Pitch accent type affects stress perception in German: Evidence from infant and adult processing

Katharina Zahner
Fachbereichskolloquium Sprachwissenschaft, Doktorandenvortrag
03.05.18

Supervisors:
Prof. Dr. Bettina Braun (University of Konstanz)
Prof. Dr. Janet Grijzenhout (Leiden University)
The Metrical Segmentation Strategy (Cutler & Norris 1988)

Behave!

I am have!

Behave → am # have

Hau auf die # Leberwurst

“Hope of deliverance”

(see Cutler & Butterfield 1992 on slips of the ear; Kentner 2015 on mondegreens (Agathe Bauer Songs))
Can anyone blame us for such auditory illusions?

**Lexical statistics for some Germanic languages** (see Cutler 2012, p. 58)

- If we include secondary stress in lexical statistics:
  - 81% of the English lexicon is stress initial
  - 89% of the German lexicon is stress initial
From a processing perspective...

Lexical stress plays a role in...

- Stressed-based segmentation strategy very relevant in second half of the first year of life

“Stress is an underlying mental phenomenon, which however must be picked up by the hearer” (Hyman 1977, p. 40)
Stress-based segmentation in infants

**Metrical Segmentation Strategy** (Cutler & Norris 1988)
- German 9-month old infants have been shown to treat lexically stressed syllables as word onsets (Höhle 2002; Bartels et al. 2009; Männel & Friederici 2013)


<table>
<thead>
<tr>
<th>Der</th>
<th>Balken</th>
<th>lag</th>
<th>quer</th>
<th>über</th>
<th>dem</th>
<th>Abgrund</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>S</td>
<td>W</td>
<td>S</td>
<td>W</td>
<td>W</td>
<td>S</td>
</tr>
</tbody>
</table>

(Höhle 2002; Bartels et al. 2009)

<table>
<thead>
<tr>
<th>Your</th>
<th>kingdom</th>
<th>is</th>
<th>in</th>
<th>a</th>
<th>far</th>
<th>a</th>
<th>way</th>
<th>place</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>S</td>
<td>W</td>
<td>W</td>
<td>W</td>
<td>S</td>
<td>W</td>
<td>W</td>
<td>S</td>
</tr>
</tbody>
</table>

(Jusczyk et al. 1999)

<table>
<thead>
<tr>
<th>Your</th>
<th>guitar</th>
<th>is</th>
<th>in</th>
<th>the</th>
<th>stu</th>
<th>di</th>
<th>o</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>S</td>
<td>W</td>
<td>W</td>
<td>W</td>
<td>S</td>
<td>W</td>
<td>W</td>
</tr>
</tbody>
</table>

(Jusczyk et al. 1999)
From a processing perspective...

Lexical stress plays a role in...

- **Stressed-based segmentation strategy** very relevant in second half of the first year of life

**Speech segmentation**

**Spoken word recognition**

- Lexical stress constraints lexical access: reduction of competitors (Ok’tober vs. ‘Oktopus) (e.g., Friedrich et al. 2004; van Donselaar et al. 2005)
- Listeners pick up on stress cues immediately (Reinisch et al. 2010)

“Stress is an underlying mental phenomenon, which however must be picked up by the hearer” (Hyman 1977, p. 40)
Overall aim of the project

How does utterance level pitch accent type affect...

Speech segmentation

Spoken word recognition

... stress-based segmentation in German 9-month olds?

... lexical activation in German adults?

Why is the interplay between intonation and lexical stress interesting at all?
The interplay between lexical stress and pitch accent type

Lexical stress can acoustically be cued by
- Duration, intensity, spectral tilt, vowel quality, $f_0$
  (see Gordon & Roettger 2017 for an overview)

F0 is special...
- F0 is determined by intonation
  - In intonation languages used to signal pragmatic and paralinguistic functions, e.g., information status of referents (e.g., Kohler 1991b, Prieto 2015)
  - In AM Phonology*, pitch accents are associated with stressed syllables
    - Alignment of f0 peak with stressed syllable varies as a function of accent type

→ High f0 / $f_0$ peak is an unreliable cue to the position of lexical stress

* Autosegmental Metrical Phonology (e.g., Ladd 2008; Pierrehumbert 1980)
Overview of the talk

- **Background**
  - Use of stress in processing mechanisms
  - The interplay between lexical stress and intonation

- **Part 1: The effect of pitch accent type on stress perception in infants**
  - **Experiment 1:** Stress-based segmentation in a head-turn preference paradigm

- **Part 2: The effect of pitch accent type on stress perception in adults**
  - **Experiment 2:** Lexical activation in a visual-world eye-tracking paradigm

- **Part 3: Underlying mechanisms: Salience vs. Frequency**
  - **Experiment 3:** Putting the frequency account to test

- **Discussion**
  - Implications
Part 1: Effect of pitch accent type on stress perception in German infants
Experiment 1: Stress-based segmentation in a head-turn preference paradigm
(Zahner, Schönhuber & Braun 2016, JCL)

Aim
- Investigate the effect of pitch accent type on stress perception and consequently stress-based segmentation in German infants

Research question
- Does the position of the pitch peak affect German infants’ segmentation of embedded SW-units?

- Side note:
  - German infants can use probabilistic cues and other word boundary cues (e.g., Höhle & Weissenborn 2003)
  - We thus tested the segmentation of a trochaic part-word (SW from WSW)
Head-turn preference paradigm (Fernald 1985; Kemler Nelson et al. 1995)

Intuitive responses
- Paradigms for studying infant speech processing rely on indirect measures that track interest, e.g.,
  - head turns → looking times (Fernald 1985; Kemler Nelson et al. 1995)

Dependent variable:
Looking time to side lights mounted at speakers that play stimuli on each trial

HPP Setup, figure taken from BSL Website, M. Schönhuber
Head-turn preference paradigm (Fernald 1985; Kemler Nelson et al. 1995)

Principle: Familiarization and recognition

Familiarization phase

- 45 seconds -

→ infants hear 2 passages (Lagune, Kasino)

Test phase

4 word lists (half familiar, half novel)


Interpretation:
Difference in looking times to familiar and novel test items

→ Segmentation!
Experiment 1: Materials

Familiarization stimuli
- Four low frequent WSW carrier words embedded in sentences
  - Kanone [ka.'noː.nə] ‘cannon’
  - Lagune [la.'guː.nə] ‘lagoon’
  - Kasino [ka.'siː.no] ‘casino’
  - Tirade [ti.'ʁaː.də] ‘tirade’


→ Targets in sentences recorded in 3 intonation conditions
Experiment 1: Materials

Example utterance in three intonation conditions

<table>
<thead>
<tr>
<th>Time (s)</th>
<th>0</th>
<th>1.686</th>
</tr>
</thead>
<tbody>
<tr>
<td>Die</td>
<td></td>
<td></td>
</tr>
<tr>
<td>La</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ne</td>
<td></td>
<td></td>
</tr>
<tr>
<td>war</td>
<td></td>
<td></td>
</tr>
<tr>
<td>trauomhaft</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WSW-CARRIER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>was</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wonderful</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Natural productions!
- WSW-Carriers closely matched for pitch range, syllable duration, vocal effort in syllables across conditions
Experiment 1: Materials

Familiarization stimuli
- Four low frequent WSW carrier words embedded in sentences
  - Kanone [ka.'noː.nə] 'cannon'
  - Lagune [la.'guː.nə] 'lagoon'
  - Kasino [ka.'siː.no] 'casino'
  - Tirade [ti.'ʁaː.də] 'tirade'

Test stimuli
- SW-part of WSW-carrier words
  - none ['noː.nə] taken from Kanone
  - gune ['guː.nə] taken from Lagune
  - sino ['siː.no] taken from Kasino
  - rade ['ʁaː.də] taken from Tirade

 12 trials (half novel and half familiarized part-words)

- Word list: 15 tokens of one SW words with falling intonation (gune, gune...)

Experiment 1: Materials

Participants
- 54 German 9-month-old infants
  - 25 female, 29 male
  - Average age: 0;9.1; range: 0;8.19 - 0;9.16

- No difference in gender or age across groups

Pitch accent type affects stress perception - Katharina Zahner - 03.05.18
Universität Konstanz
Experiment 1: Results

Average looking times to familiar and novel test lists

ANOVA: intonation condition x familiarity status
significant interaction:
F(2,51) = 3.53, p = 0.04
Experiment 1: Discussion

Looking time difference in medial-peak condition
- Extraction of *gune* (SW) from *Lagune* (WSW) only when pitch peak and metrical stress coincide

Interpretation
- Only high-pitched stressed syllables perceived as stressed and thus taken as word onsets
- *But...* could it also be the alternation of tonal targets of opposite height (i.e., the deviation) that triggers the percept of stress?
  (see Bolinger 1958, p. 112; “wide departure”)
Experiment 1b: Testing the alternative explanation

Looking time difference in medial-peak condition
- Extraction of *gune* (SW) from *Lagune* (WSW) only when pitch peak and metrical stress coincide

Peak-stress-alignment (medial-peak)  

− L – H – L

Peak-stress-misalignment (early-peak)  

X

− H – L – L

Peak-stress-misalignment (late-peak)  

X

− L – L – H

→ Flipping of intonation contour

Alternation of tonal events does not lead to the perception of stress and the extraction of a trochaic unit → High pitch is a necessary cue to stress
Part 2: Effect of pitch accent type on stress perception in German adults
Experiment 2: Lexical activation in a visual-world eye-tracking paradigm

(Zahner, Egger & Braun, Submitted, JPho)

**Aim**
- Investigate the effect of pitch accent type on stress perception and consequently on lexical activation in German **adults**

**What we know from an offline stress judgement study...** (Egger 2015)

![Waveform example](image)

- **Task:** Identification of stressed syllable
  - More errors when peak and stress do not co-occur
  - Error patterns show a strong **bias towards syllable with peak**
Experiment 2: Lexical activation in a visual-world eye-tracking paradigm
(Zahner, Egger & Braun, Submitted, JPho)

Aim
- Investigate the effect of pitch accent type on stress perception and consequently on lexical activation in German adults

Research question
- Do \( f_0 \) peaks on unstressed initial syllables (as in early-peak accents, \( H+L^* \)) lead to the temporary activation of competitors with initial stress?

Task
- Click on a word mentioned in an instruction, while eye-movements are monitored
  (Tanenhaus et al. 1995; McQueen & Viebahn 2007; Reinisch et al. 2010)
Experiment 2: Materials

Eye-tracking: Visual-world paradigm with 4 printed words on screen
(Tanenhaus et al. 1995; McQueen & Viebahn 2007; Reinisch et al. 2010)

Materials
- 64 trisyllabic cohort pairs (WSW vs. SWW)
  - segmentally identical until to at least onset consonant of second syllable

Examples:

<table>
<thead>
<tr>
<th>Target (WSW)</th>
<th>Stress Competitor (SWW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Libelle [liˈbɛlə]</td>
<td>Libero ['libəro]</td>
</tr>
<tr>
<td>Kaverne [kaˈveɐnə]</td>
<td>Kaviar ['kaviəɾ]</td>
</tr>
<tr>
<td>Albaner [alˈbaɾə]</td>
<td>Albatros ['albatɾɔs]</td>
</tr>
</tbody>
</table>

→ Matched in lexical frequency and number of characters
(Dahan & Gaskell 2007; Lavidor et al. 2001; New et al. 2006)
Experiment 2: Materials

Recordings
- Targets recorded in two intonation conditions

- Syllable durations and range of f0-movement matched across conditions (early vs. medial- peak accent)

- Targets were cross-spliced into carrier sentence “Bitte klicke # Libelle an” (‘Please click on TARGET’)

- Targets were PSOLA-resynthesized to isolate the effect of f0 → see “natural parallelism”
  i.e., syllable on which the peak is realized louder and longer
  (Kohler 1991a, Niebuhr 2007)
Experiment 2: Design

Visual-world Paradigm with 4 printed words on screen
(Tanenhaus et al. 1995; McQueen & Viebahn 2007; Reinisch et al. 2010)

64 trials
48 participants

"Bitte klicke Libelle an ".

575ms
Experiment 2: Results

Evolution of fixations to four words on screen (two intonation conditions)

Pitch accent type affects stress perception - Katharina Zahner - 03.05.18

Universität Konstanz
Experiment 2: Competitor fixations

Competitor fixations

Time from stimuli onset (auditory target WSW, e.g., Libelle)

Fixation proportion

Bitte  klicke  LIBELLE  an

Intonation condition
- Early–peak
- Medial–peak

Pitch accent type affects stress perception - Katharina Zahner - 03.05.18
Experiment 2: Competitor fixations

Competitor fixations

Window: 875-975ms
Effect of intonation condition on competitor fixations
($\beta = 0.4 \ [0.05; 0.65], \ SE = 0.15, \ df = 732, \ t = 2.31, \ p = 0.02$)
Model: $\text{lmer}(\text{elog} \sim \text{condition} + \text{control predictors} + (1|\text{item}) + (1|\text{subject}), \ data = \text{window2})$

Interpretation
- More fixations to competitor for H+L* (early), compared to L+H* (medial peak)
- F0 peak in the form of a H-leading tone on unstressed syllable (H+L*) prompts percept of stressed syllable

Pitch accent types immediately affects lexical activation in German adults
Interim Summary of Parts 1+2

Pitch accent type affects...

- ... stress-based segmentation in German infants
  - Extraction of embedded trochee in medial-peak condition, but not in misalignment (early-peak and late-peak condition) and in flipped condition
  - Only high-pitched stressed syllables interpreted as stressed

- lexical activation in German adults
  - More looks to stress competitor when WSW-target is realized with an early-peak accent compared to a medial-peak accent
  - High-pitched unstressed syllables are interpreted as stressed

F0 peak guides the perception of lexical stress in both German infants and adults
Part 3: Underlying mechanisms: Salience vs. Frequency
Underlying mechanisms

Two mechanisms that can account for f0-stress interference

Salience-Account

“bottom-up” salience
“a stimulus that stands out from a perceptual ground” (Zarcone et al. 2016, p. 6)

(see Awh et al. 2012; Blumenthal-Drame et al. 2017)

→ These two mechanisms have been shown to be beneficial for processing
   (for recent reviews, e.g., Behrens & Pfänder 2016; Diessel 2007; Ellis 2002; Zarcone et al. 2016)

Frequency-Account
Underlying mechanisms

Two mechanisms that can account for f0-stress interference

**Salience-Account**

- H* accents and rising accents perceived most prominent  
  (Baumann & Röhr 2015)

- F0 is best predictor for drumming velocity when listeners imitate syllable prominences  
  (Wagner et al. 2016)

- Use of f0 for linguistic grouping  
  H L H L H L → (SW) (SW) (SW) ...  
  (e.g., Bion et al. 2011)

**Frequency-Account**

- H* accents are most frequent in German adult-directed and infant directed speech

<table>
<thead>
<tr>
<th></th>
<th>ADS*</th>
<th>IDS*</th>
</tr>
</thead>
<tbody>
<tr>
<td>medial peak</td>
<td>56-60%</td>
<td>59%</td>
</tr>
<tr>
<td>early peak</td>
<td>13-34%</td>
<td>8%</td>
</tr>
<tr>
<td>late peak</td>
<td>11-28%</td>
<td>26%</td>
</tr>
</tbody>
</table>

*KIEL Corpus*: appointment-scheduling dialogues  
(Peters et al. 2005)

*Konstanz prosodically annotated (K)IDS Corpus*:  
16 mother-infant dyads (infants < 1 year)  
(Zahner, Schönhuber, Grijzenhout & Braun 2016)  
http://ling.uni-konstanz.de/pages/home/braun/KIDS.html
Underlying mechanisms

Two mechanisms that can account for f0-stress interference

- Salience-Account
- Frequency-Account

How can we address / test these two mechanisms?

- Can inherent salience be manipulated at all?

- Manipulation of distribution frequency of H* in the immediate input (see e.g., accent adaptation studies)

- Other languages in which stressed syllables are low-pitched, e.g., Indian English (Pickering & Wilthshire 2000)
Experiment 3: Putting the frequency account to test...

1) Exposure phase (~3 mins)

- alternative questions (L*+H H+L* L-%)
- contrastive topics (L*+H H+L* L-%)
- polar questions (L* H-H%)

2) Test phase ( = Exp. 2)

Frequency account predicts...
- ... lower or no competitor activation for early-peak accents when L* accents are more frequent in the immediate input

Bitte klicke Libelle an.
Experiment 3: Putting the frequency account to test...

Effect of intonation condition on competitor fixations?

**Effect** of intonation condition: Window 875-975ms ($p = 0.02$)

**No effect** of intonation condition: All 100ms windows ($p > 0.3$)

**Interaction** between condition x exposure (window 875-975): $p = 0.16$

The amount of exposure to high-pitched stressed syllables is one of the mechanisms that explain the f0-stress-interference!

Yet: the salience also plays a role - it cannot be fully overridden by exposure
Discussion: Implications

F0 as a cue to stress?
- Different pitch accent types render f0 an unreliable cue to stress (see Intro)
- For the infant and adult listener, however, its perceptual salience and the high occurrence frequency of high-pitched stressed syllables promote f0 to a strong predictor for stress (Experiments 1-3)

Nature of f0 processing
- There are two theoretical options of f0 processing as a stress cue
  - direct path
    - high f0 \(\rightarrow\) lexical stress
  - mediated through phonological accent categories
    - high f0 \(\rightarrow\) H* \(\rightarrow\) lexical stress

(in case of a direct path, we would not have seen an effect of the frequency manipulation in Experiment 3)
Discussion: Implications

Infant speech processing from a developmental perspective
- Co-use of strategies to solve word segmentation problem
  - Hase? vs. Hase!
- Future work: Unravel development of cue weighting in stress perception
  - see role of intonation in
    - word recognition in 2.5- to 5-year-old English children
      (Quam & Swingley 2014; Fikkert & Chen 2011; Song et al. 2010)
    - prosodic phrase boundary perception in German pre-schoolers
      (Männel & Friederici 2016)

Adult speech processing – beyond the lexical level
- Pitch accent type may also affect processing at the phrase level
  - Changes in the interpretation of focus structure
  - He flew to Paris.  
    H+L*  L-%
  - He flew to Paris.  
    H*  L-%
Thank you for your attention!

Special thanks to...
- Bettina Braun
- and Janet Grijzenhout for their great support
- Muna Schönhuber and Sophie Egger for collaboration, discussion and support
- Clara Huttenlauch and Jana Neitsch for recording the stimuli
- Achim Kleinmann for technical support
- Monika for sharing the office with me
- my “peer group” (I cannot name all of you here) and all student assistants
- the Cusanuswerk for funding
- All adult participants, but especially all infants and their parents for coming all the way from the city (and beyond) to the babylab
A brief summary...

Pitch accent type affects...

- ... stress-based segmentation in German infants (Experiment 1)
  - Extraction of embedded trochee in medial-peak condition, but not in misalignment (early-peak and late-peak condition) and in flipped condition
  - Only high-pitched stressed syllables interpreted as stressed

- ... lexical activation in German adults (Experiment 2)
  - More looks to stress competitor when WSW-target is realized with an early-peak accent compared to a medial-peak accent
  - High-pitched unstressed syllables are interpreted as stressed

Why does the f0 peak matter so much in stress perception? (Experiment 3)

Salience-Account ✔

Frequency-Account ✔
References


References (continued)


References (continued)

References (continued)


