

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

This study investigates how 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 saw four printed words displayed on screen: one that was a contrastive alternative to the subject noun (e.g., *diver*), one that was non-contrastively related (e.g., *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 noun (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 may trigger alternatives itself. It associated with the accented subject. 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*) 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 the processing of 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 intonational, morphological and syntactic means (e.g., Krifka, 2008; Vallduví & Engdahl, 1996). The same string of words can therefore signal different focus-background structures, as shown in the German examples (1A) and (2A). In (1A), the subject will be intonationally marked as narrow focus (typically with a nuclear accent) and in (2A) as part of a broad focus domain (typically with a prenuclear accent).

- (1) Q: Wer aß die Kekse?
Who ate the cookies?
A: [Der Junge]_F aß die Kekse
[The boy]_F ate the cookies.
- (2) Q: Was ist passiert?
What is happened?
'What happened?'
A: [Der Junge aß die Kekse]_F
[The boy ate the cookies]_F

The past approximately 20 years have shown an increasing interest in how intonational information is processed online as the utterance unfolds over time (Chen, et al., 2007; Dahan, et al., 2002; Dennison & Schafter, 2010; Esteve-Gibert, et al., 2016; Husband & Ferreira, 2012; Ito & Speer, 2008; Watson, et al., 2008; Weber, Braun, et al., 2006). In a seminal study, Dahan, et al. (2002) investigated the effect of accentuation on reference resolution using the visual world eye tracking paradigm. Participants heard two instructions: In the first instruction, they were asked to move an object in a display (e.g. the *candle* in *Put the candle above the triangle*); according to a second instruction they had to move either the same object again (*candle*) or a lexical cohort competitor (*candy*). Object and competitor were either accented (H* or L+H*) or unaccented, resulting in four conditions. Before the cohort competitors were disambiguated segmentally, participants fixated the competitor *candy* more when the noun was accented, suggesting that listeners immediately exploited the relation between pitch accents and discourse for referent resolution. In the psycholinguistics literature, accented referents have received particular attention, since they seem to be interpreted as contrastive, leading to a temporary activation of contrastive alternatives (see below for details). The current study investigates whether this holds true for different pitch accent types

(in particular, nuclear L+H*, a rising-falling contour, and nuclear H+L*, a falling contour with an early peak), whether the activation of alternatives is driven immediately by these pitch accents, and whether the interpretation of pitch accents is affected by the presence of additive particles, which lexically induce alternatives.

Background

Focus vs. contrast

Focus is generally defined as a constituent that corresponds to the open proposition in the question, e.g., *the boy* in (1A) and *the boy ate the cookies* in (2A). In the semantic literature, narrow focus is defined in terms of the presence of alternatives that are relevant for interpretation (e.g., Krifka, 2008; Rooth, 1992), e.g., {*the boy, the girl, Marina,....*} in 1A and {*the boy ate the cookies, the girl studied algebra, ...*} in 2A. This semantic definition is close to what many researchers associate with the notion of "contrast". In fact, the relation between the categories focus and contrast is not entirely clear (Krahmer & Swerts, 2001; Molnár, 2001; Repp, 2010). For some authors, the notion of contrast is synonymous to (narrow) focus (e.g., Bolinger, 1961; Jackendoff, 1972; Krifka, 2008; Lambrecht, 1994; Rooth, 1992; Wong & Diehl, 2003). In the prosodic literature on the other hand, 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, et al., 2006; Kiss, 1998; Neeleman, et al., 2009; Pierrehumbert & Hirschberg, 1990; Selkirk, 2002; Sudhoff, 2010; Watson, et al., 2008 for studies on German and English). For instance, the German example (1A) can be uttered with a rising-falling accent on the subject noun (nuclear L+H* accent) to signal that the referent is contrastive (Baumann, et al., 2006; Kügler & Gollrad, 2015) or with a falling accent with an early-peak (nuclear H+L*) to signal that it is inferable (Baumann & Grice, 2006; Kohler, 1991), see *Figure 1*. In both cases, the subject noun is narrowly focused (it is the sole accent in the phrase and it is suitable as an answer to a question focusing on the subject). In broad focus, the subject is produced with a prenuclear L+H*, followed by a nuclear accent later in the utterance.

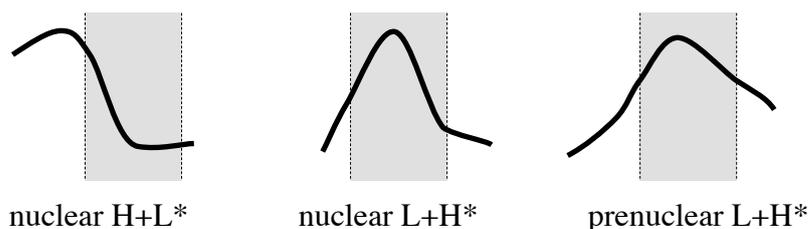


Figure 1. Stylization of nuclear H+L*, nuclear L+H* and prenuclear L+H* accents. The shaded area indicates the stressed syllable.

Experimental evidence from online processing speaks in favour of the distinction between contrastive and non-contrastive focus, based on differences in pitch accent type (Braun, 2005, 2006; Braun & Tagliapietra, 2010; Chen, et al., 2007; Husband & Ferreira, 2012; Kügler & Gollrad, 2015; Watson, et al., 2008). For instance, Braun and Tagliapietra (2010) conducted a cross-modal priming experiment in Dutch to test the activation of alternatives to words in narrow focus (as in Example 1) and in broad focus contexts (as in Example 2). Information structure was signalled prosodically, using different pitch accent types. The prime occurred always in utterance-final position. In one condition, the prime word was realized as a contrastive focus (nuclear H*+L), preceded by another focal accent (prenuclear H*+L) earlier in the utterance (resulting in a perceived information structure like [HENK]_F photographed a [FLAMINGO]_F); in the other it was realized as the focus exponent of a broad focus utterance (nuclear !H+L*), preceded by a prenuclear (L+H*) accent on the subject (resulting in the information structure [Henk photographed a FLAMINGO]_F). At the offset of the prime, participants saw visual target words that were contrastively related to the utterance-final prime (e.g., *pelican*) or a non-contrastively related word (e.g., *pink*). Compared to an unrelated auditory control condition, lexical decision times to the contrastive alternatives were faster when the prime had a contrastive focus accent (nuclear H*+L) and there was no difference to the control condition when the prime was part of a broad focus and realized with a nuclear !H+L*. Reaction times to the contrastively related words were always faster than in the control condition, independent of intonational realization. These priming data suggest that Dutch listeners indeed activated contrastive alternatives to narrowly focused prime words but not to the same words in a broad focus realization.

However, experimental processing studies of this kind allow alternative interpretations. First, studies collecting participants' responses at the offset of the utterance bear the risk that the responses are also influenced by other accents in the clause, for instance accents that come later (cf. Braun, 2005) or earlier (cf. Kügler and Gollrad, 2015; Braun and Tagliapietra, 2010; Watson, et al., 2008) than the accents the studies focus on. To illustrate this 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 the cross-modal priming

experiment discussed above among others (cf. also Husband & Ferreira, 2012), and for the visual-world eye-tracking studies. The above-mentioned study by Dahan et al. (2002), for instance, used utterances such as *Now move the candle/candy above the triangle*. However, prior research has shown that the intonation of the utterance-initial *now* is sensitive to discourse structure (Hirschberg & Litman, 1993) and affects the interpretation of the subsequent referent, at least for Dutch (Braun & Chen, 2012)¹, which is why it is conceivable that Dahan et al.'s (2002) results may have been affected by the intonation of *now* in addition to the intonation of the target words.

It is an open question whether certain accents alone may indeed render a referent contrastive. This can be tested using online studies, in which the respective accents appear in phrase-initial position. Based on the experimental prosodic evidence, we take the assumption that there is a distinction between contrastive and non-contrastive focus as our working hypothesis for the experiment. The first goal of this paper is to test whether a contrastive interpretation is also available based on the accent type alone, i.e., without prosodic information that is available before or after the (contrastively or non-contrastively) accented referential expression (Experiment 1). We used the visual world eye-tracking paradigm (Dahan, et al., 2002; Huettig, et al., 2011; Ito & Speer, 2008; Tanenhaus, et al., 1995; Weber, Braun, et al., 2006), but with printed words instead of objects (Braun & Chen, 2012; McQueen & Viebahn, 2007). Participants heard neutral German declarative sentences (e.g., *Der Turner hatte Blasen bekommen*, 'The gymnast had gotten blisters') and saw four printed words on a screen: (i) a noun that referred to a referent which was contrastively related to the referent of the subject noun (e.g., *Tänzer* 'dancer', henceforth contrastive associate), (ii) a noun that referred to a concept which was non-contrastively related to the subject referent (e.g., *Sport* 'sports', henceforth non-contrastive associate), (iii) the grammatical object of the utterance (e.g., *Blasen* 'blisters', which had to be clicked in all trials) and (iv) an unrelated distractor. All utterances were presented in isolation, i.e., listeners had to rely on the intonational realization of the utterances to decode the information structure and the contrastive or non-contrastive interpretation of the subject accent. 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 contrastive associates while they were processing the utterances. The critical analysis window was the time period during which participants processed the subject noun, because this was the time window during which they integrated the intonational

¹ Note that Ito and Speer (2008) did not find an effect of the intonation of the discourse particle *now* in English (their Experiment 3), a discrepancy that has been discussed in some detail in Braun and Chen (2012).

information and the lexical information. As control condition, we used a broad focus realization (comparable to example 2A) of the same sentence.

In Experiment 1, we also investigated whether the accent type realized on a narrow focus constituent affects the activation of contrastive alternatives (and hence the perception of contrast). Note that the term "activation" is understood here as shorthand for "consider as lexical or conceptual alternatives" and is operationalized by more fixations to the contrastive alternatives in the narrow focus condition compared to the broad focus control condition. In Experiment 1a, the subject noun was produced with a nuclear rising-falling accent on the subject (L+H* L- according to GToBI, cf. Grice, et al., 2005), followed by deaccentuation, a configuration that has been argued or shown to be contrastive in German (Baumann, et al., 2006; Kügler & Gollrad, 2015) and English (see Chafe, 1976; Couper-Kuhlen, 1984; Ito & Speer, 2008; Krahmer & Swerts, 2001; Pierrehumbert & Hirschberg, 1990; Selkirk, 2002; Watson, et al., 2008). 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).² The H+L* accent is not as frequent as the L+H* accent in German, but it is not infrequent either: There is corpus work on German appointment-making dialogues (KIEL corpus), which has quantitative data on the occurrence of different pitch accent types (Peters, et al., 2005). In utterances with a single pitch accent (as the ones we deal with here), the early-peak accent (H+L*) accent occurred in 30% of the accents that were realized turn-internally and in 50% of the accents that were realized in turn-final position. The medial-peak accent (L+H*), on the other hand, occurred in 47% of the accents in turn-medial position and in 49% of the accents in turn-final position.

Additive particles

The third 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 the utterance-initial additive particle *auch* ('also'), which associates with the accented constituent to its right (Büring & Hartmann, 2001; König, 1991; Reis & Rosengren, 1997 for German). Interestingly, particles like *auch* have been argued to presuppose alternatives to the accented noun (e.g., *Peter* in Example 3) and are hence expected to activate alternatives (Krifka, 1991).

² Note that H+L* may signal contrast as well, but only when this accent follows a contrastive topic constituent (Braun, 2005).

(3) Auch [PETER]_F hat ein Bier getrunken.

Also Peter has a beer drank

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

In a series of production and perception studies in German, Sudhoff (2010) investigated whether focus sensitive particles, specifically *auch* ('also') and *sogar* ('even'), induce contrastive focus, or whether they interact with the focus-background structure specified by the context. He analyzed the prosodic realization of nouns which were embedded in sentences with or without focus particles, the sentence in turn being embedded in contrastive and non-contrastive contexts. Sudhoff showed that in production experiments, the prosodic realization was affected by context, but not by presence or absence of a focus-sensitive particle. In perception, listeners did not assign more contrastive interpretations to utterances with focus-sensitive particles. Based on these data, he argues that focus particles interact with the information-structure of the utterance, but do not impose contrast on their own.

In an online processing study on English, Schwarz (2015) showed that additive particles facilitate reference resolution. In his Experiment 1, participants saw visual displays depicting two boys and two girls, each of whom held one or two objects in their hands (e.g., a knife and a plate, a spoon). Participants heard instructions like '*One of the boys is holding a spoon. Click on the girl who is also holding a spoon*' (with an unaccented 'also') and had to click on the respective persons. In the critical displays only the target person held the same object as the person mentioned in the preceding context sentence (in addition to holding another object), in control displays two other persons held the same object as the person mentioned in the preceding context sentence (in addition to each holding another object). In this setting, the particle 'also' disambiguated the referent in critical displays but not in control displays. Analysis of fixations showed that the participants fixated the target picture earlier in critical trials compared to control trials. Unfortunately, not much information is provided about the prosodic realization of the auditory stimuli.

In a truth-value judgement task in German, Gotzner (2016) presented contexts as in (4) and had participants judge whether they had heard sentences such as (5, here translated into English) in the experiment or not. They manipulated the presence of focus-sensitive particles (*also*, *only*, \emptyset) and the pitch accent type of the associated constituent:

(4) Der Richter und der Zeuge folgten dem Argument. Nur/Auch/ \emptyset der Richter_(H*/L+H*)

The judge and the witness followed the argument. Only/Also/∅ the judge_(H*/L+H*)

glaubte dem Angeklagten. Er verkündete das Urteil.

believed the defendant. He announced the verdict.

(5) Der Zeuge glaubte dem Angeklagten.

The witness believed the defendant.

Their results showed that for sentences with 'also', the accent type of the associated constituent did not affect the judgments. Participants correctly indicated that sentence (5) was true in about 75% of the cases (judging from their Figure 3). For sentences without focus particles, accent type affected the truth-value judgments: participants more often falsely reported that (5) was true for H* accents (ca. 35%) than for L+H* accents (ca. 20%). In other words, accent type only played a role when there was no additive particle.

Taken together, previous results are slightly contradictory: while Sudhoff (2010) argues on the basis of perception data that accent type but not the presence of focus particles matters for the perception of contrast, Gotzner et al. (2016) conclude on the basis of a truth value judgment task that focus particles affect their results more than accent type. The only online processing study (Schwarz, 2015) shows that information on focus particles is used immediately when it is helpful for disambiguation, but there is not much information about the prosodic realization of the auditory stimuli. Regarding the additive particle, our question is whether in online processing, *auch* ('also') adds to the distinctive force of the pitch accent, or whether the particle may overwrite potential effects of pitch accents. For Experiment 2, we used the same materials as in Experiment 1, but this time, the unstressed additive particle *auch* directly preceded the subject constituent in the narrow focus conditions. Based on the literature review, we formulate the following hypotheses for Experiments 1 and 2.

Hypothesis 1: If the focus-semantic account holds (narrow focus equals contrast), we expect more fixations to the contrastive associate in both narrow focus conditions (i.e., in the nuclear L+H* condition in Experiment 1a and in the nuclear H+L* condition in Experiment 1b compared to the control condition). We expect no differences in fixations to the non-contrastive associate and the distractor.

Hypothesis 2: If the prosodic account holds (only 'contrastive' narrow focus equals contrast), we expect more fixations to the contrastive associate in the nuclear L+H* condition (Exp. 1a)

compared to the control condition, but no differences in the H+L* condition (Exp. 1b) compared to the control condition. We expect no differences in fixations to the non-contrastive associate and the distractor.

Hypothesis 3: Based on the semantic literature and Schwarz' (2015) processing data, we predict that the additive particle will lead to the activation of alternatives to the accented subject constituent, irrespective of the type of accent on the subject. We therefore expect more fixations to the contrastive associate compared to the broad focus control condition without a particle in both Experiments 2a (with a nuclear L+H* accent) and 2b (with a nuclear H+L* accent). We expect no differences in fixations to the non-contrastive associate and the distractor.

Hypothesis 4: Based on Sudhoff's (2010) findings, we do not expect the particle to contribute to the contrastive or non-contrastive interpretation of these sentences. Instead, the interpretation is based on the accent type on the subject, leading to more fixations to the contrastive associate in the nuclear L+H* condition (Exp. 2a) compared to the broad focus control condition, but no differences in Exp. 2b. We expect no differences in fixations to the non-contrastive associate and the distractor.

Hypotheses 1 and 2, as well as Hypotheses 3 and 4, are in part contradictory, which is due to the contradictory results in previous literature outlined above. These contradictions will be resolved by our experimental results and discussed accordingly. Hypotheses 1 and 2 are addressed in Experiment 1, Hypotheses 3 and 4 in Experiment 2.

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 and whether this effect happens immediately upon processing the accented word. Focus condition was manipulated between-subjects: In Experiment 1a, the subject was realized with a contrastive accent (nuclear L+H*), in Experiment 1b with a non-contrastive accent (nuclear H+L*). We monitored participants' fixations to the contrastive associate, while they heard context-free utterances in one of the three focus conditions mentioned above. Note that the use of printed words instead of object drawings allowed 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 Experiment 1a or 1b. In Experiment 1a (contrastive narrow focus accent L+H*), participants were between 18 and 29 years of age (average 21.5 years, 34 female, 6 male); in Experiment 1b (non-contrastive narrow focus accent H+L*), 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, consisting of 24 experimental and 24 filler trials. All experimental utterances started with a subject constituent (see Table A1 in the Appendix for subject nouns used in experimental trials; they were preceded in the experimental utterances by definite articles), followed by a disyllabic auxiliary (*wollte* 'wanted', *hatte* 'had', *konnte* 'was able to', *sollte* 'should'), an object noun (without definite article) 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 related 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 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 first 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 making sure that 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, et al., 2001; Kliegl, et al., 2004). In cases in which the contrastive and non-contrastive associates were too different from one another in lexical frequency or number of characters (resulting in significant differences across groups), 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, see Table 1. For each experimental trial, one contrastive and one non-contrastive associate, the grammatical object that had to be clicked, and a semantically unrelated distractor were used. The four words in any given experimental trial had comparable length and lexical frequency..

	Contrastive associate	Non-contrastive associate	t-value (df), p-value
Association strength (percentage)	30.3 (<i>SD</i> = 14.9)	27.9 (<i>SD</i> = 16.6)	t(23) = 0.5, p = 0.6
Lexical frequency (occurrences per million)	1.5 (<i>SD</i> = 2.1)	4.6 (<i>SD</i> = 5.5)	t(23) = 1.6, p = 0.1
Number of characters	6.8 (<i>SD</i> = 1.5)	5.9 (<i>SD</i> = 1.9)	t(23) = 1.6, p = 0.1

Table 1. Average association strength, lexical frequency and number of characters (and standard deviations) of contrastive and non-contrastive associates to the subject nouns. The last column shows the results of a paired t-test.

Please insert *Table 1* about here

The filler trials were aimed at distracting participants from the presence of contrastive and non-contrastive associates to the subject noun. Therefore, in filler trials (e.g., 'The pensioner had blocked the train beforehand. '), the display showed two related words (e.g., *book* and *page*) which were unrelated to the words in the sentence, the target that had to be clicked (e.g., *train*) and an unrelated distractor (e.g., *laundry*). As in the target trials, the four words in the filler trials had comparable length and lexical frequency.

Recordings

The utterances were spoken by a phonetically trained female speaker of German and recorded in a sound-attenuated cabin at the University of Konstanz (44.1 kHz, 16 Bit). The speaker recorded all experimental sentences as triplets: (i) with a contrastive focus accent and subsequent deaccentuation (nuclear L+H* L-, see *Figure 2*), (ii) with a non-contrastive focus accent and subsequent deaccentuation (nuclear H+L* L-, see *Figure 3*), and (iii) with a non-contrastive prenuclear rise (prenuclear L+H*) with a slowly falling following contour and a nuclear pitch accent (L+H*) on

the object (see Figure 4).³ The recordings were cued by appropriate context questions (as in Examples 1Q and 2Q).

Please insert Figures 1-3 about here.

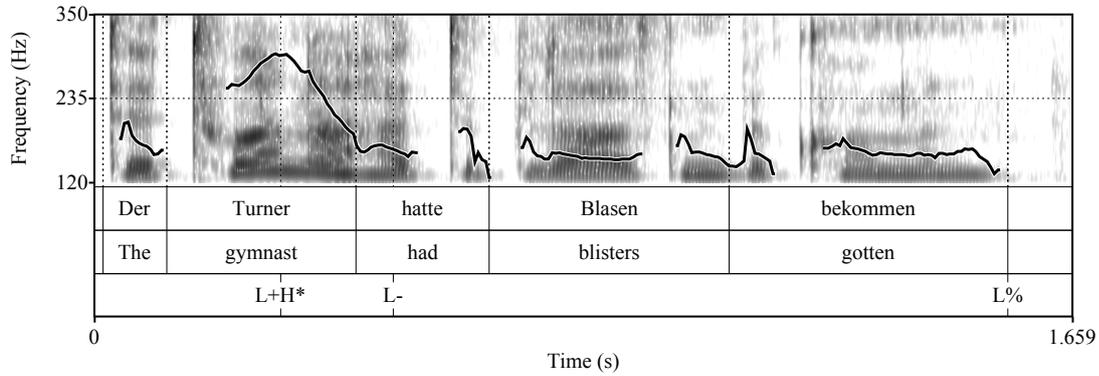


Figure 2. Representative example realization of an item in the narrow focus condition in Experiment 1a (with nuclear L+H* L- on the subject noun). Note that the L- phase accent has a double association with the right boundary and the stressed syllable following the nuclear accent.

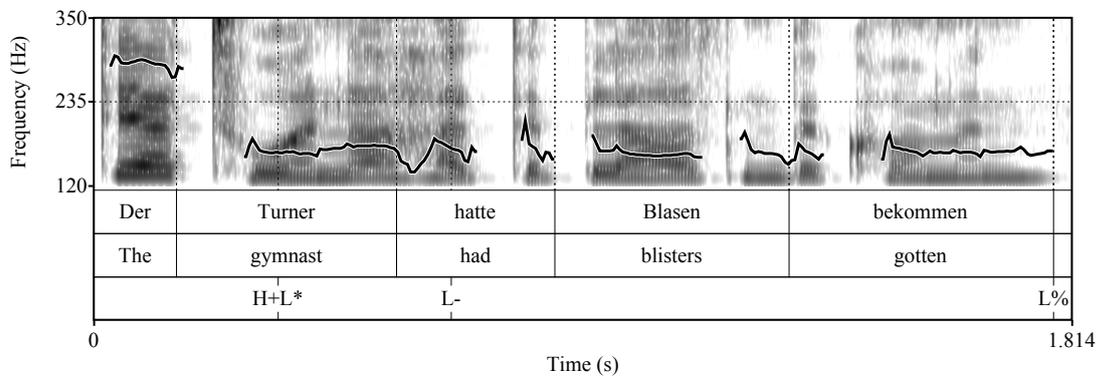


Figure 3. Representative example realization of an item in the non-contrastive focus condition in Experiment 1b (with nuclear H+L* on the subject noun).

³ The sound files are available on ling.uni-konstanz.de/braun-zinn/publications.

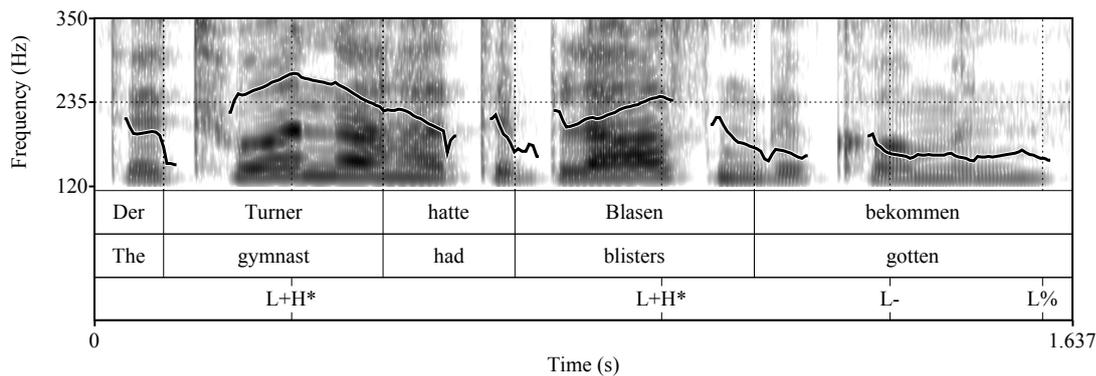


Figure 4. Representative example realization of an item in the broad focus control condition in all experiments (with a prenuclear L+H* on the subject noun).

Note that the autosegmental-metrical description of the pitch accent types suggests that two conditions are very similar: the broad focus control condition (prenuclear L+H*) and the contrastive focus condition (nuclear L+H*). Indeed, they both have 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 'Und ich habe gehört' (*And 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.⁴

Please Figure 5 about here

⁴ The spliced experimental materials were pre-tested for their naturalness in a web-based judgment task. Participants had to judge all 24 sentences but each sentence in only one of the three intonation conditions, resulting in eight lists, whose trials were randomized. Each list started with four familiarization trials. Participants judged each item on a scale from 1 (very unnatural) to 7 (very natural). Average ratings of 24 participants (average age 23.5 years, 13 female, 11 male), eight for each list were 5.1 for the control condition (SD = 1.4), 4.9 for the nuclear H+L* condition (SD = 1.9) and 5.4 for the nuclear L+H* condition (SD = 1.5). Results of a linear-mixed effects regression model with intonation condition as fixed factor and participants and items as crossed random factors showed no effect of intonation condition ($F(2, 23) = 1.1, p = 0.35$). These results suggest that none of the contours is judged as sounding significantly less unnatural than others to native speakers of German.

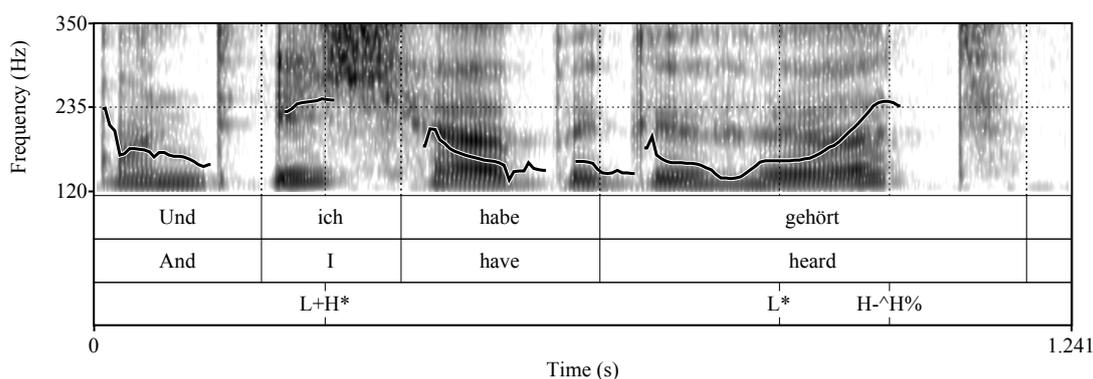


Figure 5. Realization of the prelude.

To validate the intonational realizations of the productions and the differences across conditions, we manually annotated the start and the end of the subject noun, the start and the end of its 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), and the f₀-range in semitones (st), as well as the duration of the stressed syllable, the duration of the f₀-rise and the total duration of the subject noun phrase. 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 did not differ in terms of f₀-range of the accentual movement (Experiment 1a: 9.0st rise excursion, Experiment 1b: 9.1st fall excursion, $p > 0.9$). The f₀-ranges in the narrow focus conditions were significantly larger than the f₀-rise in the broad focus control condition (Experiment 1a: 9.0st rise excursion in the narrow focus condition vs. 5.9st in the broad focus control condition, $t(23) = 8.9$, $p < 0.0001$, Experiment 1b: 9.1st fall excursion in the narrow focus condition vs. 5.9st in the broad focus control condition, $t(23)=10.9$, $p < 0.0001$).⁵ The range of the f₀-fall was significantly larger in the nuclear (narrow focus) L+H* compared to the prenuclear L+H* control condition (10.9st vs. 6.8st, $t(23) = 4.8$, $p < 0.0001$).

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

⁵ The absolute f₀ maximum was 286.4Hz in the narrow focus condition of Experiment 1a and 268Hz in Experiment 1b ($t(23) = 5.2$, $p < 0.0001$). Likewise, the f₀ minimum was higher in Experiment 1a than in Experiment 1b (169.6Hz vs. 159.2Hz, $t(23) = 3.5$, $p < 0.005$).

possible. They were given an example to make sure that they all understood what was meant by 'object' (e.g., *blisters* in the sentence *The gymnast had gotten blisters*).

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 plus system (pupil and corneal reflection at a sampling rate of 500 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 experimental 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 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) were 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 black fixation cross on white screen, which was shown until participants clicked on it. In all trials, the same token of the prelude (*Und ich habe gehört* 'And I have heard'; duration: 1240ms) was played. This was followed by 1000ms of silence, after which the target utterance was presented auditorily. After participants had clicked on the object, as instructed, 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 2ms 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 four words on screen over time is shown in *Figure 6*. These fixation proportions were calculated over consecutive 80ms bins, using the R-package VWPre

(Porretta, et al., 2016).

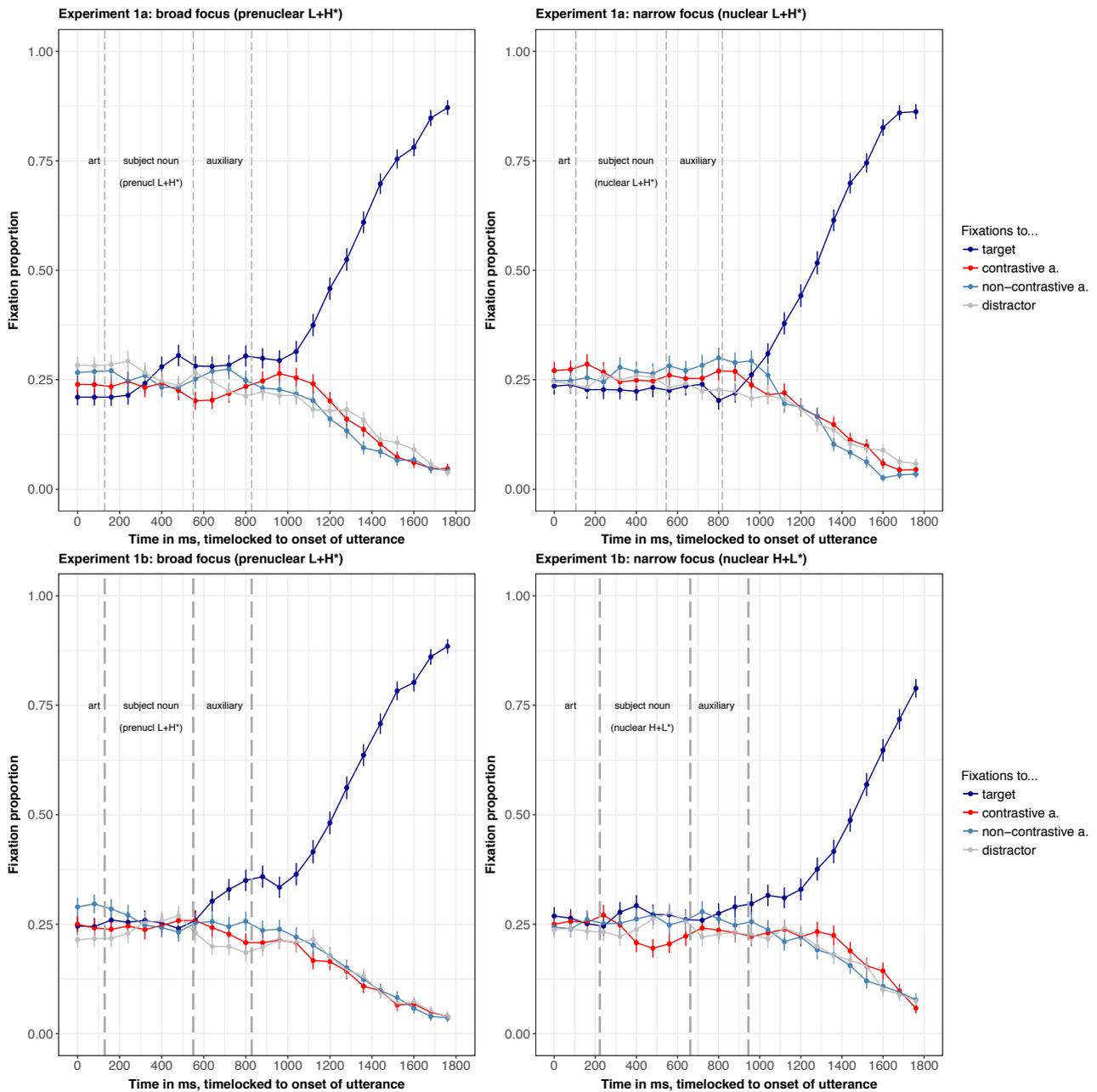


Figure 6. Fixation proportion over time for Experiment 1a (top panel) and Experiment 1b (bottom panel). Left panels show the broad focus control condition, right panels the narrow focus conditions. Fixations were sampled at 80ms bins. Error bars areas indicate standard error.

Please insert *Figure 6* about here

Figure 6 shows that fixations to the target (the object in the utterance, indicated by the dark blue line) started to be higher than fixations to the other words at 1000ms after the onset of the utterance the latest (approximately 200ms after the onset of the object noun), except for the broad focus condition in Experiment 1b (left bottom figure), in which there were already a lot of correct

anticipatory fixations to the object during the processing of the auxiliary. The interesting line for our research question is the red line in the time window from about 380ms to 780ms, which shows fixations to the contrastive associate while participants were processing the subject noun. The prosodic account predicted that there are more fixations in the narrow focus condition than in the broad focus condition in Experiment 1a but not in Experiment 1b. A direct visual and statistical comparison of the fixations to the contrastive associates in the narrow and broad focus conditions of each experiment demonstrates that this is borne out, see *Figure 7*.

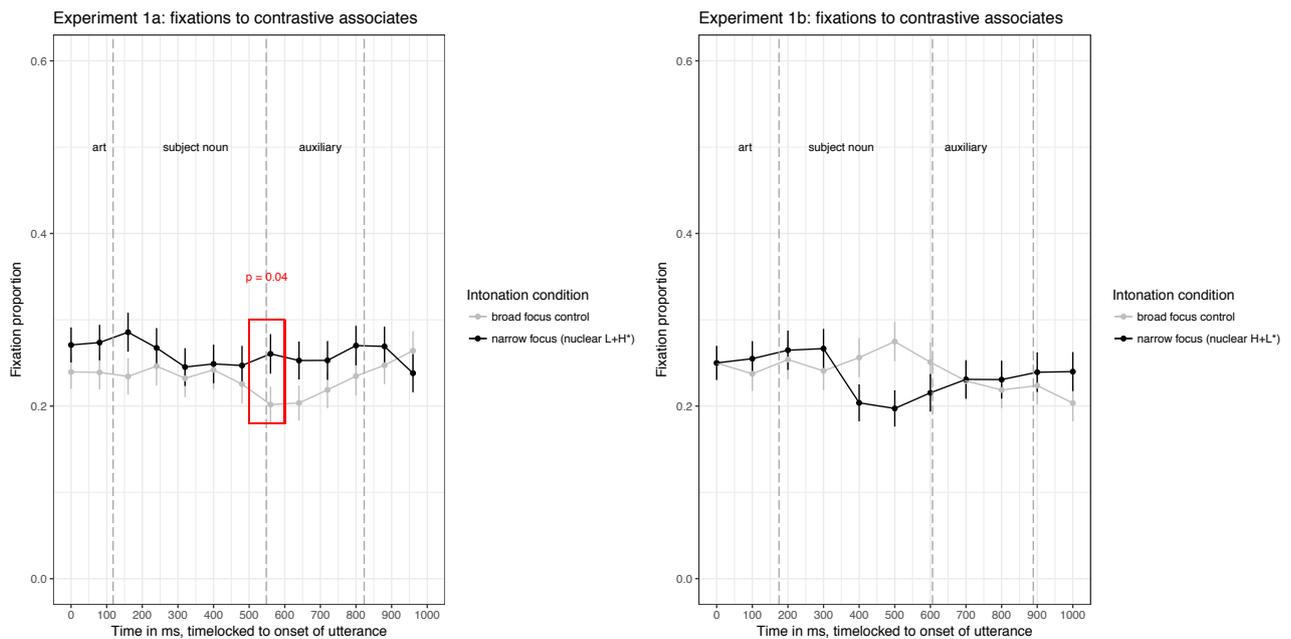


Figure 7. Comparison of fixations to the contrastive associate in the two conditions in Experiment 1a (left panel) and Experiment 1b (right panel). Vertical lines indicate average landmarks of the broad focus and narrow focus conditions. The red box indicates the analysis window with a significant difference across conditions.

Please insert *Figure 7* about here

For statistical analysis we analysed participants' fixations to the contrastive referent in consecutive 100ms steps (cf. McQueen & Viebahn, 2007). The statistical analyses largely follow the proposal in Barr, et al. (2011). We calculated the empirical logits of fixations to the contrastive associate in consecutive 100ms windows starting from 100 ms after the onset of the utterance until 800ms after its onset, 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 (dummy coded) and random intercepts for

participants and *items*. The model further included random slopes for the two within-group factors when this improved the fit of the model, as determined by LogLikelihood comparisons, using the R-function `anova()`. P-values were calculated using the Satterthwaite approximation of degrees-of-freedom in the R-package `lmerTest` (Kuznetsova, et al., 2016), which is based on `lme4` (Bates, et al., 2014).

Experiment 1a (nuclear L+H vs. prenuclear L+H*).*

In the time window 500-600ms after the onset of the utterance, there were significantly more fixations to the contrastive associate in the narrow focus condition (average logits = -1.8) than in the broad focus control condition (average logits = -2.2, $\beta = 0.4$, 95% CI: [0.01;0.78], $SE = 0.19$, $t = 2.0$, $p = 0.04$) see first row in Table 2. Given the typical time needed to plan a saccade (between 150-250ms, cf. Matin, et al., 1993), this difference is well within the time during which participants were processing the subject noun. In both intonation conditions, there were also many fixations to the non-contrastive associate, but these fixations to the non-contrastive associate were not affected by focus condition; for a summary of p-values, see Table A3 in the Appendix. Fixations to the distractor were not affected by intonation condition either (see Table A4 in the Appendix). Fixations to the target showed an effect of intonation condition in the time window starting at 500ms after target onset (see Table A5): there were more fixations to the target in the broad focus control condition than in the narrow focus condition. This is the opposite pattern as for the fixations to the contrastive associate, which suggests that target fixations are reduced in the narrow focus condition because of increased fixations to the contrastive associate.

	100- 200ms	200- 300ms	300- 400ms	400- 500ms	500-600ms	600- 700ms	700- 800ms
Exp 1a	p = 0.2	p = 0.1	p = 0.6	p = 0.8	p < 0.04	p = 0.2	p = 0.3
Exp 1b	p = 0.5	p = 0.2	p = 0.5	p = 0.9	p = 0.2	p = 0.1	p = 0.1

Table 2. Summary of p-values of comparisons to fixations to the contrastive associate across conditions in consecutive 100ms analysis windows. The subject noun starts on average 200ms after the onset of the sentence; it ends on average 600ms after the onset of the sentence (averages over all conditions).

Please insert Table 2 about here.

Experiment 1b (nuclear H+L vs. prenuclear L+H*).*

Unlike in Experiment 1a, 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). Fixations to the non-contrastive associate, the distractor and the target were not affected by focus condition, either (see Tables A3, A4 and A5 in the Appendix). To corroborate the differences in fixations to the

contrastive associate between Experiments 1a and 1b statistically, we combined the two data sets for the analysis window between 500 and 600ms and added *experiment* as a between-subjects factor in the data table. We then calculated a mixed-effects logistic regression model with *focus condition* (narrow vs. broad) and *experiment* (Experiment 1a and 1b) as fixed factors (both dummy coded) and *participants* and *items* as crossed random factors. The final model included random slopes for *focus condition* for *participants* and *items* and random slopes for *experiment* for *items*.⁶ The results showed a statistically significant interaction between focus condition and experiment ($\beta = 0.64$, $SE = 0.27$, $t = 2.34$, $p = 0.02$).

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-). These results show that not every narrow focus constituent is interpreted as contrastive, contra the view that narrow focus equals contrast (Hypothesis 1). Instead, the data lend support to Hypothesis 2 and provide further processing evidence for theories that distinguish the concepts of narrow focus and contrastive focus (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 in the same phrase is needed to activate contrastive alternatives.⁷ 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 reported effects of pitch accent type 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, et al. (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

⁶ both.glmer=glmer(eolog ~ cond*exp + (1+cond|vp)+(1+cond*exp|item), data =both, family = "binomial")

⁷ One reviewer correctly pointed out that the carrier phrase "And I have heard" may have provided context for intonational normalization, similar to what Wong and Diehl (2003) showed for perceptual normalization for Cantonese level tones. However, in the absence of a carrier phrase, normalization would also have been possible based on preceding experimental items, unless each item had been produced by a different speaker.

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, et al., 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, et al., 2014; Ito & Speer, 2008; Sedivy, et al., 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 statistically significant. 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 analysis window appears to coincide approximately with the end of the target words in Huettig and McQueen's study). In the current experiment, this number is 0.57 in the broad focus control condition and 0.62 in the narrow focus condition, i.e., of similar magnitude. In Yee and Sedivy (2006), the differences in fixation proportions to semantically related pictures (*key* upon hearing *lock*) and unrelated words was approximately 7%, i.e., 0.07 (judging from Figure 2 in Yee & Sedivy, 2006, p. 6). 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 in the on-ramp (the rising part) of the accent. 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 more or less pronounced subsequent fall) and not in the

conditions with more distinct accentual realizations across conditions (Experiment 1b with nuclear H+L* vs. prenuclear L+H*). It is hence unlikely that differences in the salience of an intonational contrast explain the data. In German intonational phonology (and its labelling system GToBI, cf. Grice, et al., 2005), both accents in Experiment 1a 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 located just before the object noun. However, the mere status of an accent as a nuclear accent (signalling a narrow focus on the subject) is not sufficient to trigger fixations to the contrastive alternative; otherwise we would have expected the same findings in Experiments 1a and 1b.

We have now established the contribution of pitch accent type towards the activation of alternatives. In Experiment 2 we turn to the processing of the additive particle *auch* ('also'), a lexical trigger for alternative sets and its interaction with the narrow focus realizations.

Experiment 2: Additive particles

Experiment 2 tested the effect of the unstressed phrase-initial additive particle *auch* ('also') on the interpretation of prenuclear and nuclear accents (e.g., *Auch der Turner hatte Blasen bekommen*; lit: Also the gymnast had blisters gotten; 'The gymnast, too, 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. Of them, forty were assigned to Experiment 2a (contrastive accent condition, aged between 18 and 32 years, average 23.3 years, 29 female, 11 male), forty to Experiment 2b (non-contrastive accent condition, aged between 19 and 33 years, average 25.7 years, 28 female, 12 male). They had not taken part in any of the web experiments and were unaware of the purpose of the experiment. They had normal (or corrected to normal) hearing and vision.

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 splice

the same token of *auch* in front of the recorded utterances from Experiment 1 because the transition sounded very unnatural for some items. 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 8*) and a non-contrastive accent (nuclear H+L* L-, see *Figure 9*). The displays and the experimental lists were the same as in Experiment 1, but the recordings including *auch* were used.

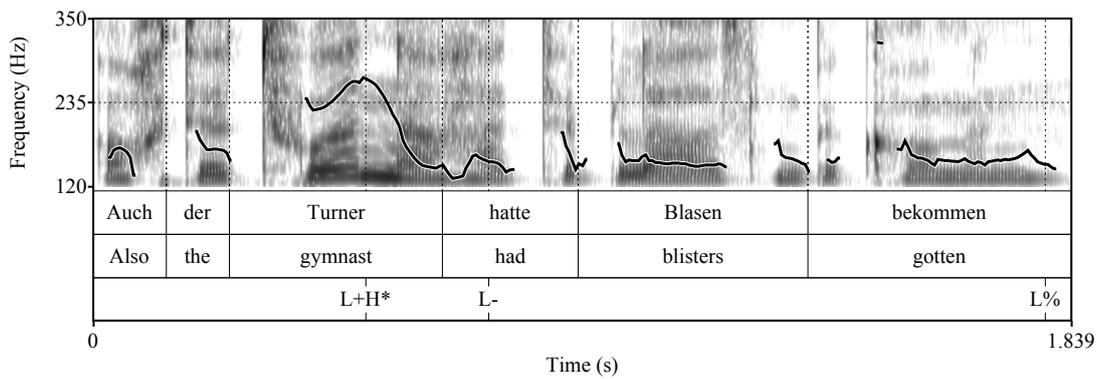


Figure 8. Example f0-track of the *auch*-condition in Experiment 2a (with a nuclear L+H* accent on the subject)

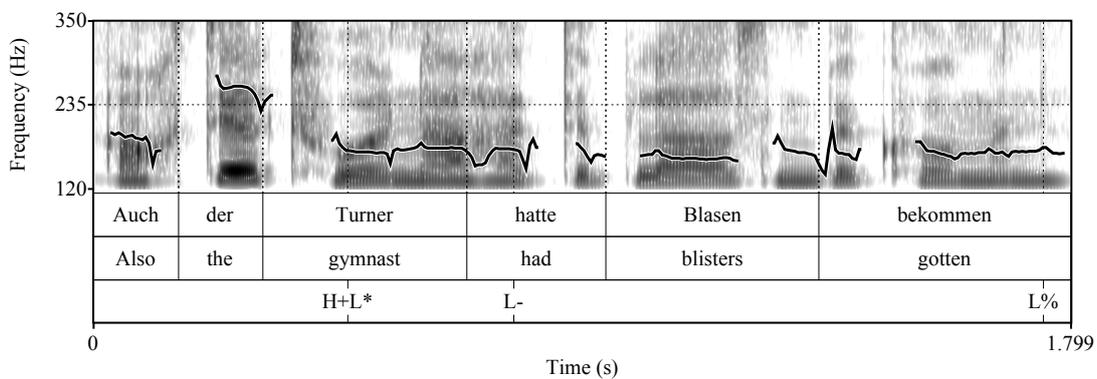


Figure 9. Example f0-track of the *auch*-condition in Experiment 2b (with a nuclear H+L* accent on the subject)

Insert *Figure 8* and *Figure 9* about here

Procedure

The testing procedure was identical to Experiments 1a and 1b.

Results

The evolution of fixations to the four words in the two conditions is shown in *Figure 10*, the fixations to the contrastive alternative across conditions are compared more directly in *Figure 11*.

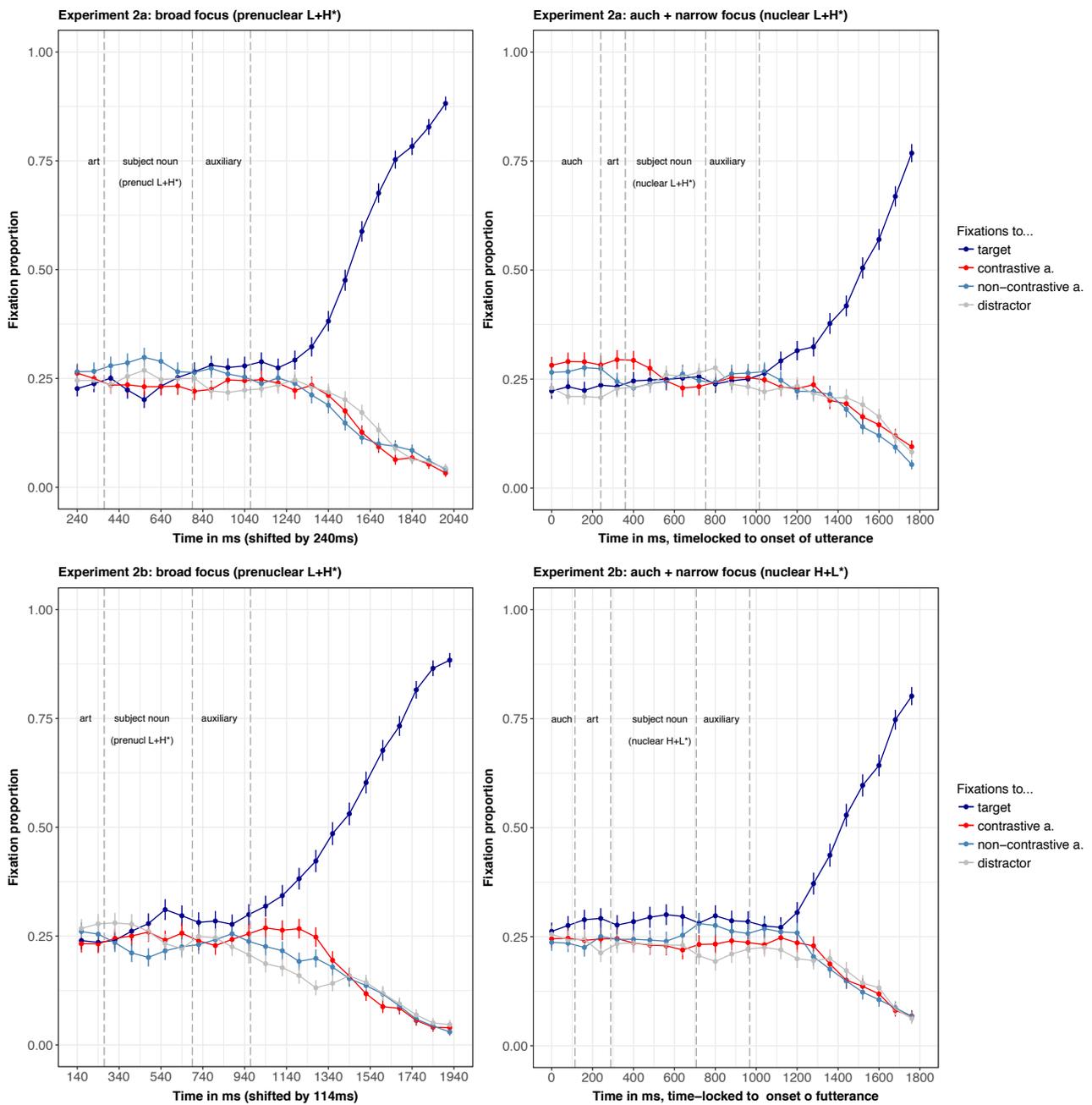


Figure 10. Evolution of fixations of Experiment 2a (upper panel) and 2b (lower panel). Left panels show the broad focus control condition without the particle, right panels the narrow focus conditions with the particle

Please insert *Figure 10* about here

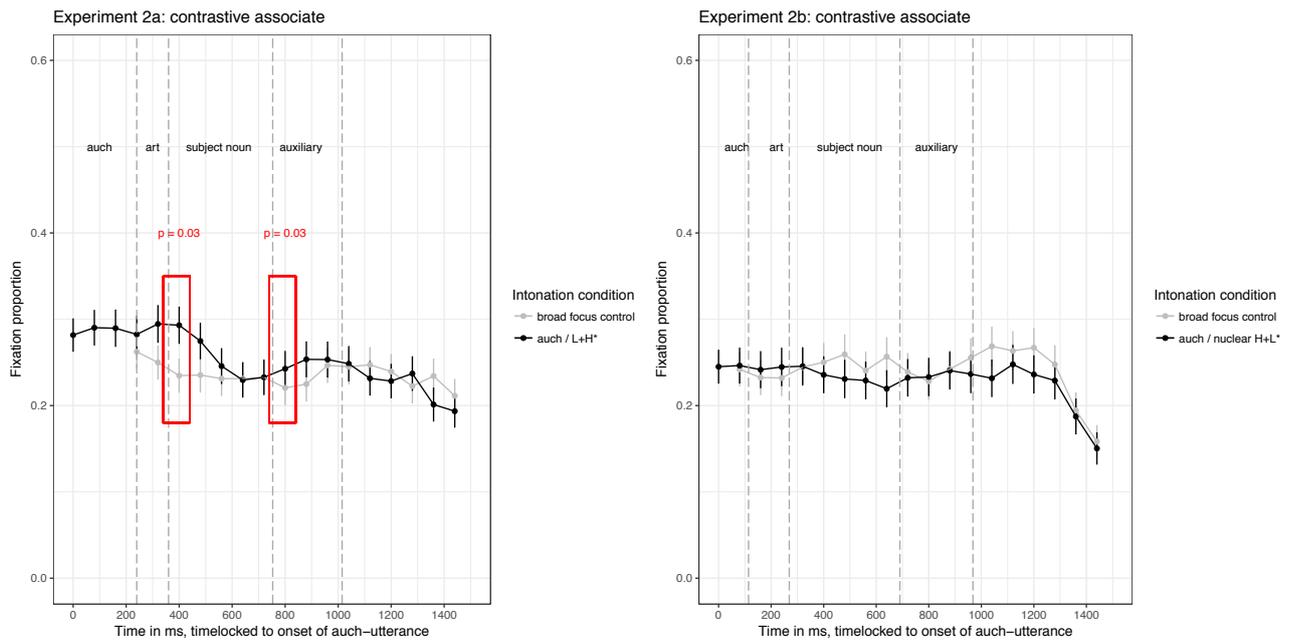


Figure 11. Comparison of fixations to the contrastive associate in the two conditions in Experiment 2a (left panel) and Experiment 2b (right panel). Fixations are plotted relative to the onset of the *auch*-utterance. Vertical lines refer to average acoustic landmarks of the *auch*-utterance. The red boxes indicate analysis windows with significant differences across conditions.

Please insert *Figure 11* about here

Experiment 2a (auch + nuclear L+H vs. prenuclear L+H*)*

For statistical comparison of fixations to the contrastive associate across conditions, we time-locked the fixations relative to the start of the definite article (shifting the fixations of the broad focus condition by 240ms, the average duration of the start of the article in the experimental condition) and compared fixations in successive 100ms steps. Results (see first row in Table 3) showed that participants fixated the contrastive associate more often in narrow focus utterances containing *auch* (average logits: -1.64) than in broad focus utterances without *auch* in the following time windows

- i. between 100 and 200 after the onset of the article ($\beta = 0.43$, $SE = 0.21$, $t = 2.1$, $p = 0.03$)
- ii. between 500 and 600ms after the onset of the article ($\beta = 0.41$, $SE = 0.19$, $t = 2.1$, $p = 0.03$)

Given the time it takes to program a saccade, these time windows are (i) before participants processed the article and (ii) while they were processing the subject noun.⁸

⁸ Fixations to the non-contrastive associate (see Table A3 in Appendix) showed more fixations to the non-contrastive associate in the broad focus control condition than the narrow focus condition, but these differences merely approached significance. Even if they were significant, the effect goes in the opposite direction than predicted by the semantics of the accents or the additive particle.

	100- 200ms	200- 300ms	300- 400ms	400- 500ms	500- 600ms	600- 700ms	700- 800ms
Exp 2a	p = 0.03	p = 0.2	p = 0.4	p = 0.4	p = 0.03	p = 0.2	p = 0.9
Exp 2b	p = 0.3	p = 0.3	p = 0.3	p = 0.3	p = 0.4	p = 0.6	p = 0.7

Table 3. Summary of p-values for comparisons of fixations to the contrastive associate across conditions in consecutive 100ms analysis windows for Experiments 2a and 2b, **time-locked to the onset of the article** (i.e., the window 100-200ms starts at 240ms of the auch-condition in Experiment 2a and at 140ms of the auch-condition in Experiment 2b)

Please insert Table 3 about here

Experiment 2b (auch + nuclear H+L vs. prenuclear L+H*)*

For statistical analysis, the fixations in the broad focus condition were also time-locked to the onset of the article (shifting fixations by 140ms, the average duration of the particle *auch*). Results showed no effects 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 experiment (contrastive accent in Experiment 2a vs. non-contrastive pitch accent in Experiment 2b) in the time window 100-200ms after the onset of the article ($\beta = 0.62$, $SE = 0.27$, $t = 2.24$, $p = 0.02$) and the interaction approached significance in the time window 500-600ms after the onset of the particle ($\beta = 0.50$, $SE = 0.27$, $t = 1.84$, $p = 0.06$).

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*). This is consistent with Hypothesis 4. Contra Hypothesis 3, the presence of the additive particle did not lead to the activation of contrastive 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 is in line with the findings by Sudhoff (2010), who

argued that the accent type but not the particle plays a role for the perception of contrast. The differences in fixations manifested themselves in two time windows. There was an effect of focus condition that came very early, while participants were still processing the article (considering that it takes at least 150ms to process the auditory input). This early effect could not have been caused by the intonational realization of the subject noun itself and we have no explanation for it. The expected effect of focus condition became apparent after the acoustic offset of the subject noun and is hence driven directly by the processing of the subject noun (and its pitch accent). The presence of effects of focus condition in Experiment 2a and their absence in Experiment 2b suggest that the additive particle did not affect fixations on its own.

General discussion and conclusion

The results show that (i) sentence-initial narrow focus constituents did not generally lead to more fixations to a word that was contrastively related to the subject noun (e.g., *dancer* upon hearing the sentence *The gymnast had gotten blisters*), but that (ii) accent type was crucial. In particular, there were more fixations to the contrastive associate when the subject noun was presented with a nuclear L+H* accent, but not when it was produced with a nuclear H+L* accent, when compared to the broad focus control condition. This supports Hypothesis 2, which predicted different effects depending on pitch accent type. The result goes against Hypothesis 1, which was based on semantic accounts maintaining that each narrow focus is contrastive. This effect of accent type was found for bare sentences without additive particles (e.g., *The gymnast had gotten blisters*, see Experiment 1), as well as for sentences with a sentence-initial particle (e.g., lit: *Also the gymnast had gotten blisters*, see Experiment 2). The online processing of narrow focus constituents is hence affected by pitch accent type. These data extend earlier eye tracking experiments, which reported differences in fixations for the salient contrast between accented and unaccented referents (Dahan, et al., 2002; Ito & Speer, 2008; Weber, Braun, et al., 2006), to a less salient contrast between different accent types (cf. Watson, et al., 2008).

Our data speak to the issue of the relation between the concepts of narrow focus and contrast. The semantic conception is that each narrow focus can be formalized by a set of alternatives (e.g., Bolinger, 1961; Jackendoff, 1972; Krifka, 2008; Lambrecht, 1994; Rooth, 1992; Wong & Diehl, 2003), while the prosodic conception is that there is a difference between narrow focus and contrastive focus, which is signalled by different accent types (e.g., Bartels & Kingston, 1994; Baumann, et al., 2006; Kiss, 1998; Neeleman, et al., 2009; Pierrehumbert & Hirschberg, 1990; Selkirk, 2002; Sudhoff, 2010; Watson, et al., 2008). Previous experimental research has already lent

support to the prosodic view, but it has not yet been shown whether contrast can be signalled by the narrow focus constituent alone, or whether the preceding and/or following prosodic context is necessary for interpretation. Our results show 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 (relative to a broad focus control condition), nuclear H+L* accents did not. Our data are therefore difficult to reconcile with the semantic view that "focus in general indicates the presence of alternatives for interpretation" (Krifka, 2008, p. 25) and instead support the view that a narrow focus with a nuclear L+H* accent leads to a contrastive focus, while a narrow focus constituent with a nuclear H+L* is a non-contrastive focus, 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 (2008), too, discusses the category "contrastive focus", but he reserves this category for corrections and some additive uses ("John wants coffee. [Mary]_{Focus} wants coffee, too", p 259). His argument is that, cross-linguistically, these contrastive foci lead to particular marking strategies. We have shown that in German, accent type is such a particular marking strategy. We therefore argue that a terminological distinction between contrastive and non-contrastive narrow focus is justified. One other way to reconcile the current data with Krifka's framework is to assume that – when heard out of context – the nuclear L+H* accents may receive a corrective interpretation ("The gymnast - and not someone else - had gotten blisters"), an interpretation that is hard to generate when the subject noun is realised with H+L*. However, one would then have to give reasons for why a corrective interpretation, which should lead to the exclusion of alternatives, leads to *more* fixations to alternatives. In any case, Krifka's (2008: 257) 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 hence be more appropriate to use the term 'contrastive' when alternatives are actually activated (in the case of nuclear L+H*), instead of assuming alternative sets for all narrow focus constituents.

What is troublesome for the formulation of the prosody-semantics interface (i.e., the meaning contribution of particular pitch accent types) is that, phonologically speaking, the subject noun in the contrastive narrow focus condition (nuclear L+H* in GToBI) has the same accent type as the subject noun in the broad focus control condition (prenuclear L+H* in GToBI). However, the two conditions differ in the position of these accents within the prosodic phrase, such that nuclear L+H*, but not prenuclear L+H*, is followed by a low phrase accent (L-). This suggests that 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 GToBI.⁹ Rather, the low phrase accent (L-) (and the sudden drop in f₀ by which it is characterised) also plays a crucial role in connection with contrastive accents, as has been noted before (e.g., Chafe, 1976; Couper-Kuhlen, 1984; Kügler & Gollrad, 2015). However, note that in our experiments, the processing differences mainly result from the nuclear accent type (H+L* vs. L+H*), and cannot be put down to the presence of an L-phrase boundary tone (or the status of the relevant accent as the nucleus) plus following deaccentuation, because L- follows the accent in both narrow focus conditions. It seems, therefore, that neither the L+H* nor the low phrase accent (L-) alone can account for contrastive interpretation. Instead, the accent together with its immediate tonal environment (and/or its position in the prosodic phrase) seem crucial for interpretation (see Selkirk, 2002). The current data show that nuclear H+L* does not lead to the activation of alternatives, which makes it a good candidate for a genuinely 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 (or: position)-dependency of intonational meaning. Note also that our discussion has concentrated on phonological pitch accent types, ignoring phonetic realization. However, parallel to the variation in the phonetic realization of segments in speech, the phonetic realization of a given pitch accent category may vary depending on factors such as prosodic context and speaker. Given that the current materials were naturally produced, it is conceivable that phonetic differences in intonation affect processing, a matter which will have to be explored in future research.

An alternative interpretation is that the differences in fixations to the contrastive associate are caused by differences in markedness between the more unmarked nuclear L+H* accent in Experiment 1a and the more marked H+L* accent in Experiment 1b.¹⁰ The less frequent nuclear H+L* may make it harder for participants to imagine a discourse context than for the more frequent nuclear L+H* accent. While it is true that one of the accents is more frequent than the other, we do not believe that the size of the difference in frequency (17% in turn-medial position and no difference in turn-final position, cf. Peters, et al., 2005) explain the data as well as the prosodic interpretation. Remember also from footnote 4 above that all accents were judged equally 'natural' by the participants of the web-based judgment task. Moreover, even if it were more difficult to

⁹ Other annotation systems make the link from accent type to contrast interpretation more transparent, e.g. Peters (2014) labels the accent type on the subject noun in the narrow focus condition as H*L (Experiments 1a and 2b) and in the broad focus condition as H* (chapter 7.3 in Peters, 2014).

¹⁰ We thank an anonymous reviewer for this suggestion.

construct an appropriate context for the less frequent accent – maybe because this accent type poses more constraints on a context – if participants still constructed the correct contexts (even taking a little bit longer), the results would be the same. This is because this accent is not a contrastive accent. Therefore, participants would not fixate the contrastive associate.

We will now discuss our data in light of the question of whether the presence of additive particles affects fixation behaviour. The fact that the presence of the additive particle *auch* ('also') did not lead to more fixations to the contrastive associate than the control condition for both narrow focus conditions is surprising from a semantic perspective, since additive particles presuppose the existence of alternatives to the accented item the particle associates with (Dimroth, 2004; Jacobs, 1983; König, 1991; König, 1994). Furthermore, recent online processing data from English suggest that listeners immediately exploit additive particles for reference resolution (Schwarz, 2015) and that the use of additive particles is only felicitous for the listener if there is a salient alternative to the accented item in the context or common ground. Hypothesis 3 therefore hypothesized that additive particles would lead to more fixations to the contrastive word than the control condition without particles, but this was not the case. From a processing point of view, the particle was useful for disambiguating the referent in the eye-tracking study by Schwarz (2015), while the additive particle was not beneficial for the task in our experiments (neither was the mere presence of an accent on the subject noun). The discrepancy in findings between Schwarz (2015) and our study may hence be due to the informativeness of the particle in a given situation. On the other hand, our data support Hypothesis 4, which hypothesized that the contrastive interpretation is not directly driven by the context. In our experiments, the contrastive interpretation came from the intonational realization of the subject noun. In German perception data by Sudhoff (2010), the contrastive interpretation did not depend on the presence of focus particles either (there, it was depended on the context).

We see two possible explanations for why *auch* ('also') did not lead to more fixations to the contrastive associate in the two narrow focus conditions alike. First, it is possible that alternatives evoked by contrastive accents have a different mental representation from 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 by contrastive accents are available earlier than alternatives due to additive particles. An argument in support of the latter explanation 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 may not be the case for contrastive pitch accents. In any case, the observed differences in the processing of contrastive pitch accents and focus particles are in line with Gotzner (2016). She tested the recognition of probes that were contrastively related to a word produced with a contrastive L+H* accent or a non-contrastive H* accent and found that contrastively related words were recognized faster in the contrastive accent condition than the non-contrastive accent condition. Probe recognition was slowed down by adding the particles *also* and *only*, however. She argued that contrastive accentuation increases the salience of contrastive alternatives, while focus-sensitive particles increase 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 which 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 better understand the relevance of those aspects of a pitch accent (slope of the rise and fall, relative or absolute alignment of tonal targets, intensity) for its interpretation as a contrastive accent. Furthermore, we will compare the processing of prenuclear L*+H accents in German, which have been found in contrastive topics German. This comparison involves two kinds of prenuclear accents (contrastive (topic) vs. non-contrastive (neutral) rising accents), an intonational contrast that is hard to represent phonologically (Braun, 2004, 2005, 2006; Zellers & Post, 2012; Zerbian, et al., 2012). Another direction for future research is to study how listeners cope with variability in the marking of 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, 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, et al., 2015 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 Neitsch for support in stimulus preparation and testing, Katharina Zahner and Vincent Porretta for support with the fixation plots, 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

- Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of "segmental anchoring" of F0: evidence from German. *Journal of Phonetics*, 32(3), 177-197.
- Barr, D. J., Gann, T. M., & Pierce, R. S. (2011). Anticipatory baseline effects and information integration in visual world studies. *Acta Psychologica*, 37, 201-207.
- Bartels, C., & Kingston, J. (1994). Salient pitch cues in the perception of contrastive focus. In P. Bosch & R. van der Sandt (Eds.), *Focus and Natural Language Processing* (Vol. 1).
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-6. <http://CRAN.R-project.org/package=lme4>.
- Baumann, S., & Grice, M. (2006). The intonation of accessibility. *Journal of Pragmatics*, 38, 1636-1657.
- Baumann, S., Grice, M., & Steindamm, S. (2006). *Prosodic marking of focus domains - categorical or gradient?* Paper presented at the 3rd International Conference on Speech Prosody, Dresden, Germany.
- Bolinger, D. (1961). Contrastive accent and contrastive stress. *Language*, 37, 83-96.
- Braun, B. (2004). *Phonetics and phonology of contrast marking in German: Issues for the Prosody-Semantics Interface*. Paper presented at the KONVENS 2004 (ÖGAI Schriftenreihe), Wien, Austria.
- Braun, B. (2005). *Production and Perception of Thematic Contrast in German*. Frankfurt am Main: Peter Lang Publishing.
- Braun, B. (2006). Phonetics and phonology of thematic contrast in German. *Language and Speech*, 49(4), 451-493.
- Braun, B. (2007). *Effects of dialect and context on the realisation of German prenuclear accents*. Paper presented at the XVIth International Congress of the Phonetic Sciences, Saarbrücken.
- Braun, B., & Asano, Y. (2013). *Double contrast is signalled by prenuclear and nuclear accent types alone, not by f0-plateaux*. Paper presented at the 14th Annual Conference of the International Speech Communication Association.
- Braun, B., & Chen, A. (2012). Now for something completely different: Anticipatory effects of intonation. In O. Niebuhr & H. Pfitzinger (Eds.), *Understanding Prosody: Context, Function, and Communication* (pp. 289-311). Berlin: deGruyter.
- Braun, B., & Tagliapietra, L. (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes*, 25, 1024-1043.
- Büring, D., & Hartmann, K. (2001). The syntax and semantics of focus-sensitive particles in German. *Natural Language & Linguistic Theory*, 19, 221-281.
- Chafe, W. L. (1976). Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View. In C. N. Li (Ed.), *Subject and Topic* (pp. 25-55). New York: Academic Press.
- Chen, A., Den Os, E., & De Ruiter, J. P. (2007). Pitch accent type matters for online processing of information status: Evidence from natural and synthetic speech. *Linguistic Review*, 23(2-3), 317-344.
- Couper-Kuhlen, E. (1984). A new look at contrastive intonation. In R. Watts, Weideman, U. (Ed.), *Modes of Interpretation: Essays presented to Ernst Leisi* (pp. 137-158). Tübingen: Gunter Narr Verlag.
- Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: evidence from eye movements. *Cognitive Psychology*, 42(4), 317-367.
- Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, 47(2), 292-314.

- Dennison, H. Y., & Schafter, A. (2010). *Online construction of implicature through contrastive prosody*. Paper presented at the Fifth International Conference on Speech Prosody, Chicago.
- Dimroth, C. (2004). *Fokuspunkt und Informationsgliederung im Deutschen*. Tübingen: Stauffenburg.
- Esteve-Gibert, N., Portes, C., Schafter, A., Hemforth, B., & D'Imperio, M. (2016). *Intonation in the processing of contrast meaning in French: An eye-tracking study*. Paper presented at the International Conference on Speech Prosody, Boston.
- Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, 36, 680-703.
- Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language*, 63, 367-386.
- Görs, K., & Niebuhr, O. (2012). *Hocus Focus - What the elicitation method tells us about types and exponents of contrastive focus*. Paper presented at the Proceedings of the 6th International Conference of Speech Prosody, Shanghai, China.
- Gotzner, N. (2016). *The role of focus intonation in implicature computation: A comparison with only and also*. Unpublished manuscript, https://www.researchgate.net/profile/Nicole_Gotzner/publication/299604052_The_role_of_focus_intonation_in_implicature_computation_A_comparison_with_only_and_also/links/57028ccf08ae646a9da873f4/The-role-of-focus-intonation-in-implicature-computation-A-comparison-with-only-and-also.pdf.
- Grice, M., Baumann, S., & Benz Müller, R. (2005). German Intonation in Autosegmental-Metrical Phonology. In J. Sun-Ah (Ed.), *Prosodic Typology. The Phonology of Intonation and Phrasing* (pp. 55-83). Oxford: Oxford University Press.
- Hirschberg, J., & Litman, D. (1993). Empirical studies on the disambiguation of cue phrases. *Computational Linguistics*, 19, 501-530.
- Huetting, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, 96, B23-B32.
- Huetting, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, 57(4), 460-482.
- Huetting, F., Rommers, J., & Meyer, A. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137, 151-171.
- Husband, M. E., & Ferreira, F. (2012). *Generating contrastive alternatives: Activation and suppression mechanisms*. Paper presented at the 25th Annual CUNY Conference on Human Sentence Processing, New York.
- Ito, K., Bibyk, S. A., Wagner, L., & Speer, S. R. (2014). Interpretation of contrastive pitch accent in six- to eleven-year-old English-speaking children (and adults). *Journal of Child Language*, 41(1), 84-110.
- Ito, K., & Speer, S. S. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, 58, 541-573.
- Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Jacobs, J. (1983). *Focus und Skalen: Zur Syntax und Semantik von Gradpartikeln im Deutschen*. Tübingen.
- Kim, C. S., Gunlogson, C., Tanenhaus, M. K., & Runner, J. T. (2015). Context-driven expectations about focus alternatives. *Cognition*, 139, 28-49.
- Kiss, K. É. (1998). Identificational focus versus information focus. *Language*, 74, 245-273.

- Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements on reading. *European Journal of Cognitive Psychology*, 16(1/2), 262-284.
- Kohler, K. J. (1991). 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.
- König, E. (1991). *The meaning of focus particles. A comparative perspective*. London, New York: Routledge.
- König, E. (1994). Focus Particles. In J. Jacobs, A. von Stechow, W. Sternefeld & T. Vennemann (Eds.), *Syntax. Ein internationales Handbuch der zeitgenössischen Forschung* (pp. 978 – 987). Berlin, New York: de Gruyter.
- Krahmer, E., & Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech Communication*, 34, 391-405.
- Krifka, M. (1991). *Additive particles under stress*. Paper presented at the SALT 8.
- Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55, 243-276.
- Kügler, F., & Gollrad, A. (2015). Production and perception of contrast: The case of the rise-fall contour in German. *Frontiers in Psychology*, 6.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-33. <https://CRAN.R-project.org/package=lmerTest>.
- Lambrecht, K. (1994). *Information Structure and Sentence Form. Topic, focus and the mental representations of discourse referents*. Cambridge; New York, NY, USA: Cambridge University Press.
- Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccadic overhead. *Perception & Psychophysics*, 53, 372-380.
- McQueen, J. M., & Viebahn, M. (2007). Tracking recognition of spoken words by tracking looks to printed words. *The Quarterly Journal of Experimental Psychology* 60(5), 661-671.
- Molnár, V. (2001). *Contrast from a contrastive perspective*. Paper presented at the ESSLLI Workshop on Information Structure, Discourse Structure and Discourse Semantics.
- Neeleman, A., Titov, E., van de Koot, H., & Vermeulen, R. (2009). A syntactic typology of topic, focus and contrast. In J. Van Craenenbroeck (Ed.), *Alternatives to Cartography* (pp. 15-51). Berlin: Mouton de Gruyter.
- Niebuhr, O. (2010). On the phonetics of intensifying emphasis in German. *Phonetica*, 67, 1-29.
- Peters, B., Kohler, K., & Wesener, T. (2005). Melodische Satzakkentmuster in prosodischen Phrasen deutscher Spontansprache - Statistische Verteilung und sprachliche Funktion [Melodic sentence accent patterns in spontaneous German prosodic phrases - statistical distribution and linguistic function]. In K. Kohler, F. Kleber & B. Peters (Eds.), *Prosodic Structures in German Spontaneous Speech (AIPUK 35a)* (pp. 185-201). Kiel: IPDS.
- Peters, J. (2014). *Intonation*. Heidelberg: Universitätsverlag Winter.
- Pierrehumbert, J. B., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan & M. E. Pollack (Eds.), *Intentions in Communication* (pp. 271-311). Cambridge: MIT Press.
- Porretta, V., Kyröläinen, A., van Rij, J., & Järvikivi, J. (2016). VWPre: Tools for Preprocessing Visual World Data. R package version 0.9.5.
- 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.

- Repp, S. (2010). Defining 'contrast' as an information-structural notion in grammar. *Lingua*, 120(6), 1333-1345.
- Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1, 75-116.
- Schwarz, F. (2015). Presuppositions vs. asserted content in online processing. In F. Schwarz (Ed.), *Experimental Perspectives on Presuppositions* (pp. 89-108). Cham, Heidelberg, New York, Dordrecht, London: Springer.
- Sedivy, J., Tanenhaus, M. K., Chambers, C., & Carlson, G. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71, 109-147.
- Selkirk, E. O. (2002). *Contrastive FOCUS vs. presentational focus: Prosodic evidence from right node raising in English*. Paper presented at the Proceedings of the 1st International Conference on Speech Prosody.
- Spalek, K., Gotzner, N., & Wartenburger, I. (2014). Not only the apples. Focus-sensitive particles improve memory for information-structural alternatives. *Journal of Memory and Language*, 70, 68-84.
- Sudhoff, S. (2010). Focus particles and contrast in German. *Lingua*, 120(6), 1458-1475.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632-1634.
- Vallduví, E., & Engdahl, E. (1996). The linguistic realisation of information packaging. *Linguistics*, 34, 459-519.
- Watson, D., Tanenhaus, M. K., & Gunlogson, C. A. (2008). Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science*, 32(7), 1232-1244.
- Weber, A., Braun, B., & Crocker, M. W. (2006). Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and Speech*, 49(3), 367-392.
- Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition*, 99, B64-B72.
- Wong, P. C., & Diehl, R. L. (2003). Perceptual normalization for inter- and intratalker variation in Cantonese level tones. *Journal of Speech and Hearing Research*, 46(2), 413-421.
- Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning Memory, & Cognition*, 32(1), 1-14.
- Zellers, M., & Post, B. (2012). *Fundamental frequency and other prosodic cues to topic structure*. Paper presented at the Proceedings of Workshop on the Discourse-Prosody Interface, Paris, France.
- Zerbian, S., Turco, G., Schaffler, N., Zellers, M., & Riester, A. (2012). *Contrastive topic constituents in German*. Paper presented at the Seventh International Conference on Speech Prosody, Boston, US.

Appendix

Subject noun	Contrastive associate	Non-contrastive associate
Schwimmer (swimmer)	Taucher (53) (diver)	Bad (15) (baths)
Turner (gymnast)	Tänzer (5) (dancer)	Sport (15) (sports)
Nonne (nun)	Mönch (32) (monk)	Kloster (50) (abbey)
Artistin (artist)	Clown (37) (clown)	Zirkus (55) (circus)
Italiener (Italian)	Spanier (37) (Spaniard)	Spaghetti (10) (spaghetti)
Japaner (Japanese)	Chinese (37) (Chinese)	Asien (10) (Asia)
Kunde (customer)	Verkäufer (16) (shop assistant)	Geschäft (30) (shop)
Segler (sailor)	Kapitän (21) (captain)	Boot (20) (boat)
Mieter (tenant)	Nachbar (32) (neighbour)	Wohnung (35) (apartment)
Professor (professor)	Student (58) (student)	Universität (30) (university)
Schreiner (carpenter)	Tischler (11) (cabinet maker)	Holz (40) (wood)
Direktor (director)	Sekretär (16) (secretary)	Schule (45) (school)
Züchter (breeder)	Bauer (32) (farmer)	Tiere (50) (animals)

Sänger (singer)	Techniker (68) (technician)	Lieder (30) (songs)
Maler (painter)	Zeichner (21) (draftsman)	Farben (30) (paint)
Schlagzeuger (drummer)	Gitarrist (21) (guitarist)	Band (45) (band)
Schafe (sheep)	Ziegen (21) (goats)	Herde (25) (flock)
Biene (bee)	Wespe (42) (wasp)	Honig (25) (honey)
Flamingo (flamingo)	Pelikan (16) (pelican)	Vogel (15) (bird)
Wale (whales)	Haie (16) (sharks)	Orcas (5) (orcas)
Frauchen (mistress)	Herrchen (42) (master)	Hund (45) (dog)
Tiger (tiger)	Löwe (58) (lion)	Streifen (10) (stripes)
Rehe (deer)	Hirsche (15) (stags)	Wald (20) (forest)
Geiger (violinist)	Pianist (21) (pianist)	Violine (15) (violin)

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*

	BF	NF	NF	auch + NF	auch + NF
	prenucl.	nucl L+H*	nucl	nucl L+H*	nucl. H+L*
	L+H*	(Exp 1a)	H+L*	(Exp 2a)	(Exp 2b)
	(all Exps)		(Exp 1b)		
L-alignment with respect to start of stressed syllable in ms	-25.9 (48.5)	-22.5 (40.1)	146.5 (51.1)	-17.8 (13.3.)	130.7 (32.7)
H-alignment with respect to end of stressed syllable in ms	-45.3 (33.2)	-34.2 (41.8)	-322.0 (57.5)	-35.5 (32.1)	-256.6 (48.5)
F0-range of the accentual movement in st (rise in L+H*, fall in H+L*)	5.9 (1.1)	9.0 (1.1)	9.1 (0.9)	7.3 (1.2)	6.7 (0.9)
F0 of L tone in bitonal accent in Hz	191.4 (8.1)	169.3 (4.5)	159.2 (7.6)	175.0 (8.1)	170.8 (6.8)
F0 of H tone in bitonal accent in Hz	287.3 (68.9)	287.6 (18.3)	268.6 (9.2)	278.7 (26.0)	268.7 (13.6)
Duration of the subject-NP in ms	421.0 (72.4)	439.4 (98.6)	458.9 (102.3)	431.5 (93.5)	417.9 (91.0)
F0-range from peak to end of the modal verb in st	6.0 (1.3)	10.3 (1.3)	8.8 (0.7)	9.5 (1.0)	7.1 (1.2)

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

	100- 200ms	200- 300ms	300- 400ms	400- 500ms	500- 600ms	600- 700ms	700- 800ms
Exp 1a	p = 0.3	p = 0.1	p = 0.9	p = 0.4	p = 0.7	p = 0.4	p = 0.1
Exp 2a	p = 0.3	p = 0.2	p = 0.09	p = 0.07	p = 0.13	p = 0.09	p = 0.06

Table A3. Summary of p-values of comparisons to fixations to the non-contrastive associate across conditions in consecutive 100ms analysis windows, time-locked to the onset of the article.

	100- 200ms	200- 300ms	300- 400ms	400- 500ms	500- 600ms	600- 700ms	700- 800ms
Exp 1a	p = 0.3	p = 0.2	p = 0.2	p = 0.8	p = 0.9	p = 0.9	p = 0.9
Exp 2a	p = 0.3	p = 0.4	p = 0.8	p = 0.6	p = 0.5	p = 0.7	p = 0.1

Table A4. *Summary of p-values of comparisons to fixations to the distractor across conditions in consecutive 100ms analysis windows, time-locked to the onset of the article.*

	100- 200ms	200- 300ms	300- 400ms	400- 500ms	500- 600ms	600- 700ms	700- 800ms
Exp 1a	p = 0.4	p = 0.5	p = 0.9	p = 0.1	p = 0.01	p = 0.02	p = 0.009
					(bf > nf)	(bf > nf)	(bf > nf)
Exp 2a	p = 0.09	p = 0.6	p = 0.5	p = 0.7	p = 0.8	p = 0.7	p = 0.8

Table A5. *Summary of p-values of comparisons to fixations to the target across conditions in consecutive 100ms analysis windows, time-locked to the onset of the article*