When (not) to look for contrastive alternatives: 
the role of pitch accent type and additive particles

Bettina Braun¹, Yuki Asano² and Nicole Dehé¹
¹University of Konstanz, Department of Linguistics, Germany
²University of Tübingen, English Seminar, Germany

Running head: Activation of contrastive alternatives

Corresponding author:
Bettina Braun
University of Konstanz
Postfach 186
78457 Konstanz
Germany
Bettina.Braun@uni-konstanz.de
Abstract
How do pitch accent type and additive particles affect the activation of contrastive alternatives? In Experiment 1, German listeners heard declarative utterances (e.g., The swimmer wanted to put on flappers) and watched displays that depicted four printed words: one that was an alternative to the subject (e.g., diver), one that was non-contrastively related (sports), the object (e.g., flappers), and an unrelated distractor. Experiment 1 manipulated pitch accent type, comparing a broad focus control condition to two narrow focus conditions: with a contrastive or non-contrastive accent on the subject (nuclear L+H* vs. H+L*, respectively, followed by deaccentuation). In Experiment 2, the utterances in the narrow focus conditions were preceded by the unstressed additive particle auch ('also'), which associates with the accented subject noun. Results showed that, compared to the control condition, participants directed more fixations to the contrastive alternative when the subject was realized with a contrastive accent (nuclear L+H*, prenuclear L*+H) than when it was realized with non-contrastive H+L*, while additive particles had no effect. Hence, accent type is the primary trigger for signalling the presence of alternatives (i.e., contrast). Implications for theories of information structure and additive particles are discussed.

Keynotes: contrastive alternative, contrastive accent, additive particle, contrastive focus, eye-tracking
Introduction

Information structure may be signalled by a variety of means, intonational, morphological and syntactic (e.g., Krifka, 2008; Vallduvi & Engdahl, 1996). The past years have shown an increasing interest in how intonational information is processed online as the utterance unfolds over time (Chen, Den Os, & De Ruiter, 2007; Dennison & Schafter, 2010; Esteve-Gibert, Portes, Schafter, Hemforth, & D'Imperio, 2016; Husband & Ferreira, 2012; Ito & Speer, 2008; Watson, Tanenhaus, & Gunlogson, 2008; Weber, Braun, & Crocker, 2006). The current paper investigates the interpretation of referents that are intonationally marked as narrow focus (e.g., Marina in 1A) and compares it to referents that are marked as part of a broad focus domain (e.g., Marina in 2A), all else being equal.

(1) Q: Who ate the cookies?  
A: [Marina]$_F$ ate the cookies.

(2) Q: What happened?  
A: [Marina ate the cookies.]$_F$

Narrow focus is semantically defined in terms of the presence of alternatives that are relevant for interpretation (e.g., Krifka, 2008; Rooth, 1992) - a definition that is close to what many researchers associate with the notion of "contrast". In fact, it is not entirely clear what the relation between these categories is (Krahmer & Swerts, 2001; Molnár, 2001; Repp, to appear). For some authors, the notion of contrast is synonymous to (narrow) focus (e.g., Bolinger, 1961; Jackendoff, 1972; Lambrecht, 1994; Rooth, 1992). For others, contrast represents a specific kind of narrow focus, resulting in a distinction between contrastive focus and non-contrastive (presentational, newness or information) focus (e.g., Bartels & Kingston, 1994; Baumann, Grice, & Steindamm, 2006; Kiss, 1998; Neeleman, Titov, van de Koot, & Vermeulen, 2009; Pierrehumbert & Hirschberg, 1990; Selkirk, 2002; Watson, et al., 2008).

Experimental evidence speaks in favour of the latter view, but leaves degrees-of-freedom for alternative interpretations. First, studies collecting participants' responses regarding the interpretation of pitch accents at the offset of the utterance bear the risk that the responses are also influenced by other accents in the clause (i.e., accents that occur later, cf. Braun (2005), or earlier, cf. Kögler and Gollrad (2015), Braun and Tagliapietra (2010), and Watson, et al. (2008)). To illustrate the point: Braun (2006), who used an appropriateness rating task between context and
target utterance, showed that prenuclear accents with a higher and later peak were judged as better suited in a contrastive context than accents with a lower and earlier peak. However, since the nuclear accent in the utterance remained the same, it may not only be the actual shape of the prenuclear accent that affected the results but the relation between prenuclear and nuclear accent. The same caveat holds for cross-modal priming experiments which show that visual targets are responded to faster when they are contrastively related to the prime, but only when the prime is produced with a contrastive intonation contour (Braun & Tagliapietra, 2010; Husband & Ferreira, 2012), as well as for visual-world eye-tracking studies. For instance, the study by (Watson, Tanenhaus, & Gunlogson, 2008) used utterances such as Now move the candle/candy above the triangle (where the context rendered candle accessible, contrastive information and candy new information) but only controlled for the intonation of the target words (candle or candy, which were produced as L+H* or H*). However, prior research has shown that the intonation of the utterance-initial now is sensitive to discourse structure (Hirschberg & Litman, 1993) and affects fixation behaviour, at least for Dutch (Braun & Chen, 2012), which is why it is conceivable that the results by (Watson, Tanenhaus, & Gunlogson, 2008) may have been affected by the intonation of now in addition to the intonation of the target words.

The first goal of the current study was to design an experiment that allows us to attribute contrast interpretation to the referent in question (void of earlier or later prosodic information), cf. Experiment 1. We used the visual world eye-tracking paradigm with printed words (Braun & Chen, 2012; Dahan, Tanenhaus, & Chambers, 2002; Huettig, Rommers, & Meyer, 2011; Ito & Speer, 2008; McQueen & Viebahn, 2007; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Weber, Braun, et al., 2006). Participants heard neutral German declarative sentences (e.g., Der Turner hatte Blasen bekommen, 'The gymnast (subject) had gotten blisters') and saw four printed words on screen: a word that referred to a referent that was contrastively related to the subject referent (e.g., Tänzer 'dancer', henceforth contrastive associate), one that referred to a concept that was non-contrastively related to the subject referent (e.g., Sport 'sports', henceforth non-contrastive associate), the grammatical object of the utterance (e.g., Blasen 'blisters', which had to be clicked in all trials) and an unrelated distractor. In order to eliminate context effects, all utterances were presented in isolation, i.e., listeners had to rely on the intonational realization of the utterances to decode their information structure. In this experimental situation, the contrastive associate is visually available and identifiable but not explicitly mentioned in the prior context or inferable from it. We measured participants' fixations towards these referents while they processed the utterances. The decisive analysis window was the time period during which participants processed the subject
noun, because this is the time window during which they integrated the intonational information and the lexical information. As control condition, we used a broad focus realization (example (2A)) of the same sentence. We claim that participants activate alternatives when they fixate the contrastive associate more in an experimental condition compared to the control condition. Only in this case we can speak of contrast.

In Experiment 1, we investigated whether the accent type realized on a narrow focus constituent modulates the activation of alternatives (and hence the perception of contrast). Note that the term "activation" is understood here as shorthand for "consider as lexical or conceptual alternatives". In Experiment 1a, the subject noun was produced with a nuclear rising-falling accent on the subject (L+H* L- according to GToBI), followed by deaccentuation, a configuration that has been argued or shown to be contrastive in English (see Chafe, 1976; Couper-Kuhlen, 1984; Ito & Speer, 2008; Krahmer & Swerts, 2001; Pierrehumbert & Hirschberg, 1990; Selkirk, 2002; Watson, et al., 2008) and German (Baumann, et al., 2006; Kügler & Gollrad, 2015). In Experiment 1b, the subject noun was produced with a nuclear fall with an early peak (H+L* L- in GToBI), an accent that is well suited in conditions in which the referent is discourse-given and/or inferable from the context (Baumann & Grice, 2006; Kohler, 1991).\(^1\) If narrow focus equals contrast, we expect more fixations to the contrastive associate in both narrow focus conditions compared with the control condition. If, on the other hand, contrastive and non-contrastive focus constitute distinct categories, we expect more fixations to the contrastive associate in the nuclear L+H* condition (Exp. 1a) compared to the H+L* condition (Exp. 1b).

The second goal of the study is to compare the processing of intonational contrast to the processing of lexical items that presuppose the presence of alternatives. To this end, Experiment 2 tested the impact of utterance-initial additive particles, which associate with the accented constituent to their right in German (Büring & Hartmann, 2001; König, 1991; Reis & Rosengren, 1997). Interestingly, these particles have been argued to presuppose alternatives to the accented noun (e.g., Peter in Example (3)). The question is whether in online processing, auch (‘also’) leads to more fixations to contrastive alternatives, similar to the effect of contrastive accents and whether this effect is independent of the type of pitch accent of the subject noun. Preliminary evidence for this view comes from probe recognition experiments in German (Gotzner & Spalek, in press; Spalek, Gotzner, & Wartenburger, 2014): In utterances with focus-sensitive particles (also, only), participants are better at recognizing probes that are contrastive associates.

\(^1\) Note, however, that H+L* may signal contrast as well, but only when this accent follows a contrastive topic constituent (Braun, 2005).
Also Peter has a beer drank

(Somebody else drank a beer and) Peter drank a beer, too.

Based on the literature on auch, we predict that the particle will associate with the accented subject constituent, which leads to an increase in fixations to the contrastive associate, irrespective of the intonational realization of the subject. We used the same materials as in Experiment 1, but this time, the unstressed additive particle auch directly preceded the subject in initial position.

**Experiment 1**

Experiment 1 investigated whether different kinds of narrow focus realizations (contrastive vs. non-contrastive accent on the subject) lead to the same increase in fixations to contrastive associates as found in a broad focus control condition. In Experiment 1a, the subject was realized with a supposedly contrastive accent (nuclear L+H*), in Experiment 1b with a supposedly non-contrastive accent (nuclear H+L*). We monitored participants' fixations to the contrastive associate, while they heard isolated utterances in one of the three focus conditions mentioned above. Note that the use of printed words instead of object drawings allows us to test semantic associations without interference from visual similarity effects (Huettig & McQueen, 2007).

**Methods**

**Participants**

Eighty native speakers of German participated for a small fee. They were randomly assigned to Experiments 1a and 1b. In Experiment 1a (contrastive narrow focus accent), participants' age ranged between 18 and 29 years (average 21.5 years, 34 female, 6 male); in Experiment 1b (non-contrastive narrow focus accent), the age ranged between 19 and 33 years (average 25.7 years, 28 female, 12 male). The participants were unaware of the purpose of the experiment and had not taken part in experiments involving similar materials or intonation contours or in the web experiments that were used to construct the materials (see below). All participants reported to have normal hearing and had normal or corrected-to-normal vision.

**Materials**

Sentences and visual displays
The experiment was comprised of 48 trials, 24 experimental and 24 filler trials. All experimental utterances started with a subject constituent (see Table A1 in the Appendix), followed by a disyllabic auxiliary (wollte 'wanted', hatte 'had', konnte 'was able to', sollte 'should'), an object noun and a non-finite verb (e.g. Der Turner hatte Blasen bekommen 'The gymnast had gotten blisters'). All subject nouns had penultimate stress and between two and four syllables.

The words for the display in experimental trials were prepared as follows. For each subject noun, we selected one noun that was contrastively and one that was non-contrastively related to it (see Table A1). These data were gathered in two web experiments. The non-contrastive associate was collected in a free association task. Nineteen participants saw one noun at a time (e.g., gymnast), printed on screen, and they had to type in the first word that came to their mind (e.g., sports). To collect the contrastive associate, 24 participants (different from those of the other web experiment) saw a sentence fragment with a negated subject noun (e.g., 'Not the gymnast had gotten blisters but the...') and had to type in the most plausible continuation. For both the contrastive and the non-contrastive associates we chose the most frequent responses, if they differed from each other, were not onset competitors and had similar word lengths and lexical frequencies (factors that are known to affect fixation behaviour, cf. Dahan, Magnuson, & Tanenhaus, 2001; Kliegl, Grabner, Rolfs, & Engbert, 2004). In cases where the contrastive and non-contrastive associates were too different from one another in lexical frequency or number of characters, we chose a less frequently named associate as visual target. Average association strength, lexical frequency and number of characters of the selected contrastive and non-contrastive associates were matched across the two types of associates (see Table 1); grammatical gender could not be matched. Each experimental trial depicted the contrastive and non-contrastive associate, the grammatical object that had to be clicked, and a semantically unrelated distractor. The four words in any given experimental trial had comparable length and lexical frequency.

Filler trials served to distract participants from the presence of both contrastive and non-contrastive associates to the subject noun. Therefore, in filler trials, a noun that was non-contrastively related to the object noun of the utterance was used instead of a word which was non-contrastively related to the subject noun. In filler trials, the display thus showed the contrastive associate of the subject, the grammatical object that had to be clicked, a word that was non-contrastively related to the object noun, and an unrelated distractor. Like in target trials, the four words in filler trials had comparable length and lexical frequency.
<table>
<thead>
<tr>
<th></th>
<th>Contrastive associate</th>
<th>Non-contrastive associate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association strength (percentage)</td>
<td>30.3 (SD = 14.9)</td>
<td>27.9 (SD = 16.6)</td>
</tr>
<tr>
<td>Lexical frequency (occurrences per million)</td>
<td>1.5 (SD = 2.1)</td>
<td>4.6 (SD = 5.5)</td>
</tr>
<tr>
<td>Number of characters</td>
<td>7.0 (SD = 1.5)</td>
<td>6.0 (SD = 2.1)</td>
</tr>
</tbody>
</table>

Table 1. Average association strength, lexical frequency and number of characters (and standard deviations) of contrastive and non-contrastive associates to the subject nouns

Please insert Table 1 about here

Recordings.

The utterances were recorded by a phonetically trained female speaker of German in a sound-attenuated cabin at the University of Konstanz (44.1 kHz, 16 Bit). The speaker recorded all experimental sentences as triplets: with a contrastive focus accent and subsequent deaccentuation (nuclear L+H* L-, see Figure 1), with a non-contrastive focus accent and subsequent deaccentuation (nuclear H+L* L-, see Figure 2), and with a non-contrastive prenuclear rise (prenuclear L+H*, see Figure 3) with a subsequent slowly falling contour and a nuclear pitch accent (L+H*) on the object. The recordings were cued by appropriate context questions (as in Examples 1Q and 2Q). Note that the autosegmental-metrical description of the pitch accent types (prenuclear L+H* in the broad focus control condition and nuclear L+H* in the contrastive focus condition) suggests that the two conditions are very similar. Indeed they both constitute rising accents with the high tone being associated with the stressed syllable. Phonetically, however, there are very clear differences: the nuclear L+H* generally has a steeper pitch rise than the prenuclear L+H*, its peak is reached slightly earlier in the stressed syllable, and the fall from the peak is much steeper (see Table A2 for more details). For the sake of clarity in the text, we will use the modifiers 'prenuclear' (for the broad focus control condition) and 'nuclear' (for the contrastive narrow focus condition).

All filler trials were recorded in the broad focus condition, which is the most neutral and frequent contour in all-new contexts.

All sentences in the experiment were preceded by the prelude ‘Ich habe gehört’ (I have heard), to increase the preview time for the words in a natural way. This prelude was recorded once and spliced in front of all sentences with a pause of 1000ms between prelude and target.
Figure 1. Example realization of an item in the narrow focus condition in Experiment 1a (with nuclear L+H* L- on the subject noun). Here and in the other figures, F0 is shown between 100 and 350 Hz.

Figure 2. Example realization of an item in the non-contrastive focus condition in Experiment 1b (with nuclear H+L* on the subject noun).

Figure 3. Example realization of an item in the broad focus control condition in all experiments (with a prenuclear L+H* on the subject noun).
To validate the intonational realizations of the productions and the differences across conditions, we manually annotated the end of the subject-NP, the start and the end of the stressed syllable, as well as the low tonal target before the f0-rise and the f0-peak (Braun, 2006). From these annotations, we automatically extracted the alignment of the f0-minima (relative to the start of the stressed syllable), f0-maxima (relative of the end of the stressed syllable), the f0-range in semitones, as well as the duration of the stressed syllable, the duration of the f0-rise and the total duration of the subject-NP. The mean values and standard deviations for each of these measurements in the two intonation conditions are listed in Table A2 in the Appendix. The accents in the two narrow focus conditions had a consistently larger f0-excursion of the rise and the fall than accents in the broad focus condition.

Procedure
Participants were tested individually in a sound attenuated room at the University of Konstanz. They were instructed to listen to the utterances and to click on the object noun as quickly as possible. Participants sat at a distance of approximately 70 cm from a 20 inch LCD screen, so that they could freely move the computer mouse. They rested their chin on the provided chin rest. Their dominant eye was calibrated with an SMI Eyelink 1000 system (pupil and corneal reflection at a sampling rate of 250 Hz). The same sampling rate was used during trials. Auditory stimuli were presented via headphones (Beyerdynamics DT 990 PRO, 250 ohm) at a comfortable loudness. Each participant was presented with all 24 filler trials and all 24 experimental trials, but each experimental sentence was presented in only one of the two focus conditions (totalling in 12 trials with narrow focus and 12 with broad focus). Half of the participants heard the narrow focus with a contrastive accent, half with a non-contrastive accent. In sum, we hence had four basic lists. There were four pseudo-randomized versions of each basic list, such that there were at most three experimental trials in a row (but at most two of the same intonation condition). After each block of five trials, an automatic drift correction was initiated. In total, we had sixteen experimental lists, to which participants were randomly assigned (five participants for each list). The positions of the different types of printed words (contrastive and non-contrastive associate, object, distractor) was
varied such that across the experiment, each of the different types of printed words occurred equally frequently in each position (upper left and right, lower left and right).
Every trial started with a fixation cross which was shown until participants clicked on it. In all trials, the same token of the prelude (Ich habe gehört 'I have heard'; duration: 897ms) was used. This was followed by 1000ms of silence, after which the target utterance was auditorily presented. After participants had clicked on the respective object, there was a 1000ms inter-trial interval. Eye-movement data (fixations, blinks, saccades) were recorded throughout the experiment.

Results
The eye-tracking data were extracted in 4ms steps. The eye movement record was automatically parsed into saccades, fixations and blinks by the EyeLink software (using normal saccade sensitivity). Only fixations were further processed. They were automatically coded as pertaining to a given word if they fell within a rectangle of 100 x 100 pixels, centred on the middle of that word. The evolution of fixations to the words on screen over time is shown in Figure 4.
Figure 4. Fixation proportion over time for the two narrow focus conditions (top panel Experiment 1a, middle panel Experiment 1b) and the broad focus control condition (lower panel). The fixations to the broad focus control condition stem from Experiment 1a, but look identical for Experiment 1b. Gray shaded areas indicate standard error.

Please insert Figure 4 (a,b,c) about here.

Figure 4 indicates that the contrastive associate (red line) is fixated relatively more in the narrow focus condition with a nuclear L+H* accent than both in the narrow focus condition with a nuclear H+L* accent and in the broad focus control condition. For statistical analysis we analysed participants’ fixations in three lexically determined analysis windows. The first analysis window included fixations before participants processed the subject noun (from the start of the target sentence until 150ms after the start of the subject noun), the second while they processed the subject noun (150ms after the start of the subject noun until 150ms after the end of the subject noun) and the third while they processed the auxiliary (from 150ms after the start of the auxiliary until 150ms after the end of the auxiliary). The 150ms delay added to each acoustic landmark reflects the time that is needed to plan a saccade following auditory input (Matin, Shao, & Boff, 1993). The start and end of each analysis window was calculated for each item individually, based on manual annotations of the respective acoustic landmarks (start and end of the subject as well as the end of the auxiliary). The statistical analyses largely follow the proposal in Barr, Gann, and Pierce (2011). We calculated the empirical logits of fixations to the contrastive associate for each of the two analysis windows, dividing the fixations to that word by fixations that were directed elsewhere. A constant of 0.5 was added to both the denominator and the numerator. Empirical logits were analysed for Experiments 1a and 1b separately, using linear mixed effects regression models with focus condition (broad vs. narrow focus) as fixed factor and random intercepts for participants and items. The model further included random slopes for within-group factors. In case the model did not converge, the random slope for items was removed. P-values were calculated using the Satterthwaite approximation in the R-package lmerTest (http://cran.r-
For the visual displays, we further calculated fixation proportions (dividing fixations to a certain word type by all fixations in that time window).

<table>
<thead>
<tr>
<th></th>
<th>while processing the prelude</th>
<th>while processing the subject noun</th>
<th>while processing the auxiliary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1a (nuclear L+H*)</td>
<td>n.s.</td>
<td>p = 0.03</td>
<td>n.s.</td>
</tr>
<tr>
<td>(contrastive narrow focus accent vs. broad focus control condition)</td>
<td>(p = 0.57)</td>
<td>(p = 0.50)</td>
<td></td>
</tr>
<tr>
<td>Experiment 1b (nuclear H+L*)</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>(non-contrastive narrow focus accent vs. broad focus control condition)</td>
<td>(p = 0.41)</td>
<td>(p = 0.95)</td>
<td>(p = 0.56)</td>
</tr>
</tbody>
</table>

*Table 2. Overview of effects of focus condition on fixations to contrastive associate (elog) in the three analysis windows, split by Experiment*

Please insert Table 2 about here

*Experiment 1a (contrastive narrow focus accent: nuclear L+H*).

As predicted, there was an effect of focus condition on fixations to the contrastive associate in the analysis window in which participants processed the subject noun (see first row in Table 2): There were significantly more fixations to the contrastive associate in the narrow focus condition (average logits = -2.4) than in the broad focus control condition (average logits = -2.8, β = 0.36, 95% CI: [0.03;0.68], SE = 0.16, t = 2.15, p = 0.03). In the analysis windows before and after the subject noun, there was no effect of focus condition.

*Experiment 1b (non-contrastive narrow focus accent: nuclear H+L*).

There was no effect of focus condition on fixations to the contrastive associate, in any of the analysis windows (see second row in Table 2).

When combining the data of the two sub-experiments, we find an interaction between focus condition (narrow vs. broad) and experiment (β = 0.51, SE = 0.26, t = 1.95, p = 0.05). This interaction is visualized in Figure 5.
Discussion

The analysis of fixations showed that contrastive associates were fixated more when the subject noun was realized as a narrow focus than when it was part of a broad focus constituent, but only when the subject was produced with a contrastive pitch accent (nuclear L+H* L-) and not when it was produced with a non-contrastive accent (nuclear H+L* L-). Clearly, these results show that not every narrow focus is interpreted as contrastive. This finding provides strong evidence for theories that distinguish the concepts of narrow focus and contrast (e.g., Bartels & Kingston, 1994; Baumann, et al., 2006; Kiss, 1998; Neeleman, et al., 2009; Pierrehumbert & Hirschberg, 1990; Selkirk, 2002; Watson, et al., 2008). For German, a nuclear L+H* L- accent on a narrow focus constituent appears to be a necessary condition for the activation of contrastive alternatives (cf. Kügler & Gollrad, 2015).

This is the first study to show that neither preceding nor following prosodic information is needed to activate contrastive alternatives. In prior cross-modal priming studies by Braun and Tagliapietra (2010) and Husband and Ferreira (2012) and in earlier eye-tracking studies (Watson et al. 2008), the effects may also have been caused by prosodic features preceding the focused constituent; the same is true for the semantic congruency task used by Kügler and Gollrad (2015). In the memory studies by Fraundorf, Watson, and Benjamin (2010) and Spalek, et al. (2014), which had accents in phrase-
initial position, the effects may also have been caused by the prosodic realization of the remaining utterance. Our data show that the accentual information on the accented constituent alone is sufficient for the activation of alternatives; prior and following prosodic indicators may strengthen this effect but are clearly not mandatory. The activation of alternatives already occurred while participants were processing the subject nouns, which provides additional evidence for the immediate integration of visual, segmental-lexical and intonational information (Braun & Chen, 2012; Chen, et al., 2007; Dahan, et al., 2002; Ito & Speer, 2008; Watson, et al., 2008; Weber, Braun, et al., 2006; Weber, Grice, & Crocker, 2006). Note also that the alternatives tested in this paper were different from those in earlier eye-tracking studies, in which the critical contrast set had been mentioned in the prior discourse (as e.g., in Watson, et al., 2008). In our study, the contrast pair consisted of members of the same semantic category, in many other studies, the contrasts lie in colour or size only (as in Ito, Biblyk, Wagner, & Speer, 2014; Ito & Speer, 2008; Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Weber, Braun, et al., 2006).

In the present study, the differences in fixations across conditions were not large (on average 5% in the analysis window in which participants were processing the subject-NP), but consistent. The small effect size may be explained by the fact that we monitored fixations to words that were neither mentioned in the auditory input nor had a similar visual shape as the mentioned subject noun (Huettig & McQueen, 2007). The design hence resembles eye-tracking studies with target-absent trials (Huettig & Altmann, 2005; Huettig & McQueen, 2007) and we find similar effect sizes. For instance, in Huettig and McQueen (2007), the average proportion of fixations to a visually presented semantic competitor (relative to fixations to the competitor and the unrelated distractor) ranged from 0.52 (in the analysis window 200ms to 299ms after word onset) to 0.59 (between 600 and 699ms after word onset – given the length of the auditory target words in Appendix A, this appears to coincide approximately with the end of their target words). In the current experiment, this number is 0.57 in the broad focus control condition and 0.62 in the narrow focus condition. In Yee and Sedivy (2006), the differences in fixation proportions to semantically related pictures (key upon hearing lock) and unrelated words was approximately 7% (judging from their Figure 2). These comparisons show that the observed difference of 5% in the proportion of fixations across fixations is small, but comparable to related studies.

We would finally like to point out that the semantic contribution of the pitch accents appears to play a larger role than phonetic similarity. After all, the differences in fixations to the contrastive associate occurred in the conditions in which the accentual realizations of the subject were most similar to each other (nuclear L+H* L- vs. prenuclear L+H*, both realized with a pitch rise and a
subsequent fall) and not in the conditions with distinct realizations across conditions (Experiment 1b with nuclear H+L* vs. prenuclear L+H*). In German intonational phonology (and its labelling system GToBI, cf. Grice, Baumann, and Benzmüller (2005)), both accents are labelled with an L+H*; in the narrow focus condition, the accent is immediately followed by a low phrase accent (L-), while in the broad focus control condition, there is no phrase accent after the rising accents and f0 therefore slowly declines towards the low tonal target that is located just before the object noun. It is tempting to conclude that the part following the accent (off-ramp) has a stronger effect than the part preceding the accent (on-ramp) but since the accentual realizations differ in other aspects besides the off-ramp, the question of which phonetic cues are most relevant for the perception of contrast remains for future research (but see Kügler and Gollrad (2015) and Ritter and Grice (2015)).

We have now established the contribution of pitch accent type towards the activation of alternatives. In Experiment 2 we turn to the additive particle auch 'also', a lexical trigger for alternative sets.

**Experiment 2: Additive particles**

Experiment 2 tested the effect of the phrase-initial additive particle auch 'also, too' on the interpretation of prenuclear and nuclear accents (e.g., Auch der Turner hatte Blasen bekommen 'Also the gymnast had gotten Blisters'). Experiment 2a combined auch 'also' with a nuclear L+H* accent on the subject, Experiment 2b with a nuclear H+L* accent on the subject (in analogy to Experiments 1a and 1b).

**Participants**

Another group of eighty participants participated for a small fee. Half were assigned to Experiment 2a (contrastive accent condition, aged between 18 and 32 years, average 23.3 years, 29 female, 11 male), half to Experiment 2b (non-contrastive accent condition, aged between 19 and 33 years, average 25.7 years, 28 female, 12 male).

**Materials**

The sentences were the same as in Experiment 1, but the sentences with a narrow focus on the subject (contrastive and non-contrastive) and half of the filler items were recorded again to add the additive particle auch 'also' at the beginning of the sentence. Note that it was impossible to copy the same token of auch in front of the recorded utterances from Experiment 1 because the transition
sounded very unnatural. As in Experiment 1, the narrow focus was realized in two different ways, with a contrastive accent on the subject (nuclear L+H* L- , see Figure 6) and a non-contrastive accent (nuclear H+L* L-, see Figure 7). The displays and the experimental lists were the same as in Experiment 1, but the recordings including auch were used.

![Figure 6](image6.png)

**Figure 6. Example f0-track of the auch-condition in Experiment 2a (with a nuclear L+H* accent on the subject)**

![Figure 7](image7.png)

**Figure 7. Example f0-track of the auch-condition in Experiment 2b (with a nuclear H+L* accent on the subject)**

Insert Figure 6 and Figure 7 about here

**Procedure**

The testing procedure was identical to Experiments 1a and 1b. For the statistical analysis, the first analysis window (prior to the processing of the subject noun) was the same for the control trials and included the time from the onset of auch until 150ms after the onset of the subject noun for the sentences starting with auch. As a consequence the duration of the first analysis window is on average 100ms longer in the experimental conditions than in the broad focus control condition.

**Results**

The evolution of fixations in the two conditions is shown in Figure 8.
Figure 8. Evolution of fixations in the experimental condition of Experiment 2a (upper panel) and 2b (lower panel).

**Experiment 2a.**

While processing the subject-noun, participants fixated the contrastive associate more in narrow focus utterances containing *auch* than in broad focus utterances without it (β = 0.48, 95%CI: [0.01;0.78], SE = 0.2, t = 2.0,  p = 0.04). There were no effects of focus condition in the analysis window before the subject-DP (*auch* vs. silence) or during the auxiliary (see first row in Table 3).

<table>
<thead>
<tr>
<th></th>
<th>while processing the prelude</th>
<th>while processing the subject noun</th>
<th>while processing the auxiliary</th>
</tr>
</thead>
<tbody>
<tr>
<td><em><em>Experiment 2a (nuclear L+H</em>)</em>*</td>
<td>n.s.</td>
<td>p = 0.04</td>
<td>n.s.</td>
</tr>
<tr>
<td>(contrastive narrow focus accent vs. broad focus control condition)</td>
<td>(p = 0.91)</td>
<td></td>
<td>(p = 0.81)</td>
</tr>
<tr>
<td><em><em>Experiment 2b (nuclear H+L</em>)</em>*</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>(non-contrastive narrow focus accent vs. broad focus control condition)</td>
<td>(p = 0.52)</td>
<td>(p = 0.37)</td>
<td>(p = 0.17)</td>
</tr>
</tbody>
</table>

Table 3. Overview of effects of focus condition on fixations to contrastive associate in three analysis windows, split by Experiment

Please insert Table 3 about here
Experiment 2b (non-contrastive accent condition)

There was no effect of intonation condition on fixations towards the contrastive associate, in any of the analysis windows (see second row in Table 3).

To corroborate the differences in fixations to the contrastive associate in the analysis window in which participants were processing the subject noun between Experiments 2a and 2b statistically, we combined the data sets (as in Experiment 1). In the combined data set, there was a significant interaction between focus condition (broad focus control condition without particle and narrow focus condition with particle) and contrast condition (contrastive vs. non-contrastive pitch accent): β = 0.51, SE = 0.26, t = 1.95, p = 0.05.

![Figure 9. Average fixations to the contrastive associate during the processing of the subject noun (in elog), split by focus condition (broad vs. narrow) and Experiment. Smaller negative values indicate more fixations. Whiskers represent standard error.](image)

Discussion

The results of Experiment 2 (in which the narrow focus utterances were preceded by the unaccented additive particle auch 'also') are strikingly similar to the results of Experiment 1 (which did not contain the particle): Compared to the broad focus control condition, participants directed more fixations to the contrastive associate when they were processing a subject noun with a contrastive accent (nuclear L+H*), but not when the subject noun was produced with a non-contrastive accent (nuclear H+L*). The additive particle had little to no impact on the activation of referential alternatives. On the other hand, the semantic contribution of the accent type appears to be decisive for the activation of alternatives: nuclear L+H* does, and nuclear H+L* does not activate...
alternatives, irrespective of the presence or absence of the additive particle. This finding suggests that alternatives that are triggered by contrastive pitch accents are processed differently than alternatives due to the additive particle *auch* 'also'.

**General discussion and conclusion**

Compared to a broad focus control condition, listeners showed more fixations to contrastive associates to a mentioned referent if the referent was produced with a nuclear L+H* accent, but not when it was produced with a nuclear H+L* accent. The online processing of narrow focus constituents is hence strongly affected by pitch accent type. This finding is difficult to reconcile with the view that "focus in general indicates the presence of alternatives for interpretation" (Krifka, 2008, p. 25). What our data show is that alternatives do not play the same role in the processing of the two kinds of narrow focus constituents tested in this paper. While nuclear L+H* accents led to the activation of alternatives, nuclear H+L* accents did not. It would hence appear that nuclear L+H* is a prototypical contrastive pitch accent and that nuclear H+L* is non-contrastive, at least when realized on the phrase-initial subject noun (see below for more discussion on the context-dependency of intonational meaning). Note that Krifka also discusses a category "contrastive focus", which he reserves for corrections and some additive uses ('Peter wants coffee. [Mary] *Focus wants coffee, too"), since cross-linguistically, these contrastive foci lead to particular marking strategies. One way to reconcile the current data with his framework is to assume that – when heard out of context – the nuclear L+H* accents may have had a corrective interpretation ('The gymnast - and not someone else - had gotten blisters'), an interpretation that is hard to generate for the subject nouns realised with H+L*. This suggests that Krifka's difference between contrastive focus and (ordinary) focus is relevant for speech processing. Nevertheless, the claim that "focus in general indicates the presence of alternatives for interpretation" does not necessarily correspond to what listeners experience. From a speech processing perspective, it would be more appropriate to use the term 'contrastive' when alternatives are actually activated (in the case of nuclear L+H* and prenuclear L*+H accents), instead of assuming alternative sets for any focus constituents.

Regarding the prosody-semantics interface, i.e., the meaning contribution of particular pitch accent types, both experiments show that the same accent type (L+H*) is processed differently, depending on its position in the prosodic phrase (in particular, depending on whether it is followed by a low phrase accent (nuclear L+H*) or not (prenuclear L+H*)). Therefore, the general association between L+H* accents and contrast (e.g., Pierrehumbert & Hirschberg, 1990; Watson, et al., 2008) is difficult to maintain, at least for German. The influence of the following low phrase accent (and
the accompanying sudden drop in f0) has been noted before in connection with contrastive accents (e.g., Chafe, 1976; Couper-Kuhlen, 1984; Kügler & Gollrad, 2015). Note that the processing differences result from the accent type and not from its nuclear/prenuclear status and cannot be reduced to the dichotomy between nuclear and prenuclear accents alone, since not every nuclear accent had the potential to activate alternatives. It seems that neither the L+H* nor the low phrase accent (L-) alone can account for the contrastive interpretation. Instead, the accent together with its immediate tonal environment (or its position in the prosodic phrase) seem crucial for interpretation; see Selkirk (2002). This positional effect on interpretation applies for both accent types: L+H* and H+L*. The current data show that H+L* does not lead to the activation of alternatives, which makes it a good candidate for a non-contrastive pitch accent. However, we are not in a position to fully generalize this finding to other utterance positions. For instance, German speakers often produce (and expect) this accent in contexts in which both the topic and the focus are contrasted (Braun, 2006; Braun & Asano, 2013). Clearly, follow-up experiments are necessary to get a better insight into the context-dependency of intonational meaning.

The fact that the presence of the additive particle auch 'also' has no effect is surprising, at first, since given its semantic meaning, its use in production presupposes the existence of alternatives to the accented item (Dimroth, 2004; Jacobs, 1983; König, 1991; König, 1994). Likewise, an additive particle is only felicitous for the listener, if there is a salient alternative to the accented item in the context or common ground. We see two possible explanations for why auch did not affect fixations over and above the effects of pitch accent type reported above. First, it is possible that alternatives evoked by contrastive accents have a different mental representation than alternatives triggered by additive particles, which in turn leads to differences in online processing. Second, intonational information may be processed more immediately than lexical information, so that alternatives triggered contrastive accents are available earlier than alternatives due to additive particles. An argument in support of the latter option is that the interpretation of additive particles is dependent on the immediate linguistic context (which may or may not contain the associated constituent) and on a local semantic representation of the clause, while this is not the case for contrastive pitch accents. In any case, the observed differences in processing contrastive pitch accents and focus particles are in line with Gotzner, Spalek, and Wartenburger (2013). They tested the recognition of probes that were contrastively related to word produced with contrastive L+H* accent or non-contrastive H* accent and found that contrastively related words were recognized faster in the contrastive accent condition. Probe recognition was slowed down by adding the particles also and
only, however. They argued that contrastive accentuation increases the salience of contrastive alternatives, while focus-sensitive particles increased processing costs. The current paradigm has provided us with insights on the activation of alternatives in different information structure configurations as the utterance unfolds over time. In future studies we plan to use this paradigm to study intonational meaning contrasts that are signalled less categorically than in the current series of experiments, but are characterized instead by more fine-grained (gradient) phonetic differences in tonal alignment, scaling and intensity (Baumann, et al., 2006; Braun, 2006; Féry & Kügler, 2008; Görs & Niebuhr, 2012; Niebuhr, 2010). This will help us to get a better understanding of the aspects of an accent (slope of the rise and fall, relative or absolute alignment of tonal targets, intensity) that are relevant for its interpretation as contrastive accent. Furthermore, we will investigate the processing of contrastive topic accents in German, which are particularly interesting because it involved the comparison of two kinds of prenuclear accents (contrastive vs. non-contrastive rising accents), with an intonational contrast that is hard to represent phonologically (Braun, 2004, 2005, 2006; Zellers & Post, 2012; Zerbian, Turco, Schauffler, Zellers, & Riester, 2012). Another direction for future research is to study how listeners cope with variability in marking information-structural differences, such as variability due to differences in regional variation or speaker idiosyncracies (Atterer & Ladd, 2004; Braun, 2007; Zerbian, et al., 2012). Furthermore, in future studies it would be worthwhile to use this paradigm to investigate other kinds of focus-sensitive particles (exclusive 'only' or inclusive 'also', as tested in Kim, Gunlogson, Tanenhaus, & Runner, in press using the visual-world paradigm; Spalek, et al., 2014 with memory tasks) to document the time-course of alternative activation and suppression.
Acknowledgements
We thank Sophie Egger, Clara Huttenlauch, Angela James, Verena Köppel, Jana Schlegel and Katharina Zahner for support in stimulus preparation and testing as well as Nicole Gotzner, Filippo Domaneschi and Giusy Turco for valuable discussion and comments on methodology and earlier versions of this paper.

Funding
This work was supported by a grant from the German Research Foundation (DFG) awarded to the first author [BR-3428/1-1].
References


### Appendix

<table>
<thead>
<tr>
<th>Subject noun</th>
<th>Contrastive associate</th>
<th>Non-contrastive associate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schwimmer (swimmer)</td>
<td>Taucher (53) (diver)</td>
<td>Bad (15) (baths)</td>
</tr>
<tr>
<td>Turner (gymnast)</td>
<td>Tänzer (5) (dancer)</td>
<td>Sport (15) (sports)</td>
</tr>
<tr>
<td>Nonne (nun)</td>
<td>Mönch (32) (monk)</td>
<td>Kloster (50) (abbey)</td>
</tr>
<tr>
<td>Artistin (artist)</td>
<td>Clown (37) (clown)</td>
<td>Zirkus (55) (circus)</td>
</tr>
<tr>
<td>Italiener (Italian)</td>
<td>Spanier (37) (Spaniard)</td>
<td>Spaghetti (10) (spaghetti)</td>
</tr>
<tr>
<td>Japaner (Japanese)</td>
<td>Chinese (37) (Chinese)</td>
<td>Asien (10) (Asia)</td>
</tr>
<tr>
<td>Kunde (customer)</td>
<td>Verkäufer (16) (shop assistant)</td>
<td>Geschäft (30) (shop)</td>
</tr>
<tr>
<td>Segler (sailor)</td>
<td>Kapitän (21) (captain)</td>
<td>Boot (20) (boat)</td>
</tr>
<tr>
<td>Mieter (tenant)</td>
<td>Nachbar (32) (neighbour)</td>
<td>Wohnung (35) (apartment)</td>
</tr>
<tr>
<td>Professor (professor)</td>
<td>Student (58) (student)</td>
<td>Universität (30) (university)</td>
</tr>
<tr>
<td>Schreiner (carpenter)</td>
<td>Tischler (11) (cabinet maker)</td>
<td>Holz (40) (wood)</td>
</tr>
<tr>
<td>Direktor (director)</td>
<td>Sekretär (16) (secretary)</td>
<td>Schule (45) (school)</td>
</tr>
<tr>
<td>Züchter (breeder)</td>
<td>Bauer (32) (farmer)</td>
<td>Tiere (50) (animals)</td>
</tr>
<tr>
<td>Subject</td>
<td>Contrastive Associate</td>
<td>Non-Contrastive Associate</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Sänger</td>
<td>Techniker (68)</td>
<td>Lieder (30)</td>
</tr>
<tr>
<td>(singer)</td>
<td>(technician)</td>
<td>(songs)</td>
</tr>
<tr>
<td>Maler</td>
<td>Zeichner (21)</td>
<td>Farben (30)</td>
</tr>
<tr>
<td>(painter)</td>
<td>(draftsman)</td>
<td>(paint)</td>
</tr>
<tr>
<td>Schlagzeuger</td>
<td>Gitarist (21)</td>
<td>Band (45)</td>
</tr>
<tr>
<td>(drummer)</td>
<td>(guitarist)</td>
<td>(band)</td>
</tr>
<tr>
<td>Schafe</td>
<td>Ziegen (21)</td>
<td>Herde (25)</td>
</tr>
<tr>
<td>(sheep)</td>
<td>(goats)</td>
<td>(flock)</td>
</tr>
<tr>
<td>Biene</td>
<td>Wespe (42)</td>
<td>Honig (25)</td>
</tr>
<tr>
<td>(bee)</td>
<td>(wasp)</td>
<td>(honey)</td>
</tr>
<tr>
<td>Flamingo</td>
<td>Pelikan (16)</td>
<td>Vogel (15)</td>
</tr>
<tr>
<td>(flamingo)</td>
<td>(pelican)</td>
<td>(bird)</td>
</tr>
<tr>
<td>Wale</td>
<td>Haie (16)</td>
<td>Orcas (5)</td>
</tr>
<tr>
<td>(whales)</td>
<td>(sharks)</td>
<td>(orcas)</td>
</tr>
<tr>
<td>Frauchen</td>
<td>Herrchen (42)</td>
<td>Hund (45)</td>
</tr>
<tr>
<td>(mistress)</td>
<td>(master)</td>
<td>(dog)</td>
</tr>
<tr>
<td>Tiger</td>
<td>Löwe (58)</td>
<td>Streifen (10)</td>
</tr>
<tr>
<td>(tiger)</td>
<td>(lion)</td>
<td>(stripes)</td>
</tr>
<tr>
<td>Rehe</td>
<td>Hirsche (15)</td>
<td>Wald (20)</td>
</tr>
<tr>
<td>(deer)</td>
<td>(stags)</td>
<td>(forest)</td>
</tr>
<tr>
<td>Geiger</td>
<td>Pianist (21)</td>
<td>Violine (15)</td>
</tr>
<tr>
<td>(violinist)</td>
<td>(pianist)</td>
<td>(violin)</td>
</tr>
</tbody>
</table>

Table A1. **Subject noun together with contrastive and non-contrastive associate. The number in brackets refers to the percentage of participants that named this associate in the web experiment**
<table>
<thead>
<tr>
<th></th>
<th>BF prenucl. L+H*</th>
<th>NF nucl L+H* (Exp 1a)</th>
<th>NF nucl H+L* (Exp 1b)</th>
<th>auch + NF nucl L+H* (Exp 2a)</th>
<th>auch + NF nucl. H+L* (Exp 2b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-alignment with respect to start of stressed syllable in ms</td>
<td>-25.9 (48.5)</td>
<td>-22.5 (40.1)</td>
<td>146.5 (51.1)</td>
<td>-17.8 (13.3)</td>
<td>130.7 (32.7)</td>
</tr>
<tr>
<td>H-alignment with respect to end of stressed syllable in ms</td>
<td>-45.3 (33.2)</td>
<td>-34.2 (41.8)</td>
<td>-322.0 (57.5)</td>
<td>-35.5 (32.1)</td>
<td>-256.6 (48.5)</td>
</tr>
<tr>
<td>F0-range of the pitch rise in semitones</td>
<td>5.9 (1.1)</td>
<td>9.0 (1.1)</td>
<td>9.1 (0.9)</td>
<td>7.3 (1.2)</td>
<td>6.7 (0.9)</td>
</tr>
<tr>
<td>F0-minimum before the pitch rise in Hz</td>
<td>191.4 (8.1)</td>
<td>169.3 (4.5)</td>
<td>159.2 (7.6)</td>
<td>175.0 (8.1)</td>
<td>170.8 (6.8)</td>
</tr>
<tr>
<td>Duration of the stressed syllable in ms</td>
<td>247.5 (37.3)</td>
<td>257.8 (53.8)</td>
<td>267.1 (49.8)</td>
<td>247.0 (36.0)</td>
<td>237.9 (41.2)</td>
</tr>
<tr>
<td>Duration of the subject-NP in ms</td>
<td>421.0 (72.4)</td>
<td>439.4 (98.6)</td>
<td>458.9 (102.3)</td>
<td>431.5 (93.5)</td>
<td>417.9 (91.0)</td>
</tr>
<tr>
<td>F0-range from peak to end of the modal verb in st</td>
<td>6.0 (1.3)</td>
<td>10.3 (1.3)</td>
<td>8.8 (0.7)</td>
<td>9.5 (1.0)</td>
<td>7.1 (1.2)</td>
</tr>
</tbody>
</table>

Table A2. Mean values and standard deviations of acoustic measures for the conditions in all experiments