

## RESEARCH

# The incremental processing of focus, givenness and prosodic prominence

Stefan Baumann<sup>1</sup> and Petra B. Schumacher<sup>2</sup><sup>1</sup> IfL Phonetik, University of Cologne, Köln, DE<sup>2</sup> Department of German Language and Literature I, Linguistics, University of Cologne, Köln, DECorresponding author: Stefan Baumann ([stefan.baumann@uni-koeln.de](mailto:stefan.baumann@uni-koeln.de))

This study on German investigates the real-time comprehension of items in *First Occurrence Focus* (focused and new), *Second Occurrence Focus* (focused and given), *Quasi Second Occurrence Focus* (derogatory expressions that are referentially given and lexically new) and *Background* (non-focused and given), which are marked by different levels of prosodic prominence. While previous electrophysiological research tested mismatches between prosody and information structure, the present study assessed contextually licensed, appropriate prosodic realizations of stimuli. Our EEG experiment revealed distinct topographic profiles for information structure and prosody. As to prosody, we found a biphasic pattern over anterior brain regions for (secondarily prominent) phrase accents (marking *Second Occurrence Focus*) and deaccentuation (marking *Background*) but not for pitch accents (marking *First Occurrence Focus*), indicating an inverse relation between processing effort and the level of perceived prominence. The event-related potentials for *Quasi Second Occurrence Focus* items resembled *First Occurrence Focus* items although the former were deaccented. As to information structural contrasts, *First Occurrence Focus* engendered a pronounced negativity over posterior sites relative to *Second Occurrence Focus* and *Background*. *Quasi Second Occurrence Focus* showed an intermediate negativity. These differences can probably be accounted for by (lexically) *new* rather than *focused* information. In general, the data indicate that both prosodic cues and information structural categories influence the incremental processing of spoken language and that pitch accents and newness fulfill independent prominence-lending functions.

**Keywords:** prominence; phrase accent; pitch accent; second occurrence focus; information status; ERP; focus; German

## 1 Background and motivation

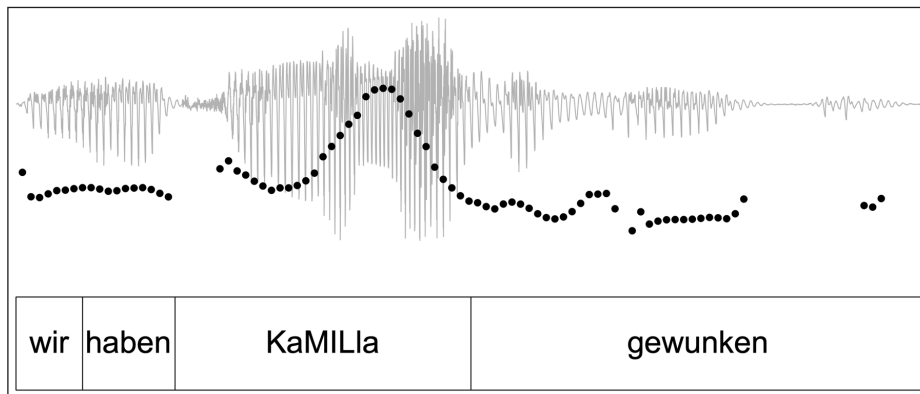
Speakers use various linguistic means to structure their utterances and draw the hearers' attention to specific aspects of an utterance. The current research investigates the interplay of prosodic cues with information structural aspects, such as focus and givenness. While much research has targeted the contribution and classification of pitch accents (defined here as primary prominences), the current paper is additionally concerned with the processing of secondary prominences in postnuclear position, which are not marked by pitch movement. Using electrophysiological measures, we contrast the processing of expressions (in German) that are contextually licensed as *First Occurrence Focus*, *Second Occurrence Focus* and *Background*.

### 1.1 Prosodic prominence

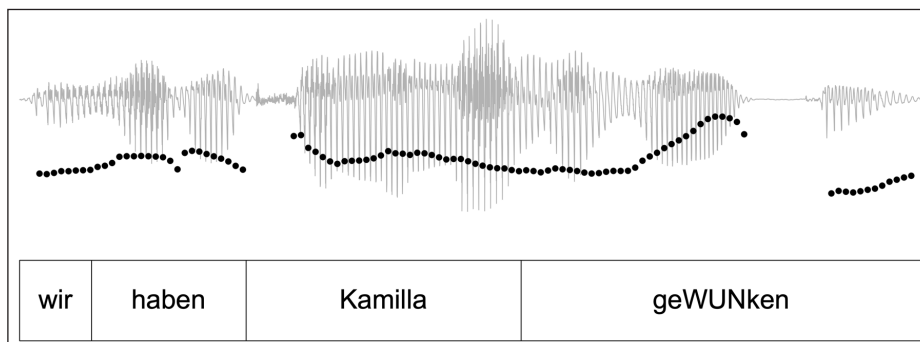
One of the most essential aspects of spoken communication is an appropriate interpretation of an utterance's *prosody*. That is, we only understand all of the intended meaning of an utterance in a discourse context if we are able to process and judge a speaker's accent

placement and intonation correctly. A central function of prosody is highlighting, i.e. making elements *prominent* in relation to neighbouring elements. Acoustically, the cues fundamental frequency (F0) movement, increased duration and intensity as well as spectral emphasis and vowel quality serve to lend prominence to a syllable or word (at least in Germanic languages and other intonation languages; see e.g. Fry 1955; 1958; Sluijter & van Heuven 1996; Kochanski et al. 2005). The illustrated German examples show the difference between a (prominence-lending) pitch accent on *Kamilla* (Figure 1) and lack of accent on the same word (Figure 2). Accentuation is cued here primarily by a rising local pitch movement on the lexically stressed syllable *-mil-* (vs. lack of movement in the unaccented version). In the example in Figure 2, the final accent has been “shifted” to the second syllable of *gewunken* (‘waved’), which now carries the most decisive pitch movement (and is longer and louder than in the first utterance).

The final pitch accent in an intonation unit is commonly referred to as the *nucleus* or *nuclear accent*, which has a special status in most phonological theories (most explicitly stated in the British School; see e.g. Crystal 1969). Structurally, the nucleus is most prominent, since it is the only obligatory element in its unit, often carrying the most distinct tonal movement. More formally oriented approaches define the nucleus as the *head* of an (intermediate) intonation phrase, assigning it a central role in the prosodic hierarchy (e.g. Beckman & Edwards 1990; 1994). Semantic-pragmatically, the nuclear accent marks the most important element, in the sense that its position determines the interpretation of an



**Figure 1:** Oscillogram and superimposed F0 contour for the utterance *Wir haben Kamilla gewunken* (‘we waved to Kamilla’) with a nuclear pitch accent on *Kamilla*. The accented syllable *-MIL-* is printed in capital letters.



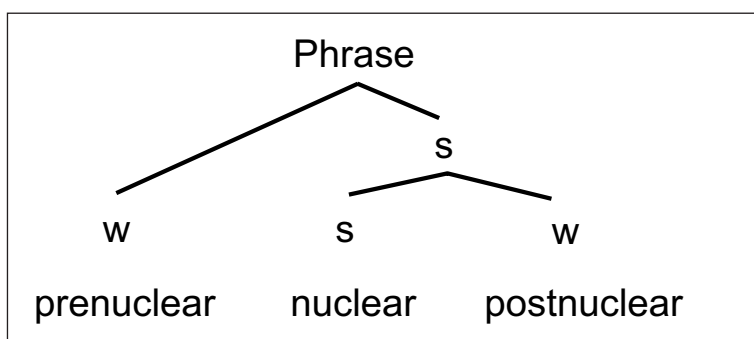
**Figure 2:** Oscillogram and superimposed F0 contour for the utterance *Wir haben Kamilla gewunken* (‘we waved to Kamilla’) with a nuclear pitch accent on *gewunken* (‘waved’) and lack of accent on *Kamilla*. The accented syllable *-WUN-* is printed in capital letters.

utterance's information structure. Applied to the utterance in Figures 1 and 2, e.g., this means that the first utterance could be interpreted as broad focus, since its prosody makes it an appropriate answer to a question like "What happened?", whereas the accent placement in the second utterance suggests a narrow focus context, in which *Kamilla* is treated as background information, i.e. as derivable from the previous context. (The relation between prosody and information structure will be discussed in more detail in section 1.2.)

Nuclear accents can thus be considered primary in terms of phonological strength, while prosodically prominent words or syllables before and after the nucleus – *pre-* and *postnuclear* elements, respectively – may be regarded as secondary in nature. This prominence relation (which is syntagmatic in nature) is illustrated in the metrical tree in Figure 3.

Nevertheless, there is an important difference between the two types of secondary prominences: While prenuclear prominences can surface as fully-fledged pitch accents (usually being indicated by a combination of longer duration, higher intensity and local pitch movement), postnuclear prominences cannot, since the nuclear accent is by definition the last pitch accent in the phrase. Postnuclear prominences are not marked by pitch movement but mostly by increased duration and intensity. We will refer to this type of secondary prominences as *phrase accents* as proposed by Grice, Ladd & Arvaniti (2000). They define a phrase accent as an edge tone that is additionally associated with a lexically stressed syllable, lending a certain degree of prosodic prominence to the constituent in question. Several examples will be given in the course of the paper.

There is a growing body of behavioural studies investigating the *perception* of prosodic prominence. It has been shown, in particular for Germanic languages, that even untrained listeners are able to judge prominence consistently. One of the elicitation methods currently used is the so-called *Rapid Prosody Transcription* (RPT) method developed by Cole and colleagues (see overview in Cole & Shattuck-Hufnagel 2016) in which naive listeners have to indicate all words on a transcript which they feel to stand out in a given utterance.<sup>1</sup> The method records patterns of inter-transcriber agreement since it calculates the prominence score (*p-score*) for each word and can at the same time capture inter-transcriber differences in the prosodic annotation. Furthermore, although the prominence judgments are binary in nature, the resulting p-scores are (quasi-)continuous-valued since they can be translated into percentages, which not only indicate the probability of a word to be marked as prominent but also – however only indirectly – suggest different degrees of perceived prominence. In fact, the "graded prosodic labels [...] can be used to test the contribution of individual acoustic cues or other non-acoustic predictors to the perception



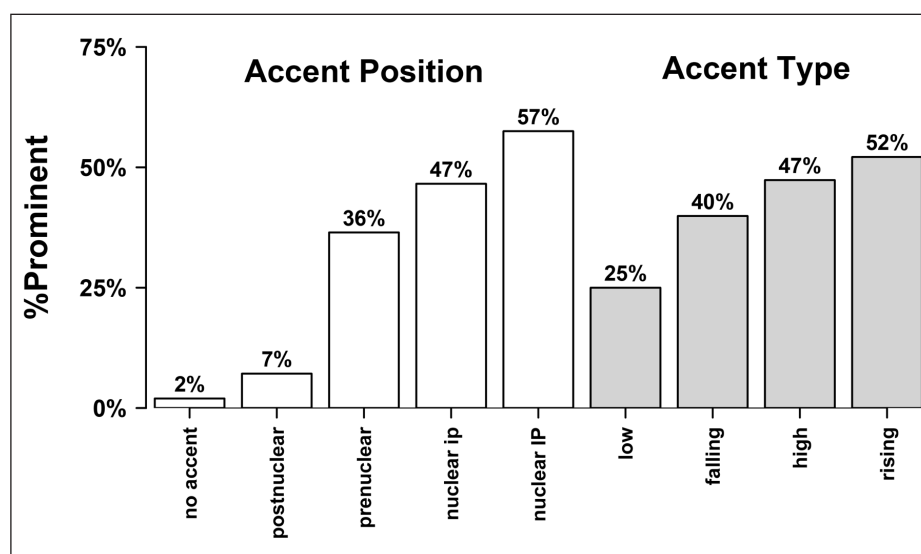
**Figure 3:** Metrical prominence relation (s = stronger, w = weaker) between prosodically defined elements (adopted from Liberman 1975).

<sup>1</sup> RPT has already been applied to a wide variety of languages, among them Hindi (Jyothi et al. 2014), Russian (Luchkina & Cole 2014), French (Smith 2011; 2013) and Spanish (Hualde et al. 2016).

of prominence” (Cole & Shattuck-Hufnagel 2016: 10), which has been done in several previous studies. For spontaneous American English, Cole and colleagues (Cole, Mo & Baek 2010; Cole, Mo & Hasegawa-Johnson 2010; Mahrt et al. 2012) showed that duration and overall intensity (RMS) were particularly important for prominence perception. For German, Baumann & Winter (2018) found that pitch movement in the vicinity of the stressed syllable was the most important factor. Figure 4 indicates that the prominence scores for fully-fledged pitch accents (in prenuclear and nuclear position) were higher than the scores for postnuclear prominences. As can be seen in Figure 4 (on the left), phrase accents (referred to as “postnuclear accents” in the figure) were identified only as slightly more prominent than “no accents” by most listeners (mean: 7.4% vs. 2.1%), whereas prenuclear accents had a much higher prominence score (mean: 36.5%).

The same study revealed a systematic ranking of pitch accent *types* affecting prominence perception (irrespective of their position in the utterance; see Figure 4 on the right). The classification was based on another metalinguistic prominence rating task (see Baumann & Röhr 2015) accounting for the factors *direction of pitch movement* (on)to the accented syllable (rising > falling), *degree of pitch excursion* (steep > shallow) and *height of the accentual tone* (high > mid/downstepped > low).

However, it has been shown that it is not only *acoustic* highlighting that leads to the perception of an element’s prominence but also *expectations* derived from the listener’s knowledge about the linguistic structure of a language. Cole, Mo & Hasegawa-Johnson (2010) found that word frequency and textual givenness influence a listener’s judgment of prominence as well (both standing in an inverse relation with perceived prominence). New information is typically processed faster when it is accented and given information shows processing advantages when being deaccented (e.g. Bock & Mazzella 1983; Terken & Nootboom 1987; Birch & Clifton 1995). These effects are further modulated by information structural notions like focus, contrast and the presence of licit accents elsewhere in a clause (e.g. Nootboom & Kruyt 1987; Sedivy et al. 1999; see Cutler et al. 1997; Birch & Clifton 2002 for a comprehensive overview). The interplay of acoustic cues and expectations has further been observed during the processing of intonational boundaries and in segmentation (e.g. Brown et al. 2011; Buxó-Lugo & Watson 2016). In addition, the more *attentional resources* are required in processing a signal that stands out or is otherwise unexpected (e.g. by encountering strong acoustic/prosodic cues and



**Figure 4:** Percentages of average prominence marks for different accent positions (left) and accent types (right) in a *Rapid Prosody Transcription* task on German (Baumann & Winter 2018: 31).

by the cognitive inaccessibility – i.e. newness – of lexical items), the higher is the probability that a word or syllable will be perceived as prominent (for attention allocation see Corbetta & Shulman 2002; Ranganath & Rainer 2003).

Actually, the expectation evoked by the discourse context may modulate a listener's speech perception so that the information conveyed by the acoustic signal is reinterpreted. This has been shown by Bishop (2012), who conducted a prominence rating study on American English SVO constructions where *acoustically identical* target sentences with a potentially ambiguous focus structure (nuclear accent on the object in sentences such as ...*because I bought a MOTOrcycle*) were presented with context questions that induced varying focus expectations. Although previous (production) studies have shown that the prosodic differences between broad and narrow focus are often only subtle, and at the same time subject to speaker-specific variation (see Snedeker & Trueswell 2003), it was still revealing that the context-induced expectations led to systematic differences in prominence perception: the object in a narrow focus structure was judged as more prominent (and the verb as less prominent) than the object in a broad focus sentence (where the verb was judged as more prominent than under narrow focus). The results were consistent across individual subjects, although they can be assumed to differ, among other things, in their pragmatic skills (Bishop 2017). Bishop (2012) concludes that listeners are actually aware of the subtle differences in prosody-meaning mapping that speakers employ but that these differences usually do not surface in laboratory settings but only in more natural contexts utilizing a purposeful communicative goal (as e.g. in Breen et al. 2010).

In the current research, we are particularly interested in the real-time processing of secondary prominences and how they relate to two information structural notions, focus and newness. To this end, we turn to cases of *Second Occurrence Focus*, which will be described in more detail in the next section.

### 1.2 The relation between prosodic prominence and information structure

Following the logic of the allegedly universal *effort code* (Gussenhoven 2004) – i.e. the more important an item is for a speaker, the more articulatory effort s/he will spend to produce it – we may assume a more or less linear relation between linguistic importance and prosodic prominence. In fact, there is some evidence, at least for Germanic languages, that the information structurally relevant concepts of *focus* and *newness* are marked by greater prosodic prominence whereas *given* elements in the *background* are produced in a prosodically less prominent manner. To be more concrete, in English, German and Dutch, an item that is at the same time discourse-new and (narrowly) focused is often produced with a high or rising pitch accent on that item (as *movies* in (1B)),<sup>2</sup> while a discourse-given item is deaccented (indicated by “Ø”), since it is predictable from the context, either textually or inferentially (as *movies* in (2B) which is lexically given). Deaccentuation implies lack of pitch movement, generally accompanied by reduced duration and intensity (see the description of Figs. 1 and 2 above).

- (1) A: Where did you go?  
 B: To the MOVies.  
       L + H\*
- (2) A: Did you go to the movies?  
 B: No, I don't LIKE movies.  
       Ø


<sup>2</sup> Nuclear accents are indicated by capital letters. Note that the notation of accent types throughout the whole paper follows GToBI (German Tones and Break Indices; Grice et al. 2005), which is based on the framework of autosegmental-metrical phonology (see Ladd 2008).

More interesting than information-structurally and prosodically clear-cut examples like these are hybrid cases whose (prosodic) encoding and (cognitive) decoding is far less well understood. A case in point is the so-called *Second Occurrence Focus* (SOF), which is defined, in its classic form, as a contextually given expression that is at the same time morpho-syntactically focused by virtue of a focus-sensitive particle (Büring 2013; Baumann 2016).<sup>3</sup> In the famous example (3) the second mention of *vegetables* is both focused (due to *only*) and textually given, while the first mentions of both *vegetables* and *Paul* are focused and new (and are thus called *First Occurrence Focus*, FOF).

- (3) Partee (1999: 215)  
 A: Everyone knew that Mary only eats [VEgetables]<sub>FOF</sub>.  
 B: If even [PAUL]<sub>FOF</sub> knew that Mary only eats [VEgetables]<sub>SOF</sub>,  
 then he should have suggested a different restaurant.

The assumption of standard “association with focus” theories (e.g. Jackendoff 1972) is that a focus particle like *only* is associated with a syntactic constituent which contains a semantic focus. This semantic focus in turn contains at least one element which is marked by prosodic prominence. Acoustic and articulatory production studies have shown that FOF elements are generally marked by fully-fledged pitch accents (indicated in (3) by capital letters on the accented syllables), whereas SOF elements are marked by phrase accents, i.e. postnuclear prominences expressed by increased duration and intensity – in comparison with Background elements – but not by tonal movement (indicated in (3) by small capitals) (see Rooth 1996; Bartels 2004; Beaver et al. 2007 for American English; Féry & Ishihara 2009 and Baumann et al. 2010 for German).

As mentioned in section 1.1 above, phrase accents can be regarded as *secondary prominences*, here reflecting the combination of *boosting* (FOCUS) and *inhibiting* (GIVENNESS) information structural factors (see Féry & Ishihara 2009). Following up on this line of argumentation, FOF elements can be claimed to be marked by two boosting factors (FOCUS and NEWNESS), leading to a *primary prominence*, and Background elements (as e.g. *Mary* in example (3)) to be marked by two inhibiting factors (NON-FOCUS and GIVENNESS), resulting in complete *lack of prominence*. Equivalent weighting procedures with factors that have a “positive/amplifying” or a “negative/weakening” influence on an element’s surface prominence have been proposed by Selkirk (2008) and Beaver & Velleman (2011). What all these approaches have in common is the assumed relevance of three distinct levels of prosodic prominence that mirror three distinct levels of information structural weight or importance. (4) translates this relation into a system of binary features.

- |     |  |   |  |
|-----|--|---|--|
| (4) | <i>Information structural importance</i><br>FOF (+ focus, + new)<br>SOF (+ focus, –new)<br>Background (–focus, –new) | <div style="border: 1px solid black; padding: 2px; display: inline-block;">increase</div>  | <i>Prosodic prominence</i><br>Pitch accent (+ pitch, + dur)<br>Phrase accent (–pitch, + dur)<br>No accent (–pitch, –dur) |
|-----|--|---|--|

In the largest empirical study on SOF to date, Beaver et al. (2007) investigated not only the production but also the *perception* of SOF elements, in comparison with Non-Focus items (here: background elements that are not in the scope of a focus particle, see examples (5) and (6), which are adopted from Beaver et al. (2007));<sup>4</sup> FOF elements were not tested in

<sup>3</sup> See Büring (2013) and Baumann (2016) for an overview and a discussion of variants of the traditional view on Second Occurrence Focus.

<sup>4</sup> Note that the semantic (first occurrence) foci – namely *Sid* in (5B) and *court* in (6B) – are determined by the (B) sentences, since they select an alternative presented in the (A) sentences. Only these semantic foci of the (B) sentences count as SOF elements when repeated in the (C) sentences, whereas the elements that

the perception part of the study). Since it has been claimed that SOF is an “inaudible” focus (see Krifka 2004), it was important to test whether the acoustic and articulatory correlates found for SOF actually are perceptible.

- (5) A: Both Sid and his accomplices should have been named in this morning’s court session.  
 B: But the defendant only named [SID]<sub>FOF</sub> [in court]<sub>NON-FOCUS</sub> today.  
 C: Even [the state PROsecutor]<sub>FOF</sub> only named [SID]<sub>SOF</sub> [in court]<sub>NON-FOCUS</sub> today.
- (6) A: Defense and Prosecution had agreed to implicate Sid both in court and on television.  
 B: Still, the defense attorney only named [Sid]<sub>NON-FOCUS</sub> [in COURT]<sub>FOF</sub> today.  
 C: Even [the state PROsecutor]<sub>FOF</sub> only named [Sid]<sub>NON-FOCUS</sub> [in COURT]<sub>SOF</sub> today.

As stimuli for the perception test the authors selected 40 minimal sentence pairs like (5C) and (6C) from the production data taken from the same speaker. Subjects had to judge in which of the two isolated sentences a target word (here: *Sid*) was more prominent than a competitor (here: *court*). The experiment revealed that SOF targets were judged as more prominent than Non-Focus elements in 63% of the cases (all 14 subjects performed above chance). This result can be taken as support for the assumption that the secondary prominence of SOF does not only manifest itself in acoustic features (increased duration and relative energy<sup>5</sup>) but also has perceptual relevance.

There has been some debate about whether only the “classic” SOF cases with their verbatim repetitions of first occurrence expressions are marked by secondary prominences or also elements that are accessible from the context but which are expressed by lexically new material (or at least *not* by identical copies of the first occurrence expressions). Krifka (2004: 203) called this specific group of cases “quasi second occurrence expressions”, and we will borrow this term for the present study. In fact, we will refer to elements which display a combination of boosting and inhibiting factors but which are not morpho-syntactically marked as focus as *Quasi-SOF* items.

An example from Buring (2007) for this type of information structural hybrid is given in (7). Here, *the butcher* constitutes an *epithet* as described by Clark (1977), namely as a bridging inference that adds some derogatory information about an antecedent.

- (7) A: Did you see Dr. Cremer to get your root canal?  
 B: Don’t remind me. I’d like to STRANgle [the butcher]<sub>QUASI-SOF</sub>.

That is, *the butcher* is identical with the previously mentioned *Dr. Cremer* (and thus referentially *given* or coreferential) but at the same time consists of a discourse-new expression (i.e. it is lexically *new*). This distinction between a referential and a lexical level of givenness<sup>6</sup> leads to a more fine-grained differentiation of the information structure categories discussed so far, refining the strictly binary model in (4) above.

This is illustrated in (8). Note that for the purposes of the present study, “±focus” refers to the presence or absence of a focus particle, which will be restricted here to *only*. In the framework of alternative semantics (based on Rooth 1992), *only* is an (exclusive)

were not part of the alternative set are classified as “non-focal”, i.e. they are not regarded as being in the scope of the focus particle *only*.

<sup>5</sup> *Relative energy* is derived by multiplying duration by root-mean-square intensity, a measure Beckman (1986) called “total amplitude”, which she found to be the most important postlexical stress marker in English.

<sup>6</sup> For a detailed discussion of the relevance of this distinction, including the proposal of an annotation scheme differentiating between a referential and a lexical level of givenness (*RefLex*) see Baumann & Riester (2012; 2013). See also Lam & Watson (2014, and the references therein) who conducted psycholinguistic production experiments on American English disentangling the two levels.

expression whose interpretation is sensitive to focus (see also the stimuli description in 2.2, Beck 2016, as well as Beaver & Clark 2008 for an elaborate definition of *only* and its semantic effects). That is, *only* associates with the target words as the semantic foci in FOF and SOF structures, which makes them “important”, but they differ in whether they are at the same time contextually new (more “important”, FOF) or given (“less important”, SOF). Nevertheless, SOF is ranked above Quasi-SOF because the former is focused by the exclusive expression *only*, which may be considered stronger than standing in “free focus” (i.e. only contextually focused but not associated with a focus-sensitive operator, see Büring 2013) like Quasi-SOF expressions. Furthermore, Quasi-SOF items are only *lexically* but not *referentially* new, indicated by the combined  $\pm$  feature for newness, which is considered weaker than the  $+$  feature for focus in SOF.<sup>7</sup> Finally, Background elements are neither contextually nor morpho-syntactically focused and additionally given information, which clearly makes them least “important”.

(8) *Information structural importance*

|            |   |
|------------|---|
| FOF        | + focus, + new (referentially and lexically new)          |
| SOF        | + focus, –new (referentially and lexically given)         |
| QUASI-SOF  | –focus, $\pm$ new (referentially given but lexically new) |
| Background | –focus, –new (referentially and lexically given)          |



Actually, the degree of prosodic prominence (or rather lack of it) of Quasi-SOF items like *the butcher* in (7) is disputed and probably subject to a large amount of variation. In German and English, such expressions will either be deaccented (as proposed by Büring 2007) or carry some kind of secondary prominence (as suggested e.g. by Rooth 1996 and Krifka 2004). However, these proposals are based on individual intuitions rather than an extended data analysis. In any case, though, a fully-fledged pitch accent on *the butcher* is clearly prohibited, since it would rule out the intended coreference reading. For our experiment, we chose to produce Quasi-SOF items as deaccented, following Büring’s (2007) original intuition.

While previous research on SOF has employed behavioral measures, we investigate the real-time correlates of SOF and Quasi-SOF using electrophysiological measures during language comprehension. This will allow us to assess the contribution of prosody, focus and newness to the processing of SOF elements.

### 1.3 Neurocognitive processing of prosody and information structure

Event-related brain potentials (ERPs) allow for a fine-grained characterization of the time-course of the underlying processes (see Bornkessel-Schlesewsky & Schumacher 2016). They represent synaptic changes that are time-locked to a cognitive event and that are recorded from electrodes placed on the listeners’ scalp.

From a neurocognitive perspective, prominence-lending cues such as pitch movement or increased duration are computed as the sensory input unfolds, and expectations are incrementally built up for upcoming entities, including their prosodic realization (see Hickok & Poeppel 2015; Bornkessel-Schlesewsky & Schumacher 2016 for overviews). Input that mismatches these expectations results in a prediction error (see Friston 2010; Bornkessel-Schlesewsky & Schlewsky 2019), which has, for instance, been shown to yield a pronounced N400 – a negative potential peaking around 400 ms after the onset of a

<sup>7</sup> We refrained from including a condition combining a free focus with fully new information, since it would have further increased the complexity of the design. Nevertheless, it is certainly an interesting research question whether a focus derived from the context alone – and marked by a pitch accent – is processed differently from a focus marked by *both* a focus particle *and* an accent.



critical entity. The N400 has been observed during semantic processing: the less expected a word, the more pronounced the amplitude of the N400 (e.g. Kutas & Federmeier 2011; inter alia); it has also been observed during referential processing: the less accessible a discourse entity, the more pronounced the amplitude of the N400 (e.g. Burkhardt 2006; Schumacher & Hung 2012). In prosody, inappropriate and thus unexpected accent types have also been shown to elicit an N400 (e.g. Heim & Alter 2006; Toepel et al. 2007; Baumann & Schumacher 2012).

Previous neurophysiological studies on the processing of information structure and prosody further reported on another relevant ERP marker, known as *expectancy negativity* (EN). This potential has been observed in anterior regions and as early as 200 ms after the onset of a target entity. It is commonly interpreted as being triggered by the expectation of a focused, and accented, constituent in a given discourse context. The contextual licensors in earlier studies consisted either of a focus-eliciting question (e.g. Hruska & Alter 2004; Toepel et al. 2009) or a focus particle assigning focus to its right-adjacent constituent (e.g. Heim & Alter 2006; 2007). These findings indicate that different cues (focus particles or context questions) can lead to similar processing responses to accented constituents.

Prominence-lending cues further serve as attention orienting signals. The mechanism of attention orienting requires listeners to update their mental model (e.g. when expressing a topic shift or introducing new information), which has been claimed to engender a late positivity (henceforth “LP”), i.e. a positive-going potential with a peak latency around 600 ms after the critical entity (see Burkhardt 2006; Schumacher & Hung 2012; Brouwer & Hoeks 2013; Wang & Schumacher 2013 – see also Bornkessel et al. 2003 for earlier onset latencies). This LP has further been shown to occur with prosodic cues: It was found for prosodically marked new information and focused constituents (e.g. Hruska & Alter 2004; Toepel et al. 2007) as well as with deaccentuation (Baumann & Schumacher 2012). Thus, the results for prosodic cues are mixed. Unexpected accents, too, may cause a LP (in addition to an N400), which may be interpreted as an instance of mental model repair, since unexpected accents can conflict with information structural cues and this conflict must be resolved (e.g. Magne et al. 2005; Toepel et al. 2007; Schumacher & Baumann 2010; Dimitrova et al. 2012; Brouwer & Hoeks 2013).

Baumann & Schumacher (2012) have shown for German that prosody and information structure are processed by evoking similar combinations of N400 and LP; that is, less expected cues engendered a more enhanced N400 and attention-drawing cues showed a pronounced LP. They crossed information structure (discourse-new vs. discourse-given) and accent type (L + H\* vs. deaccentuation), as illustrated in (9)–(12), and observed effects of information structure in both the N400 and the LP time windows (new > given) as well as a biphasic N400-LP response to prosody (deaccentuation > L + H\* accent). In this study, crucially, the mismatches – i.e. missing accent on the noun (here: *Winzer*, ‘winegrower’) in (10) and superfluous accent on the noun in (11) – did not pattern together relative to the matching conditions. Rather, prosodic cues (deaccentuation > pitch accent) and information structural cues (new > given) grouped together. This was observed both at the noun position (representing referential givenness) and at the position of the adjective in the target sentence (constituting lexical givenness; see Baumann & Riester 2012; 2013 for different types of givenness).

(9) **Discourse-new & L + H\* accent on noun (match)**

Context: FRAUke meinte, dass der HOLZfäller nicht sehr HEIter war.  
‘Frauke said that the lumberjack was not very cheerful.’

Target: Sie erWÄHNte, dass **der WINzer** sehr heiter war.  
‘She mentioned that the winegrower was very cheerful.’

- (10) **Discourse-new & deaccented noun (mismatch)**  
 Context: FRAUke meinte, dass der HOLZfäller nicht sehr HEIter war.  
 ‘Frauke said that the lumberjack was not very cheerful.’  
 Target: Sie erWÄHNte, dass **der Winzer** sehr HEIter war.  
 ‘She mentioned that the winegrower was very cheerful.’
- (11) **Discourse-given & L + H\* accent on noun (mismatch)**  
 Context: Vivian berichtete von einem WINzer in BAden.  
 ‘Vivian talked about a winegrower in Baden.’  
 Target: Sie erWÄHNte, dass **der WINzer** sehr heiter war.  
 ‘She mentioned that the winegrower was very cheerful.’
- (12) **Discourse-given & deaccented noun (match)**  
 Context: VIvian berichtete von einem WINzer in BAden.  
 ‘Vivian talked about a winegrower in Baden.’  
 Target: Sie erWÄHNte, dass **der Winzer** sehr HEIter war.  
 ‘She mentioned that the winegrower was very cheerful.’

Interestingly, the biphasic patterns for prosody and information structure mapped onto *different topographical distributions* on the scalp: Prosodic cues led to ERP effects with an *anterior* maximum while information structural cues evoked ERP profiles that were most pronounced over *posterior* regions. This led the authors to suggest that prosodic and information structural cues are subject to the same general mechanisms (expectation-based processing and mental model updating) but that they may be processed by discrete underlying networks that surface on the scalp in an anterior-posterior divide. This latter claim is partly supported by previous studies on information structure in the written modality that show effects with a posterior maximum (e.g. Burkhardt 2006; Schumacher & Hung 2012) while research on ambiguity resolution associates discourse complexity with effects over anterior electrode sites (e.g. Kaan & Swaab 2003; van Berkum et al. 2007). Prosodic effects like the *expectancy negativity* have surfaced over anterior regions, but findings from the N400 or LP have been mixed.

Previous ERP research on prosody has usually employed a mismatch paradigm and contrasted the processing of appropriate and inappropriate prosodic realizations as a function of context (e.g. Hruska & Alter 2004; Magne et al. 2005; Heim & Alter 2006; 2007; Toepel et al. 2007; 2009; Dimitrova et al. 2012; Baumann & Schumacher 2012). In the following, we want to investigate the processing of *prosodic prominence* in a more direct way. Usually, two clearly different prominence values occur in matching vs. mismatching contexts (e.g. expressed by a steeply rising “contrastive” accent on the one hand and complete lack of accent on the other, see Toepel et al. 2007; Baumann & Schumacher 2012). A first step to investigate different degrees of accentuation was Schumacher & Baumann’s (2010) investigation of two accent types (H\*, H+L\*) plus deaccentuation, representing three degrees of prominence. H\* stands for a high pitch on an accented syllable and H+L\* for a falling accent with high pitch on the pre-accentual syllable and low pitch on the accented syllable. They investigated the comprehension of inferential relations in German (e.g. *Sabine repairs an old shoe. In doing so, she cuts the sole.*) and varied the accent types on the sentence-final inferentially linked definite expression. Previous research has identified an H+L\* accent to be the most appropriate accent for this whole-part relation (Baumann & Grice 2006). And indeed, deviations from the expected prosodic realization were reflected in N400 effects, which were further modulated by the severity of the deviation (H+L\* < H\* < deaccentuation) and followed by a LP. Nevertheless, the aim of that study was once more to examine the contextual appropriateness of the target items and not the processing of their prosodic prominence.

In the present study, we are not concerned with mismatches but exclusively with contextually appropriate prosody, expressing three different degrees of prominence. In particular, we investigate the processing of primary and secondary prominences by comparing *First Occurrence Focus*, *Second Occurrence Focus* and *Background* in the response to short dialogue sequences. As outlined in (4) and (8) above, these conditions differ in terms of their information structural importance ( $\pm$ focus,  $\pm$ new) and prosodic prominence ( $\pm$ pitch,  $\pm$ dur). In addition, we will include a condition reflecting *Quasi Second Occurrence Focus*, which differs from SOF by being referentially given yet lexically new.

#### 1.4 Research questions

The studies discussed so far provide evidence that secondary prosodic prominences, e.g. in the form of (postnuclear) phrase accents, may serve as markers of linguistically meaningful distinctions. Here, we presented *Second Occurrence Focus* (SOF) as a case in point. Investigations on SOF have shown that an intermediate level of prominence has a phonetic basis, and that it can also be perceived as different from higher as well as lower levels of prominence. Furthermore, different levels of prosodic prominence have been shown to be perceived differently in a wide range of behavioural studies, both metalinguistic and related to a linguistic task.

The present study is looking for evidence for the assumption that secondary prominence (here: phrase accents) is also *neurocognitively* processed in a different way than primary prominence (here: pitch accents) on the one hand and lack of prominence (here: deaccentuation) on the other. SOF is an ideal testbed for this research question, since it additionally provides clearly defined information structural categories such as focus, newness and (textual) givenness whose role in cognitive processing has been tested before but not in a sufficiently controlled way.

## 2 The present study

To examine the processing of primary and secondary prominences, we compare four types of critical items. We contrast *First Occurrence Focus* (FOF), *Second Occurrence Focus* (SOF), *Quasi Second Occurrence Focus* (Quasi-SOF) and *Background* (BG) in the answer of a mini dialogue to investigate the neural correlates of prosodic and information structural cues in contextually licensed exchanges. Hypotheses 1–2, formulated below, apply to the comparison of FOF, SOF and BG, while hypothesis 3 addresses Quasi-SOF.

### 2.1 Hypotheses

Our hypotheses are concerned with the processing of prominence and with the potential difference between the processing of information structural and prosodic cues. Critically, the present study does not intend to investigate mismatches between the two levels of linguistic description but measures *contextually licensed prosody*. In other words, we claim that all prosodic realizations used in the test materials are appropriate in their given contexts – a claim which is based on the results of previous production and perception studies discussed above.

The processing of prominence affects both expectation-based mechanisms and mental model updating. Cues given by the context are used to generate expectations for upcoming entities and incoming prosodic or information structural cues modulate the amplitude of the N400. Prominence-lending cues further have the capacity to initiate updates of the mental representation, which is reflected in a more enhanced LP.

The current research seeks to tease apart effects of prosody (primary prominence, secondary prominence, no prominence) and information structure (newness and focus) by looking at ERP effects associated with expectation-based processing and mental model

updating that have been shown in previous research. Accordingly, we formulate separate hypotheses for prosody (H1) and information structure (H2) below.

*Hypothesis 1:* As to the prosodic contrasts, we are primarily interested in the status of phrase accents as secondary prominences, i.e. we predict that they represent an intermediate level of prominence between pitch accents and deaccentuation.

The three levels of prosodic prominence tested may trigger stepwise differences in ERPs. Based on the findings of Baumann & Schumacher (2012) we expect decreasing N400 and LP amplitudes from deaccentuation ( $\emptyset$ ) through phrase accents to fully-fledged pitch accents. The ranking in (13) uses the phonetic parameters PITCH MOVEMENT ( $\pm$ pitch) and DURATION ( $\pm$ dur) somewhat simplified as binary features. In combination, they reflect the three levels NO PROMINENCE, SECONDARY PROMINENCE and PRIMARY PROMINENCE.

(13) **H1:**  $\emptyset$  ( $-$ pitch,  $-$ dur) > phrase accent ( $-$ pitch, + dur) > pitch accent (+ pitch, + dur)

Translated into the information structural conditions which carry the three degrees of prominence, H1 can be stated as:

(14) BG ( $\emptyset$ ) > SOF (phrase accent) > FOF (pitch accent)

Building on the observation that prosodic and information structural mechanisms show distinct topographical distributions (see Baumann & Schumacher 2012), the prosodic effects are predicted to surface over anterior brain regions.

*Hypothesis 2:* Regarding the information structural contrasts, we investigate whether information structural effects are driven by  $\pm$ newness or  $\pm$ focus or a combination of them. If  $\pm$ newness is the driving force, then SOF and BG should pattern together relative to FOF. If the distinction  $\pm$ focus is primarily used, then SOF and FOF on the one hand should differ from BG on the other hand. If both newness and focus contribute to the underlying processes, a graded distribution is predicted.

In this latter case, we would expect increasing N400 and LP amplitudes from BG through SOF to FOF. The ranking in (15) uses the binary feature FOCUS ( $\pm$ focus) and the simplified feature NEWNESS ( $\pm$ new). The combined features indicate three levels of *importance* ranging from low for BG to high for FOF.

(15) **H2:** BG ( $-$ foc,  $-$ new) < SOF (+ foc,  $-$ new) < FOF (+ foc, + new)

Following Baumann & Schumacher (2012), the information structural differences should have a maximum over posterior regions.

H1 and H2 work in opposite directions according to the results found in previous studies. If the two levels of investigation, i.e. prosodic prominence and information structural importance, operate independently and assuming different topographical distributions, we should observe the patterns predicted in (13)–(15). If the two levels interact with each other and are subserved by the same neural networks, we might observe that one cue outranks the other (reflected in ERP signatures in line with *either* H1 *or* H2). Alternatively, prosodic and information structural cues may interact with each other in an unweighted manner; since we are testing contextually licensed utterances, FOF and BG continuations are more predictable than secondary prominences, and the latter might thus evoke more processing effort with respect to expectation-based mechanisms and mental model updating.

*Hypothesis 3:* As to the exploratory investigation of the processing of Quasi-SOFs, we have the following predictions: With respect to prosody, Quasi-SOF items will be processed in a similar fashion as BG items, since both are deaccented. With respect to information

structure, the Quasi-SOF condition can shed further light on whether newness relies on referential or lexical information: Quasi-SOF elements are lexically new (but referentially given) and should be processed similarly as SOF elements because they show attenuating (= coreference) and boosting (= lexical newness) effects just like SOF items (which are given but focused). Hence Quasi-SOF should pattern with SOF but lexical differences may evoke a moderate N400 amplitude, however less pronounced than (referentially and lexically new) FOF.

- (16) **H3:** prosody: BG ( $\emptyset$ ) = Quasi-SOF ( $\emptyset$ ) > SOF (phrase accent) > FOF (pitch accent)  
 information structure: BG (-foc, -new) < SOF (+foc, -new) = Quasi-SOF (-foc,  $\pm$ new) < FOF (+foc, +new)

## 2.2 Methods

### 2.2.1 Participants

Twenty-four right-handed, monolingual native speakers of German from the University of Cologne participated in the ERP experiment after giving written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the ethics committee of the *German Linguistic Society* (No. 2016-09-160914). Three participants had to be discarded from the analysis due to excessive ocular and movement artifacts. The age of the remaining twenty-one participants (17 women, 4 men) ranged from 19 to 27 years (mean-age: 23.1 years). None of them reported any auditory, visual or neurological deficits.

### 2.2.2 Stimuli

Stimuli were created for four conditions, with the target word as a FOF, SOF, BG or Quasi-SOF item integrated in the answer of a mini dialogue (see examples (17) to (20)).<sup>8</sup> All target items consist of monosyllabic German words. In the FOF, SOF and BG conditions, the target words denote beverages or food, whereas the Quasi-SOF items constitute more general expressions with a derogatory connotation. We made sure that the frequencies of the target items were comparable.<sup>9</sup> Focus is marked morpho-syntactically by the exclusive *nur* ('only'), which has been classified by Beaver and Clark (2008) as an expression that has a conventional association with focus. It is kept constant in the FOF and SOF conditions, i.e. we exclusively use the focus particle *only* (in a non-scalar reading) before the target word to avoid potential meaning differences caused by other types of focus operator. In the SOF and Quasi-SOF conditions, the subject is further indicated by the scalar additive *sogar* ('even'), another focus-sensitive particle introducing the element that carries the nuclear accent of the respective utterance. This structure is kept constant throughout the whole stimuli set.

The target words in the examples below are printed in bold face. Capitals (in FOF condition) indicate fully-fledged (nuclear) pitch accents, small capitals (in SOF condition) mark phrase accents, and lack of capitalisation indicates complete lack of prominence (in BG and Quasi-SOF conditions).

- (17) **FOF**  
 Context: Was gibt's Neues? ('What's new?')  
 Target: Karl hat nur **BIER**<sub>FOF</sub> getrunken. ('Karl only drank BEER.')

<sup>8</sup> Initially, the Quasi-SOF condition, which is treated as an exploratory contrast, was compared to SOF and BG in an independent analysis. Following the suggestion of two independent reviewers, however, we included all conditions in a single analysis.

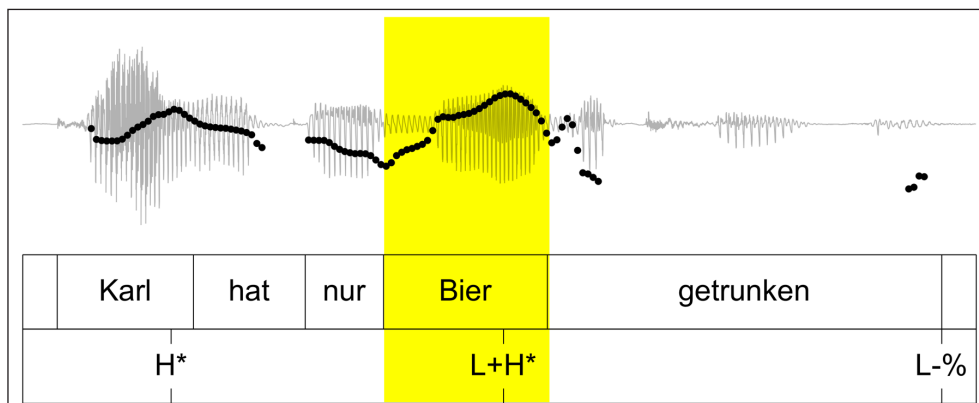
<sup>9</sup> The frequency values were checked on <http://wortschatz.uni-leipzig.de/> which is based on the *Leipzig Corpora Collection* (LCC, Goldhahn et al. 2012).

- (18) **SOF**  
 Context: Eva hat nur Bier getrunken. ('Eva only drank beer.')
- Target: Sogar THOMAS hat nur **BIER**<sub>SOF</sub> getrunken.  
 ('Even THOMAS only drank BEER.')
  
- (19) **BG**  
 Context: Wer hat Bier getrunken? ('Who drank beer?')
- Target: HANS hat **Bier**<sub>BG</sub> getrunken. ('HANS drank beer.')
  
- (20) **Quasi-SOF**  
 Context: Maria hat ein Bier getrunken. ('Maria drank a beer.')
- Target: Sogar MELANIE hat das **Zeug**<sub>QUASI-SOF</sub> getrunken.  
 ('Even MELANIE drank that stuff.')

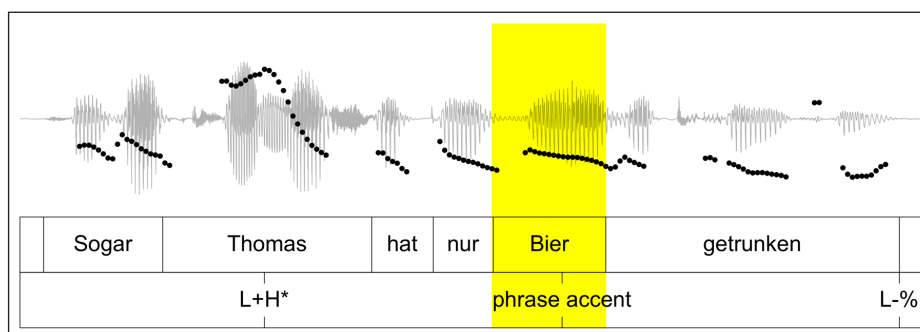
Forty stimuli and the accompanying mini dialogues were created per condition, plus 120 dialogues with a similar structure which served as filler items and which were neither controlled for the number of syllables nor the semantic field of the object in the second sentence (e.g. *Saskia hat einen ROMAN gelesen. Auch PHILLIP hat den Roman gelesen.* 'Saskia read a NOvel. Also PHILLip read the novel.').

All stimuli were read by a trained male phonetician and recorded in a sound-attenuated cabin with a sampling rate of 44100 Hz and 16 bit resolution (mono). Examples of the experimental conditions showing the acoustic differences of the same target word (*Bier*, 'beer'; see (17)–(19) above) as well as the Quasi-SOF condition (see (20)) are given in Figures 5 (FOF), 6 (SOF), 7 (BG) and 8 (Quasi-SOF). Furthermore, the phonological description of the intonation contour using GToBI (see Grice et al. 2005) is shown, which is the same for all forty target sentences per condition. The summarizing plot in Figure 9 indicates that the contours in each of the four conditions were produced in a very stable manner.

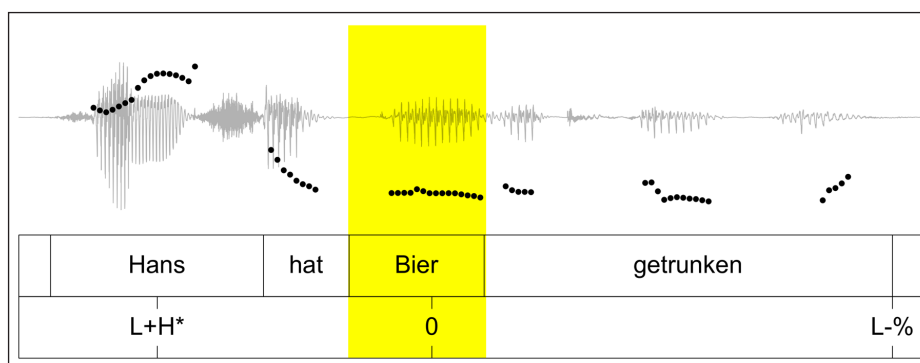
For most stimuli, the originally read version entered the experiment. However, we decided to adjust a small number of target words (in *Praat*; Boersma & Weenink 2013) if the values for duration and pitch range exceeded the mean for a specific condition by more than one standard deviation (SD). Table 1 shows the means and SDs for pitch range (i.e. the difference between the minimum and maximum pitch in semitones (st)) and duration (in milliseconds (ms)) for the target words in each experimental condition. In line with Hypothesis 1 (see (13)) and the actual examples in Figures 5–8, the values show that FOF target words are marked by a larger pitch range than the target words in the



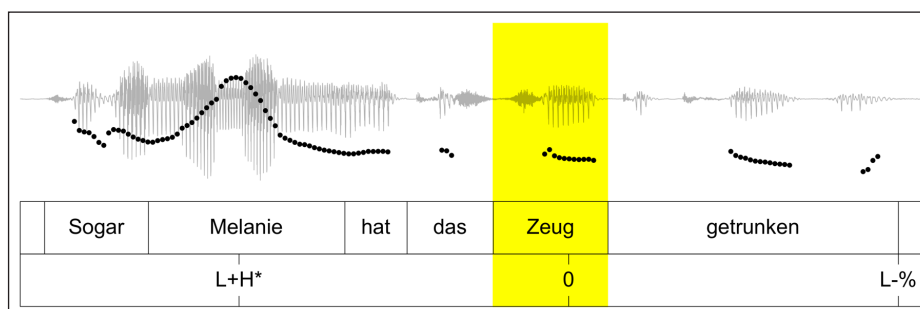
**Figure 5:** Oscillogram and pitch contour of the FOF target utterance *Karl hat nur Bier getrunken* (see (17) above). The GToBI annotation indicates a (high) prenuclear accent on *Karl* and a (rising) nuclear accent on *Bier*, plus a (low) final boundary tone of the intonation phrase.



**Figure 6:** Oscillogram and pitch contour of the SOF target utterance *Sogar Thomas hat nur Bier getrunken* (see (18) above). The GToBI annotation indicates a (rising) nuclear accent on *Thomas* and a phrase accent on *Bier*, plus a (low) final boundary tone of the intonation phrase.



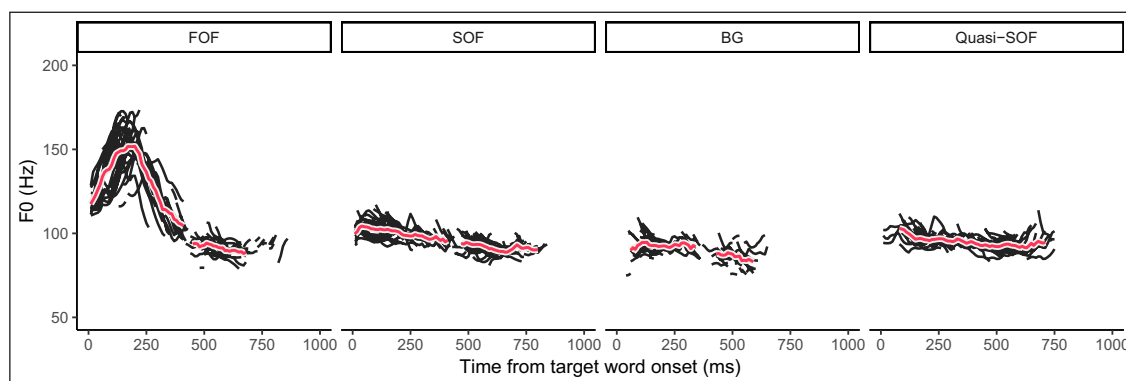
**Figure 7:** Oscillogram and pitch contour of the BG target utterance *Hans hat Bier getrunken* (see (19) above). The GToBI annotation indicates a (rising) nuclear accent on *Hans* and a (low) final boundary tone of the intonation phrase. The “0” on *Bier*, which is not part of the GToBI inventory, is used to indicate complete deaccentuation of the target word *Bier*.



**Figure 8:** Oscillogram and pitch contour of the Quasi-SOF target utterance *Sogar Melanie hat das Zeug getrunken* (see (20) above). The GToBI annotation indicates a (rising) nuclear accent on *Melanie* and a (low) final boundary tone of the intonation phrase. The “0” on *Zeug* indicates complete deaccentuation of the target word.

other conditions. Furthermore, the values for duration are higher for SOF and FOF than for the other two conditions.

After each stimulus, participants performed a word recognition task. For this task, stimuli were matched with a correct and an incorrect probe word, which were then distributed across different lists. Recognition items from the critical conditions represented either verbs or participants in the events; the target noun was never used to probe word recognition. Fillers asked for all content words in the recognition task. Critical and filler



**Figure 9:** Spaghetti plot showing the intonation contours of each of the four experimental conditions with the items superimposed on each other and with an average contour (in red; gained by *Loess* smoothing in R; R Core Team 2018) from the onset of the target word until the end of the target sentence.

**Table 1:** Means and standard deviations (in brackets) for pitch range (in semitones, st) and duration (in milliseconds, ms) for the target words in the four experimental conditions.

|                  | <b>pitch range</b> | <b>duration</b> |
|------------------|--------------------|-----------------|
| <b>FOF</b>       | 4.0 st (1.8 st)    | 268 ms (47 ms)  |
| <b>SOF</b>       | 1.7 st (0.7 st)    | 280 ms (43 ms)  |
| <b>BG</b>        | 1.7 st (1.0 st)    | 235 ms (34 ms)  |
| <b>Quasi-SOF</b> | 1.6 st (0.7 st)    | 234 ms (54 ms)  |

items were pseudo-randomized and presented in seven different blocks with short breaks in between. Participants heard all versions of a set. To counter repetition effects, the members of a set were assigned to different experimental blocks, and in order to avoid systematic order effects in the exposure to the stimuli, the set members were presented in different condition sequences across sets. Each participant saw one of two lists of the 280 items with different randomizations.

### 2.2.3 Procedure

After electrode application, participants were seated in a sound-attenuating cabin. They were instructed to look at the computer monitor in front of them and to focus on a fixation star while the auditory stimuli were presented over loudspeakers. The subjects' task was a word recognition task: After each mini text a word was presented and the participants had to decide whether this word had occurred in the previous item or not. Answers were given by pressing one of two buttons on a game controller.

Each trial began with the presentation of a fixation star in the center of the screen. After 500 ms, the auditory stimulus was presented while the fixation star remained on the screen. After the end of the auditory stimulus, there was a 500 ms blank screen before the probe word was presented visually for the recognition task. Maximum response times to this question were set to 4000 ms, and the inter-trial interval lasted 1000 ms. The experiment consisted of six blocks and participants individually determined the duration of the pauses between blocks. The recording session started with a short practice block during which participants were familiarized with the experimental procedure.

### 2.2.4 Data recording and preprocessing

The electroencephalogram (EEG) was recorded and digitized (500 Hz) by means of 24 Ag/AgCl electrodes placed according to the standard 10-20 system (*BrainVision Brain-Amp*



*amplifier*). EEGs were referenced online to the left mastoid. The ground electrode was placed at AFz. To control for eye-movement artifacts, the electrooculogram (EOG) was recorded by two pairs of electrodes. For horizontal eye movements, these were placed at the outer canthus of each eye, and for vertical eye movements, electrodes were placed above and below the left eye. Electrode impedances were kept below 5 k $\Omega$ .

During preprocessing, data were rereferenced offline to linked mastoids. Instead of applying a baseline correction, the EEG was filtered with a 0.3–20 Hz bandpass filter to remove unsystematic pre-stimulus differences caused by slow signal drifts (see Schumacher & Hung 2012; Widmann et al. 2015; Maess et al. 2016).<sup>10</sup> Trials with eye movements, muscular or amplifier-saturation artifacts were removed automatically (EOG cutoff of  $\pm 40$   $\mu$ V) as well as manually. This resulted in the rejection of 19.34% of the data points over all conditions, with no differences between conditions (minimum rejection per participant and condition: 0, maximum rejection: 20). Average ERPs time-locked to the onset of the target noun were first computed per condition and participant before grand-averaging was performed over all participants.

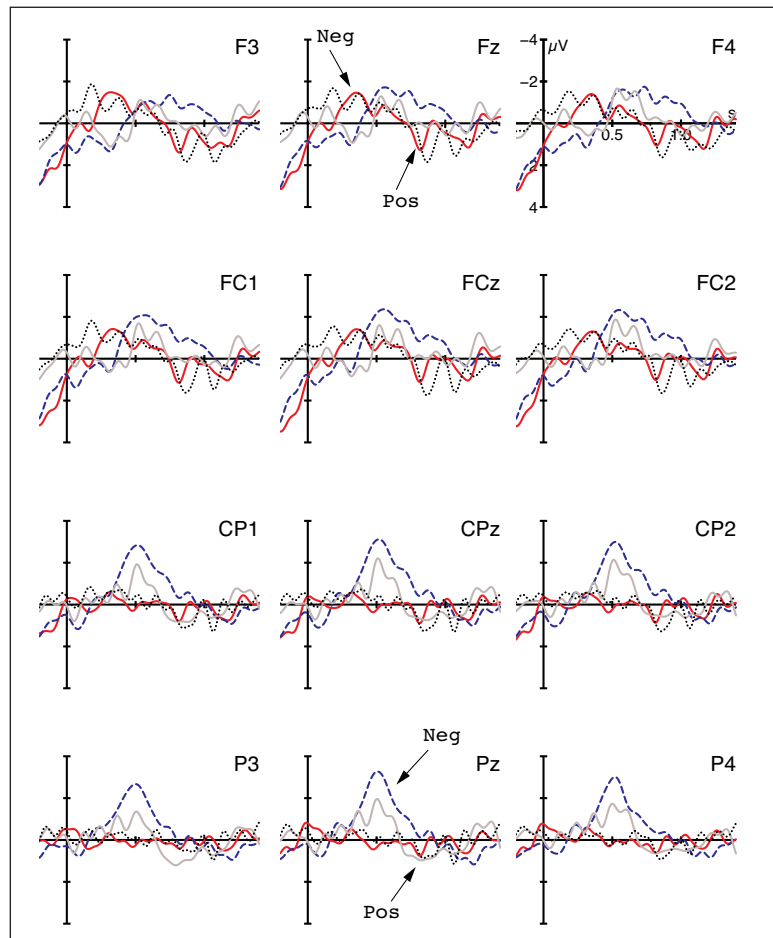
### 2.2.5 Data analysis

Repeated-measures analyses of variance (ANOVAs) using the *ez*-package (Lawrence 2016) in R were computed for the mean amplitude per condition in pretermed time windows with the factor CONDITION and four levels (FOF vs. SOF vs. BG vs. Quasi-SOF). Statistical analyses further included the topographical factor region of interest (ROI) with two levels, which is in line with our topographical hypotheses for prosody and information structure, respectively: anterior (F3/F4/F7/F8/Fz/FC1/FC2/FC5/FC6/FCz/Cz) and posterior (CP1/CP2/CP5/CP6/CPz/P3/P4/P7/P8/Pz/POz). All analyses were carried out hierarchically (i.e., only reliable CONDITION  $\times$  ROI interactions of  $p < .05$  were resolved and followed up by planned comparisons). Huynh–Feldt corrections were applied to counter violations of sphericity (Huynh & Feldt 1970). All possible pair-wise comparisons between the conditions were carried out. The threshold for the p-value of pairwise comparisons was adjusted to  $p < .025$  (Keppel 1991). Analyses were calculated for temporal windows determined by visual inspection.

### 2.3 Results

Figure 10 depicts the grand-average ERPs time-locked to the onset of the critical word (e.g. *Bier* resp. *Zeug* in the examples above). It illustrates different ERP effects for anterior vs. posterior electrode sites. *Over anterior regions*, BG (red solid line) and SOF (black dotted line) show a biphasic pattern compared to FOF (blue dashed line), reflected in a more pronounced negative deflection between 250–400 ms that is followed by a positive-going wave between 750–950 ms. Crucially, BG and SOF do not differ from each other. Note also that there appears to be a very early effect for SOF; however, since this difference already emerges before the onset of the critical word, we refrain from interpreting the early negativity for SOF. We suspect that the pre-stimulus differences arise from length variation between the target sentences of the different conditions. The Quasi-SOF condition differs from SOF and BG: BG (red solid line) and SOF (black dotted line) evoke more pronounced negativities between 250 and 400 ms relative to Quasi-SOF (grey solid line).

<sup>10</sup> We used a bandpass filter (0.3–20 Hz) instead of baseline correction since we consider it a better method to deal with potential pre-stimulus differences, which were unavoidable given our experimental design (the critical word in FOF and SOF is immediately preceded by the focus particle *nur* ('only'); BG and Quasi-SOF are preceded by an auxiliary; in addition, different focus structures are anticipated from the onset of the target sentence). The dangers of using a baseline correction have been made particularly clear by advances in ERP research that have shown that at least certain ERP effects should be viewed as a reorganisation of pre-stimulus activity (e.g., Makeig et al. 2002). Accordingly, we decided to apply a filtering procedure (see Maess et al. 2016 on advantages of appropriate filters over baseline correction).



**Figure 10:** Grand-average ERPs at selected electrodes for the contrast BG (red solid line) vs. SOF (black dotted line) vs. FOF (blue dashed line) vs. Quasi-SOF (grey solid line). Negativity is plotted upwards. Time course on horizontal axis spans from 200 ms before until 1400 ms after the onset of the critical word. An 8 Hz low pass filter was applied for visual presentation.

Subsequently, Quasi-SOF reveals a reduced positivity in the 750–950 ms time-window relative to BG and SOF.

Over posterior electrode sites, the FOF condition (blue dashed) shows a pronounced negative amplitude roughly around 400–650 ms after the onset of the critical word relative to the other two conditions. Regarding the Quasi-SOF condition, electrodes over posterior sites reveal an enhanced negativity between 400 and 650 ms in contrast to SOF and BG, which is less pronounced than the FOF amplitude. In the window from 750–950 ms the FOF condition shows the least positive-going trend relative to the other conditions.

Statistical analyses for the time-window between 250 and 400 ms revealed a  $\text{CONDITION} \times \text{ROI}$  interaction [ $F(3,60) = 21.83, p < .001$ ]. Resolution of this interaction by ROI registered significant effects of  $\text{CONDITION}$  over anterior regions [ $F(3,60) = 13.78, p < .001$ ] but not over posterior regions [ $F(3,60) = 1.27, p > .29$ ]. Pairwise comparison in the anterior ROI yielded a significant effect of  $\text{CONDITION}$  for the contrast BG vs. FOF, SOF vs. FOF, BG vs. Quasi-SOF and SOF vs. Quasi-SOF but no effect between BG and SOF and between Quasi-SOF and FOF (see Table 2 for statistical details).

Analyses for the 400–650 ms window registered a main effect of  $\text{CONDITION}$  [ $F(3,60) = 9.80, p < .001$ ] and an interaction of  $\text{CONDITION}$  BY ROI [ $F(3,60) = 17.45, p < .001$ ]. Resolution of the interaction by ROI showed an effect of  $\text{CONDITION}$  over posterior sites [ $F(3,60) = 25.37, p < .001$ ] and no effect over anterior sites [ $F(3,60) = 1.09, p > .35$ ]. Pairwise comparisons in the posterior ROI revealed differences between all contrasts except for BG vs. SOF.

**Table 2:** Pairwise comparisons for BG, FOF, SOF and Quasi-SOF. Grey shades indicate non-significant differences (which result in non-resolution of the overall CONDITION × ROI interaction).

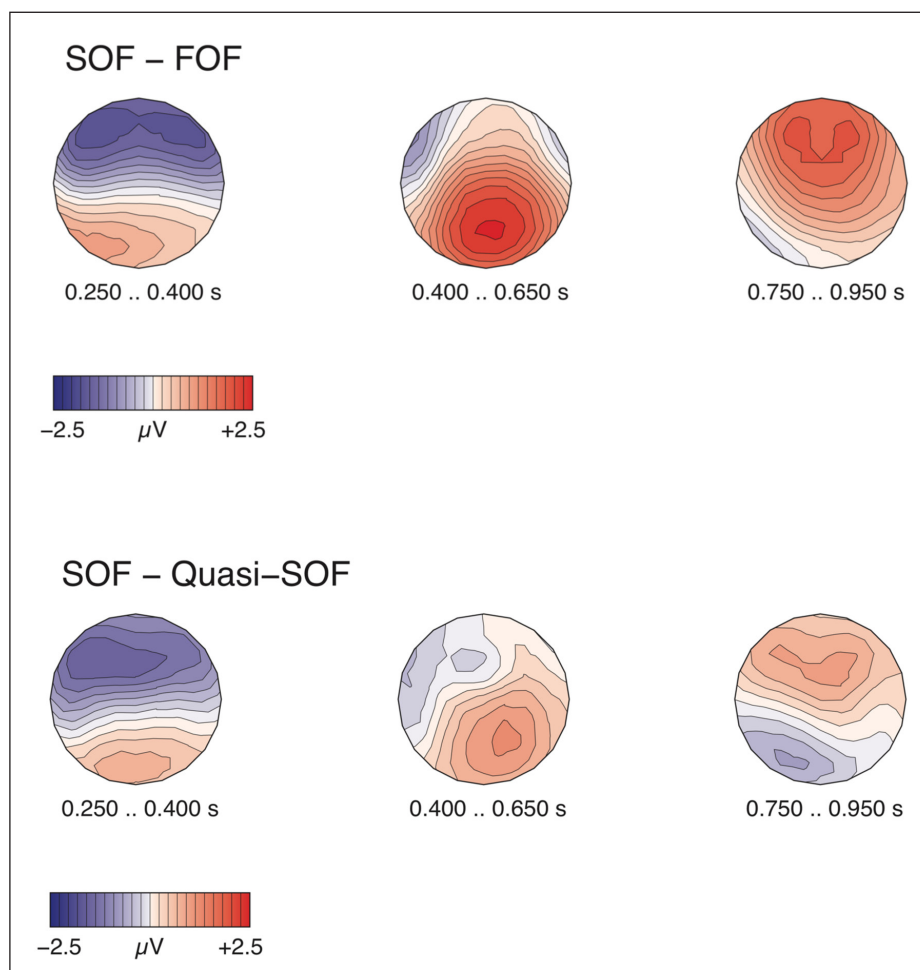
|             | COND × ROI | BG vs. FOF            | BG vs. SOF          | FOF vs. SOF           | BG vs. Quasi-SOF      | SOF vs. Quasi-SOF     | FOF vs. Quasi-SOF     |
|-------------|------------|-----------------------|---------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 250–400 ms  |            |                       |                     |                       |                       |                       |                       |
| - anterior  | $p < .001$ | $F = 20.98, p < .001$ | $F = 0.18, p > .67$ | $F = 30.00, p < .001$ | $F = 13.33, p < .002$ | $F = 16.06, p < .001$ | $F = 0.04, p > .841$  |
| - posterior | $p > .29$  |                       |                     |                       |                       |                       |                       |
| 400–650 ms  |            |                       |                     |                       |                       |                       |                       |
| - anterior  | $p > .35$  |                       |                     |                       |                       |                       |                       |
| - posterior | $p < .001$ | $F = 60.66, p < .001$ | $F = 0.31, p > .58$ | $F = 37.42, p < .001$ | $F = 17.34, p < .001$ | $F = 11.31, p < .004$ | $F = 37.42, p < .001$ |
| 750–950 ms  |            |                       |                     |                       |                       |                       |                       |
| - anterior  | $p < .001$ | $F = 12.15, p < .003$ | $F = 1.58, p > .22$ | $F = 22.03, p < .001$ | $F = 3.76, p > .06$   | $F = 7.56, p < .02$   | $F = 5.36, p < .04$   |
| - posterior | $p < .01$  | $F = 2.16, p > .16$   | $F = 1.51, p > .23$ | $F = 6.47, p < .02$   | $F = 3.80, p > .06$   | $F = 0.63, p > .43$   | $F = 6.47, p < .02$   |

In the 750–950 ms time-window, there was a main effect of CONDITION [ $F(3,60) = 7.41, p < .001$ ] and a CONDITION × ROI interaction [ $F(3,60) = 8.77, p < .001$ ]. Resolving the interaction by ROI registered CONDITION effects over anterior [ $F(3,60) = 10.06, p < .001$ ] and posterior regions [ $F(3,60) = 4.25, p < .01$ ]. Planned comparisons in this window yielded an anterior difference between BG and FOF, and SOF and Quasi-SOF, a posterior difference between FOF and Quasi-SOF as well as differences in the anterior and posterior ROIs for FOF vs. SOF.

### 3 Discussion

The comparison of *Second Occurrence Focus* with *First Occurrence Focus* and *Background* elements proved to be a fruitful testbed for teasing apart aspects of the neurocognitive processing of information status on the one hand and (morpho-syntactic) focus on the other, as well as their prosodic marking. The ERP data indicate that the processing of prosody and information structure show distinct profiles on the surface of the scalp (see also Baumann & Schumacher 2012), which we are using in the following to discuss the contribution of prosodic and information structural cues separately. Figure 11 shows the topographical distribution in the three different time windows for the contrast between SOF and FOF (top panel) as well as between SOF and Quasi-SOF (bottom panel). The other contrasts can be accessed at [osf.io](http://osf.io). The figure indicates that the effects in the windows between 250–450 ms and 750–950 ms have anterior maxima, while the effect in the 400–650 ms window has a posterior maximum. Pronounced effects are observable for the first contrast between SOF and FOF and weaker effects in the topographical map for the SOF vs. Quasi-SOF comparison.

As to prosody, our data show that *pitch accents* are processed differently than *lack of accent*. That is, we did not find a difference between phrase accents (SOF) and complete deaccentuation (BG), which only differ in the duration of the target word. Neither of them displays a tonal movement, and the data suggest that this lack of movement triggers



**Figure 11:** Topographical maps for the three time windows, comparing SOF with FOF (upper map) and SOF with Quasi-SOF (bottom map). Anterior electrodes are at the top of each map. Note that the second condition is always subtracted from the first condition, i.e. colour coding (in the 400–650 ms window) does not necessarily correspond to polarity.

a negativity (in the region between 250 and 400 ms after target word onset), which we consider an instance of a N400,<sup>11</sup> plus a late positivity over anterior brain regions. Thus, our result does *not* confirm an intermediate status of phrase accents in terms of prosodic prominence, due to the very similar pattern with deaccentuation. Nevertheless, the hypothesis *is* confirmed to the extent that it shows a clear difference in the ERPs for pitch accents (FOF) on the one hand and phrase accents (SOF) and deaccentuation (BG) on the other. In fact, this result is in line with a recent behavioural prominence rating study on German mentioned above (Baumann & Winter 2018, see section 1.1 and Figure 4) in which phrase accents received only slightly higher prominence scores than deaccentuations. The prosodic variation between the two conditions in the present study (a difference in duration of 45 ms on average, see Table 1) was probably too subtle to be perceptible and thus too subtle to impact processing in most of the participants. Furthermore, previous research suggesting that prosodically prominent constituents following focus particles trigger a particular electrophysiological signal – the *expectancy negativity* (EN) – could not be confirmed.

<sup>11</sup> We consider both negativities (between 250–400 ms and 400–650 ms) members of the N400 family (Bornkessel-Schlesewsky & Schlewsky 2019). The differences in latency and topography may reflect domain-specific differences as well as distinct sources of the underlying operations (prosody vs. information structure).

As to information structure, which shows processing correlates over posterior scalp regions, BG and SOF pattern together as well, in contrast to FOF. There is a pronounced negativity (N400) for FOF (in the region between 400 and 650 ms after target word onset), which is missing in the other conditions. These results thus partly confirm our hypotheses, although there is no three-way distinction between the conditions. The fact that we failed to find a difference between BG and SOF items unlike Beaver et al. (2007) may have a methodological explanation: While Beaver et al. conducted an offline study with subjects paying attention to the contrast in question in a forced-choice task (see section 1.2 above), our ERP patterns reflect automatic effects of real-time processing.

Apart from such task-related differences, however, it could generally be claimed that the construct of *Second Occurrence Focus* may not be a unified phenomenon, which also has focus theoretic implications. We already stated that previous studies have shown that the difference between the prosodic marking of SOF and BG elements in both production and perception in German and English is only subtle – much more so than the difference between FOF and BG, which is generally marked by presence versus absence of a pitch accent. The present investigation finds no difference in neurocognitive processing between SOF and BG, a result which is incompatible with “association with focus” theories (e.g. Jackendoff 1972) mentioned above. These theories claim that a focus particle or “operator” (such as *only*) is associated with a focused constituent, and this association is indicated by some degree of prosodic prominence – irrespective of whether the constituent is given or new. Our results rather support a contextual account of focus, as proposed by Rooth (1992) and von Stechow (1994) and discussed by Krifka (2004), overcoming the need for an obligatory one-to-one relation between a semantic focus of a focus operator and its marking by prosodic prominence. In fact, such a pragmatic approach assumes that a focus operator is either context-sensitive or focus-sensitive introducing the relevant contextual features (Krifka 2004: 196). Importantly, the contextual approach allows for operators that are not associated with a focus in the first place, and consequently do not have to be marked by prosodic prominence. This approach thus seems to offer an explanation for the finding that the participants in our EEG study did not process SOF elements differently from BG elements: Both are given in the context (since they are direct repetitions of previous text) and as such not *interpreted* to be prosodically prominent, so that the subtle durational prominence of SOF items was not perceived.

Crucially, the information structural divide between the three conditions discussed so far lies in the factor *±newness*, i.e. only FOF represents new information while BG and SOF display given items. If *focus* were the crucial factor, we would have found a similar ERP pattern (increased N400) for FOF and SOF, with only BG showing an attenuated negativity. In earlier studies, new and focussed information has not been distinguished (e.g. Hruska & Alter 2004; Toepel et al. 2007). Our study is the first one that investigates independent effects of newness and focus by comparing FOF and SOF – and it clearly shows that it is *newness* that leads to increased processing effort (over posterior brain regions) and not focus – defined here as the item that is in the scope of a focus particle.

This result may come as a surprise, since focus marking is commonly associated with increased effort. However, the data may point to a distribution of effort between speaker and hearer. For focus marking, the effort is on the side of the *speaker*, which may rather ease the processing on the side of the *listener*. That is, the complex and less economical morpho-syntactic construction with the focus particle (more costly for the speaker) does not trigger an ERP effect, since the focus reading is readily available (less costly for the listener). Similarly, the production of a pitch accent (more costly for the speaker) makes it easy for a listener to detect a FOF, leading to reduced processing costs over anterior brain regions (less costly for the listener).

The crucial role of *newness* (or, more generally, “information status”) for speech processing also becomes obvious when comparing Quasi-SOF with the other conditions. Quasi-SOF triggers an ERP effect over posterior regions (negativity between 400 and 650 ms after target word onset) which is more pronounced than for BG and SOF but less pronounced than for FOF. The reason for this intermediate status may lie in the fact that the Quasi-SOF target words are *lexically new* (see section 1.2) but at the same time referentially given. BG and SOF items are both referentially and lexically given whereas FOF items are new on both levels. This result is in line with previous research on bridging inferences that reported a three-way amplitudinal modulation of the N400 with an increase from lexically given information (e.g. *beer – beer*) through indirectly given or bridged information (which is lexically new, e.g. *picnic – beer*) to fully new information (Burkhardt 2006). It is further supported by a finding from a study on the processing of set-superset relations that indicate a more enhanced N400 for reference via a superset term (e.g. *carp – fish*) compared to coreference via repetition (e.g. *carp – carp*) (see Schumacher & Weiland 2014). Finally, the results are not only backed by production studies on American English differentiating between lexical and referential givenness (see Lam & Watson 2014) but also by Almor’s (1999) *Information Load Hypothesis* claiming that the renaming of a given referent by new lexical material provides additional information which increases the cognitive load. In sum, the information structural part of Hypothesis 3 could be confirmed for Quasi-SOF but not for SOF, since we did not find the expected boosting effect due to focus marking in SOF. In other words, the “newness effect” for Quasi-SOF was stronger than the “focus effect” for SOF.<sup>12</sup>

We already discussed the ERPs for BG and SOF items in anterior brain regions (negativity between 250–400 ms plus late positivity), triggered by lack of accent. Since Quasi-SOF is marked by deaccentuation, we would expect the same brain potentials in this condition. Surprisingly, however, the ERP pattern rather indicates that Quasi-SOF differs from BG and SOF (resembling the pattern for a pitch accent, i.e. in the FOF condition). Thus, the part of Hypothesis 3 that deals with prosody is clearly disconfirmed. Three tentative lines of explanation are conceivable here. The first one is derived from the relation between prosody and information structure: The lexical newness effect may have led to an interpretation of Quasi-SOF items as being also *prosodically* more prominent than SOF and BG elements (and as more prominent than they actually are), as a reflex of their semantic-pragmatic importance. The second explanation is semantic in nature and suggests that the negative valence of the target words in the Quasi-SOF condition has a prominence effect in comparison with BG and SOF. Finally, a third explanation could lie in the segmental setup of the target words in Quasi-SOF: Most of them contain several voiceless obstruents (*Zeug* [tsɔɪk] ‘stuff’, *Quatsch* [kvatʃ] ‘nonsense’, *Schrott* [ʃrɔtʰ] ‘scrap’, *Schund* [ʃuntʰ] ‘trash’) which sound more prominent by virtue of their large amount of aperiodic energy in high frequency regions of the speech signal. That is, the articulatory and acoustic strength of these words may have led to a perception of prosodic prominence that is comparable to a pitch accent, despite of and to some extent compensating for the lack of both pitch movement and increased duration.<sup>13</sup> In fact, the large number of voiceless obstruents (including affricates) in these words might be regarded as somewhat onomatopoeic, adding to the negative connotation of the derogative common nouns which the Quasi-SOF target words are composed of.

<sup>12</sup> We refrain from interpreting the effects for the later time window over posterior electrode sites due to component overlap from the pronounced N400 between 400-650 ms of the FOF condition.

<sup>13</sup> Although this explanation resembles Kohler’s (2005) description of a *force accent*, we do not claim this category to be applicable to our data: While a force accent implies some extra effort in articulation, our Quasi-SOF stimuli are produced in an attenuated manner.

## 4 Conclusions

The present study is the first neurolinguistic investigation which directly links the processing of different levels of prosodic prominence to corresponding levels of information structure. More concretely, we tested the processing of items which are contextually licensed as *First Occurrence Focus*, *Second Occurrence Focus*, *Quasi Second Occurrence Focus* and *Background* as well as their appropriate prosodic realizations (primary (pitch) accent, secondary (phrase) accent and deaccentuation, respectively). Thus, no mismatches between information structural categories and their prosody were investigated but the specific contributions of different levels on both dimensions to the incremental processing of spoken language. In particular, the setup of our study makes it possible to tease apart the independent contributions of *focus* (defined here morpho-syntactically) on the one hand and *newness* on the other. This is done by investigating SOF items (which are both focused and given) in comparison with FOF (focused and new) and BG items (non-focused and given).

The main result with respect to **information structure** is that the increased processing effort in posterior brain regions that has been found in previous studies can probably be attributed to *new* rather than *focused* information. Evidence can be gained from the fact that SOF elements pattern together with BG rather than FOF elements, i.e. the divide is between *given* and *new* and not between *focused* and *non-focused* information. Actually, the additional analysis of Quasi-SOF items suggests that it is *lexical* newness which triggers the negative ERP, rather than newness at the *referential* level. As discussed above, our results further support a contextual account of focus (e.g. Rooth 1992; von Stechow 1994; Krifka 2004).

As to **prosody**, our results indicate an inverse relation between processing effort and the level of perceived prominence: We find a clear difference in anterior brain regions between the processing of *pitch accents* (to be found in FOF contexts), which are prosodically prominent due to tonal movement in the vicinity of a stressed syllable, and *no pitch accents* (comprising phrase accents and deaccentuation in SOF and BG contexts), which lack this tonal movement. Since *increased* processing effort is only found for *lack* of accents, we assume by implication that the production of a pitch accent, which is more costly for the *speaker*, reduces the processing costs on the side of the *listener*. An intermediate status of phrase accents in terms of processing effort and, in turn, prominence perception could not be confirmed. Interestingly, however, Quasi-SOF items, which were deaccented in the stimuli presented, might have been perceived as rather prominent, since their ERPs resemble FOF items. A tentative explanation can be derived from the derogative meaning of the target words and/or the segmental setup, which involves a relative increase of acoustic and articulatory strength.

## Abbreviations

BG = Background, EEG = electroencephalography, EN = expectancy negativity, EOG = electrooculogram, ERP = event-related brain potential, F0 = fundamental frequency, FOF = First Occurrence Focus, GToBI = German Tones and Break Indices, H = high tone, L = low tone, N400 = negative brain potential 400 ms after onset of critical entity, LP = late positivity, RMS = root mean square (intensity measure), ROI = region of interest, RPT = Rapid Prosody Transcription, SOF = Second Occurrence Focus

## Supplementary files

EEG raw data, audio stimuli and textgrids as well as topographical maps are available via [osf.io/6gzqc/](https://osf.io/6gzqc/).

## Acknowledgements

We would like to thank Janina Kalbertodt for her invaluable help with stimulus creation and preparation, Jane Mertens for help with stimulus preparation, Claudia Kilter for assistance with data collection and Simon Rössig for the spaghetti plot.

## Funding Information

This research has been funded by the German Research Foundation (DFG) as part of grant BA 4734/1-2 and the CRC 1252 “Prominence in Language” (project number 281511265) in the project A01 “Intonation and attention orienting: Neurophysiological and behavioral correlates” at the University of Cologne, Germany.

## Competing Interests

The authors have no competing interests to declare.

## References

- Almor, Amit. 1999. Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review* 106(4). 748–765. [PubMed: 10560327]. DOI: <https://doi.org/10.1037//0033-295X.106.4.748>
- Bartels, Christine. 2004. Acoustic correlates of ‘second occurrence’ focus: Towards an experimental investigation. In Hans Kamp & Barbara Partee (eds.), *Context-dependence in the analysis of linguistic meaning*, 345–361. Amsterdam: Elsevier.
- Baumann, Stefan. 2016. Second Occurrence Focus. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 483–502. Oxford: Oxford University Press.
- Baumann, Stefan & Arndt Riester. 2012. Referential and lexical givenness: Semantic, prosodic and cognitive aspects. In Gorka Elordieta & Pilar Prieto (eds.), *Prosody and meaning* (Interface Explorations 25), 119–162. Berlin, New York: Mouton De Gruyter.
- Baumann, Stefan & Arndt Riester. 2013. Coreference, lexical givenness and prosody in German. *Lingua* 136. 16–37. (Special Issue “Information Structure Triggers”, edited by Jutta Hartmann, Janina Radó & Susanne Winkler). DOI: <https://doi.org/10.1016/j.lingua.2013.07.012>
- Baumann, Stefan & Bodo Winter. 2018. What makes a word prominent? Predicting untrained German listeners’ perceptual judgments. *Journal of Phonetics* 70. 20–38. DOI: <https://doi.org/10.1016/j.wocn.2018.05.004>
- Baumann, Stefan & Christine T. Röhr. 2015. The perceptual prominence of pitch accent types in German. *Proceedings 18th ICPHS*. Paper number 298, 1–5. Glasgow: University of Glasgow.
- Baumann, Stefan, Doris Mücke & Johannes Becker. 2010. Expression of second occurrence focus in German. *Linguistische Berichte* 221. 61–78.
- Baumann, Stefan & Martine Grice. 2006. The intonation of accessibility. *Journal of Pragmatics* 38(10). 1636–1657. DOI: <https://doi.org/10.1016/j.pragma.2005.03.017>
- Baumann, Stefan & Petra B. Schumacher. 2012. (De-)accentuation and the processing of information status: evidence from event-related brain potentials. *Language and Speech* 55(3). 361–381. DOI: <https://doi.org/10.1177/0023830911422184>
- Beaver, David & Brady Clark. 2008. *Sense and sensitivity. How focus determines meaning*. Chichester: Wiley & Sons.
- Beaver, David, Brady Clark, Edward Flemming, T. Florian Jaeger & Maria Wolters. 2007. When semantics meets phonetics: Acoustical studies of Second Occurrence Focus. *Language* 83(2). 245–276. DOI: <https://doi.org/10.1353/lan.2007.0053>



- Beaver, David & Dan Velleman. 2011. The communicative significance of primary and secondary accents. *Lingua* 121. 1671–1692. DOI: <https://doi.org/10.1016/j.lingua.2011.04.004>
- Beck, Sigrid. 2016. Focus sensitive operators. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 227–250. Oxford: Oxford University Press.
- Beckman, Mary E. 1986. *Stress and non-stress accent*. Dordrecht: Foris. DOI: <https://doi.org/10.1515/9783110874020>
- Beckman, Mary E. & Jan Edwards. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In John Kingston and Mary E. Beckman (eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511627736.009>
- Beckman, Mary E. & Jan Edwards. 1994. Articulatory evidence for differentiating stress categories. In Pat Keating (ed.), *Phonological structure and phonetic form: Papers in laboratory phonology III*. Cambridge: Cambridge University Press.
- Birch, Stacy & Charles Clifton. 1995. Focus, accent, and argument structure: Effects on language comprehension. *Language and Speech* 38(4). 365–391. DOI: <https://doi.org/10.1177/002383099503800403>
- Birch, Stacy & Charles Clifton. 2002. Effects of varying focus and accenting of adjuncts on the comprehension of utterances. *Journal of Memory and Language* 47(4). 571–588. DOI: [https://doi.org/10.1016/S0749-596X\(02\)00018-9](https://doi.org/10.1016/S0749-596X(02)00018-9)
- Bishop, Jason. 2012. Information structural expectations in the perception of prosodic prominence. In Gorka Elordieta & Pilar Prieto (eds.), *Prosody and meaning* (Interface Explorations 25), 239–270. Berlin, New York: Mouton De Gruyter.
- Bishop, Jason. 2017. Focus projection and prenuclear accents: Evidence from lexical processing. *Language, Cognition and Neuroscience* 32(2). 236–253. DOI: <https://doi.org/10.1080/23273798.2016.1246745>
- Bock, J. Kathryn & Joanne R. Mazzella. 1983. Intonational marking of given and new information: Some consequences for comprehension. *Memory & Cognition* 11(1). 64–76. DOI: <https://doi.org/10.3758/BF03197663>
- Boersma, Paul & David Weenink. 2013. *Praat: doing phonetics by computer* (Computer program). Version 5.3.80, retrieved from <http://www.fon.hum.uva.nl/praat/>.
- Bornkessel, Ina, Matthias Schlesewsky & Angela D. Friederici. 2003. Contextual information modulates initial processes of syntactic integration: The role of inter-versus intrasentential predictions. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29(5). 871–882. DOI: <https://doi.org/10.1037/0278-7393.29.5.871>
- Bornkessel-Schlesewsky, Ina & Matthias Schlesewsky. 2019. Towards a neurobiologically plausible model of language-related, negative event-related potentials. *Frontiers in Psychology* 10. 298. DOI: <https://doi.org/10.3389/fpsyg.2019.00298>
- Bornkessel-Schlesewsky, Ina & Petra B. Schumacher. 2016. Towards a neurobiology of information structure. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 581–598. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780199642670.013.22>
- Brouwer, Harm & John C. J. Hoeks. 2013. A time and place for language comprehension: Mapping the N400 and the P600 to a minimal cortical network. *Frontiers in Human Neuroscience* 7. 758. DOI: <https://doi.org/10.3389/fnhum.2013.00758>
- Brown, Meredith, Anne Pier Salverda, Laura C. Dilley & Michael K. Tanenhaus. 2011. Expectations from preceding prosody influence segmentation in online sentence processing. *Psychonomic Bulletin and Review* 18. 1189–96. DOI: <https://doi.org/10.3758/s13423-011-0167-9>

- Büring, Daniel. 2007. Intonation, semantics and information structure. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 445–474. Oxford: Oxford University Press.
- Büring, Daniel. 2013/15. A theory of Second Occurrence Focus. *Language and Cognitive Processes/Language, Cognition and Neuroscience* 30(1–2). 73–87. DOI: <https://doi.org/10.1080/01690965.2013.835433>
- Burkhardt, Petra. 2006. Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain & Language* 98(2). 159–168. DOI: <https://doi.org/10.1016/j.bandl.2006.04.005>
- Buxó-Lugo, Andrés & Duane G. Watson. 2016. Evidence for the influence of syntax on prosodic parsing. *Journal of Memory and Language* 90. 1–13. DOI: <https://doi.org/10.1016/j.jml.2016.03.001>
- Cole, Jennifer & Stefanie Shattuck-Hufnagel. 2016. New methods for prosodic transcription: Capturing variability as a source of information. *Laboratory Phonology* 7. 1–29. DOI: <https://doi.org/10.5334/labphon.29>
- Cole, Jennifer, Yoonsook Mo & Mark Hasegawa-Johnson. 2010. Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology* 1. 425–452. DOI: <https://doi.org/10.1515/labphon.2010.022>
- Cole, Jennifer, Yoonsook Mo & Soondo Baek. 2010. The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes* 25(7). 1141–1177. DOI: <https://doi.org/10.1080/01690960903525507>
- Corbetta, Maurizio & Gordon L. Shulman. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience* 3(3). 201. DOI: <https://doi.org/10.1038/nrn755>
- Crystal, David. 1969. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- Cutler, Anne, Delphine Dahan & Wilma Van Donselaar. 1997. Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 40(2). 141–201. DOI: <https://doi.org/10.1177/002383099704000203>
- Dimitrova, Diana, Laurie A. Stowe, Gisela Redeker & John C. J. Hoeks. 2012. Less is not more: Neural responses to missing and superfluous accents in context. *Journal of Cognitive Neuroscience* 24(12). 2400–2418. DOI: [https://doi.org/10.1162/jocn\\_a\\_00302](https://doi.org/10.1162/jocn_a_00302)
- Féry, Caroline & Shinichiro Ishihara. 2009. The phonology of second occurrence focus. *Journal of Linguistics* 45. 285–313. DOI: <https://doi.org/10.1017/S0022226709005702>
- Friston, Karl. 2010. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience* 11(2). 127–138. DOI: <https://doi.org/10.1038/nrn2787>
- Fry, D. B. 1955. Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America* 27. 765–768. DOI: <https://doi.org/10.1121/1.1908022>
- Fry, D. B. 1958. Experiments in the perception of stress. *Language and Speech* 1. 126–152. DOI: <https://doi.org/10.1177/002383095800100207>
- Goldhahn, Dirk, Thomas Eckart & Uwe Quasthoff. 2012. Building large monolingual dictionaries at the Leipzig corpora collection: From 100 to 200 Languages. *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*. 759–765. Luxembourg: European Language Resources Association (ELRA).
- Grice, Martine, D. Robert Ladd & Amalia Arvaniti. 2000. On the place of phrase accents in intonational phonology. *Phonology* 17. 143–185. DOI: <https://doi.org/10.1017/S0952675700003924>

- Grice, Martine, Stefan Baumann & Ralf Benzmüller. 2005. German intonation in autosegmental-metric phonology. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, 55–83. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199249633.003.0003>
- Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511616983>
- Heim, Stefan & Kai Alter. 2006. Prosodic pitch accents in language comprehension and production: Erp data and acoustic analyses. *Acta Neurobiologiae Experimentalis* 66(1). 55–68.
- Heim, Stefan & Kai Alter. 2007. Focus on focus: The brain's electrophysiological response to focus particles and accents in German. In Andreas Späth (ed.), *Interfaces and interface conditions*, 277–298. Berlin: De Gruyter.
- Hickok, Gregory & David Poeppel. 2015. Neural basis of speech perception. In Michael J. Aminoff, François Boller & Dick F. Swaab (eds.), *Handbook of Clinical Neurology* 129, 149–160. London: Elsevier. DOI: <https://doi.org/10.1016/B978-0-444-62630-1.00008-1>
- Hruska, Claudia & Kai Alter. 2004. Prosody in dialogues and single sentences: How prosody can influence speech perception. In Anita Steube (ed.), *Information structure: Theoretical and empirical aspects*, 221–226. Berlin: De Gruyter.
- Hualde, José I., Jennifer Cole, Caroline L. Smith, Christopher Eager, Timothy Mahrt & Ricardo Napoleão de Souza. 2016. The perception of phrasal prominence in English, Spanish and French conversational speech. *Proceedings of Speech Prosody* 8. 459–463. [http://www.isca-speech.org/archive/SpeechProsody\\_2016/](http://www.isca-speech.org/archive/SpeechProsody_2016/). DOI: <https://doi.org/10.21437/SpeechProsody.2016-94>
- Huynh, Huynh & Leonard S. Feldt. 1970. Conditions under which mean square ratios repeated measurements designs have exact f distributions. *Journal of the American Statistical Association* 65(332). 1582–1589. DOI: <https://doi.org/10.1080/01621459.1970.10481187>
- Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Jyothi, Preethi, Jennifer Cole, Mark Hasegawa-Johnson & Vandana Puri. 2014. An investigation of prosody in Hindi narrative speech. *Proceedings of Speech Prosody* 7. 623–627. Dublin: Trinity College Dublin. DOI: <https://doi.org/10.21437/SpeechProsody.2014-112>
- Kaan, Edith & Tamara Y. Swaab. 2003. Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience* 15. 98–110. DOI: <https://doi.org/10.1162/089892903321107855>
- Keppel, Geoffrey. 1991. *Design and analysis: A researchers handbook*. Englewood Cliffs, NJ: Prentice Hall.
- Kochanski, Greg, Esther Grabe, John Coleman & Burton Rosner. 2005. Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America* 118. 1038–1054. DOI: <https://doi.org/10.1121/1.1923349>
- Kohler, Klaus. 2005. Form and function of non-pitch accents. *AIPUK* 35a. 97–123.
- Krifka, Manfred. 2004. Focus and/or context: A second look at second occurrence expressions. In Hans Kamp & Barbara Partee (eds.), *Context-dependence in the analysis of linguistic meaning*, 187–207. Amsterdam: Elsevier.
- Kutas, Marta & Kara D. Federmeier. 2011. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology* 62. 621–647. DOI: <https://doi.org/10.1146/annurev.psych.093008.131123>

- Ladd, D. Robert. 2008. *Intonational Phonology* (2nd Ed.). Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511808814>
- Lam, Tuan Q. & Duane G. Watson. 2014. Repetition reduction: Lexical repetition in the absence of referent repetition. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 40(3). 829–43. DOI: <https://doi.org/10.1037/a0035780>
- Lawrence, Michael A. 2016. ez: *Easy analysis and visualization of factorial experiments*. R package version 3.5.0. Available online at: <http://CRAN.R-project.org/package=ez>.
- Lieberman, Mark. 1975. *The intonational system of English*. New York: Garland.
- Luchkina, Tatiana & Jennifer Cole. 2014. Structural and prosodic correlates of prominence in free word order language discourse. *Proceedings of Speech Prosody* 7. 1119–1123. Dublin: Trinity College Dublin. DOI: <https://doi.org/10.21437/Speech-Prosody.2014-213>
- Maess, Burkhard, Erich Schröger & Andreas Widmann. 2016. High-pass filters and baseline correction in M/EEG analysis. Commentary on: “How inappropriate high-pass filters can produce artefacts and incorrect conclusions in ERP studies of language and cognition”. *Journal of Neuroscience Methods* 266. 164–165. DOI: <https://doi.org/10.1016/j.jneumeth.2015.12.003>
- Magne, Cyrille, Corine Astesano, Anne Lacheret-Dujour, Michel Morel, Kai Alter & Mireille Besson. 2005. On-line processing of “pop-out” words in spoken french dialogues. *Journal of Cognitive Neuroscience* 17(5). 740–756. DOI: <https://doi.org/10.1162/0898929053747667>
- Mahrt, Timoty, Jennifer Cole, Margaret Fleck & Mark Hasegawa-Johnson. 2012. Modeling speaker variation in cues to prominence using the Bayesian information criterion. *Proceedings of Speech Prosody* 6. 322–325. Shanghai: Tongji University Press.
- Makeig, Scott, Marissa Westerfield, Tzyy-Ping Jung, Sigurd Enghoff, Jeanne Townsend, Eric Courchesne & Terrence J. Sejnowski. 2002. Dynamic brain sources of visual evoked responses. *Science* 295(5555). 690–694. DOI: <https://doi.org/10.1126/science.1066168>
- Nooteboom, Sieb G. & Johanna G. Kruyt. 1987. Accents, focus distribution, and the perceived distribution of given and new information: An experiment. *The Journal of the Acoustical Society of America* 82(5). 1512–1524. DOI: <https://doi.org/10.1121/1.395195>
- Partee, Barbara. 1999. Focus, quantification, and semantics-pragmatics issues. In Peter Bosch & Rob van der Sandt (eds.), *Focus – linguistic, cognitive, and computational perspectives*, 213–232. Cambridge: Cambridge University Press.
- Ranganath, Charan & Gregor Rainer. 2003. Cognitive neuroscience: Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience* 4(3). 193. DOI: <https://doi.org/10.1038/nrn1052>
- R Core Team. 2018. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75–116. DOI: <https://doi.org/10.1007/BF02342617>
- Rooth, Mats. 1996. On the interface principles for intonational focus. In Teresa Galloway & Justin Spence (eds.), *Proceedings of the 6th Semantics and Linguistic Theory Conference*, 202–226. New Brunswick: Rutgers University. DOI: <https://doi.org/10.3765/salt.v6i0.2767>
- Schumacher, Petra B. & Hanna Weiland-Breckle. 2014. Referential properties of definite NPs and salience spreading. In Ana Aguilar-Guevara, Bert Le Bruyn & Joost Zwarts (eds.), *Weak referentiality*, 365–388. Amsterdam/Philadelphia: John Benjamins.

- Schumacher, Petra B. & Stefan Baumann. 2010. Pitch accent type affects the N400 during referential processing. *Neuroreport* 21(9). 618–622. DOI: <https://doi.org/10.1097/WNR.0b013e328339874a>
- Schumacher, Petra B. & Yu-Chen Hung. 2012. Positional influences on information packaging: Insights from topological fields in German. *Journal of Memory and Language* 67(2). 295–310. DOI: <https://doi.org/10.1016/j.jml.2012.05.006>
- Sedivy, Julie C., Michael K. Tanenhaus, Chraig G. Chambers & Gregory N. Carlson. 1999. Achieving incremental semantic interpretation through contextual representation. *Cognition* 71(2). 109–147. DOI: [https://doi.org/10.1016/S0010-0277\(99\)00025-6](https://doi.org/10.1016/S0010-0277(99)00025-6)
- Selkirk, Elisabeth. 2008. Contrastive focus, givenness and the unmarked status of “discourse-new”. In Caroline Féry & Gisbert Fanselow (eds.), *Acta Hungarica Linguistica* 55(3–4). 331–346. (Special Issue on Information Structure.) DOI: <https://doi.org/10.1556/ALing.55.2008.3-4.8>
- Sluiter, Agaath M. C. & Vincent J. van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America* 100. 2471–2485. DOI: <https://doi.org/10.1121/1.417955>
- Smith, Caroline. 2011. Perception of prominence and boundaries by naïve French listeners. *Proceedings 17th ICPHS*. 1874–1877. Hong Kong: City University of Hong Kong.
- Smith, Caroline. 2013. French listeners’ perceptions of prominence and phrasing are differentially affected by instruction set. *Proceedings of Meetings on Acoustics* 19(1). 60191. Melville, NY: Acoustical Society of America. DOI: <https://doi.org/10.1121/1.4799041>
- Snedeker, Jesse & John Trueswell. 2003. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language* 48. 103–130. DOI: [https://doi.org/10.1016/S0749-596X\(02\)00519-3](https://doi.org/10.1016/S0749-596X(02)00519-3)
- Terken, Jacques & Sieb G. Nooteboom. 1987. Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes* 2(3–4). 145–163. DOI: <https://doi.org/10.1080/01690968708406928>
- Toepel, Ulrike, Ann Pannekamp & Elke van der Meer. 2009. Fishing for information: The interpretation of focus in dialogs. In Kai Alter, Merle Horne, Magnus Lindgren, Mikael Roll & Janne von Koss Torkildsen (eds.), *Papers from braintalk. The 1st Birgit Rausing language program conference in linguistics*, 1–22. Lund: Lund University.
- Toepel, Ulrike, Ann Pannekamp & Kai Alter. 2007. Catching the news: Processing strategies in listening to dialogs as measured by erps. *Behavioral and Brain Functions* 3. ARTN 53. DOI: <https://doi.org/10.1186/1744-9081-3-53>
- van Berkum, Jos J. A., Arnout W. Koornneef, Marte Otten & Mante S. Nieuwland. 2007. Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research* 1146. 158–171. DOI: <https://doi.org/10.1016/j.brainres.2006.06.091>
- von Fintel, Kai. 1994. *Restrictions on Quantifier Domains*. Amherst, MA: University of Massachusetts at Amherst dissertation.
- Wang, Luming & Schumacher, Petra B. 2013. New is not always costly: Evidence from online processing of topic and contrast in Japanese. *Frontiers in Psychology* 4. 363. DOI: <https://doi.org/10.3389/fpsyg.2013.00363>
- Widmann, Andreas, Erich Schröger & Burkhard Maess. 2015. Digital filter design for electrophysiological data – a practical approach. *Journal of Neuroscience Methods* 250. 34–46. DOI: <https://doi.org/10.1016/j.jneumeth.2014.08.002>

**How to cite this article:** Baumann, Stefan and Petra B. Schumacher. 2020. The incremental processing of focus, givenness and prosodic prominence. *Glossa: a journal of general linguistics* 5(1): 6.1–30. DOI: <https://doi.org/10.5334/gjgl.914>

**Submitted:** 08 February 2019

**Accepted:** 15 October 2019

**Published:** 10 January 2020

**Copyright:** © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiq Press.

OPEN ACCESS 