

Double contrast is signalled by prenuclear and nuclear accent types alone, not by f0-plateaux

Bettina Braun, Yuki Asano

¹ Department of Linguistics, University of Konstanz, Germany
(bettina.braun, yuki.asano)@uni-konstanz.de

Abstract

In two speeded acceptability experiments we tested which combination of prenuclear accent, nuclear accent and f0-interpolation between them is best suited to signal a double contrast in German (i.e., a contrastive topic followed by a contrastive focus). The experimental utterances differed in the prenuclear accent (medial- vs. late-peak, i.e., L+H* vs. L*+H), the nuclear accent (early- vs. medial-peak, i.e., H+L* vs. H*) and the f0-interpolation between them (high or dipping). All utterances were judged for their acceptability in a contrastive (Experiment 1) and a non-contrastive context (control Experiment 2). Our results showed that the combination of a late-peak prenuclear accent (L*+H) and an early-peak nuclear accent (H+L*) is best suited to signal a double contrast, independent of the f0-interpolation. The reaction time data also support the view that the f0-interpolation is not necessary for the interpretation of a double contrast.

Index Terms: contrast, perception, intonation, information structure

1. Introduction

In this paper we investigate which intonational realization is best suited to signal a double contrast [1], i.e., a contrastive topic followed by a contrastive focus as shown in (1). The italic capitals signal contrastive topics, the non-italic capitals contrastive foci.

- (1) What did the popstars wear?
 - a. The *FEMALE* popstars wore KAFTANS.
 - b. The *MALE* popstars wore SUITS.

The intonational realization of such utterances has been described rather differently in the literature, especially for German [2-6]. One prominent realization of such double contrast sentences in German is the so-called hat pattern [see 7]. This contour consists of an initial pitch rise (prenuclear accent or *accent1*, in (1) realized for instance on 'male'), a sustained high pitch (f0-interpolation), and a pitch fall (nuclear accent or *accent2*, in (1) realized for instance on 'suits').

Interestingly, the intonational form of the German hat pattern appears to vary a lot. Féry [5], for instance, describes two types of hat patterns, H* followed by H*L (type 1), and L*H followed by H*L (type 2). Hat pattern 2, which is the contour that can signal a double contrast, is annotated as two intermediate phrases, so the rising first accent has a nuclear status [5, p 128ff]. Wunderlich [6] and Kohler [8], on the other hand, consider the f0-interpolation as a phonological entity: Wunderlich [6], for instance, annotates the hat pattern as H* H L*. The middle H-tone without diacritics signals a floating tone, which is associated with all unspecified intervening syllables. In the Kiel Intonation Model (KIM), Kohler [8]

distinguishes between two types of f0-interpolations, dipped and non-dipped. The hat pattern contains a non-dipped interpolation and can be combined with prenuclear accents whose peaks are either medial or late (with respect to the stressed syllable) and with nuclear accents whose peaks are either early or medial.

In summary, there is large disagreement regarding the phonological form of hat patterns in German. The only similarity across descriptions is the sustained high pitch between accents (which is not even considered phonological in all models of intonation). However, these intonational descriptions are mostly based on introspection and not on a representative sample of production or perception data.

Two production studies in German shed more light on the phonological realization of double contrast utterances: They analysed the intonational realization of identical utterances in either contrastive or non-contrastive contexts and showed differences in the realization of the prenuclear accent [3, 4] and the nuclear accent [4]: Contrastive topics are preferably signalled with later and higher peaks than non-contrastive topics. The nuclear accents are more often realized with an early-peak (H+L*) accent in such contrastive context [4]. Results on the presence of a high f0-interpolation are mixed, however. While [3] reported them as typical realization, [4] showed that they were not consistently produced in contrastive contexts. Therefore, it is still a matter of debate what role the high f0-interpolation between accents plays in signalling a double contrast, as compared to the contribution of prenuclear and nuclear accent types. Offline perception studies that test which contours are more appropriate in a double contrast context have focused either on the contribution of the prenuclear pitch accent [3, 4] or on the perceptual importance of the f0-transition between the accents [e.g., 9]. Therefore it is unclear what the relevance of the different parts of the hat pattern is.

Recently, we investigated the online interpretation of utterances with a sentence-initial contrastive or non-contrastive topic accent in a visual-world eye-tracking paradigm [10]. Participants heard a context sentence that set up a potential double contrast (contrast on the subject-NP and the object-NP, e.g., *the mason wanted to climb onto the table*). Participants then heard utterances starting with a subject-NP that contrasted lexically with that in the context sentence (e.g., *the carpenter wanted to...*). Intonationally, the subject-NP was realized as contrastive or non-contrastive topic. The visual display showed four objects, one of which was a visual object that contrasted with the grammatical object of the context sentence (e.g., *chair*, contrasting with *table*). We investigated whether participants fixate that contrastive object *chair* more when they perceive the subject-NP as contrastive topic. Results confirmed this hypothesis, but the anticipatory effects emerged only when participants processed the first part of the f0-interpolation (and hence knew whether there was a high or dipping pattern) and not during the processing of the subject-

NP. These results are compatible with 3 explanations. 1) it is not the prenuclear accent type but the f0-interpolation that drives the effect, 2) it is the combination of prenuclear accent and f0-interpolation and 3) it is only the prenuclear accent that causes the anticipatory effect but it is delayed in time and therefore only shows during the f0-interpolation.

Our experiments fill this gap by providing data on the intonational realization that is best suited to signal a double contrast (contrastive topic followed by a contrastive focus). Participants performed speeded acceptability judgments in contrastive (Exp. 1) and non-contrastive (Exp. 2) contexts; these binary responses were complemented by reaction times as a measure of task difficulty.

2. Experiment 1: contrastive context

Experiment 1 tests the perception of utterances with different intonational contours in a double contrast context.

2.1. Methods

2.1.1. Participants

Sixteen native German participants (5 male, 20-29 years) took part for a small fee or course credit. They were unaware of the purpose of the experiment and had not participated in other experiments with similar materials and had no prior training in intonational phonology.

2.1.2. Materials

Eight sentence pairs with a double contrast were constructed. These pairs had a parallel syntactic structure and differed only in the grammatical subject and object. The first member of the sentence-pair appeared visually as context sentence, the second auditorily as target sentence (see examples (2) and (3)). The auditory sentences contained 5-6 unaccented syllables between the accented ones; to avoid tonal crowding, the accented-syllables did not appear phrase-initially or -finally. The auditory stimuli (3) were produced by a female speaker of Standard German who was trained in intonational phonology. She recorded each sentence in eight different versions, crossing the prenuclear accent (medial peak accent L+H* or late peak accent L*+H, henceforth *accent1*), the nuclear accent (medial peak accent H* or early peak accent H+L*, henceforth *accent2*) and the f0-interpolation between the two accents (high or dipping) on a digital recorder (44.1kHz, 16Bit).

Acoustic measurements revealed that the productions were as intended. Paired t-tests showed that in the prenuclear accent, the peak was significantly later in the late peak condition (148ms after the offset of the stressed syllable) than in the medial peak condition (21ms before the offset of the stressed syllable, $t(31) = 10.4$, $p < 0.0001$). In the nuclear accent, the peak was significantly earlier in the early peak condition (35ms before the start of the stressed syllable) than in the medial peak condition (138ms after the start of the stressed syllable, $t(31) = 57$, $p < 0.0001$). Furthermore, the minimum-f0 for the high plateau condition was significantly higher in the high plateau condition (221Hz) than in the dipping condition (155Hz, $t(31) = 19.4$, $p < 0.0001$).

2.1.3. Procedure

Eight pseudo-randomized experimental lists were constructed by presenting all eight sentences in all eight intonation

contours (64 trials in total). In each list, two identical sentences or two identical intonation contours were separated by at least four intervening items with a different content or intonation.

Participants were tested individually in a silent room. The experiment was controlled using presentation. Each trial started with a fixation cross of 800ms. Then, the visual display of the context appeared in white Arial 30pt characters on black background centered on screen and remained there for 3 seconds (see example (2)). Then, the auditory stimulus was played in stereo via headphones at a comfortable loudness. Response times were measured relative to the end of the auditory stimulus for a duration of two seconds.

(2) Visual display of context:

Jetzt geht es um einen Sohn und eine Tochter. Der Sohn beschäftigt sich mit Latein und...

‘The next story is about a son and a daughter. The son is occupying himself with Latin and...’

(3) Auditory stimulus:

Die Tochter beschäftigt sich mit Mathe.

‘The daughter is occupying herself with mathematics.’

Participants received written instructions to press the right button when the intonational realization of the auditory stimulus was acceptable as a continuation to the visual context and the left button otherwise. Left-handed participants had to press the left button for ‘yes’-responses. Participants were randomly assigned to one of these lists.

2.2. Results

Thirty data points were missing due to timeout (4.6% of the data).

2.2.1. Response data

The remaining responses were summed by condition (*accent1*, *accent2*, *f0-interpolation*) and were subjected to separate by-subjects and by-items general linear logistic regression models [11]. The percentages of acceptable responses in the eight conditions are shown in Table 1. Results of the by-subjects analysis showed a main effect of *accent1* with a preference for late peak accents ($\beta = 0.8$, $SE = 0.15$, $p < 0.0001$), as well as main effects for *accent2* and *f0-interpolation* as well as an interaction between them. If *accent2* had an early peak, utterances with a high f0-interpolation were judged worse than those with a dipping pattern ($\beta = -0.5$, $SE = 0.2$, $p < 0.05$). This negative effect of a high f0-interpolation was significantly stronger when *accent2* had a medial peak ($\beta = -1.3$, $SE = 0.2$, $p < 0.0001$). Results for the by-items analysis were identical.

Table 1. Average percentage of acceptable responses in Experiment 1.

Dipping f0-interpolation		
	nuclear accent (<i>accent2</i>)	
prenuclear accent (<i>accent1</i>)	medial peak (H*)	early peak (H+L*)
medial peak (L+H*)	57.4%	77.4%
late peak (L*+H)	70.4%	89.3%

High f0-interpolation		
	nuclear accent (<i>accent2</i>)	
prenuclear accent (<i>accent1</i>)	medial peak (H*)	early peak (H+L*)
medial peak (L+H*)	25.5%	55.6%
late peak (L*+H)	34.9%	81.6%

2.2.2. Reaction times

Reaction time data were analysed as an indicator for the difficulty of the decision. Reaction times for responses in which participants indicated the intonational realization was acceptable were log-normalized and analysed using linear mixed effects regression models with *accent1*, *accent2* and *f0-interpolation* as fixed factors and *items* and *participants* as crossed random factors [11]. The model included random slopes for the fixed factors [12, 13]. Model selection was performed using backward elimination and log likelihood ratio tests (function *anova()* in R). The best model was validated by removing data points with residuals that lie beyond 2.5sd from the mean and the model was refitted. A t-value > |2| indicates a significant effect at $\alpha = 0.05$.

Results showed a main effects of *accent2* ($\beta = 0.43$, SE = 0.09, $t = 4.5$) with a preference for early-peak accents and a significant interaction between *accent1* and *accent2* ($\beta = 0.32$, SE = 0.12, $t = 2.6$). As shown in Table 1, utterances with an early-peak *accent2* (right columns) were reacted to faster than utterances with a medial-peak *accent2* (left columns). Furthermore, utterances with a late-peak *accent1* and an early-peak *accent2* (lower-right corners) were reacted to fastest.

Table 2. Log- reaction times for 'acceptable' responses in Experiment 1, as calculated from the statistical model.

Dipping f0-interpolation		
	nuclear accent (<i>accent2</i>)	
prenuclear accent (<i>accent1</i>)	medial peak (H*)	early peak (H+L*)
medial peak (L+H*)	6.32	6.23
late peak (L*+H)	6.53	6.02

High f0-interpolation		
	nuclear accent (<i>accent2</i>)	
prenuclear accent (<i>accent1</i>)	medial peak (H*)	early peak (H+L*)
medial peak (L+H*)	6.31	6.24
late peak (L*+H)	6.36	6.06

2.3. Discussion

The data of the response analysis showed that contours with an L*+H prenuclear accent followed by an H+L* nuclear accent were best suited to signal a double contrast. A high f0-interpolation between accents was judged less acceptable, particularly when it was followed by a nuclear accent with a medial peak (H*). These data suggest that neither prenuclear nor nuclear accent type alone are able to signal a double contrast. Instead, it is a combination of prenuclear and nuclear accent types that is relevant, while the f0-interpolation only had a small effect (especially, when comparing the two best contours, L*H H+L* with a high or dipping interpolation). These data stand in contrast to most intonational descriptions

of German data, which mostly highlighted the presence of an f0-plateau between the two accents.

These response data were supported by the reaction time analyses: Participants were fastest to judge the two most appropriate contours (L*+H followed by H+L*, with high or dipping interpolation) as acceptable in the double contrast conditions.

3. Experiment 2: non-contrastive context

Experiment 2 was a control experiment to ensure that the results of Experiment 1 were not caused by inherent qualitative differences in the stimuli. To avoid that the auditory stimuli in subsequent trials are not perceived as contrasting with one another, we did not use a broad-focus question like "What happened?" [14]. Rather, the visual context followed the auditory stimulus in Experiment 2.

3.1. Methods

3.1.1. Participants

Another 16 native German participants (1 male, 20-27 years), different from those in Experiment 1, took part for a small fee.

3.1.2. Materials

The auditory materials were the same as in Experiment 1. The visual stimuli were as follows: Eight contextually neutral dependent clause sentences were constructed for each of the eight auditory sentences used in Experiment 1, see (4) and (5).

(4) Auditory stimulus:

Die Tochter beschäftigt sich mit Mathe.

'The daughter is occupying herself with mathematics.'

(5) Visual stimulus:

..., weil sie morgen eine Klausur schreibt.

'... because she will have a test tomorrow.'

3.1.3. Procedure

The testing procedure and order of trials in the experimental lists was identical to Experiment 1. Each trial started with a fixation cross, which was displayed for 800ms. Then, the visual stimulus appeared at the centre of the screen (white Arial 30pt on black background, remaining for 3 seconds); the auditory stimulus started simultaneously. Reaction times were measured relative to the offset of the auditory stimulus.

Participants were instructed to press the right button when the intonational realization of the auditory stimulus fitted the context and the left button otherwise. Left-handed participants had to press the left button for 'yes'-responses.

3.2. Results

Thirty-five points were missing due to timeout (3.4%). The remaining data were analysed as described in 2.2.

3.2.1. Responses

Results of the by-subjects and by-items analyses showed a three-way-interaction between *accent1*, *accent2* and *f0-interpolation* ($\beta = 1.2$, SE = 0.5, $p < 0.05$). Participants judged utterances with a medial-peak *accent1*, a medial-peak *accent2* and a high f0-interpolation as best suited in the presented non-contrastive context (see Table 4).

Table 4. Average percentage of acceptable responses in Experiment 2.

Dipping f0-interpolation		
	nuclear accent (<i>accent2</i>)	
prenuclear accent (<i>accent1</i>)	medial peak (H*)	early peak (H+L*)
medial peak (L+H*)	68.0%	61.2%
late peak (L*+H)	66.3%	54.3%

High f0-interpolation		
	nuclear accent (<i>accent2</i>)	
prenuclear accent (<i>accent1</i>)	medial peak (H*)	early peak (H+L*)
medial peak (L+H*)	69.0%	53.3%
late peak (L*+H)	59.8%	61.9%

3.2.2. Reaction times

There were no main effects and no interactions (all t-values < |0.3|).

3.3. Discussion

Importantly, this control experiment showed different responses than Experiment 1, suggesting that participants reacted to the semantic felicity of the intonational realizations and not to inherent qualitative differences in the stimuli. In the non-contrastive context, participants preferred contours with medial-peak prenuclear and nuclear accents, connected by a high f0-interpolation (average acceptability 69%), a contour that only had an average acceptability rating of 25.5% in a contrastive context. Overall, there was much less variation in acceptability scores in Experiment 2 and scores were above 50% for all contours, suggesting that no contour sounded particularly inappropriate in the non-contrastive context. This latter interpretation is supported by reaction time data, which did not differ across contours.

4. General Discussion

We tested the contribution of prenuclear and nuclear accent type as well as the f0-interpolation between accents in the perception of German utterances signalling a double contrast in a speeded acceptability judgement task. Previous phonological descriptions gave various competing analyses on the hat pattern in German [1, 2, 5, 6], which differed in the prenuclear and nuclear accent types and only agreed on the high f0-interpolation between the accents. Our experimental data stand in contrast to these analyses and show that the high f0-interpolation does not contribute much to the perception of a double contrast (on the contrary, a high f0-interpolation was suitable in a non-contrastive context). Rather, it was a late-peak prenuclear accent (L*+H) in combination with an early-peak nuclear accent (H+L*) that was best suited to signal a double contrast. Experiment 2 further showed that this intonational realization is not appropriate in non-contrastive contexts, which demonstrates that it was not the general auditory quality of this contour, but its semantic contribution that drove the effect. The fact that the high f0-interpolation was not decisive for a double contrast interpretation lets us see the online eye-tracking data [10] in a new light. There, anticipatory effects only emerged while participants processed

the f0-interpolation between accents (and not when they processed the accent on the sentence-initial subject-NP). In the light of the current findings, the late occurrence of these anticipatory effects appears to result from the time needed to establish the contrastive inference caused by the prenuclear accent, rather than from the meaning contribution of the f0-interpolation itself (explanation 3 of the introduction).

5. Conclusion

In order to signal a double contrast (contrastive topic followed by contrastive focus), the prenuclear accent needs to have a late peak (L*+H) and the nuclear accent needs to have an early peak (H+L*). The f0-interpolation between these accents – often thought to be a constituting part of this contour – is merely phonetic in nature and does not contribute to the interpretation of such contours.

6. Acknowledgements

This work was supported by a research grant from the German Research Foundation to the first author (BR-3428/1-1). We thank Katja Häuser for recording the stimuli.

7. References

- [1] Büring, D., "The Meaning of Topic and Focus: The 59th Street Bridge Accent", London: Routledge, 1997.
- [2] Büring, D., "On D-trees, beans, and B-accents", *Linguistics & Philosophy*, 26(5):511-545, 2003.
- [3] Mehlhorn, G., "Produktion und Perzeption von Hutkonturen im Deutschen [Production and perception of hat pattern in German]", *Linguistische Arbeitsberichte*, 77:31-57, 2001.
- [4] Braun, B., "Phonetics and phonology of thematic contrast in German", *Language and Speech*, 49(4):451-493, 2006.
- [5] Féry, C., "German Intonational Patterns", Tübingen: Niemeyer, 1993.
- [6] Wunderlich, D., "Intonation and contrast", *Journal of Semantics*, 8(239 - 251), 1991.
- [7] Cohen, A. and 't Hart, J., "On the anatomy of intonation", *Lingua*, 19:177-192, 1967.
- [8] Kohler, K., "Terminal intonation patterns in single-accent utterances of German: phonetics, phonology and semantics", *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK)*, 25:115-185, 1991.
- [9] Hertrich, I., "Perceptual categorization in terminal intonation patterns of two-accent utterances of German", *Arbeitsberichte IPdS Universität Kiel*, 25:243-294, 1991.
- [10] Braun, B. and Asano, Y., "Contrastive topics and foci: From the activation of alternatives to anticipatory effects", submitted.
- [11] Baayen, H.R., "Analyzing linguistic data. A practical introduction to statistics using R", Cambridge: Cambridge University Press, 2008.
- [12] Barr, D.J., et al., "Random-effects structure for confirmatory hypothesis testing: Keep it maximal", *Journal of Memory and Language*, 68(3):255-278, 2013.
- [13] Cunnings, I., "An overview of mixed-effects statistical models for second language researchers", *Second Language Research*, 28(3):369-382, 2012.
- [14] Ladd, D.R., "Intonational phonology", 2nd ed. Cambridge studies in linguistics. Vol. 119. Cambridge: Cambridge University Press. 349, 2008.