PRODUCTION AND PERCEPTION OF CONTRAST IN GERMAN

Frank Kügler & Anja Gollrad

Department of Linguistics, University of Potsdam, Germany
kuegler@uni-potsdam.de; agollrad@googlemail.com

ABSTRACT

In German intonational phonology two different approaches to model nuclear falling accents exist (cf. [6, 8] vs. GToBI [11]). While [6] and [8] assume only left-headed pitch accents, the GToBI system allows for left- and right headedness. Consequently, in the former approach one simple falling accent (H*L) is represented in the tonal grammar while GToBI distinguishes between two falling accents, a simple falling and a rising-falling one (H* L- and L+H* L-). The present production and perception studies aim at providing phonetically based evidence in order to argue for one of the two phonological approaches of representing falling accents in German.

Keywords: falling accents, focus, contrast, tonal perception, German

1. INTRODUCTION

The concept of contrastiveness has received much attention in psycholinguistic research e.g., [4, 6, 13, 14, 18]. However, the phonetic and/or phonological nature of accents realized under a contrastive context compared to a neutral accentuation is still a matter of debate, i.e., there is a diversity of interpretations with respect to the phonological category falling accent types belong to in German.

According to GToBI there is a difference between two types of falling accents, which is accompanied by a difference in intonational function, i.e. a difference in focus structure [10]. According to [6], however, there is no tonal distinction between narrow and wide focus in German. Both [6] and [8] claim that in case of falling accents the pitch shape prior to the pitch peak is phonetic in nature, thus no further low leading tone is necessary to present falling accents. [8] proved this claim by means of a discrimination test showing that grouping of her data according to two different tonal categories failed. All studies provide poor evidence in favour of their claims: [6] argues phonologically not providing detailed phonetic evidence, [8] argues perceptually not referring to information structure, thus not considering the distinction assumed by GToBI, and [9] argue functionally without further phonetic evidence except for inter-transcriber consistency [12]. The present research is approaching the debatable issue by a combined production and perception study to examine how falling accents are represented in the tonal grammar of German.

2. EXPERIMENTS

In speech production the realization of wide vs. narrow focused sentences were compared investigating the phonetics of falling accents in nuclear position in German, and how the realization can be represented phonologically.

2.1. Production

2.1.1. Speech materials

The experimental sentences contain a SAuxOV word order with target words embedded as objects in non sentence final position in order to avoid any intonational phrase boundary effects. The following factors were manipulated:

- The number of accents in the sentence: one and two accents.
- The number of syllables of the target word varied between one (Wal ‘whale’), two (Roman ‘novel’) and three (Admiral ‘admiral’), all with ultima word stress.
- The length of the sentence: sentences were gradually lengthened by adding one of the two adverbials (gestern ‘yesterday’), and (glücklicherweise ‘luckily’) or a combination of both prior to the target word to increase the inter-accentual distance (between a sentence initial, and an accent on the target word. We expected that a larger inter-accentual distance would increase the chance that speakers realise two single peak accents instead of a hat pattern, which is a frequent pattern in German.

(1a) illustrates an example of a wide focus target sentence, (1b) illustrates a target sentence realized under contrastive focus. In both sentences, the target word is monosyllabic.

(1a) Erzähl mir bitte, was passiert ist.
‘Please tell me what happened.’
Martin hat den Wal gesehen.
‘Martin has seen the whale.’

(1b) Hat Martin den Frosch gesehen?
‘Has Martin seen the frog?’
Nein, Martin hat den Wal gesehen.
‘No, Martin has seen the whale.’

The experimental sentences are highly sonorant to allow for a maximally accurate f0 analysis. Sentences were interspersed with numerous fillers and fed into DMDX [5]. The experimental sentences were pseudo-randomized for each subject so that sentences of the same condition did not appear adjacently and corresponding sentences had a maximal distance.

2.1.2. Speakers

8 speakers participated in the experiment. All were female undergraduate students at the University in Potsdam. All were native speakers of standard German spoken in the Berlin-Brandenburg region and reported no speech or hearing impairment. They either received course credit or were paid for participation.

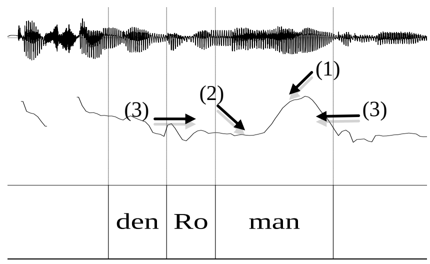
2.1.3. Recording procedure

For each sentence, a context question eliciting wide focus (1a) and contrastive focus (1b), spoken by a male voice, had been previously recorded. The contexts were presented together with a target sentence both visually on screen and auditorily over headphones. Speakers were asked to read and listen to the context and then to say the answer displayed on the screen as a response to the question. Subjects were familiarized with the task through written and verbal instructions. In case of hesitations or false starts, participants were asked to repeat the sentence. Recordings took place in a sound-proof chamber equipped with an AT4033a audio-technica studio microphone, using a C-Media Wave soundcard at a sampling rate of 44.1 kHz with 16 bit resolution. Presentation flow was controlled by the experimenter, and participants were allowed to take a break at any point.

A total of 486 target sentences had been recorded (discarding 11% due to creaky voice and mispronunciations). The resulting 432 sentences were submitted to an extensive perceptual inspection which revealed four distinct phonological contours of the target sentences realized by the speakers. The following patterns were categorized: (a) 299 non-downstepped nuclear falling accents, (b) 25 downstepped nuclear falling accents, (c) 36 hat patterns, and (d) 72 other types of nuclear accents, such as early falls. The phonetic examination of falling accents was based on group (a). Since GToBI assumes rising-falling accents we were particularly interested in contours that contain a nuclear peak accent, hence hat patterns were excluded as well as downstepped accents. The distribution over the broad- and contrastive focus cases of the 133 tonal accents, excluded from further analysis, was as follows: 79 broad focus cases and 54 contrastive focus cases (for similar figures see [11]).

The experimental sentences were hand-annotated and subjected to phonetic analysis using Praat software [3]. The following phonetic measurements were conducted: (1) the pitch peak of the target words in Hertz (Hz) and the corresponding time of the peak (t_h), (2) a low turning point in pitch prior to the peak in Hz (l) and the corresponding time of the low point (t_l), (3) the beginning and the end of the accented syllable, (4) the excursion between the low pitch point and the peak, (5) the velocity of the rise, and (6) the alignment of the low turning point in relation to the accented syllable.

Figure 1: Phonetic measurements of the target word.



Pitch analysis was conducted using a Hanning window of 0.4 seconds length with a default 10 ms analysis frame. The pitch contour was smoothed using the Praat [3] smoothing algorithm (frequency band 10 Hz) to diminish microprosodic perturbations.

2.1.4. Results

We fitted a multilevel model [2] using crossed random factors subject and item, and focus condition (wide focus (WF), contrastive focus (CF)) as fixed factor. The acoustic analysis revealed no significant difference between the focus conditions of the factor alignment of the low turning point. The analysis revealed only a marginal significant difference between the focus conditions for the f0 of the low turning point (WF: 188.07 Hz, CF: 182.45 Hz, t=2.73), for the f0 of the pitch peak (WF: 232.87 Hz, CF: 233.90 Hz, t=2.93) and for the alignment of the peak (WF: -0.055, CF: -0.065, t=2.08). The pitch excursion (WF: 44.79, CF: 51.45, t=-4.592) and the velocity of the rise (WF: 254.28, CF: 297.80, t=-5.19) yield significant differences between the two focus conditions.

The small differences between the two focus conditions with respect to (a) the alignment and (b) the Hz-values of the low turning point and the peak raises the question if and to what extend listeners would perceive a difference of the targets words as a function of focus.

2.2. Perception

A semantic congruency task [15] was conducted to investigate whether German listeners use the small phonetic differences of (a) the low target prior to the high tone and (b) of the high tone itself in the nuclear falling accent as a primary perceptual cue to distinguish accents in contrastive focus from accents in wide focus. Semantic congruity tests have been successfully used to explore the perception of intonation contrasts [16, 17]. The test allows us to evaluate the degree of perceived appropriateness of target intonation patterns to different pragmatic contexts.

2.2.1. Material

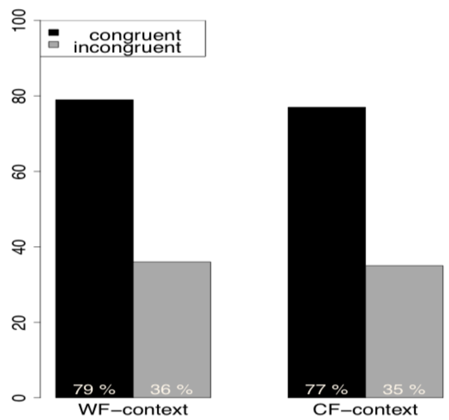
Stimulus materials for the perception experiment were taken from one of the speakers of the production study who produced the most prominent difference from the mean value of the low turning point in the two focus conditions. Context questions and target sentences were concatenated to one sound chain at a sampling frequency of 48000 Hz and were scaled at an intensity of 70db. The perception experiment consisted of 12 target sentences where intonation was coherent with the pragmatic context (6 WF-WF dialogs, 6 CF-CF dialogs) and 12 target sentences where intonation was not coherent with the pragmatic context (6 CF-WF dialogs, 6 WF-CF dialogs). Each dialog was presented 3 times which resulted in a total of 72 dialogs. The stimuli were auditorily presented over headphones with the MFC Praat software [3].

16 listeners were asked to listen to each dialog carefully and then evaluate whether they regard the intonation of the target sentence as “congruent” (by clicking on the “congruent box” visible on the screen) or as “incongruent” (by clicking on the “incongruent box”). After written and verbal instructions, a test run of 3 dialogs was carried out before the experiment started. The experiment lasted approximately 20 minutes.

2.2.2. Results

Figure 2 displays the rate of congruent responses to all dialog types, separated into WF-context (left bars) and CF-context (right bars). The results revealed that the appropriateness of the target intonation pattern to a context were rated higher for congruent (WF-WF and CF-CF) than for incongruent dialog types (CF-WF and WF-CF). Linear mixed-effects models [2] revealed that the difference of the rating response between the two conditions (congruous and incongruous) were significant for the WF-context dialogs (WF-context: t=11.75) as well as for the CF-context dialogs (CF-context: t=11.25). These results indicate that listeners are extremely sensitive to congruous and incongruous intonation of a context-target sequence.

Figure 2: Rate of congruous responses to all dialog types, separated by WF-context (right) and CF-context (left).



3. CONCLUSION

A production experiment had been carried out testing the influence of wide focus and contrastive focus contexts on the realization of nuclear fallings accents in German. The analysis revealed that the phonetic difference of the target words is only marginal with respect to (a) the low turning point prior to the peak, (b) the f0-value of the peak itself and (c) the alignment of the peak in both focus conditions.

The fact that contrastive focus raises nuclear H* accents in German is well-known, e.g. [7]. If there are two distinct falling accents in German [9], we would expect a difference in phonetic realization regarding the low turning point prior to the high accent. This is indeed the case for scaling. The low turning point is realized lower in contrastive focus than in wide focus. However, there is no significant difference in alignment of this low turning point. Given the marginal acoustic differences the question remains whether listeners perceive a difference between the wide and contrastive focus elicitations at all.

A semantic congruency experiment using congruous (WF-WF and CF-CF) and incongruous dialogs (WF-CF and CF-WF) was conducted aiming to investigate whether listeners are able to perceive the phonetic difference as a function of focus. Interestingly, the results of the perception study show that listeners are able to distinguish between a congruous and an incongruous dialog, in the WF-context as well as in the CF-context condition (see Fig. 2). In sum, the accents realized under contrastive focus which have a lower f0 turning point and a later as well as a higher f0 peak compared to wide focus accents can be distinguished from each other perceptually.

In order to answer the question of tonal representation, the question remains whether the lower scaling of the low turning point causes the perceptual impression of contrastive focus, or whether higher scaling of H* accents is a sufficient phonetic cue. The fact that listeners differentiate between marginally distinct intonation contours as recorded in our production data allows for a further perception study manipulating the individual cues.

4. ACKNOWLEDGEMENTS

This research was supported by the DFG-grant (KU 2323/1-2) to the project "Prosody in Parsing" at Potsdam University, principle investigators: Frank Kügler, Caroline Féry and Shravan Vasishth. We are grateful to Tobias Günther who provided considerable technical assistance.

5. REFERENCES

[1] Alter, K., Mleinek, I., Umbach, C., Rohe, T., Steube, A. 2002. Kontrastprosodie in Sprachproduktion und –perzeption. *Linguistische Arbeitsberichte* 77, 59-81.
[2] Bates, D., Sarkar, D. 2007. *Lme4: Linear Mixed-Effects Models Using S4 Classes*. R package version 0.9975-11.
[3] Boersma, P., Weenink, D. Praat-doing phonetics by computer, online resource: *http://www.praat.org*
[4] Carlson, K. 2001. The effects of parallelism and prosody on the processing of gapping structures. *Language and Speech* 44, 1-26.
[5] DMDX, online resource: *http://www.u.arizona.edu/ /~jforster/dmdx.htm*
[6] Féry, C. 1993. *German Intonational Patterns*. Tübingen: Niemeyer.
[7] Féry, C., Kügler, F. 2008. Pitch accent scaling on given, new and focused constituents in German, *Journal of Phonetics* 36, 680-703.
[8] Grabe, E. 1998. *Comparative Intonational Phonology. English and German*. Ph.D. thesis, Nijmegen.
[9] Grice, M., Baumann, S. 2002. Deutsche intonation und GToBI. *Linguistische Berichte* 191, 267-298.
[10] Grice, M., Baumann, S., Benzmüller, R. 2005. German intonation in autosegmental-metrical phonology. In Jun, S.A. (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: OUP.
[11] Grice, M., Baumann, S., Jagdfeld, N. 2009. Tonal association and derived nuclear accents: The case of downstepping contours in German. *Lingua* 119, 881-905.
[12] Grice, M., Reyelt, M., Benzmüller, R., Mayer, J., Batliner, A. 1996. Consistency in transcription and labelling of German intonation with GToBI. *Proc. ICSLP* 96 Philadelphia.
[13] Hanssen, J., Peters, J., Gussenhoven, C. 2008. Prosodic effects of focus in Dutch declaratives. *Speech Prosody*-2008, 609-612.
[14] Kügler, F., Drenhaus. H. 2006. Negation, prosody and replacives: An auditory acceptability judgement study. *Journal of Cognitive Neuroscience* 46.
[15] Prieto, P. Submitted. Experimental methods and paradigms for prosodic analysis. In Cohn, A., Fougeron, C., Huffman, M. (eds.), *Handbook of Laboratory Phonology*. Oxford University Press.
[16] Prieto, P., Torres-Tamarit, F., Vanrell, M.M. Submitted. The role of tonal scaling in encoding intonational contrast in Catalan. *Journal of Phonetics*.
[17] Rathcke, T., Harrington, J. 2010. The variability of early accent peaks in Standard German. In Fougeron, C., Kühnert, B., d'Imperio, M., Vallée, N. (eds.), *Laboratory Phonology* Berlin / New York: Mouton de Gruyter, 10, 533-555.
[18] Toepel, U. 2005. Contrastive topic and focus information in discourse. *MPI Series in Human Cognitive and Brain Science* 66. Dresden: Sächsisches Digitaldruck Zentrum.