Can you give me a ride to the school? I need to pick up the flyers there." → (> 600 ms) → B: "Sure."

Low perceived willingness of B
The affirmative answer is **not** sincere

- The **"tolerance threshold" of 600 ms** is related to the minimum time that it takes to respond to a new unexpected stimulus
- Speakers (questioners) are aware of this response time
- → Affirmative responses later than 600 ms are obviously not given spontaneously…
- Interlocutors had to think twice before they said "yes" = **reluctant**



Pauses after requests (VI questions)

- **Roberts and Francis (2013)**
  - A: "Can you give me a ride to the school? I need to pick up the flyers there."

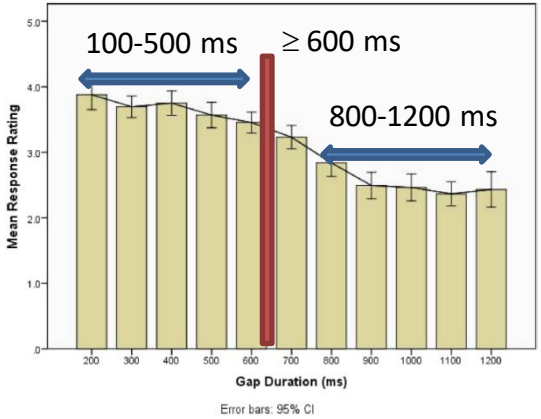
> 600 ms

B: "Sure."

Low perceived willingness of B
The affirmative answer is **not** sincere

Are these results replicable for German?

- What about context factors, esp. those related to ecological validity?
  - Speaking rate?
  - Inhalation/exhalation noises?
  - Speech in noise? (100% silent studio conditions are rare in natural every-day conversation)



Pauses after requests (VI questions)

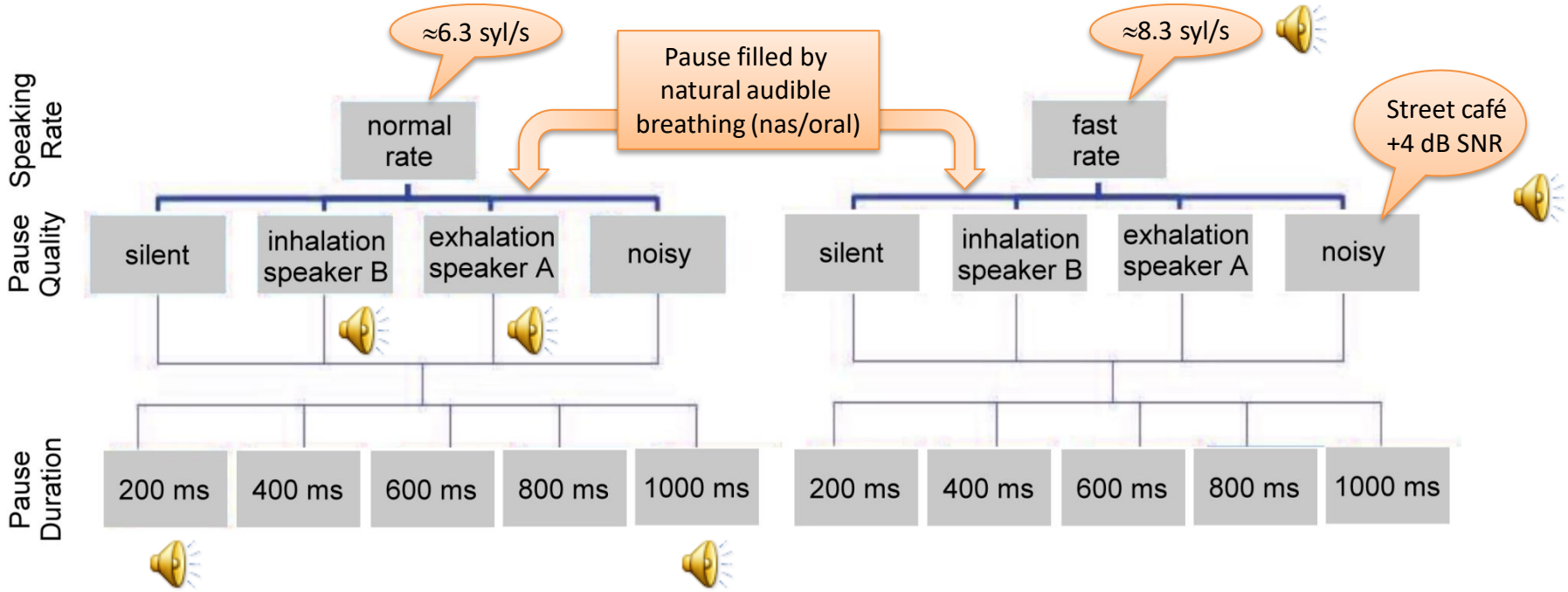
H* L*+H H-%

- A: **“Kannst Du mich nachher in die Uni fahren?”**
  (Can you take me to the university later on?)

H* L+H* L-%

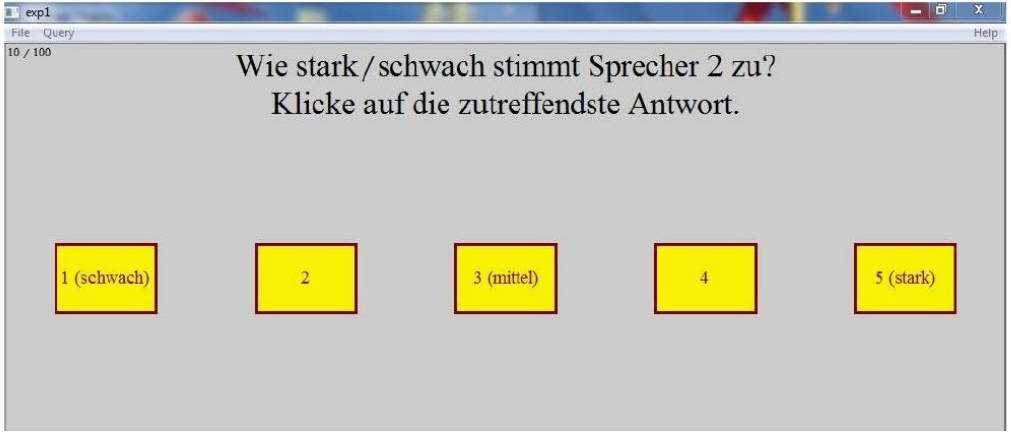
B: **“Ja, natürlich”**
(Yes, of course)

- 3 independent variables: **Rate, Quality, Duration** → 2 x 4 x 5 = 40 AB dialogue stimuli



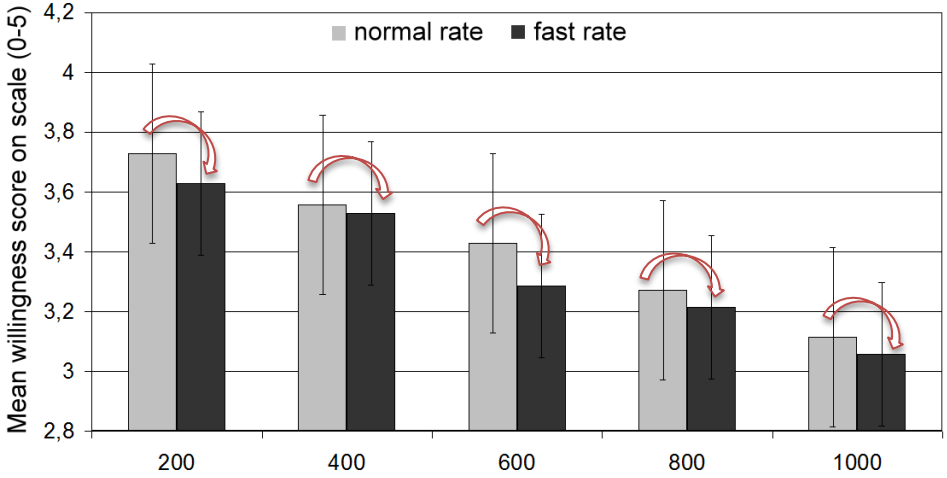
Pauses after requests (VI questions)

- **28 participants** (L1 German) took part in the perception experiment (ø 22 years old)
- Conducted with **PRAAT-MFC** on individual desktop PCs
- Stimuli presented **five times in randomized order** (different for each participant)
  = 200 stimuli



- **Instruction:** *"Read between the lines" and judge after each dialogue on a scale from 1 (weak) to 5 (strong) how willing B actually is to comply with A's request, i.e. how sincere B's "Ja, natürlich" (B) sounds to you.*

Pauses after requests (VI questions)



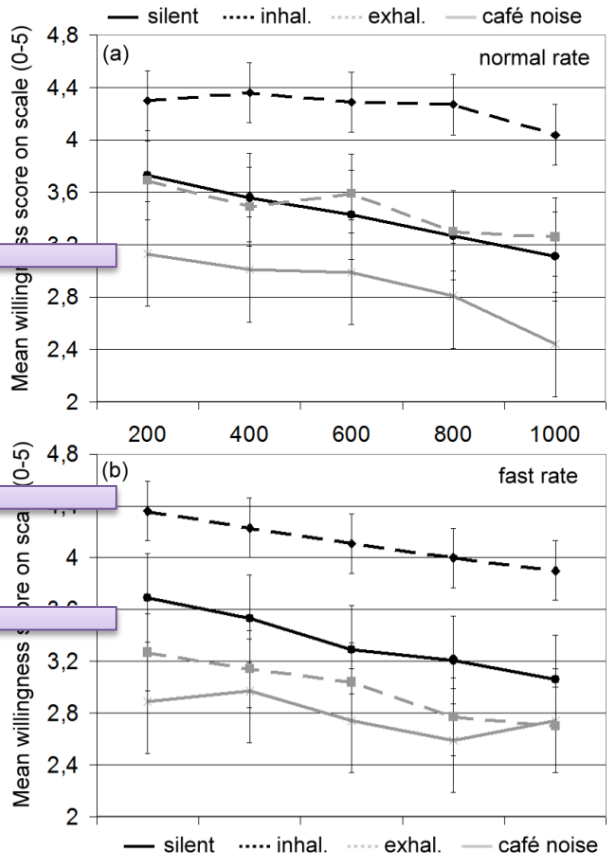
→ For each pause duration: Faster speaking rate = lower perceived willingness

Pauses after requests (VI questions)

Affirmative answers in AB dialogues overlaid with street café noise yielded lowest perceived willingness scores

Audible inhalation of B prior to her affirmative answer supports B's perceived willingness ↔ compared to audible exhalation of A after the request

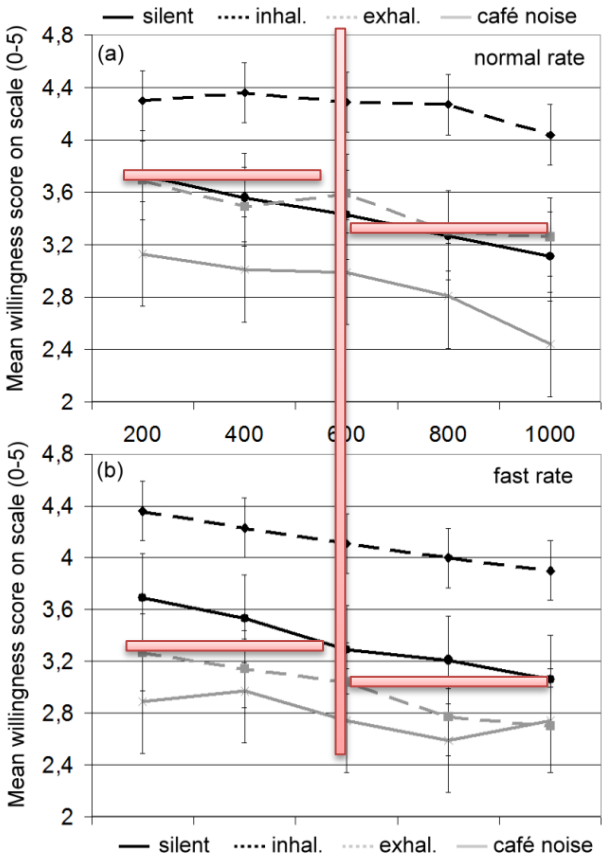
The silent pauses fall in between inhalation and café noise



Pauses after requests (VI questions)

The affirmative answer's perceived willingness decreases with increasing pause duration

The only significant decrease **across all three independent variables** is that **from 600 ms to 800 ms**



Pauses after requests (VI questions)

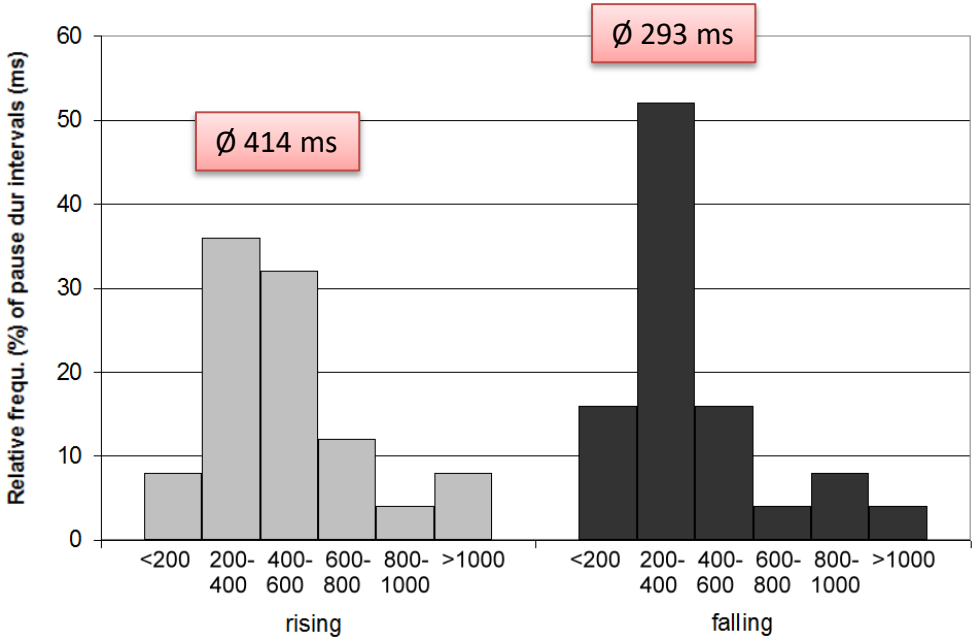
- Findings of Roberts and Francis (2013) **were replicated**
  - If a speaker waits too long to give an affirmative answer to a request, then the perceived willingness of this answer is negatively affected
  - **How long is too long? → Significant decrease across all tested context factors after 600 ms**

- However, **context factors play a role as well !**

Pauses after requests (VI questions)

- Further context factor to be tested: Direction of phrase-final intonation movement
- Rise/fall = the A allows B to give any answer (rise) or the answer that B expects to hear (fall)

- **Assumption: If A limits B's answer alternatives by means of a falling intonation, then A also expects B's answer to come faster**
- Initial supporting evidence from pause durations between y/n questions and subsequent answers in a German speech corpus (KIESEL)



Pauses after requests (VI questions)

- What about **wh-questions?** Pause durations after wh-questions play a role in many novels
  - "But the **answer came too fast**, and Tucker knew it was a lie." (Shadow of betrayal 2009:291)
  - ""Nothing," Eve said, but the **answer came too fast**. Morgan knew she was lying" (The Loner 2009: Ch.11,1)
  - ""It's not that." The **answer came too fast** to be believed. "Okay," Gabe said. (The witch who came in from he cold 2017)
  - ""What'd you do?" His **answer came too fast**. "Nothing." "Your lying skills could use some work…" (In safe hands 2014:66)

- → Suggests that pause duration has a communicative function in wh-questions as well → related to credibility or reliability
- **Unlike for y/n questions, shorter pauses after wh-questions are not better but worse for B !**

A: "What did you do yesterday evening?"
B: "I learned"

- We have conducted a perception experiment like the one presented today, but wh-question and answer
- Results support the form-function link implied in many novels. Unlike y/n questions, wh-questions have **both an upper** (> 1,000 ms) **and a lower "tolerance threshold"** (< 500 ms) !

Region of Southern Denmark

Summary

SDU MADS CLAUSEN INSTITUTE

How to ask without a question and to reply without an answer

If you wait for more than 600ms in VI or 1000ms in Wh or are faster than 500ms in Wh questions, then your answer is already given!

The right bundle of prosodic features (tempo, VQ, intonational variability/level, rhyhtm…) signals a question indicator even in declarative utterances

Region of Southern Denmark

Questions/Comments?

SDU
MADS CLAUSEN INSTITUTE