Prosodic Properties of Constituents Associated with Stressed *auch* in German

Stefan Sudhoff & Denisa Lenertová

Institut für Linguistik
Universität Leipzig, Germany
{sudhoff; lenerto}@uni-leipzig.de

Abstract

We report a production experiment and two perception studies examining the prosodic characteristics of constituents associated with the stressed variant of the German particle *auch* 'also' in potentially ambiguous constructions. The results show that these elements are marked by perceptually relevant rising pitch accents, but that there is no 1:1 mapping between the prosodic realization and the status of being associated with *auch*.

1. Introduction

It is a well known fact that the stressed variant of the German additive focus particle *auch* 'also' – in contrast to its unstressed counterpart – associates with a constituent to its left. The associated constituent (AC), which is not confined to a specific syntactic position, represents the element added to a contextually given set by means of the particle – in (1), Martin is added to the set of hungry beings.

(1)   [Martin]$_{AC}$ hat AUCH Hunger.
       Martin   has also   hunger

The syntactic, semantic, and information structural properties of constructions with stressed *auch* have been subject to much discussion, see [5] and [3], among others. At the same time, the relevance of prosodic factors did not go unnoticed. Krifka [3] argues that ACs of stressed *auch* are contrastive topics, which, however, are not obligatorily marked by a rising contrastive accent. Previous experimental work (cf. [1]) showed that the marking of contrastive topics is gradual and involves a characteristic interplay of f_0, duration, and alignment variables.

Concerning the prosodic marking of association with stressed *auch*, constructions with two potential ACs to the left of the particle provide a suitable subject for investigation, as they should call for a disambiguation by prosody. A production study was designed to explore the individual prosodic parameters (Sec. 2); subsequent perception studies (Sec. 3 and 4) tested the impact of the marking on the interpretation.[1]

2. Experiment 1: Production study

In Exp. 1, minimal pairs of utterances were elicited by contextually triggering *auch* to associate with one of two possible elements. In this way, prosodic factors relevant for the marking of ACs in two different positions (sentence initial vs. non-initial) could be identified.

2.1. Data elicitation

Participants.   7 female native speakers of German (aged 20 to 26) without any linguistic background or awareness of our analysis intentions participated in the experiment.

Materials.   5 pairs of sentences ambiguous with respect to the AC of *auch* were constructed, the potential ACs always being the subject and a temporal adverbial realized by three-syllabic constituents with penultimate stress. The two sentences of each pair only differ in the relative order of the potential ACs, located in the so-called prefield (PF) and middlefield (MF), i.e., preceding and following the finite verb, respectively (cf. (2)).

(2)   Der Rudi hat im Juni ... / Im Juni hat der Rudi ...
       the Rudi has in June   in June has the Rudi
       ... wahrscheinlich auch einen Vortrag gehalten.
         probably   also a   talk   given

In the 5 lexicalizations, the number of syllables is identical before the particle and roughly the same after it. Within each sentence, the potential ACs are kept parallel with respect to syllable structure, vowel length and vowel height (stressed and post-stressed syllable) to control the effects of the phonetic properties (e.g., intrinsic f_0, cf. [4]). For every lexicalization, two contexts were constructed, each triggering one of the possible associations with *auch*, cf. the English translations in (3) and (4).

(3)   Can you tell me who of the PhD students gave a talk in June? I heard that only Martin gave one at that time.

(4)   Can you tell me when Rudi gave talks this term? I only know of the one in May.

The combination of the two contexts with the two target sentence versions results in 4 experimental conditions, the independent variables being the position (PF vs. MF) and function (subject vs. adverbial) of the intended AC. The 20 critical items were supplemented by 20 filler items also involving the context-dependent resolution of ambiguity.

Procedure.   Subjects had to read a visually presented combination of context and target item silently, listen to the auditory presentation of the context (which had been recorded by a professional speaker) and produce the target sentence. Repeated productions (with the context replayed each time) were allowed when subjects believed their utterances to be unnatural or inappropriate. The target utterances were digitally recorded.

The 40 items were randomized anew for each subject in such a way that the minimal pairs are disguised. A short practice session (5 items) served to familiarize the subjects with the task. In the course of the actual experiment, subjects could take several short breaks. Items produced with pauses, hesitations, or slips of the tongue were repeated at the end of the session, again

randomized and interspersed with fillers. The total duration of the experiment was about 45min.

2.2. Data analysis

Data selection. The utterances of one speaker were excluded from the analysis as a post-experimental interview revealed that the sentences with stressed *auch* had attracted her attention. 13 utterances had to be excluded for various reasons (production of the unstressed variant of *auch*, change of the prosodic strategy in the middle of the utterance, hesitation, defective recording). Altogether, 107 utterances could be included in the analysis.

Annotation and measurements. On the basis of the GToBI system (cf. [2]), pitch accents were annotated. Syllable boundaries and the onsets of the stressed vowels for the two potential ACs in each utterance were labeled using information from waveform and spectrogram. f_0 was extracted,[2] and after a manual correction of artifacts, the contours were linearly interpolated and smoothed. As the f_0-contours are highly consistent within all 4 conditions, mean contours could be computed.

For both potential ACs in each utterance, the following dependent variables were determined: f_0-*min* (local f_0-minimum preceding the rise), f_0-*max* (local f_0-maximum), df_0 (difference between f_0-*max* and f_0-*min*), *dur-s23* (duration of the stressed and post-stressed syllable), *al-min* and *al-max* (alignment of the f_0-minimum and -maximum with respect to the onset of the stressed vowel). The choice of the variables was based on the observed contour shape and the hypotheses (see below).

2.3. Hypotheses

H1: ACs of stressed *auch* are marked by L*H accents.

H2: ACs of stressed *auch* show a greater f_0-rise (caused by a lower f_0-minimum or a higher f_0-peak or both), a longer duration (due to lengthening of the stressed syllable or larger domains), and a later accent alignment than non-ACs (i) in the same position (comparison between utterances) and (ii) in a different position (comparison within utterances).

2.4. Results

2.4.1. Qualitative results

ACs are always marked by rising accents: L*H (81.3%) or LH*. Non-ACs, however, also frequently (prefield: 80.8%; middlefield: 29.1%) carry an accent: L*H, LH*, or H*. The accent type of the ACs does not affect the overall contour shapes, which are illustrated by the mean contours in Fig. 1. Together with the mean values of the dependent variables, the contours suggest that the prosodic marking of the ACs is independent of their syntactic function. Consequently, no difference was made between subjects and adverbials in the statistical comparisons.

2.4.2. Quantitative results

Comparison between utterances. In the minimal pairs of lexically identical utterances, comparisons were made between associated and non-associated constituents in identical positions. The mean values of the individual variables as well as the statistical results are given in Tab. 1. The ACs show a significantly lower f_0-minimum, higher f_0-peak, longer duration[3], and later

---
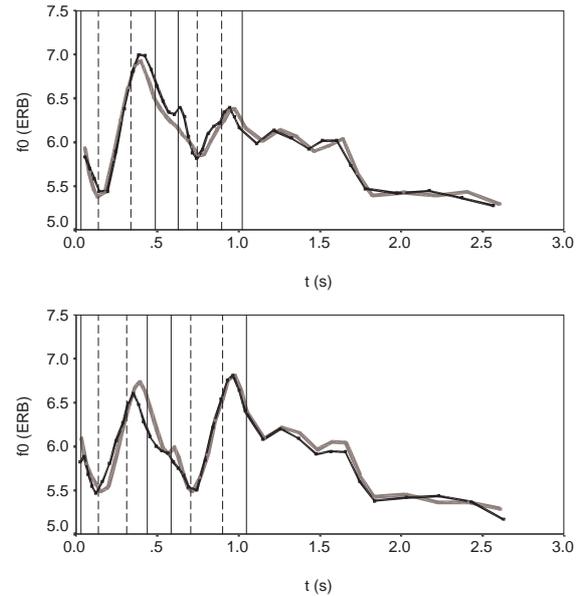
Figure 1: *Mean contours for PF (upper panel) and MF association (lower panel); gray: AC = subject, black: AC = adverbial*

peak alignment than their non-associated counterparts. There is no significant effect for the alignment of the local minimum.

Comparison within utterances. Due to the controlled materials (cf. Sec. 2.1), ACs and non-ACs could also be compared within utterances (except for f_0-*min* and *al-min*, the local minimum being mostly located in the non-controlled pre-stressed syllable). PF and MF association were evaluated separately, as their respective patterns differ considerably (cf. Fig. 1). Significant effects for the relevant variables were found for both PF and MF association, with the exception of *al-max* for MF association, cf. Tab. 2. The direction of the effects is the same as in the comparison between utterances. In addition, for all 4 variables, the differences between ACs and non-ACs are significantly greater for PF association than for MF association.[4]

2.5. Discussion

H1 was confirmed to the extent that ACs of stressed *auch* are marked by rising pitch accents. However, the identification of ACs cannot be based on their accents, as non-ACs often carry accents of the same type. Here, the quantitative analysis reveals the decisive factors: ACs and non-ACs significantly differ with respect to f_0-minimum, -maximum, and -rise, duration, and peak alignment, which substantially confirms H2.

The syntactic function of the ACs seems not to have any influence on their prosodic realization. In contrast, the prosodic marking is affected by the position of the AC: With respect to the relevant variables, the differences between ACs and non-ACs are greater for PF association. The marking of middlefield ACs is only slightly stronger than the marking of the respective non-associated prefield elements. Nevertheless, we expect the second accent to be perceived as more salient, facilitating a correct interpretation in these cases as well. This is one of the issues to be discussed in connection with the perception studies.

---

variable	pos.	assoc.	non-ass.	p
f_0-min (ERB)	PF	5.30	5.42	<.001
	MF	5.42	5.74	<.001
f_0-max (ERB)	PF	7.13	6.75	<.001
	MF	6.93	6.44	<.001
df_0 (ERB)	PF	1.83	1.33	<.001
	MF	1.51	0.70	<.001
dur-s_{23} (ms)	PF	346	300	<.001
	MF	340	281	<.001
al-min (ms)	PF	-47	-41	n.s.
	MF	-48	-44	n.s.
al-max (ms)	PF	193	173	<.003
	MF	185	139	<.001

Table 1: *Comparison between utterances: paired t-tests (one-tailed); N = 48; α = .0042 (Bonferroni adjustment)*

variable	ass. PF	non-a. MF	p	non-a. PF	ass. MF	p
f_0-max*	7.19	6.45	<.001	6.71	6.91	<.001
df_0*	1.92	0.70	<.001	1.33	1.52	<.003
dur-s_{23}**	352	277	<.001	300	343	<.001
al-max**	196	137	<.001	173	185	n.s.

* measured in ERB; ** measured in ms

Table 2: *Comparison within utterances: paired t-tests (one-tailed), α = .0063 (Bonferroni adjustment); left: PF association (N = 55); right: MF association (N = 52)*

3. Experiment 2: Perception study

By means of a completion task, Exp. 2 investigates to what extent hearers use the encoded prosodic information for the disambiguation of the utterances produced in Exp. 1.

3.1. Method

Participants. 32 native speakers of German (mean age 24.8) participated in the experiment. They were naïve with respect to the problem examined.

Materials. The original utterances of 5 speakers from Exp. 1 (except the defective or unacceptable utterances that had not been included in the analysis) served as stimuli. The 91 critical utterances were supplemented by 69 filler utterances randomly selected from all 6 speakers. Two possible continuations were constructed for each lexicalization, either of them compatible with only one interpretation (e.g., the choice for (2) would be between '... and not only in May' and '... and not only Martin').

For each utterance, 4 derived parameters expressing the clearness of the prosodic marking were calculated from the differences between the AC and the non-AC with respect to the variables f_0-max, df_0, dur-s_{23}, and al-max. The differences were transformed to scales ranging from 0 to 1 (corresponding to the smallest and biggest difference, respectively), with the intermediate values being proportionally mapped onto real numbers between the two extremes. PF and MF association were treated separately. The resulting parameters p-f_0-max, p-df_0, p-dur-s_{23}, and p-al-max can be directly related to the percentage of correct responses in the perception experiment.

Procedure. Subjects had to listen to an auditorily presented target utterance and select one of the two continuations (pre-

sented on a computer screen in a randomized left-right order). Repeated listening was possible. The 160 items were pseudo-randomized and divided into 4 blocks to enable subjects to take short breaks in between. The order of the blocks was systematically varied. The entire experiment took about 30min.

3.2. Hypothesis

H3: The percentage of correct responses for the utterances positively correlates with p-f_0-max, p-df_0, p-dur-s_{23}, p-al-max.

3.3. Results and discussion

Of the total of 2912 responses, 72.4% were correct. However, the percentage of matching answers differs between PF (84.5%) and MF association (58.3%). Crucially, both PF and MF association show a significant correlation between the proportion of correct responses and p-f_0-max, p-df_0, and p-dur-s_{23}, cf. Table 3. H3 is thus essentially confirmed.

There is, however, no significant effect for p-al-max. Still, we cannot conclude that peak alignment is perceptually irrelevant: Its possible effects could be outweighed by effects of the other parameters. Moreover, the proportion of the contribution might differ between the parameters with significant effects, especially since the parameters are probably interrelated.[5] A comprehensive examination of the interaction between the individual parameters and perception is beyond the scope of this study.

parameter	prefield ass.		middlefield ass.	
	r	p	r	p
p-f_0-max	.502	<.001	.525	<.001
p-df_0	.626	<.001	.385	<.007
p-dur-s_{23}	.253	<.05	.387	<.007
p-al-max	.223	n.s.	-.092	n.s.

Table 3: *Correlations between % of correct responses and the derived parameters for PF (N = 49) and MF assoc. (N = 42)*

4. Experiment 3: Perception study (manipulated stimuli)

Exp. 3 has a similar design, but is based on manipulated materials. This allows us to address two open issues from Exp. 2: (i) the source of the different proportions of correct responses for PF and MF association (unbalanced materials vs. general tendency) and (ii) the possibility of establishing perceptual categories on the basis of the prosodic variables examined in Exp. 1.

4.1. Method

Participants. 44 native speakers of German (mean age 24.0) participated in the experiment.

Materials. 11 versions of an utterance of (5) were produced by a systematic joint manipulation of f_0-min, f_0-max, dur-s_{23}, and al-max on both potential ACs using Praat's PSOLA algorithm.[6]

---

[5] A simple scenario is that the parameters with significant effects are equally strong perceptual cues. Although *r* differs between PF and MF association and between the parameters, the correlation with the sum of p-f_0-max, p-df_0, and p-dur-s_{23} is of similar strength for PF (r = .555, p < .001) and MF association (r = .557; p < .001).

[6] The variation of f_0-min and f_0-max is automatically accompanied by a variation of df_0; al-min was not manipulated as it did not show any significant effects in the production study.

(5) Der Wiener hat um sieben wahrscheinlich auch einen Anruf bekommen. 'The Viennese fellow probably received a call at seven, too.'

A trained female speaker produced (5), taken from the materials of Exp. 1, with a neutral intonation and a similar prominence on both potential ACs. For the two extremes – the clear cases of PF (stimulus 1) and MF association (stimulus 11) – the manipulation parameters are based on prototypical utterances of (5) in Exp. 1 that had received good results in Exp. 2, thus resembling the mean contours in Fig. 1. The target values for the 9 intermediate stimulus versions represent equal steps on the scale between the two extremes, cf. the plotted f_0-tracks in Fig. 2.[7] The resulting stimuli still sounded natural; uninformed hearers did not recognize them as being manipulated. The continuations for (5) and the fillers were taken from Exp. 2.
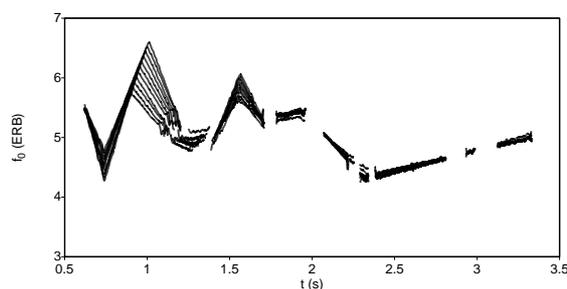
Figure 2: f_0-tracks of the 11 stimulus versions

Procedure. The task and mode of presentation were the same as in Exp. 2. Each stimulus version was included 6 times. The resulting 66 critical items were interspersed with 66 fillers and randomized anew for each subject. A filler always separated two critical items to prevent subjects from making direct comparisons between the different versions. Subjects were informed that they would have to judge many instances of the same sentence and instructed to decide afresh each time. They could take short breaks at any time; the experiment took about 30min.

4.2. Hypothesis

H4: The two extreme stimuli show the clearest interpretation preferences: Stimulus 1 receives the highest percentage of decisions for PF association and stimulus 11 for MF association. The respective percentages are similarly high.

4.3. Results and discussion

A manual inspection of the individual subjects' results with respect to the distribution of the judgements revealed two different patterns. From 1/3 of the subjects (group A), all 11 versions received about the same percentage of choices for PF and MF association. The results of the remaining subjects (group B) clearly differ across the stimulus versions. The aggregated results of the two groups are given in Fig. 3.

Group A obviously did not pay attention to the prosody of the critical items. As for group B, the results of Exp. 2 could be replicated: The interpretive preferences depend on the prosodic realization. Specifically, stimulus 1 was assigned the highest
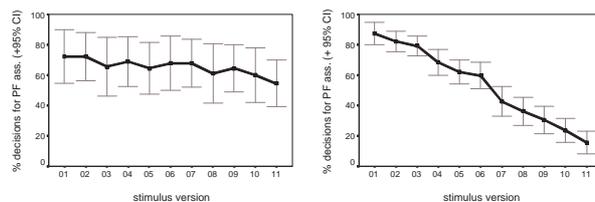
Figure 3: *% of decisions for PF association across the 11 stimulus versions; left: group A (N = 15), right: group B (N = 29)*

percentage of decisions for PF association (87.4%) and stimulus 11 for MF association (84.5%), the percentages being similar. This confirms H4. Contrary to what one might expect, the results do not show an s-shaped curve, but a rather linear relationship between the number of decisions for a particular association position and the prosodic parameters of the stimulus versions. Subjects did not assign the gradually differing stimuli to distinct perceptual categories.[8]

5. Conclusions

The production study revealed that ACs of stressed *auch* are marked by rising pitch accents, but that there is often no difference with respect to the accent type between ACs and non-ACs. However, the constituents differ in a number of prosodic properties, supporting a description in terms of continuous phonetic parameters (see Sec. 2.2) rather than categorical distinctions.[9] This is in line with conclusions drawn in [1].

The perception studies have shown that the identification of the ACs is governed by the relative magnitude of the prosodic parameters characterizing the candidates. Importantly, for the comparison between the potential ACs, their positions must be taken into account: The differences between ACs and non-ACs are greater for PF association than for MF association. This is reflected in the manipulated stimuli of Exp. 3. Nevertheless, subjects recognized the extremes equally well. This also supports the conclusion that the preference for PF association, which seems to arise when prosodic information is neglected – as for group A in Exp. 3, is not at play here.

In sum, although a 1:1 mapping between association status and prosodic realization cannot be maintained, prosody clearly affects the production and perception of constructions with stressed *auch*.

6. References

[1] Braun, B., 2005. *Production and Perception of Thematic Contrast in German.* Bern [etc.]: Peter Lang.

[2] Grice, M.; Baumann, S., 2002. Deutsche Intonation und GToBI. *Linguistische Berichte*, 191, 267-298.

[3] Krifka, M., 1999. Additive Particles under Stress. *Proceedings of SALT 8.* Cornell: CLC Publications, 111-128.

[4] Lehiste, I., 1970. *Suprasegmentals.* Cambridge: MIT Press.

[5] Reis, M.; Rosengren, I., 1997. A Modular Approach to the Grammar of Additive Particles: the Case of German *Auch*. *Journal of Semantics*, 14, 237-309.

---

[7]Note that the manipulation of *al-max* is not clearly visible in the middlefield in Fig. 2, as the lengthening in the prefield shifts the contours with early alignment on the middlefield constituent to the right.

[8]Note that the observed linear dependency does not result from variation between subjects.

[9]Interestingly, peak alignment, being the only phonetic variable at least to some extent reflected in the GToBI labels, does not correlate with the percentage of correct responses in Exp. 2.