



The effect of global melodic rhythms and pitch-accent patterns on local accent positions in German

1. Introduction

- Accented syllables are perceptually standing out, i.e. they are marked by a higher prominence than the surrounding ones.
- Basic research question:** how is this perceptual prominence created in the hearer?
- What we know:
 - (a) direct phonetic cues: local changes (esp. increases) in intensity, duration, and particularly F0 affect prominence.
 - (b) however, prominence is a perceptual construction that also involves (language-specific) top-down knowledge and expectation:
 - prominences of pitch-accents compensate for F0 declination
 - native listeners include linguistic structures and semantics in their prominence judgements of sentences and hence deviate from non-native subjects.
 - verbs evoke less prominence than adjectives and nouns.
 - visual cues (head/eyebrow movements) influence prominence.

-Rhythmic effects on local prominences and hence accent positions have not been addressed so far, although prominence is the basic constituent of rhythm and although rhythm is an important feature of speech.
 -The presented study on German investigates effects of melodic rhythm, which is created by regular patterns of pitch-accented and unaccented syllables/words. In addition, pitch-accent category is varied. The study focuses on German and uses F0 (the main prominence cue in German) as the phonetic variable.

2. Method

- 3 disyllables were used that allow for different positions of the abstract lexical stress and hence provide 3 pairs of possible target words.
- **(1) DURCH**schauen ('see through') vs. durch**SCHAUEN** ('understand')
- **(2) UM**fahren ('knock down') vs. um**FAHREN** ('drive around')
- **(3) UM**stellen ('move') vs. um**STELLEN** ('surround')

-Idea: Concrete sentence accents are bound to lexically stressed syllables => from the perceived accent positions, subjects infer the position of the lexical stress and hence the target word and its meaning.
 => easily accessible / clear judgements can be obtained on this basis.
 -One representative of each disyllable was produced (by on) in isolation with monotonous pitch and equal prominences on both syllables.

- Manipulation in praat. Aim: to create for each disyllable ambiguous and clear accent positions, according to own impression.
- 3 ambiguous accent positions/target-word identifications**
 - rising-falling F0 peaks were created that span the whole disyllable and that show maxima around the middle of the word, right after the vowel onset of the second syllable (span c. 700ms, F0 range c. 13st) => *F0 peak pattern for ambiguous accent position.*
 - Additional duration manipulations were necessary.
- Plus 6 (3*2) clear accent positions/target-word identifications**
 - Starting from the ambiguous ones, two peak-maxima shifts to the vowel offsets of the first and second syllable (= typical H* alignment in German). => *F0 patterns for clear accent positions.*

-The 9 resynthesized disyllables were integrated into **sentence frames** consisting of 7-9 mono or disyllabic words, produced by the same speaker with matching F0 range, voice quality, and speaking rate.

- 2 exp. variables for disyllables of target-word pairs (2) + (3)**
 - Position of target-word stimulus in the sentence (initial vs. final)
 - Global rhythmic pattern of sentence frame (iambic vs. trochaic); with H* in the strings of accented syllables of both rhythmic conditions
- 2 exp. variables for disyllables of target-word pair (1)**
 - Pitch-accent category (early/H+L* vs. late/L*+H) in the string of accented syllables of the sentence frame
 - Global rhythmic pattern of sentence frame (iambic vs. trochaic)
- 90 filler sentences** were added to the **26 test sentences.**
 => Experiment=Set of 116 sentences presented in 2 randomized blocks.

- Shadowing task was used to obtain indirect judgements.**
- Instruction: machine-generated speech melody needs to be improved. To this end, the natural human speech melody will be compared with the one of the machine for a number of test sentences. Subjects were asked to repeat immediately the perceived string of words in the presented 116 sentences using their own speech melody.
- For each subject, the complete session (c. 20 min.) was recorded.
- => meaning-directed subtle task, in which subjects had to focus on the whole sentences, not just on single target words. There was no direct descriptive or metalinguistic judgement of prominence, accent, or rhythm.
- 21 native German subjects;** 16 female, 5 male; ages between 22-52.
- Meaning-based labelling of the target words, done by the experimenter => determining the (re-)produced accent positions of the subjects.
- Additionally, reaction-time measurements (RTs) were taken => sentence offset to onset of subject's (re-)production of presented sentence.

3. Results and Discussion

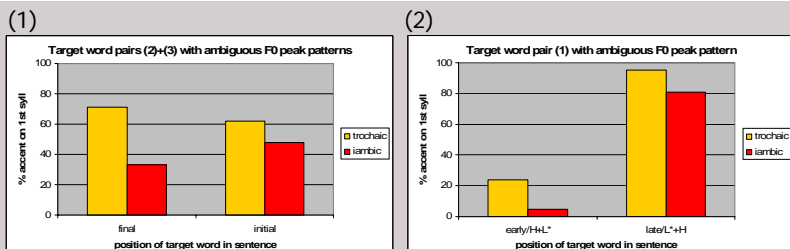


Fig. 1+2: Percentage of accent on the initial syllable = production of the first disyllabic target word in pairs 2+3 (Fig.1) and 1 (Fig.2). n=21

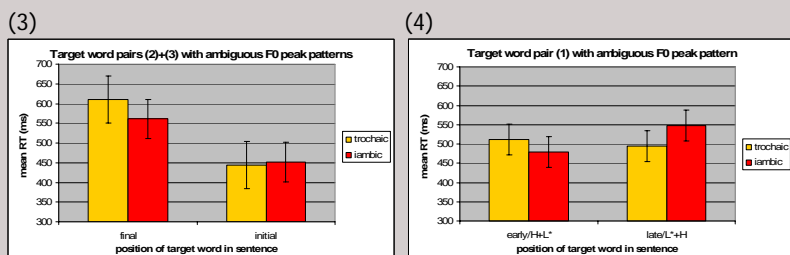


Fig. 3+4: mean RTs for the productions of test sentences of target-word pairs 2+3 (Fig.1) and 1 (Fig.2). Vertical lines show stand. deviations. n=21

- Based on Chi² tests, the findings may be summarized as follows:
 - The global melodic rhythm, represented by a preceding/following alternating string of accented/unaccented syllables, has a clear influence on the perceived position of local accents and hence on target-word identification.
 - For disyllables with ambiguous phonetic cues/F0 patterns, the perceived accent position is shifted so that the rhythmic pattern of the disyllable fits into the global melodic rhythm (e.g., trochaic=more accents on 1st syll.).
 - This effect is stronger for sentence-final disyllables (=progressive effect of preceding melodic rhythm). However, there is also a small, but significant effect for sentence-initial disyllables (=regressive effect of following melodic rhythm), see Figure 1.
 - The first finding replicates results of a previous study by Niebuhr (2008), *however with no obvious effects of musical training of the subjects.*
 - In addition of effects of global melodic rhythm, there is an even stronger effect of pitch-accent category on perceived accent position or target-word identification. The F0 patterns of the utterance-final disyllables of (1) tend to be interpreted as continuing the preceding pitch-accent categories. So, late peaks/L*+H shift the perceived accent on the first syllable/target word, while early peaks/H+L* shift the perceived accent to the second syllable/target word of pair (1), see Figure 2.
 - Different from the rhythmic effect, the effect of pitch-accent category is still observable for F0 patterns that were supposed to be a clear indicator of accent position.
 - => *one of the reasons for the empirical finding that pitch-accent categories in German IPs tend to be homogeneous?*
 - RTs were longer: for ambiguous than for clear F0 patterns, for final than for initial target-word position, and for conflicting than for matching combinations of global rhythm and local F0 peak position (cueing accent pos.).