Open Library of Humanities

Prosodic prominence does not speed up word recognition in a word monitoring task with German adult speakers

Barbara Zeyer, University of Cologne, DE, barbara.zeyer@uni-koeln.de

Martina Penke, University of Cologne, DE, martina.penke@uni-koeln.de

Prosodically prominent accent types have been proven to facilitate word recall as well as referent identification in eye-tracking studies. In the present study, we sought to investigate whether prosodic prominence also influences on-line sentence processing. We conducted a word monitoring task with 49 German speakers where either the target word or the word preceding the target word were manipulated with either high level prominent accents (*L+H**, *L*+H*), a low level prominent accent (*L**), or deaccentuation (*ø*). Based on previous findings, we expected prosodic prominence to draw attention, hence resulting in faster recognition times when the target word carried a prominent accent. At the same time, we postulated that the processing of prosodic prominence also binds processing resources, thus leading to slower recognition times when the word preceding the target word was manipulated. Specifically, we hypothesized a scale of prosodic prominence (*L+H**, *L*+H*, *L**, *ø*) leading to a stepwise increase or decrease of recognition times in word monitoring dependent on the position of the accent (on the target word or on the word preceding the target word, respectively). While previous studies have shown that prosodic prominence facilitates both word recall and referent identification, we did not find that prosodic prominence affected reaction times in the present study. Our results suggest that, when presented without further context, prosodic prominence might not exert an effect on on-line processing in a word monitoring task.

1 Introduction

In human interaction, prosody is an important tool to facilitate communication. For instance, prosody is used to mark information status (*new* vs. *given* e.g., Bolinger 1986; Baumann & Hadelich 2003; Baumann & Grice 2006; Féry & Kügler 2008; Röhr & Baumann 2010), to distinguish focus and background (e.g., Gussenhoven 1983; Mücke & Grice 2014), to indicate turn-taking (e.g., Schaffer 1983; Cutler & Pearson 1986) or to show speakers' attitudes or emotions (e.g., Bänziger & Scherer 2005; for a review see Mitchell & Ross 2013). In addition, prosody can also be used to draw a listener's attention to a particular entity within the utterance by pronouncing this entity more prominently compared to the entities in its environment, i.e., by making the entity *prosodically prominent*. Prosodically prominent entities bear a pitch accent, are produced louder and longer, or have a greater pitch excursion, i.e., the entities are hyperarticulated compared to deaccented entities (Cutler & Swinney 1987; de Jong 1995; Ladd 1996; Himmelmann & Primus 2015). Properties such as producing an entity with a greater pitch excursion, with a pitch accent, or longer and louder are defined as prominence-lending cues (Himmelmann & Primus 2015; Baumann & Winter 2018). Prosodic prominence can be defined as the "standing out" or the "highlighting" of a particular entity in relation to the neighbouring entities based on their prosodic characteristics (Steefkerk 2002; Himmelmann & Primus 2015; Cangemi & Baumann 2020). Consequently, Cangemi & Baumann (2020: 1) define prosodic prominence as a "relational property" which requires to "consider both the unit itself and the other neighbouring units within the same domain". Thus, prosodic prominence only establishes itself when the prosodically prominent entity is considered in relation to its environment. These highlighted units, as Cangemi & Baumann (2020) put it, can range from a segment or a syllable to a word (Steefkerk 2002).

In a pilot study, Baumann (2014) examined how untrained German-speaking listeners perceived prominence-lending cues. The results suggested that for German speakers tonal movement (specifically rising) and utterance-final nuclear accents are important indicators for prominence. To study the perception of prosodic prominence in German in more detail, Baumann & Röhr (2015) compared different accent types and deaccentuation with regard to their perceived prominence in phonetically untrained participants. Participants had to judge how prominent or highlighted they perceived the proper names *Lana, Lona* or *Lina* in isolated sentences such as *Sie hat mit der Lana/Lona/Lina telefoniert* ('She was on the phone with Lana/Lona/Lina', Baumann & Röhr 2015: 2). Baumann & Röhr (2015) found that high tones ($!H^*$, H^*) were judged to be more prominent than low tones (L^*) while deaccentuation ($\emptyset$) was judged as least prominent. They concluded that for German participants the most important factor to judge a word as highlighted, thus prosodically prominent, was pitch movement, with rises (L^*+H, $L+H^*$) being more prominent than falls ($H+!H^*$, $H+L^*$). Additionally, steep pitch movement ($L+H^*$, L^*+H, $H+!H^*$, $H+L^*$) was rated as more prominent than shallow pitch movement (H^*, $!H^*$, L^*) by

the participants. The results can be summed up in the following scale of perceived prosodic prominence in German: $\emptyset < L^* < H+L^* < H+!H^* < !H^* < H^* < L^*+H < L+H^*$ (Baumann & Röhr 2015: 4). While Baumann & Röhr (2015) showed how German speakers judge the prosodic prominence of pitch accents and deaccentuation in German, the study provides no indication whether and how the different accent types affect language processing, an open question which we sought to address in the current study.

1.1 Prosodic prominence and language processing

A number of studies have investigated the interplay of prosodic prominence on language processing, especially on word recall (Fraundorf et al. 2010; Savino et al. 2020; Kember et al. 2021; Koch & Spalek 2021). Fraundorf et al. (2010), for instance, compared the effects of the H^* accent to the $L+H^*$ accent on word recall in speakers of American English. Following Pierrehumbert & Hirschberg (1990), the accent type H^* is considered to introduce new information and to not show contrastive focus while the accent type $L+H^*$ is considered to not introduce new information and to show contrastive focus. The procedure was split into two parts: First, the participants had to listen to all auditorily presented experimental sentences (24 critical and 24 filler sentences), that were embedded into small discourses (one context sentence, one experimental sentence). Afterwards, participants were asked in a forced-choice recognition memory task whether a word appearing on the screen had been presented in the experimental sentences or not. The results showed that target words bearing the pitch accent $L+H^*$ were recalled better than target words accented with an H^* accent. The authors attributed the improved word recall of the accent type $L+H^*$ over H^* to the fact that the encoding of the word bearing an $L+H^*$ accent is facilitated in discourse.

A similar effect is reported by Kember et al. (2021). The authors examined speakers of Australian English or of Korean in a recall task where prominence was manipulated syntactically and prosodically. Prosodic prominence was manipulated in terms of a high pitch accent. The results showed that target words which were prosodically prominent were recalled better than target words that were syntactically prominent. Kember et al. (2021) argued that prosodic prominence facilitates lexical processing which leads to a better recall.

While the aforementioned studies measured the influence of prosodic prominence on off-line language processing, eye-tracking studies have investigated the influence of prosodic prominence on on-line language processing (see e.g., Weber et al. 2006; Chen et al. 2007; Ito & Speer 2008; Braun & Biezma 2019).[1] Ito & Speer (2008), for example, investigated anticipatory effects of

[1] Off-line and on-line effects are defined as follows: On-line tasks "measure processing as it happens", whereas off-line tasks "measure the consequences of processing, after some or all of the processing has taken place." (Warren 2012: 162).

intonation in an instructed visual search task conducted with English-speaking participants. Participants had to decorate a holiday tree with ornaments that were presented on a vision board. Different ornaments were hung up in different cells. Each cell consisted of four ornaments of the same type but differing in colour. The participants' task was to decorate a holiday tree by executing the instructions that were auditorily presented. While doing that, their eye movement was monitored with a head mounted eye-tracker. Ito & Speer (2008) manipulated the position of the contrastive $L+H^*$ accent that was either on the prenominal adjective or on the noun in instruction sentences like 'Hang the green drum. Now hang the BLUE drum./Now hang the blue DRUM.' (capitalization indicating the location of the accent; Ito & Speer 2008: 545). The authors found more early fixations and a steeper initial rise, i.e., more frequent fixations to the target cell, in the condition where the prenominal adjective was manipulated with an $L+H^*$ accent. The authors argued that the contrastive accent on the adjective facilitated fixation to the target cell because participants expected the same type of ornament than in the preceding sentence.

Similar findings were reported by Weber et al. (2006). The authors presented participants with four pictures (e.g., purple scissors (first referent), red scissors (target; contrastive referent), red vase (non-contrastive referent), and green clock (distractor)). Participants were instructed to click on the object that was specified in the sentence they heard. Participants first heard an instruction sentence such as 'Click on the purple scissors.', with *scissors* carrying an H^* accent. Then, participants heard the instruction sentence 'Click on the red scissors.' with either *red* or *scissors* carrying an $L+H^*$ accent. Weber et al. (2006) found that participants looked faster to the target referent when the preceding adjective was manipulated with an $L+H^*$ target compared to when the adjective was unaccented. The authors argued that their participants "exploited accents on preceding adjectives rapidly enough to anticipate target referents even before the referent noun was mentioned" (Weber et al. 2006: 386).

While these studies looked at prosodic prominence and its function in marking contrast, there is only sparse psycholinguistic evidence on the issue of how prosodic prominence influences on-line language processing devoid of an explicit contrasting context. Cutler & Swinney (1987) conducted a word monitoring task with English-speaking children and adults. Participants were asked to listen to pre-recorded sentences and to push a response key as soon as they recognized a target word, that was presented auditorily before the onset of the sentence. The primary sentence accent was either on the target word or appeared elsewhere in the sentences (e.g., *The family was already **at** the summer cabin*; bold indicating the accented target word; Cutler & Swinney 1987: 151). Cutler & Swinney (1987) found that their ten adult participants reacted faster when the target word carried the primary sentence accent compared to when the accent appeared elsewhere in the sentence. These findings suggest that prosodic prominence affects language processing by facilitating word identification, even if the accent is devoid of its function, i.e., does not explicitly

mark contrast.[2] The study replicated findings by Cutler & Foss (1977), who manipulated the stress of target phonemes in carrier words. Cutler & Foss (1977) asked participants to listen for previously defined target phonemes in carrier words in sentences and to push a button as soon as they heard the target phoneme. The authors found that participants recognized the target phonemes faster when they appeared in words which bore the main accent of the sentence compared to when they occurred in words that did not carry the main accent.

Further evidence that prosodic prominence might facilitate word recognition and processing comes from a study by Cole & Jakimik (1980). Cole & Jakimik (1980) examined whether stress influenced the detection of mispronunciations and asked participants to click on a button as soon as they heard a word that was mispronounced. In the mispronounced words, syllable-initial /p/ and /k/ were replaced by /b/ and /g/, respectively. The mispronounced words were either stressed, i.e., prosodically prominent, or unstressed, i.e., not prosodically prominent. Cole & Jakimik (1980) found that mispronunciations were detected significantly more often when the mispronounced word was stressed (82%) compared to when it was unstressed (47%). The authors argued that "acoustic features [were] more prominent in stressed than unstressed syllables" (Cole & Jakimik 1980: 968), making the voicing in stressed words more salient and resulting in a higher detection rate of mispronunciations in stressed compared to unstressed words.

To summarize, off-line studies have quite consistently shown facilitatory effects of prosodic prominence on word recall across studies and languages (Fraundorf et al. 2010; Kember et al. 2021; Koch & Spalek 2021; see also Savino et al. 2020 for a recall task of digits). Additionally, eye-tracking studies have revealed that prosodic prominence facilitates referent identification when the prominent entity indicates a contrast (Weber et al. 2006; Chen et al. 2007; Ito & Speer 2008; Braun & Biezma 2019). However, there is only little psycholinguistic research on the effect of prosodic prominence on on-line language processing when prosodic prominence is used devoid of context and function, for instance without signalling an explicitly mentioned contrast (Cutler & Foss 1977; Cutler & Swinney 1987; see also Cole & Jakimik 1980). Moreover, while previous research focused on the effects of one or two specific accent types, the issue whether the graded scale of prosodic prominence established by Baumann & Röhr (2015) leads to graded effects on on-line language processing has not been investigated in psycholinguistic research so far. The goal of the present study was to close these gaps by investigating the influence of prosodic prominence on language processing in German.

---

[2] Please see the section Discussion for whether processing of an accented entity may always raise an implicit contrasting context, even though this context is not explicitly mentioned.

1.2 Aims and predictions

The aim of our study was to follow-up on the research by Cutler & Swinney (1987) in investigating whether and how prosodic prominence influences on-line language processing in German when it is employed devoid of its function in a given context. In addition, we sought to examine whether the graded scale of perceptual prominence established by Baumann & Röhr (2015) would also lead to graded effects of prosodic prominence on on-line processing. To this end, we conducted a word monitoring task that was similar to that of Cutler & Swinney (1987), where participants were asked to push a response key as soon as they recognized the previously defined target word in a sentence that was auditorily presented to them. In contrast to Cutler & Swinney (1987), we did not only test the difference between accented and deaccented words but investigated the processing of three different accent types ($L+H^*$, L^*+H, L^*) and of deaccentuation (ø), each varying in perceptual prominence ratings, in prenuclear sentence position. By testing these four types, we aimed to cover the extreme ends of the prominence scale suggested by Baumann & Röhr (2015): while accents $L+H^*$ and L^*+H were rated as most prominent in the study by Baumann & Röhr (2015), the accent L^* and deaccentuation (ø) were rated as least prominent.

According to Posner (1980: 4) attention orienting is "the aligning of attention with a source of sensory input". Prosodic prominence is assumed to centre the attention of a listener to the prosodically prominent entity during language processing (e.g., Himmelmann & Primus 2015). The listener's attention is involuntarily drawn by an external stimulus, i.e., the prominence-lending prosodic cues, towards the prosodically prominent word. Thus, the orienting of attention is not under control of the listener but is an involuntary reaction to an external stimulus, the prominence-lending prosodic cues. Following James (1890: 403 f.), attention is defined as "taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought", which implies a "withdrawal from some things in order to deal effectively with others". This implies that while processing resources are drawn to the entity attention is oriented towards, information that is not in focus is neglected.

How might prosodic prominence influence word monitoring? While numerous psycholinguistic models have been put forward to explain auditory word recognition (e.g., Marslen-Wilson & Welsh 1978; McClelland & Elman 1986; Norris 1994; Cutler & Clifton 2001), most of them commonly assume the following three stages: *initial contact*, *selection* and *integration* (see e.g., Dahan & Magnuson 2006). During the initial contact stage, words that share the same onset get activated. As the auditory input unfolds, more and more activated words in this initial cohort get excluded until one candidate is selected. During the integration process, the selected lexical item is then integrated into the discourse in consideration of the semantic and syntactic structure of the input. We suggest that prosodic prominence might facilitate the selection process, resulting in a faster identification of the prosodically prominent word. Prosodically prominent words

are hyperarticulated (e.g., de Jong 1995), which might make the distinction between similarly sounding words easier. This might be due, for instance, to a smaller initial cohort for the selection process or to a quicker weeding-out of cohort candidates. As a consequence, prosodically prominent words might be more accessible for the listener and get accessed faster. This might also hold when a prosodically prominent accent type is used devoid of its pragmatic function, for example when a contrast inducing $L+H^*$ accent is used without explicitly signalling contrast. Given the assumption that prosodically prominent entities draw attention, a prosodically prominent target word should be identified quicker in a word monitoring task compared to a target word that is presented without a prosodically prominent accent. Moreover, given the graded scale of prosodic prominence proposed by Baumann & Röhr (2015), we might expect a graded effect of prosodic prominence on word monitoring times. Specifically, prosodically prominent accent types, here the accent types $L+H^*$ and L^*+H, should be more effective in drawing attention to the target word, compared to target words presented with the less prominent accent type L^* or for target words that are deaccented (ø). Hence, we expected to find the following scale for the identification times of the target word in our study: $ø > L^* > L^*+H > L+H^*$. These expectations led to the following research question:

> RQ 1: Do differently prominent accent types (i.e., $L+H^*$, L^*+H, L^*) and deaccentuation (ø) on a target word influence identification times in a word monitoring task differently?

We further assumed that the centring of attention to the prosodically prominent entity and the processing of prosodic prominence binds processing resources. In turn, these processing resources are no longer available for parsing the elements immediately following the prominent entity. Accent types that are high in prosodic prominence should bind more processing resources than accent types that are prosodically less prominent. In a word monitoring task, this would lead to the assumption that target words following a word with a prosodically prominent ($L+H^*$, L^*+H) accent type should be identified slower compared to target words that follow a word with a less prosodically prominent accent type (L^*) or that follow a deaccented word (ø). This leads to the second research question that we aimed to answer:

> RQ 2: Do differently prominent accent types (i.e., $L+H^*$, L^*+H, L^*) and deaccentuation (ø) on a word directly preceding the target word influence identification times in a word monitoring task?

In order to answer this research question, the word directly preceding the word identified as target word in the word monitoring task was either manipulated with the accent types $L+H^*$, L^*+H, L^* or was deaccented (ø). Following our line of argumentation, we assumed that the previously proposed scale of word identification times should be reversed: $ø < L^* < L+H^* < L^*+H$.

Given the predictions sketched out for our first and second research question, we also postulated that a comparison of those experimental conditions in which the target word or the word directly preceding the target word are manipulated with the same highly prominent prosodic accent types ($L+H^*$ or L^*+H) should lead to different recognition times of the target word. Assuming that prosodic prominence draws attention, a prominent accent on the target word should lead to a facilitation of the recognition of the target word. At the same time, we expect that the recognition of the target word should be hindered when the word before the target word is manipulated with that exact accent, as the processing of prosodic prominence is assumed to bind processing resources. We expect this difference in word identification times to be strongest for the two highly prominent accent types ($L+H^*$ and L^*+H), leading to an interaction between accent type and position of the accent type in the sentence, i.e., on the target word or on the preceding word. Our third research question, therefore, was:

RQ 3: Do prosodically prominent accent types (i.e., $L+H^*$ and L^*+H) influence the identification times in a word monitoring task differently when they are on the target word compared to when they are on the word directly preceding the target word?

Summarizing, to investigate the influence of different levels of prosodic prominence on identification times in a word monitoring task, either the target word or the word directly preceding the target word were presented in four different accent conditions ($L+H^*$-, L^*+H- or L^*-accent or deaccentuation, $\emptyset$).

2 Method

2.1 Participants

Participants were recruited at the University of Cologne. In total, 58 people participated in the experiment. 32 participants took part in the experiment as part of their course curriculum and were rewarded course credits. 26 more people were recruited from the university and were paid for their participation. Eight participants had to be excluded because they did not meet the inclusion criterion of having grown up with German as their only language up to the age of six years. One more participant was excluded because s/he stated to have reacted only after the sentences were finished and not during the presentation of the experimental sentences. This left 49 participants (female: 34, male: 15) for the statistical analysis. Participants' mean age was 23 years (range: 18–31 years). Two participants were left-handed. All participants reported to have normal or corrected-to-normal vision as well as normal hearing and to be neuro-typical. They furthermore reported to have grown up with German as their only language between the ages of zero and six. All participants gave written informed consent before starting the study.

2.2 Procedure

The study took place in a laboratory at the University of Cologne. The experiment was designed using OpenSesame (version 3.3.10), an open-source programme suitable for linguistic reaction time experiments (Mathôt et al. 2012). Particpants were tested individually in a quiet room and were asked to wear headphones (Sennheiser HD206) during the experiment. The loudness of the auditory stimuli was held equal for all participants. The same laptop (Dell, Latitude 3420) was used for all trials. A button response box by SR Research was used as a response device (https://www.sr-research.com).

Before the participants started the experiment, they were given information on the procedure orally and in written form, were asked to fill out a questionnaire about their personal and language background, and had to give consent for their participation. The experiment started with a practice phase of ten sentences. After the practice phase, participants had the possibility to ask questions before the actual experiment started. After each block of 28 sentences, participants could take a break. Each experimental block lasted for approximately four minutes.

Trial structure was as follows: A fixation cross appeared on the computer screen for 700 ms, followed by the target word written in caps that appeared for 1500 ms. After that, the sentence was presented auditorily while the screen was blank. Participants were asked to press a response key on the button box as soon as they identified the target word in the sentence. Timeout was set at 4000 ms after the onset of the sentence. Reaction time measurements started with the onset of the auditorily presented target word and ended either after 4000 ms or when the participant pushed the response key (see **Figure 1**).
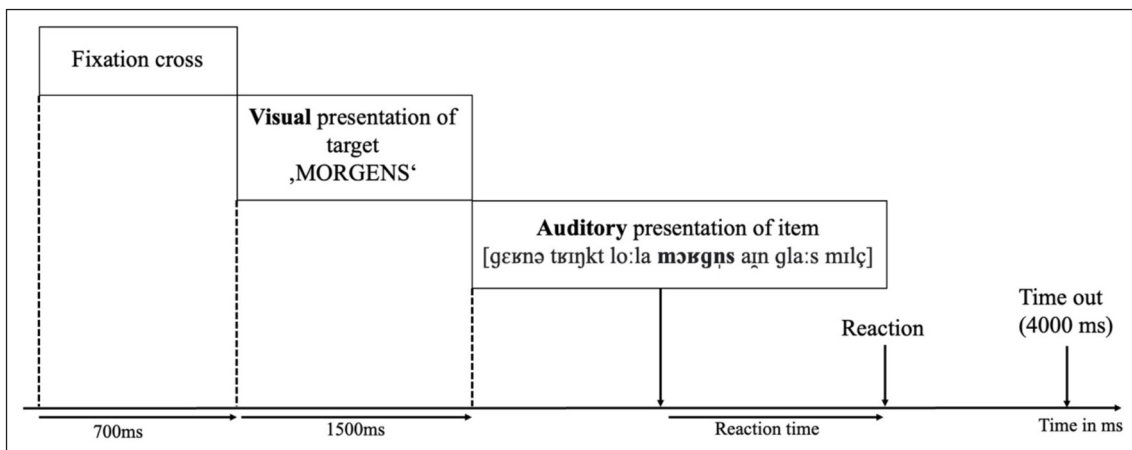

Figure 1: Experimental trial. Please note that the onset of the target word differs from sentence to sentence. This figure is supposed to exemplify the trial structure in general.

Approval for this experiment was given by the Faculty of Human Sciences of the University of Cologne (MPHF0064). The study was preregistered before data collection proceeded at OSF (https://osf.io/fea4p).

2.3 Stimuli

We tested seven different experimental conditions. In three conditions (1a-c) the target word was presented either with an $L+H^*$, $L+{}^*H$, or L^* accent, in three conditions (2a-c) the word directly preceding the target word carried either an $L+H^*$, $L+{}^*H$, or L^* accent. In the last condition, that served as baseline in this experiment, both the target word and the word preceding the target word were deaccented (see **Figure 2** for on overview of the stylized intonation contours of the experimental sentences). In total, we constructed 70 experimental sentences: ten sentences for each of the seven different experimental conditions tested. Each sentence was only presented in one of the experimental conditions, thus participants heard each sentence only once.
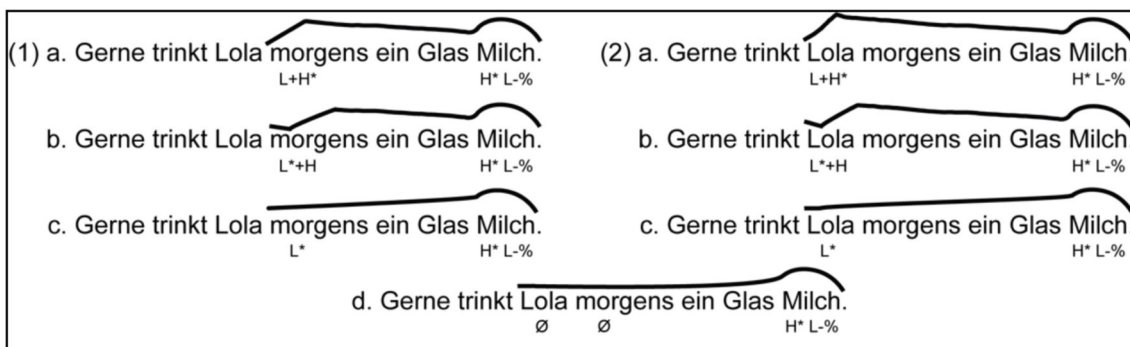

Figure 2: Stylized intonation contours of the experimental conditions exemplified here by the same experimental sentence only for ease of illustration. The example sentence translates as 'Lola likes to drink a glass of milk in the morning' In conditions (1a-c) the target word *morgens* 'in the morning' is prosodically manipulated. In conditions (2a-c) the word preceding the target word *morgens* (here *Lola*) is prosodically manipulated. In condition (d) both, the target word *morgens* and the word preceding the target word (here *Lola*) are deaccented. The experimental material including the prosodic contours can be found under https://osf.io/rsnz8.

Whereas **Figure 2** depicts the stylized intonation contours of the experimental conditions, **Figure 3** (a, b) shows the mean time-normalized f0 contours across experimental conditions. Panel 3a shows mean f0 contours when the target word was manipulated, while panel 3b shows the mean f0 contours when the word preceding the target word was manipulated. First, we extracted 20 data points per area of interest (i.e., the target word or the word preceding the target word, respectively) using PRAAT (Boersma & Weenink 2020). We then plotted the mean f0-contours in R (R Core Team 2021) for each condition using the extracted data points.
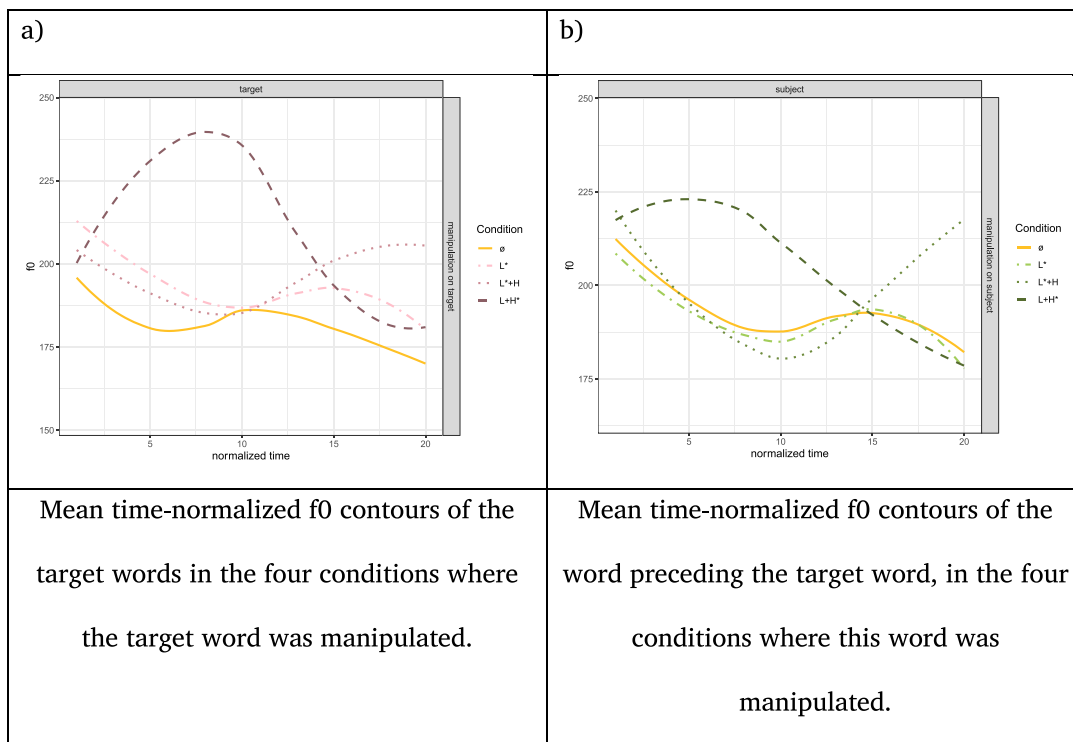
a)	b)
Mean time-normalized f0 contours of the target words in the four conditions where the target word was manipulated.	Mean time-normalized f0 contours of the word preceding the target word, in the four conditions where this word was manipulated.

Figure 3: Mean time-normalized f0 contours of the different experimental conditions.

All experimental sentences had the same syntactic structure (Adverb – Verb – Subject – target word – Object, see example (1), bold indicating the target word). The target word occurred as the fourth word of the sentence, directly after the sentence subject. This way, neither the syntactic structure of the sentence nor the position of the target word within the sentence would influence the identification of the target word. The nucleus was sentence final.

(1)   *Gerne   trinkt   Lola   **morgens**   ein Glas Milch.*
      Adv   V   S   target word O
      'Happily drinks Lola in the morning a glass of milk'

All target words were bi-syllabic, trochaic adverbials that were either adverbials of time, of reason, of modality or of space. The adverbs of the different types were equally distributed over the seven experimental accent conditions so that each experimental condition contained one modal, two causal, two local and five temporal adverbials. The frequency of occurrence of the target words was controlled for by the database WebCelex (Baayen et al. 1995). We used an R script (R Core Team 2021) to distribute the target words across the seven experimental conditions such that the mean target word frequency differed as little as possible from the mean target word frequency across experimental groups. The mean word frequency of the target words

across the seven experimental groups was 144.43 (SD = 316.29). The mean word frequency of the target words did not differ significantly between the seven different experimental conditions (F(6, 63) = .092, p = .997). We also controlled for the mean number of phonemes the target word consisted of, which was 5.9 (SD = 1.42). There was no significant difference in the mean number of phonemes of the target words across the seven experimental conditions (F(6,63) = .696, p = .654). In addition to that, we controlled for the mean duration in ms of the recorded target words, which was 461.89 ms (SD = 90.43) across experimental conditions. There was no significant difference in target word duration in the experimental conditions where the target word was accented or deaccented (conditions 1a-d in **Figure 2**, (F(3,63) = 1236, p = .311). Likewise there was no significant difference of the target word in the experimental conditions where the subject preceding the target was accented or deaccented (conditions 2a-d in **Figure 2**, F(3,63) = 2227, p = .102).

Additionally, students (n = 23) rated our written sentences with regard to how natural the sentences seemed to them on a scale from 1 (very good/natural) to 5 (very bad/unnatural). The mean rating across experimental conditions was 2.21 (SD = 1.35). The ratings indicated no significant difference between the seven experimental conditions (F(6, 1603) = 1046, p = .393). See **Table 1** for a summary of all characteristics of the target words across the seven experimental conditions.

Conditions/ Category	Word frequency of target word	Duration in ms of target word	Number of phonemes of target word	Rating of sentence
ø on target word and preceding word	200.2 (*SD* 544.25)	458.5 (*SD* 92.16)	6.4 (*SD* 1.71)	2.34 (*SD* 1.41)
$L+H^*$ on target word	164.0 (*SD* 389.78)	508.80 (*SD* 107.73)	5.8 (*SD* 1.55)	2.25 (*SD* 1.38)
L^*+H on target word	123.8 (*SD* 131.66)	524.70 (*SD* 61.8)	5.6 (*SD* 1.27)	2.17 (*SD* 1.35)
L^* on target word	108.3 (*SD* 98.54)	518.0 (*SD* 72.78)	5.5 (*SD* 1.18)	2.08 (*SD* 1.29)
$L+H^*$ on preceding word	142.1 (*SD* 358.85)	395.5 (*SD* 76.48)	6.4 (*SD* 1.90)	2.12 (*SD* 1.30)
L^*+H on preceding word	154.20 (*SD* 220.12)	389.2 (*SD* 56.28)	5.6 (*SD* 1.27)	2.27 (*SD* 1.27)
L^* on preceding word	118.40 (*SD* 217.70)	438.6 (*SD* 52.32)	6.0 (*SD* .94)	2.24 (*SD* 1.44)

Table 1: Characteristics of the target words across experimental conditions.

In addition to the 70 different experimental sentences, we created 70 filler sentences to ensure that the participants would not get accustomed to a particular sentence prosody or syntactic structure and would expect target words in a particular position in the sentence. The target words in the filler sentences were not manipulated with a particular accent and the sentences were recorded informally as if uttered in spontaneous speech. In the filler sentences, the target words were either the subject or the object of the sentence and occurred either in the second, third, fourth or seventh position of the sentence. The syntactic structure of the filler sentences was either Subject – Verb – (Prepositional Phrase) – Object or Prepositional Phrase – Verb – Subject – (Prepositional Phrase) – Object (see example (2)).

In 37 of the filler sentences the onset of the target word already occurred in a different word appearing before the target word ('onset fillers', see (3)). By including these filler sentences, we sought to ensure that the participants listened to the whole word before reacting. We furthermore created 16 'catch trials' in which the previously defined target word did not appear and where the correct reaction was to not push the response key (see (4)). This way, we wanted to ensure that the participants listened for the target word and did not react arbitrarily at some point in the sentence.

(2)   Filler
       Structure: S (= target word) V O
       Target word: *Tiger* ('tiger')
       Sentence: *Ein schneller **Tiger** jagt die langsame Beute.* ('A fast **tiger** is chasing the slow prey.')

(3)   Onset filler
       Structure: PP V S O (= target word)
       Target word: *Krone* ('crown')
       Sentence: *Auf das* Krokodil *malt Tamara die **Krone**.* ('It is on the *crocodile* that Tamara is drawing the **crown**.')

(4)   Catch trial
       Structure: S V O
       Target word: *Dackel* ('dachshund')
       Sentence: *Der starke Husky zieht den Schlitten.* ('The strong husky is pulling the sled.')

A female trained phonetician, who was a speaker of Standard German, recorded all sentences using a C520 headset and the Scarlett 2i2 3rd generation 2-in, 2-out USB audio interface in Audacity (Audacity Team 2021). The sentences were cut using PRAAT (Boersma & Weenink 2020) and adjusted for loudness across all experimental and filler sentences using Audacity (Audacity Team 2021). We normalized the data to –23.0 LUFS.

The 140 sentences were distributed across five blocks with 14 experimental and 14 filler sentences each. The order of the sentences in each block was pseudo-randomized so that no more than two experimental sentences occurred in a row. Furthermore, sentences of the same experimental accent condition did not follow each other. All participants were presented with the same previously fixed list of sentences.

2.4 Analysis

We log-transformed the reaction times prior to statistical analysis in order to normalize the data. Prior to analysis, we removed incorrect answers, where participants did not push the key although it was required, and negative reaction times, i.e., reactions where the participants pushed the button before the word appeared. We furthermore removed outliers. We identified outliers per participant, i.e., we determined the mean reaction time across all seven experimental conditions per participant. Reaction times that deviated more than two standard deviations of the mean reaction time of the participant were removed as outliers.[3] In the conditions where the target word was prosodically manipulated, 66 reaction times had to be removed (3.36%), which left us with 1894 data points. In the conditions where the word preceding the target word was prosodically manipulated 31 reactions were removed (1.02%), leaving 1909 data points for the statistical analysis. For the interaction of accent type and position, 27 reactions were removed (0.99%), leaving 1909 data points for the statistical analysis.[4]

We conducted linear mixed-effects models using the *lme4* package (Bates et al. 2015) in R (R Core Team 2021). For the first research question the independent variable was the accent type on the target word and the dependent variable was the time required for identifying the target word. As the maximally specified model did not converge, we successively simplified the structure of the model by reducing the correlation between random intercept and random slopes for participants as well as the random slopes for items until the model converged. The final model included random intercepts for participants and for items. The same procedure was adopted for the second research question and the final model also included random intercepts for participants and for items.

---

[3] To address the concerns of a reviewer, we also conducted the analysis without outlier removal. The results persisted and we did not find an effect of prosodic prominence on word identification times in any of the analyses we conducted.

[4] Please note that the conditions ø on target word and on the word preceding the target word, L^* on target word and L^* on the word preceding the target word are not included in this analysis. This is why the number of excluded trials in this sub-dataset does not equal the sum of the excluded trials of the conditions 'accent on target word' and 'accent on the word preceding the target word'.

We then conducted planned pairwise comparisons (Tukey corrected) using the *emmeans* package (Lenth 2022). We compared each of the accent types $L+H^*$, L^*+H and L^* to the baseline condition (ø) as well as the accent types with each other to assess whether the data provided evidence for the expected scale of graded prominence.

For the interaction of accent type ($L+H^*$ and L^*+H) and position of accent type (on the target words vs. on the words preceding the target words), i.e., for the third research question, the independent variables were accent type and the position of the accent type, the dependent variable was the time required for identifying the target word. As the maximally specified model for the interaction did not converge, we successively simplified the structure of the model by removing the correlation between random intercept and random slopes for the participants, the random slopes for interaction between position and type for participant, as well as the random slopes for participants until the model converged. The final model included random intercepts for items as well as random intercept for participants.[5] We then conducted planned pairwise comparisons (Bonferroni corrected) using the *emmeans* package (Lenth 2022). We compared the accent types $L+H^*$ and L^*+H in the positions *on target word* vs. *on the word preceding the target word* to assess whether the data show an interaction of accent type and its position in the sentence, i.e., whether the participants reacted faster when a prosodically prominent accent type was on the target word compared to when it was on the word directly preceding the target word.

2.5 Post-hoc study

To confirm that different accent types in the experimental stimuli were indeed perceived as prosodically prominent or not prominent, we ran an additional post-hoc study, where we implemented the design of Baumann & Röhr (2015). Participants were asked to judge how highlighted either the target word or the word preceding the target word sounded in an auditorily presented sentence on a five-point Likert scale (1 not highlighted at all, 5 very highlighted). The task (e.g., *Wie hervorgehoben klingt das Wort ‚morgens' in der folgenden Äußerung?* 'How highlighted does the word 'morgens' sound in the following utterance?') and a five-point Likert scale were presented on a computer screen while participants listened to the pre-recorded sentences that had also been used in the word monitoring task. After each auditorily presented sentence, participants were asked to respond by clicking on a value on the Likert scale. In total, participants had to listen and to respond to 30 sentences, ten sentences per condition: $L+H^*$ on the target word, $L+H^*$ on the word preceding the target word, and ø on both the target word

---

[5] To address model convergency issues, we additionally conducted Bayesian analysis. The Bayesian models confirmed our findings. The results and the models with all specifications can be found on OSF (https://osf.io/rsnz8).

and the word preceding the target word. Note that participants could listen to each presented sentence as often as they felt necessary before indicating their judgement.

In total, 36 students of the University of Cologne participated in the study as part of their course curriculum and were rewarded course credits. Four participants had to be excluded because they did not finish the task. Five more had to be excluded because they acquired German after the age of 6 years. This left 27 (23 female, 3 male, 1 diverse) participants. All of the participants grew up with German as their first language. Their mean age was 23 years (range: 19–33).

We compared the conditions $L+H*$ vs. ø on the target word, and the conditions $L+H*$ vs. ø on the word preceding the target word with a mixed effects ordinal model using the package *ordinal* (Christensen 2023) in R (R Core Team 2021). The final models included experimental group as predictor and random intercepts for items as well as for participants.

3 Results

3.1 Accent manipulation on the target word (RQ 1)

Figure 4 shows the distribution of the reaction times including the median for the conditions $L+H*$, $L*+H$, $L*$ and ø on the target word. For the condition $L+H*$ on the target word, the mean reaction time was 553 ms ($SD = 145$). In the condition $L*+H$ on the target word, the participants had a mean reaction time of 544 ms ($SD = 159$). In the condition $L*$ on the target word, the mean reaction time was 563 ms ($SD = 161$) and in the condition ø on the target word the mean reaction time averaged at 545 ms ($SD = 139$). We predicted that the mean reaction times for the accent types $L+H*$, $L*+H$ and $L*$ on the target word would be faster than for ø on the target word respectively. Furthermore, we expected that the accent type $L+H*$ would lead to the fastest word identification times with an increase of the mean reaction time from $L*+H$ and $L*$ to deaccentuation (ø) and with deaccentuation leading to the slowest word identification times. However, these predictions were not born out. In contrast to our expectations, the accent type $L*+H$ led to the fastest reaction times (544 ms), whereas the accent type $L*$ led to the longest reaction times (563 ms). Also, mean reaction times in the conditions $L+H*$ on the target word and $L*$ on target were slower than the condition ø on the target word. Only the condition $L*+H$ showed a numerically marginally faster reaction time than the baseline. For a plot of the mean reaction times, see **Figure 4**.

We ran a linear mixed-effects model with random intercepts for participant and random intercepts for items. The output of this model is given in **Table 2** and shows non-significant effects for all experimental conditions. Post hoc comparisons (Tukey corrected) showed no significant differences between experimental conditions (see **Table 3**). Thus, we found neither a graded effect of prosodic prominence nor an effect of prosodic prominence *per se.*
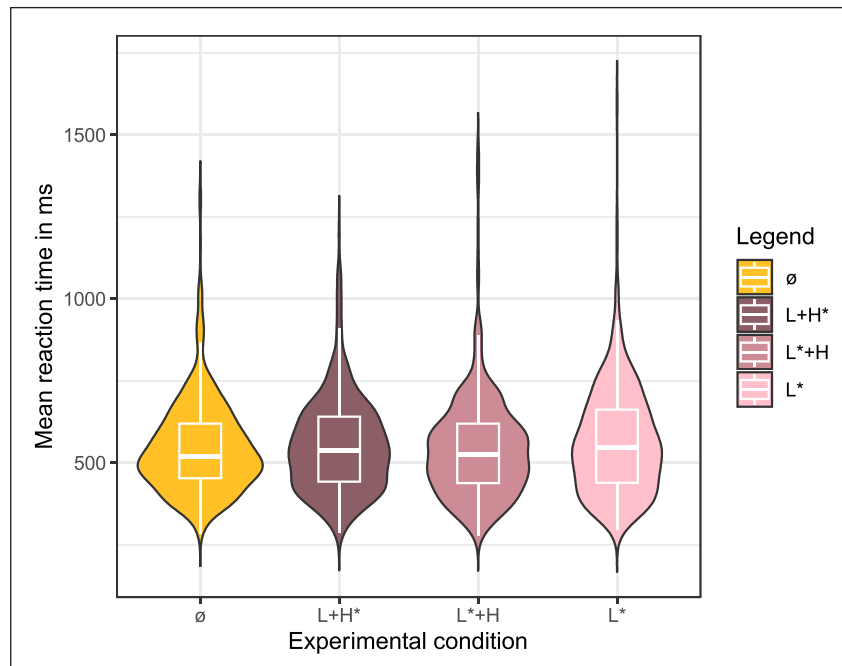
Figure 4: Violin plot of the mean reaction time in ms per experimental condition. ø serves as the baseline condition and is the condition where both the target word and the word preceding the target word are deaccented.

Accent type	Est.	SE	t	p
Intercept (ø)	6.271	.043	144.746	
L+H*	.013	.049	.278	.783
L*+H	−.008	.049	−.165	.870
L*	.036	.049	.730	.470

Table 2: Fixed effects from the linear mixed effects model in the conditions where the accent is on the target word. The table shows log-transformed reaction times.

	Comparison	Est.	SE	t.ratio	p.value
a)	ø vs. L+H*	−.01350	.0486	−.278	.9400
	ø vs. L*+H	.00804	.0486	.165	.8194
	ø vs. L*	−.035511	.0487	−.730	.9870
b)	L+H* vs. L*+H	.02154	.0486	.443	.9706
	L+H* vs. L*	−.02202	.0487	−.452	.9687
	L*+H vs. L*	−.04355	.0487	−.895	.8075

Table 3: Pairwise comparisons between a) the baseline condition (ø) vs. the three different accent types, and b) the three different accent types. Note that a '–' means that the condition mentioned second in the first column is slower than the condition mentioned first in the first column.

3.2 Accent manipulation on the word preceding the target word (RQ 2)

Figure 5 shows the distribution of the reaction times including the median for the four conditions $L+H^*$, L^*+H, L^*, or deaccentuation on the word preceding the target word (i.e., the sentential subject). In the condition $L+H^*$ on the word preceding the target word, the mean reaction time was 522 ms (SD = 143). In the condition L^*+H on the word preceding the target word the mean reaction time averaged at 515 ms (SD = 141). Participants had a mean reaction time of 549 ms (SD = 167) in the condition L^* on the word preceding the target word and a mean reaction time of 545 ms (SD = 139) in the condition ø on the word preceding the target word. We predicted that the mean target identification times for the accent types $L+H^*$, L^*+H and L^* on the word preceding the target word would be slower than for a deaccented word preceding the target word respectively, since a prominent accent on the word preceding the target word should bind processing resources. These processing resources would then be unavailable for identifying the target word. Furthermore, we anticipated that reaction times would increase as follows: ø < L^* < L^*+H < $L+H^*$, i.e., we expected that deaccentuation, ø, would lead to the fastest word identification times with an increase of the mean reaction time from L^* and L^*+H to $L+H^*$, with the accent type $L+H^*$ leading to the slowest word identification times. However, these predictions were not born out. In contrast to our expectations, the accent type L^*+H led to the fastest reaction times (515 ms), whereas the accent type L^* led to the longest reaction times (549 ms). Also, mean word identification times in the condition with the prominent $L+H^*$ *or* L^*+H accents on the word preceding the target word were faster than mean word identification times in the baseline condition (ø). For a plot of the mean reaction times, see **Figure 5**.
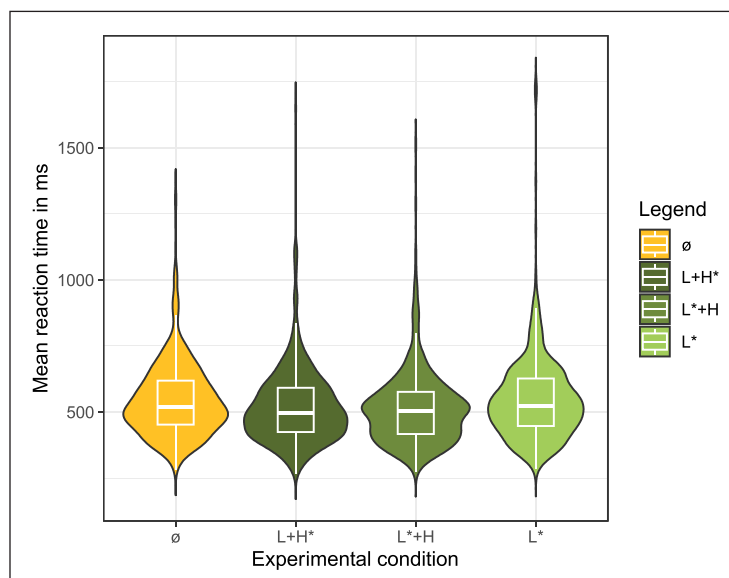

Figure 5: Violin plot of the mean reaction time in ms per experimental condition where the word preceding the target word was accentually manipulated. ø serves as the baseline condition and is the condition where both the target word and the word preceding the target word are deaccented.

We ran a linear mixed-effects model with random intercepts for participant and random intercepts for items. The output of this model is presented in **Table 4** and shows non-significant effects for all experimental conditions. Post hoc comparisons (Tukey corrected) showed no significant differences between experimental conditions (see **Table 5**). Thus, we found neither a graded effect of prosodic prominence nor an effect of prosodic prominence *per se*.

Accent type	Est.	SE	t	p
Intercept (ø)	6.271	.037	168.944	
$L+H^*$	−.044	.039	−1.137	.263
L^*+H	−.058	.039	−1.490	.145
L^*	.001	.039	.005	.996

Table 4: Fixed effects from the linear mixed effects model in the conditions where the accent is on the word preceding the target word. The table shows log-transformed reaction times.

	Contrast	Est.	SE	t.ratio	p.value
a)	ø vs. $L+H^*$	.044	.039	1.137	.345
	ø vs. L^*+H	.058	.039	1.490	.202
	ø vs. L^*	−.001	.039	−.005	.877
b)	$L+H^*$ vs. L^*+H	.014	.039	.353	.985
	$L+H^*$ vs. L^*	−.044	.039	−1.141	.667
	L^*+H vs. L^*	−.056	.039	−1.292	.452

Table 5: Pairwise comparisons between a) the baseline condition (ø) vs. the three different accent types, and b) the three different accent types. Note that a '–' means that the condition mentioned second in the first column is slower than the condition mentioned first in the first column.

3.3 Interaction of accent and position (RQ 3)

Figure 6 shows the distribution of the reaction times including the median for the conditions $L+H^*$ on the target word or on the word preceding the target word as well as for the conditions L^*+H on the target word or on the word preceding the target word.

We expected faster target word identification times when the prominent accent types ($L+H^*$ and L^*+H, respectively) were on the target word compared to when they were on the word preceding the target word. However, when the accent $L+H^*$ was on the word preceding the target word, the mean reaction time was 31 ms faster compared to when it was on the target word itself. In the condition L^*+H, the mean reaction time was 28 ms faster when on the word preceding the target word compared to when on the target word.
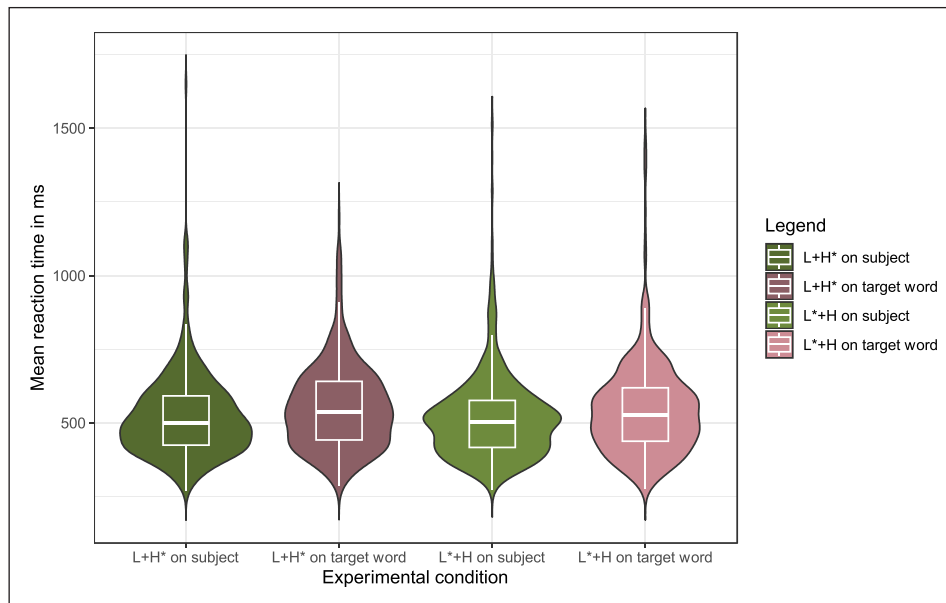
Figure 6: Mean reaction time in ms per experimental condition where either the target word or the word preceding the target word were manipulated with an *L+H* or L*+H* accent respectively.

We ran a linear mixed-effects model with random intercepts for participant and random intercepts for items. The output of this model is given in **Table 6** and shows non-significant effects for all experimental conditions. Post hoc comparisons (Bonferroni corrected) showed no significant differences between experimental conditions (see **Table 7**).

Accent type	Estimate	SE	t	p
L+H* on preceding word (Intercept)	6.227	.041	153.782	
L+H* on target word	.057	.044	1.283	.208
L*+H on preceding word	–.014	.044	–.317	.753
L*+H on target word	–.007	.063	–.117	.907

Table 6: Fixed effects from the linear mixed effects model in the conditions where the accent is on the target word or on the word preceding the target word. The table shows log-transformed reaction times.

Accent type	Contrast	Est.	SE	t.ratio	p.value
L+H*	preceding vs. target word	–.057	.045	–1.283	.416
L*+H	preceding vs. target word	–.05	. 045	–1.117	.543

Table 7: Pairwise comparisons for the accent type on the word preceding the target word vs. on target word.

3.4 Post-hoc experiment

The mean rating of how highlighted, thus prominent, the presented target words sounded was 1.75 (SD = .91) in the condition ø and 4.25 in the condition $L+H^*$ (SD = .97). We ran a mixed effects ordinal model (Christensen 2023) in R (R Core Team 2021). The model revealed a significant effect for experimental group ($ß$ = –4.8, p < .0001), which confirms that the target words manipulated with an $L+H^*$ accent type were rated as more highlighted than the target words that were deaccented. **Figure 7** shows the distribution of the ratings for the material in these two experimental conditions.

The mean rating of how highlighted, thus prominent, the word preceding the target words sounded in the experimental material was 1.75 (SD = .88) in the condition ø, and 2.3 (SD = 1.38) in the condition with an $L+H^*$ accent on the word preceding the target word. The statistical analysis confirmed a significant difference for the rating of the prosodic prominence of the preceding word in these two experimental conditions ($ß$ = –4.4, p < .0001), suggesting that the preceding words manipulated with an $L+H^*$ accent type were rated as more highlighted than the preceding words that were deaccented.
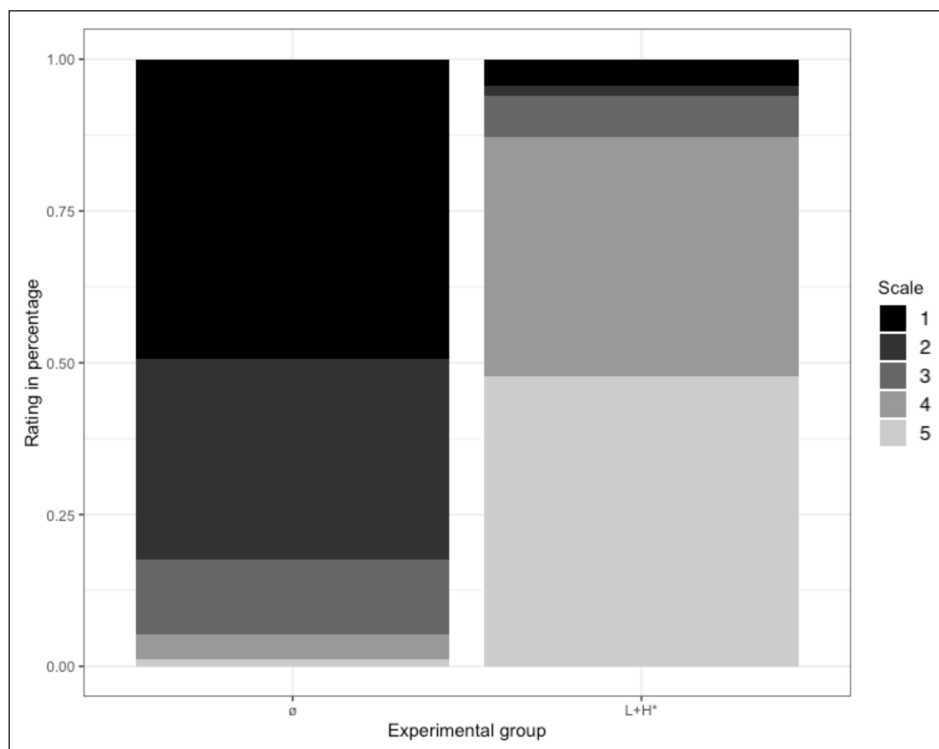

Figure 7: Distribution of the rating of the perceptual prominence in the condition ø on the target word and $L+H^*$ on the target word. The scale 1 to 5 indicates how highlighted the target word sounded, 1 being not highlighted at all, 5 being highly highlighted.

4 Discussion

Overall, our study did not show effects of prosodic prominence on language processing. Specifically, our data did not confirm the expected effect of accent type on word recognition times. Prominent accent types on the target word did not facilitate target word recognition and, likewise, prominent accent types on the word preceding the target word did not lead to hindering effects on target word recognition. We also did not find the expected interaction between accent type and position of the accent. A prosodically prominent accent type on the target did not lead to faster recognition times compared to when the same accent type was manipulated on the word preceding the target word. In the following, we will discuss these results in detail.

4.1 Accent manipulation on the target word (RQ 1)

There was no significant effect of accent type on target word recognition times. We expected prosodic prominence to draw attention to the target word, thereby facilitating word identification. More specifically, we predicted faster recognition times if the target word was manipulated with a prominent accent type (L^*+H or $L+H^*$) compared to when the target word was manipulated with a less prominent accent type (L^*) or was deaccented ($\emptyset$). By contrast, the mean reaction time in the baseline condition (deaccentuation, $\emptyset$) was numerically faster than the mean reaction times in the conditions $L+H^*$ and L^*. These results do not support the predicted scale of identification times ($L+H^* < L^*+H < L^* < \emptyset$) but rather imply that a prosodically prominent accent on the target word does not facilitate identification of these target words.

We controlled for a number of possible confounding variables that might have influenced word-identification times, such as the number of phonemes of the target words, the duration in ms of the target word and the word frequency of the target word. We furthermore had the written sentences rated before we distributed them amongst the experimental groups. Additionally, the results of a post-hoc study, where we asked the participants to rate how highlighted the target words sounded in the conditions $\emptyset$ and $L+H^*$, confirmed that participants indeed perceived the difference in prosodic prominence between these two experimental conditions. The condition $L+H^*$ was rated as significantly more highlighted, thus prominent, than the condition $\emptyset$. In the light of the finding that differences in prosodic prominence in the experimental material were clearly perceived, we conclude that the lack of an effect of prosodic prominence on language processing in our experiment is sound and reflects that differences in the prosodic prominence of a target word do not affect its identification times in a word monitoring experiment.

This result contrasts with findings by Cutler & Swinney (1987), who report a significant effect of prosody on word identification times in a word monitoring experiment. The ten tested adult participants reacted about 36 ms faster when the primary sentence accent was on the target word compared to when the primary sentence accent occurred elsewhere in the sentence. Note,

however, that Cutler & Swinney (1987) presented the same sentences once with and once without the prosodic manipulation on the target word. It is hence possible that due to the relatively small number of filler sentences (six) and experimental sentences (16), participants remembered the sentences and the position of the target word within the sentence, affecting their word identification times. In addition, it might have been the case that the participants expected the target word to be in the particular position in the sentence where it occurred because of the syntactic and prosodic structure of the experimental sentences.

Another difference between our experiment and the one by Cutler & Swinney (1987) concerns the placement of the accent in the experimental sentences. Cutler & Swinney (1987) manipulated the accent in the nuclear position of the sentence (their wording was primary sentence accent), while in our study the accentual manipulation was in the prenuclear position of the sentence. Could this difference account for the observation that we found no effect of prosodic prominence on word monitoring times contrary to the findings of Cutler & Swinney (1987)? There is an on-going discussion in the literature with regard to the status of prenuclear accents. Some researchers argue that prenuclear accents are purely ornamental without serving a function (see e.g., Büring 2007; Calhoun 2010; Jagdfeld & Baumann 2011; Kapatsinski et al. 2017; Baumann et al. 2021). For instance, Jagdfeld & Baumann (2011) found that when participants were asked to judge if a word sounded accented or not, they were less sensitive to accents in prenuclear position compared to accents in nuclear position. Similar results were reported by Kapatsinski et al. (2017) who found that adults weighed pitch peaks at the end of an utterance, equivalent to nuclear accents, as more important than in the beginning of the sentence. Against this background, one could assume that participants in our study ignored the accentual manipulations of the target words in our experiment since this manipulation occurred in the prenuclear position, thus accounting for the lack of an effect of prosodic prominence on language processing in our data. Note, however, that the data of our post-hoc study indicate that participants indeed perceived the difference in prosodic marking in the experimental material although it occurred in prenuclear position. This provides evidence against the assumption that the lack of an effect of prosodic prominence on word monitoring in our study can be attributed to the position of the accent manipulation in our experimental material. Moreover, the assumption that prenuclear accents are purely ornamental and not functional is controversial. A number of studies have argued that prenuclear accents are not purely ornamental but that they do serve a function, such as marking contrast (e.g., Féry & Krüger 2008; Braun & Biezma 2019). For instance, Braun & Biezma (2019) employed a visual-world paradigm where they auditorily presented participants with a declarative sentence (e.g., *The swimmer wanted to put on flappers*) while showing them four pictures on the screen, the target (*flappers*), an unrelated filler, a related but non-contrastive filler (*sports*) and a contrast alternative to the subject (*diver*). The prenuclear subject was either manipulated with an $L*+H$ or with an $L+H*$. Braun & Biezma (2019) found that a prenuclear

$L*+H$ accent on the subject led to higher fixation rates on the contrast alternative, i.e., more looks to the *diver* when the subject *swimmer* was accented. This led the authors to conclude that the prenuclear $L*+H$ accent serves to activate discourse alternatives. Similarly, Féry & Krüger (2008) found that the tone of the prenuclear accent changes, depending on the information status of the word in nuclear position, i.e., whether the entity in the nuclear position is given or new. These observations suggest that prenuclear accents are not purely ornamental. Given these findings and given that in our study the participants did perceive the prenuclear accent manipulation, we suggest that the lack of an effect of prosodic prominence on word monitoring cannot be attributed to the participants disregarding the accent manipulations in the prenuclear position. However, future research should specifically address the role of nuclear and prenuclear accents on word monitoring tasks, for instance by employing the word monitoring paradigm with sentences that differ in the position of the accentual manipulation (i.e., nuclear vs. prenuclear position) but are kept equal otherwise.

4.2 Accent manipulation on the word preceding the target word (RQ 2)

With respect to our second research question, we assumed that a prosodically prominent accent on the word preceding the target adverbial (i.e., on the subject) would draw attention, thereby binding processing resources to this word, leading to longer word identification times of the successive target word in comparison to a condition where the preceding word was deaccented. Specifically, we expected that the participants would react fastest when the word preceding the target word was deaccented and that identification times of the target word would increase when the word preceding the target word was presented with increased prosodic prominence ($L*, L*+H, L+H*$ respectively). However, there was no significant effect of accent type on word identification times when the word preceding the target word was prosodically manipulated. Counter to our expectations, the mean reaction times in the conditions $L+H*$ and $L*+H$ on the preceding word led to numerically faster recognition times than when the preceding word was deaccented. Moreover, pairwise comparison showed neither differences between the three accent types and the baseline condition (deaccentuation, ø), nor between the different accent types ($L+H*$ vs. $L*+H$, $L+H*$ vs. $L*$, and $L*+H$ vs. $L*$). Thus, our results do not support the initially postulated hypothesis that prosodically prominent accents bind processing resources, leading to slower identification times of the target word that directly follows the prosodically prominent word. Furthermore, the suggested identification time scale ($L+H* > L*+H > L* > ø$) was not confirmed. In contrast, these results support the findings described in section 4.1., that prosodic prominence does not influence on-line sentence processing in a word monitoring task.

Although the accent on the preceding word (here the subject) is not in the nuclear but the prenuclear position, it might still be considered more natural to have the sentential subject accented compared to having an adverbial accented. Thus, the processing of the target word

might have been slowed down in those experimental conditions where the target word (i.e., the adverbial) was prosodically manipulated (conditions 1a-d), due to a general 'unnaturalness' of prosodically marking an adverbial. This could account for the observation why we did not find an effect of prosodic prominence in these experimental conditions. Note, however, that we nevertheless would have expected an effect of prosodic prominence in those experimental conditions where the subject was prosodically manipulated (conditions 2a–d) and where prosodic manipulations might be considered more 'natural'. Contrary to this, however, our prosodic manipulations in these latter conditions did not influence word recognition times, although they were well perceived by the participants as indicated by our post-hoc study. Thus, we conclude that the position of the accent (on the subject vs. on the target adverbial) is not a decisive factor in explaining our results. Rather, the results suggest that prosodic prominence had no influence on word monitoring times in our study.

4.3 Interaction of accent and position (RQ 3)

Further evidence against the assumption that prosodic prominence draws and binds processing resources to the prosodically prominent entity, thus affecting on-line language processing, comes from the observation that we found no interaction between the type of accent (prominent $L+H^*/L^*+H$) and its position on the target word or on the word preceding it. We expected that the prominence of accent types would influence word identification times differently, depending on whether the accent was manipulated on the target word or on the preceding word. Specifically, we assumed prosodically prominent accent types on the target word to have a facilitatory effect on target word identification, while we expected prosodically prominent accent types on the word preceding the target word to bind processing resources and to slow down word identification of the following target word. However, this assumption was not supported by the statistical analyses. Neither of the two pairwise comparisons (L^*+H on target word vs. on preceding word, and $L+H^*$ on target word vs. on preceding word) was significant.

4.4 General discussion

Accents are not produced in isolation but it is argued that for instance the contrast-signalling accent type $L+H^*$ might evoke a contrastive context and activate contrast alternatives even when the context and alternatives are not mentioned explicitly (e.g., Rooth 1985; Watson et al. 2008; Braun & Tagliapietra 2010). For instance, Watson et al. (2008) compared the looks evoked by the accent types $L+H^*$ (signalling a contrast) and H^* (signalling a new referent) in an eye-tracking study, employing the visual-world paradigm. The authors created contrast pairs by first introducing two referents ("Click on A and B"), then they made one referent more salient ("Move B to the right of C") and finally they either re-mentioned one referent or introduced a new referent ("Now, move A/D below E"). The referents *A* and *D* received either the accent

$L+H^*$ or H^*. Watson et al. (2008) found that the accent type $L+H^*$ evoked more looks to the contrast alternative (i.e., *A*) whereas the accent type H^* equally triggered looks to new (i.e., *D*) and to contrast alternatives (i.e., *A*). Adding on this, Braun & Tagliapietra (2010) investigated the effect of contrastive intonation in a cross-modal priming study. They presented participants with a recorded sentence (e.g., "In Florida he photographed a flamingo", ibid.: 1041) that either included a contrastive accent on the sentence-final word or not. After each sentence, a word appeared on the screen and participants had to indicate whether it was a real word or not. The word on the screen was either a contrast alterative to the last-mentioned word of the sentence (e.g., "pelican"), non-contrastive ("pink") or a control word ("celebrity"). The results revealed that when the sentence-final entity (i.e., "flamingo") received a contrastive accent, participants reacted faster to the contrast alternative (i.e., "pelican"). The authors argued that contextual alternatives became activated and more salient, although the sentences were uttered in isolation and the contrast alternatives were not mentioned explicitly.

These studies suggest that the contrastive accent $L+H^*$ might activate contrast alternatives even when they are not explicitly mentioned in the context. These observations could imply that in our experimental conditions (1a and b), the presentation of rising accents ($L+H^*$ and L^*+H) might have evoked the activation of unmentioned contrast alternatives. The activation of unmentioned contrast alternatives to the target adverbial might lead to a slowdown in language processing because participants would have to process not just the actual word but also the co-activated contrast alternatives. This slowing down of processing might have cancelled out the potentially facilitating effects of prosodic prominence on target recognition, leading to a zero-sum situation and accounting for the lack of an effect of prosodic prominence in our word monitoring task. Evidence against this assumption, however, comes from experimental conditions (2a and b) where the word preceding the target adverbial was presented with rising accents ($L+H^*$ and L^*+H). In these two conditions we would, hence, expect an addition of two hindering effects on language processing: First, the contrast-evoking rising accents on the word preceding the target word might co-activate unmentioned contrast alternatives, which would slow down processing. In addition, processing the prosodic prominence of the word preceding the target word should also slow down recognition times of the target word, as resources are bound to processing the prosodic prominence of the preceding word and are, hence, lacking for parsing the target word immediately following the prominent word. Consequently, these two experimental conditions should not only lead to the slowest recognition times in our word monitoring task but should also result in significantly slower recognition times compared to the baseline condition (ø) where neither co-activation of unmentioned contrast alternatives nor the processing of prosodic prominence would hinder the recognition of the target word. Note, however, that the data display no evidence for a significant difference in word recognition times of these two experimental conditions against the baseline condition. Rather, recognition times in both experimental conditions (2a and b) were numerically shorter than recognition times in

the baseline condition. Likewise, the comparison between experimental conditions with rising accents on the target word and conditions with rising accents on the word preceding the target word (RQ 3) indicated no differences in word recognition times. Such differences, however, would have been expected under the assumption that rising accents lead to the activation of unmentioned contrast alternatives. This should have resulted in an accumulation of processing costs in experimental conditions where the word preceding the target word was prosodically manipulated. We, therefore, conclude that the possible activation of lexical competitors was not decisive for the lack of an effect of prosodic prominence on word recognition times in our study.

While we found no evidence that a prominent accent affects word identification times in an on-line word monitoring task, studies using the off-line measure of word recall have quite consistently reported an effect of prosodic prominence on the recall of accented words both in sentences (e.g., Fraundorf et al. 2010; Kember et al. 2021) and serial lists (e.g., Savino et al. 2020), indicating that prosodic prominence serves to anchor words in working memory more deeply. Furthermore, eye-tracking studies, for example, have repeatedly shown facilitatory effects of prosodic prominence in visual search tasks (e.g., Weber et al. 2006; Chen et al. 2007; Ito & Speer 2008; Braun & Biezma 2019), i.e., participants found target referents earlier when they were presented with a prosodically prominent accent. However, in these studies prosodic prominence served its function of indicating contrast to an element given in a preceding context sentence or in a visual display. This suggests that it was not prosodic prominence alone that facilitated the search for the target referent but it was the prediction generated from the interplay of the accent type used and the contrast relationship between the referents. However, in this present study, we sought to investigate whether prosodic prominence *per se* affects on-line language processing, i.e., when it is presented without an explicitly contrast-inducing context. In contrast to findings reported by Cutler & Swinney (1987) who also conducted a word monitoring task to investigate the influence of prosodic prominence on on-line language processing, we found no evidence that word identification times were affected by prosodic prominence. We assumed that prosodic prominence leads to hyperarticulation and speculated that this might result in a faster lexical selection of the prosodically prominent word and, hence, in faster word recognition times in our word monitoring task. However, our results provide no indication for such a facilitatory effect of prosodic prominence on word recognition times. This might suggest that the process of selecting the correct lexical item is a stage during language processing that is not influenced by prosodic cues. More research using different on-line language processing measures is needed to determine whether this suggestion holds true and to gain a more detailed view on the influence of prosodic prominence on on-line language processing.

5 Conclusion

We conducted a word monitoring task with speakers of German in order to find out whether and how prosodic prominence influences language processing on-line. We postulated that

prosodic prominence draws attention, and thereby binds processing resources, to the prominent entity, thus affecting language processing. In order to test this hypothesis, we designed a word monitoring task where either the target word itself or the word directly preceding the target word was prosodically manipulated with accent types of high ($L+H^*$, L^*+H) and low (L^*, $\emptyset$) prosodic prominence. We assumed that identification of a target word would proceed quicker when it was prosodically prominent ($L+H^*$, L^*+H) than when it was not (L^* or $\emptyset$). At the same time, we expected that identification of a target word would be slowed down when the word directly preceding the target word was prosodically prominent, because we expected the parsing of prosodic prominence to bind processing resources that would, hence, not be available for the identification of the following element.

The results, however, did not confirm these hypotheses. Linear mixed regression models yielded non-significant differences in identification times between conditions where the prosodic prominence of the target word was manipulated, as well as between conditions where the prosodic prominence of the word preceding the target word was manipulated. In addition, neither direct comparisons between conditions where the target word or the word preceding the target word were presented with the same prominent accent type nor direct comparisons between the different accent conditions indicated significant differences in word identification times. Our findings therefore suggest that prosodic prominence does not affect on-line language processing when it is presented without a context indicating the information status of the prosodically marked element.

While prosodic prominence has been shown to deeper anchor words in our memory, as repeatedly shown in word recall tasks, its influence on on-line sentence processing seems to be less clear-cut. On the one hand, eye-tracking studies have shown that prosodic prominence facilitated the search for referents in visual world tasks, suggesting that prosodic prominence influences on-line language processing. On the other hand, we did not find an effect of prosodic prominence on word identification times in our word monitoring task. Further research is certainly needed to confirm our results that prosodic prominence does not influence word identification times in a word monitoring task and to also address the issue of how prosodic prominence influences language processing across different experimental on-line methodologies.

Data availability

Data, materials and scripts for analyses can be found here: https://osf.io/fea4p.

Ethics and consent

Approval for this experiment was given by the Faculty of Human Sciences of the University of Cologne (MPHF0064).

Competing interests

The authors have no competing interests to declare.

Authors' contributions

Barbara Zeyer collected the data and developed the concept for this paper. This concept work was supported by Martina Penke. Barbara Zeyer computed all analyses which were controlled by and discussed with Martina Penke. This paper was written by Barbara Zeyer and revised by Martina Penke.

Barbara Zeyer: https://orcid.org/0000-0002-7240-6240

Martina Penke: https://orcid.org/0000-0003-4686-7673

References

Audacity Team. 2021. *Audacity(R): Free audio editor and recorder*, version 2.4.2.

Baayen, Harald & Piepenbrock, Richard & Gulikers, Leon. 1995. *The CELEX lexical database (WebCelex)*. Philadelphia: University of Philadelphia, Linguistic Data Consortium.

Bänziger, Tanja & Scherer, Klaus. 2005. The role of intonation in emotional expressions. *Speech Communication* 46(3–4). 252–267. DOI: https://doi.org/10.1016/j.specom.2005.02.016

Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steven. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Baumann, Stefan. 2014. The importance of tonal cues for untrained listeners in judging prominence. *Proceedings of the 10th ISSP*. 21–24.

Baumann, Stefan & Grice, Martine. 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 & Hadelich, Kerstin. 2003. Accent type and givenness: An experiment with auditory and visual priming. *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*. 1–4.

Baumann, Stefan & Mertens, Jane & Kalbertodt, Janina. 2021. The influence of informativeness on the prosody of sentence topics. *Glossa: A Journal of General Linguistics* 6(1). 1–28. DOI: https://doi.org/10.16995/glossa.5871

Baumann, Stefan & Röhr, Christine. 2015. The perceptual prominence of pitch accent types in German. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*. 1–5.

Baumann, Stefan & Winter, Bodo. 2018. What makes a word prominent? Predicting untrained German listeners' perceptual judgements. *Journal of Phonetics* 70. 20–38. DOI: https://doi.org/10.1016/j.wocn.2018.05.004

Boersma, Paul & Weenink, David. 2020. *Praat: Doing phonetics by computer*, version 6.1.16.

Bolinger, Dwight. 1986. *Intonation and its part: Melody in spoken English.* Stanford: Stanford University Press. DOI: https://doi.org/10.1515/9781503622906

Braun, Bettina & Biezma, María. 2019. Prenuclear L*+H activates alternatives for accented words. *Frontiers in Psychology* 20. 1–22. DOI: https://doi.org/10.3389/fpsyg.2019.01993

Braun, Bettina & Tagliapietra, Lara. 2010. The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes* 25. 1024–1043. DOI: https://doi.org/10.1080/01690960903036836

Büring, Daniel. 2007. Intonation, semantics and information structure. In Ramchand, Gillian & Reiss, Charles (eds.), *The Oxford handbook of linguistic interfaces*, 445–474. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199247455.001.0001

Calhoun, Sasha. 2010. The centrality of metrical structure in signaling information structure: A probabilistic perspective. *Language* 86(1). 1–42. DOI: https://doi.org/10.1353/lan.0.0186

Cangemi, Francesco & Baumann, Stefan. 2020. Integrating phonetics and phonology in the study of linguistic prominence. *Journal of Phonetics* 81. 1–6. DOI: https://doi.org/10.1016/j.wocn.2020.100993

Chen, Aoju & den Os, Els & de Ruiter, Jan Peter. 2007. Pitch accent type matters for online processing of information status: Evidence from natural and synthetic speech. *The Linguistic Review* 24. 317–344. DOI: https://doi.org/10.1515/TLR.2007.012

Christensen, Rune H. B. 2023. *Regression Models for Ordinal Data*, version 2023.12-4.

Cole, Ronald & Jakimik, Jola. 1980. How are syllables used to recognize word? *Journal of the Acoustical Society of America* 67(3). 965–970. DOI: https://doi.org/10.1121/1.383939

Cutler, Anne & Clifton, Charles. 2001. Comprehending spoken language: A blueprint of the listener. In Brown, Colin & Hagoort, Peter (eds.), *The neurocognition of language*, 123–166. Oxford. Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780198507932.003.0005

Cutler, Anne & Foss, Donald. 1977. On the role of sentence stress in sentence processing. *Language and Speech* 20(1). 1–10. DOI: https://doi.org/10.1177/002383097702000101

Cutler, Anne & Pearson, Mark. 1986. On the analysis of prosodic turn-taking cues. In Johns-Lewis, Catherine (ed.), *Intonation in discourse*, 139–155. London: Routledge. DOI: https://doi.org/10.4324/9780429468650

Cutler, Anne & Swinney, David. 1987. Prosody and the development of comprehension. *Journal of Child Language* 14(1). 145–167. DOI: https://doi.org/10.1017/S0305000900012782

Dahan, Delphine & Magnuson, James. 2006. Spoken word recognition. In Traxler, Matthew & Gernsbacher, Morton (eds.), *Handbook of psycholinguistics*, 249–283. DOI: https://doi.org/10.1016/B978-012369374-7/50009-2

de Jong, Kenneth. 1995. The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustic Society of America* 97(1). 491–504. DOI: https://doi.org/10.1121/1.412275

Féry, Caroline & Kügler, Frank. 2008. Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics* 36. 680–703. DOI: https://doi.org/10.1016/j.wocn.2008.05.001

Fraundorf, Scott & Watson, Duane & Benjamin, Aaron. 2010. Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language* 63(3). 367–386. DOI: https://doi.org/10.1016/j.jml.2010.06.004

Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of Linguistics* 19(2). 377–417. DOI: https://doi.org/10.1017/S0022226700007799

Himmelmann, Nikolaus & Primus, Beatrice. 2015. Prominence beyond prosody – a first approximation. In de Dominicis, Amedeo (ed.), *pS-prominenceS: Prominences in linguistics. Proceedings of the International Conference*, 38–58. Viterbo: Dicum Press.

Ito, Kiwako & Speer, Shari. 2008. Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language* 58. 541–573. DOI: https://doi.org/10.1016/j.jml.2007.06.013

Jagdfeld, Nils & Baumann, Stefan. 2011. Order effects on the perception of relative prominence. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)*. 958–961.

James, William. 1890. *The principles of psychology*. New York: H. Holt and Company. DOI: https://doi.org/10.1037/10538-000

Kapatsinski, Vsevolod & Olejarczuk, Paul & Redford, Melissa. 2017. Perceptual learning of intonation contour categories in adults and 9- to 11-year- old children: Adults are more narrow-minded. *Cognitive Science* 41. 383–415. DOI: https://doi.org/10.1111/cogs.12345

Kember, Heather & Choi, Jiyoun & Yu, Jenny & Cutler, Anne. 2021. The processing of linguistic prominence. *Language and Speech* 64(2). 413–436. DOI: https://doi.org/10.1177/0023830919880217

Koch, Xaver & Spalek, Katharina. 2021. Contrastive intonation effects on word recall for information-structural alternatives across the sexes. *Memory & Cognition* 49. 1312–1333. DOI: https://doi.org/10.3758/s13421-021-01174-1

Ladd, Robert. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

Lenth, Russel. 2022. emmeans: estimated marginal means, aka least-squares means. https://CRAN.R-project.org/package=emmeans, version 1.7.2.

Marslen-Wilson, William & Welsh, Alan. 1978. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychogology* 10(1). 29–63. DOI: https://doi.org/10.1016/0010-0285(78)90018-X

Mathôt, Sebastiaan & Schreij, Daniel & Theeuwes, Jan. 2012. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods* 44(2). 314–324. DOI: https://doi.org/10.3758/s13428-011-0168-7

McClelland, James & Elman, Jeffrey. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18(1). 1–86. DOI: https://doi.org/10.1016/0010-0285(86)90015-0

Mitchell, Rachel & Ross, Elliott. 2013. Attitudinal prosody: What we know and directions for future study. *Neuroscience and Biobehavioral Reviews* 37(3). 471–479. DOI: https://doi.org/10.1016/j.neubiorev.2013.01.027

Mücke, Doris & Grice, Martine. 2014. The effect of focus marking on supralaryngeal articulation – Is it mediated by accentuation? *Journal of Phonetics* 44(1). 47–61. DOI: https://doi.org/10.1016/j.wocn.2014.02.003

Norris, Dennis. 1994. Shortlist: A connectionist model of continuous speech recognition. *Cognition* 52(3). 189–234. DOI: https://doi.org/10.1016/0010-0277(94)90043-4

Pierrehumbert, Janet & Hirschberg, Julia. 1990. The meaning of intonation in the interpretation of discourse. In Cohen, Philip & Morgan, Jerry & Pollack, Martha (eds.), *Intentions in communication*, 271–311. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/3839.003.0016

Posner, Michael. 1980. Orienting of attention. *Quarterly Journal of Experimental Psychology* 32. 3–25. DOI: https://doi.org/10.1080/00335558008248231

R Core Team. 2021. *R: A language and environment for statistical computing*, version 1.4.1106. R Foundation for Statistical Computing.

Röhr, Christine & Baumann, Stefan. 2010. Prosodic marking of information status in German. In *Proceedings of the Fifth International Conference on Speech Prosody 2016*, 1–4. Chicago: USA. DOI: https://doi.org/10.21437/SpeechProsody.2010-203

Rooth, Mats. 1985. *Association with focus*. Unpublished PhD dissertation. Amherst: University of Massachusetts.

Savino, Michaelina & Winter, Bodo & Bosco, Andrea & Grice, Martine. 2020. Intonation does aid serial recall after all. *Psychonomic Bulletin and Review* 27(2). 366–372. DOI: https://doi.org/10.3758/s13423-019-01708-4

Schaffer, Deborah. 1983. The role of intonation as a cue to turn taking in conversation. *Journal of Phonetics* 11(3). 243–257. DOI: https://doi.org/10.1016/S0095-4470(19)30825-3

Steefkerk, Barbertje. 2002. *Prominence. Acoustic and lexical/syntactic correlates* (Doctoral thesis). Available online from UvA-DARE (UMI No. 1.198694).

Warren, Paul. 2012. *Introducing Psycholinguistics*. Cambridge: CUP. DOI: https://doi.org/10.1017/CBO9780511978531

Watson, Duane & Tanenhaus, Michael & Gunlogson, Christine. 2008. Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science* 32. 1232–1244. DOI: https://doi.org/10.1080/03640210802138755

Weber, Andrea & Braun, Bettina & Crocker, Matthew. 2006. Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language an Speech* 49(3). 367–392. DOI: https://doi.org/10.1177/00238309060490030301