THE PROCESSING OF PROSODIC PROMINENCE IN GERMAN

Barbara Zeyer[1], Martina Penke[1]

[1] University of Cologne
bzeyer2@uni-koeln.de, martina.penke@uni-koeln.de

ABSTRACT

The aim of our study was to investigate how prosodic prominence affects language processing. For that, a word-monitoring task was conducted where the targets either carried a high level prominent accent (*LH**, *L*H*), an accent that is lower in prominence (*L**) or no (prominent) accent at all (*ø*). Based on the as yet sparse findings on prosodic prominence and its effect on language processing in the literature, we postulated that prosodic prominence draws attention and will lead to faster recognition times in a word-monitoring task. Specifically, we assumed that target words carrying a prominent accent type (*LH**, *L*H*, *L**) are recognized faster in a word-monitoring task than target words that are deaccented (*ø*). While previous evidence has found prosodic prominence to affect word recall, it did not speed up word identification in a word-monitoring task, suggesting that prosodic prominence might not exert an immediate influence on on-line language processing.

Keywords: Prosodic prominence, language processing, word-monitoring, German

1. INTRODUCTION

Prosody is used to mark information status [1], to distinguish focus and background [2], or to indicate turn-taking [3]. Also, speakers can use prosodically prominent accent types to draw the listeners' attention to a specific part of the utterance. A prosodically prominent entity stands out from its environment because of the prosodic characteristic it holds, such as loudness, duration or pitch accent [4].

First psycholinguistic studies have found that prosodic prominence affects language processing. Recall-tasks revealed a facilitatory effect of intonation and prosodic manipulation (see e.g., [5]), indicating that prosodically more prominent words get recalled easier than prosodically less prominent words. A word-monitoring task, conducted in English, showed that both, adults and children above six years of age identified the target word faster when it bore the primary sentence accent compared to when the primary sentence accent appeared elsewhere, either before or after the target word [6].

The reported results suggest that prosodic prominence serves to draw the attention to the accented entity, thus enabling a better recall of this entity. Here we wanted to build up on these findings, by investigating whether prosodic prominence also exerts a more immediate processing effect in speeding up language processing. We also wanted to explore the effects of different levels of prosodic prominence on language processing.

2. METHODOLOGY

2.1. Research Questions

In the current study, we wanted to investigate how different levels of prosodic prominence draw the attention of a hearer and affect language processing. For that, we conducted a word-monitoring task, similar to that of Cutler & Swinney [6], where participants were asked to push a key as soon as they recognized the previously defined target word.

To investigate the influence of different levels of prosodic prominence, we manipulated the accent on the target words. In German, different accent types are perceived differently with regards to their perceived prosodic prominence. Baumann & Röhr [7] found that, inter alia, the accent *LH** is perceived as most prominent by native speakers of German, followed by the accent *L*H*. The study furthermore showed that deaccentuation (*ø*) in German is perceived as least prominent and the accent *L** as second least prominent. To investigate the influence of different levels of prosodic prominence on identification times in a word-monitoring task, we presented target words in four different accent conditions: target words either carried *LH**-, *L*H*- or *L**-accent or were deaccented (*ø*).

We predicted that target words having a prominent accent type (*LH**, *L*H*, and *L**) would lead to faster recognition times of the target word compared to the deaccentuation condition (*ø*, the baseline condition) as the prominent accent draws the hearers attention. Moreover, we expected identification times to decrease with increasing prosodic prominence of the target word. The more prominent a word is, the more attention it should draw, leading to shorter identification times. Thus, based on the findings by Baumann & Röhr [7], we expected reaction times to display the following scale: *LH* < L*H < L* < ø*.

The study was preregistered at OSF (https://osf.io/fea4p).

2.2. Stimuli

In total, we constructed 40 experimental (10 sentences for each of the four different accent conditions *LH**, *L*H*, *L** and *ø* on the target word) and 100 filler sentences. Both, the experimental and the filler sentences consisted of seven words each.

In the experimental sentences, the target word always occurred in the fourth position after the subject. The target words always bore the prenuclear accent and the nucleus was sentence-final. All experimental sentences had the same syntactic structure (Adv V S target word O; see example (1), bold indicating the target word). This way, we ensured that neither the position of the target word nor the syntactic structure of the sentence would be a confounding variable for the results.

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

The target words were bi-syllabic, trochaic adverbials that were either adverbials of time, of reason, of modality, or of space. Adverbs of different type were equally distributed over the four experimental accent conditions so that each experimental condition contained one modal, two causal, two local, and five temporal adverbials.

We controlled for the word frequency of the target words using WebCelex [8]. The mean word frequency of the target words across the four experimental conditions was 149.08 ($SD = 328.94$). The mean word frequency of the target words did not differ significantly between the different experimental conditions ($F(3, 36) = .144$, $p = .933$).

Additionally, we had students (n = 23) rate our written sentences with regards to how natural the sentences seemed to them on a scale from 1 (very good/natural) to 5 (very bad/unnatural). Mean rating across experimental conditions was 2.21 ($SD = 1.36$). The ratings indicated no significant difference between the four experimental conditions ($F(3, 916) = 1596$, $p = .189$).

Moreover, we controlled for the mean duration in ms of the recorded target words as well as the number of phonemes the target words consisted of. The mean duration in ms of the spoken target words was 502.5 ms ($SD = 85.15$). The mean duration in ms of the target words did not differ significantly across experimental conditions ($F(3, 36) = 1.236$, $p = .311$). The mean number of phonemes the target words consisted of was 5.83 ($SD = 1.43$). There was no significant difference in the mean number of phonemes of the target words across experimental conditions ($F(3, 36) = .781$, $p = .512$).

Conditions / Category	Word frequency	Duration in ms	No. of phonemes	Rating
*LH** 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)
ø on target word	200.2 (*SD* 544.25)	458.5 (*SD* 92.16)	6.4 (*SD* 1.71)	2.34 (*SD* 1.41)

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

In addition, we created 100 sentences that served as fillers for this investigation. The filler sentences were recorded informally as in spontaneous speech. Target words in the filler sentences were not prosodically manipulated with a particular accent. In these sentences the target words were either the subject or the object and occurred either in the second, third, fourth or seventh position in the sentence. The syntactic structure was either PP V S (= possible target word) (Adv) O (= possible target word) or S (= possible target word) V (Adv) O (= possible target word). The filler sentences contained 37 sentences where the onset of the target word already appeared in a different word before the target word ('onset fillers' see (2)). We included these fillers to ensure that the participants listened to the whole word before reacting. We additionally created 16 'catch trials' which did not contain the previously defined target word and where the correct reaction was not to push the button (see (3)). With this kind of filler, we wanted to ensure that the participants did not arbitrarily push the button at some point in the sentence but listened for the target word.

(2) Onset filler
Syntax: PP V S 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.')

(3) Catch trial
Syntax: SVO
Target word: *Dackel* ('dachshund')
Sentence: *Der starke Husky zieht den Schlitten.* ('The strong husky is pulling the sled.')

All sentences were recorded by a female trained phonetician who was a native speaker of Standard German. The items were recorded in Audacity [9], using a C520 headset and the Scarlett 2i2 3rd generation 2-in, 2-out USB audio interface. We cut the items using PRAAT [10] and adjusted loudness using Audacity [9].

The experimental material was presented in five blocks. Each block consisted of 7 experimental and 21 filler sentences. The order of the sentences was pseudo-randomized so that not more than two experimental sentences occurred in a row. Furthermore, we made sure that sentences of the same experimental accent condition did not follow each other. This way we wanted to ensure that the participants would not get accustomed to a particular sentence prosody or would expect target words in a particular position in the sentence.

2.3. Participants

The participants were students at the University of Cologne. Participation in the experiment was voluntary and participants were either rewarded course credits or were reimbursed for their participation.

In total, 58 people participated in the experiment. We excluded eight participants who did not grow up monolingually with German between the age of zero and six years. One more participant had to be excluded because s/he reported to have systematically pushed the key only after the sentence had finished. This left 49 participants for statistical analysis, 15 of which reported to identify as male. The mean age was 23.3 years (range: 18–31 years). All participants reported to be neuro-typical, to have normal hearing abilities and normal or corrected-to-normal visual abilities.

2.4. Procedure

Participants were asked to identify a visually presented target word in a sentence that was subsequently presented auditorily. Trial structure was as follows: A fixation cross appeared for 700 ms on the computer screen, 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. Subjects were asked to press a 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 target word and ended either after 4000 ms or with the participant pushing the key (see Figure 1).

Participants were tested individually in a quiet room and were required to wear headphones (Sennheiser HD260) during the experiment. The experiment was designed and ran in OpenSesame (version 3.3.10) [11] and was run on a Dell, Latitude 3420, laptop. As a response device, we used the button box by SR Research.

Prior to the experiment, the participants were given oral and written information about what they were asked to do in the experiment, they were asked to fill out a questionnaire on their personal and language background and gave written consent for their participation.

After a practice phase of six sentences, the participants started the experiment. After each block of 28 sentences, they had the possibility to take a break for as long as they wanted to. Each block roughly lasted four minutes.

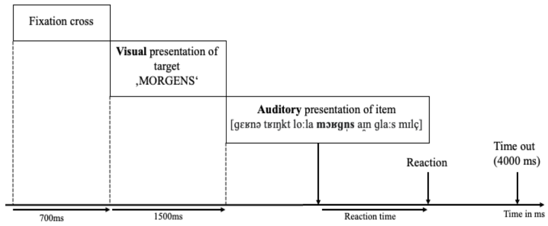

Figure 1: Experimental trial.

3. RESULTS

Reaction times of each participant that differed more than two SDs from the mean reaction time of this particular participant across experimental conditions were removed as outliers. Moreover, we removed the reaction times from the trials where the participants failed to push the button resulting in a wrong answer. In total, 66 reaction times were removed (3.36 %), leaving 1894 data points for statistical analysis.

The mean identification time for target words carrying the prominent *LH** accent was 552.78 ms (SD = 144.63 ms). In the condition *L*H*, the mean word identification time was 543.75 ms (SD = 159.75 ms). The mean identification time for the condition *L** was 562.89 ms (SD = 161.42 ms); the mean reaction time for the condition *ø* was 544.96 ms (SD = 139.01 ms) (see Figure 2).

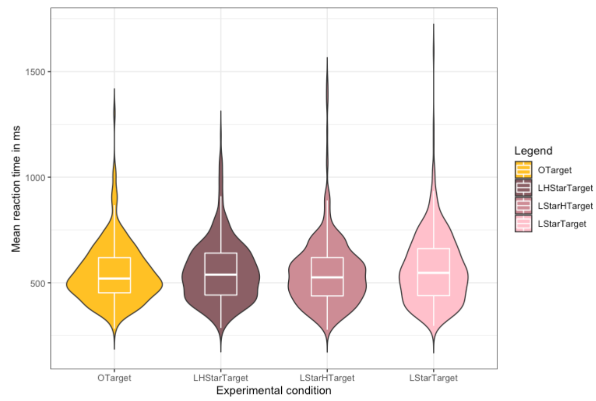

Figure 2: Violin plot of the mean reaction times in ms per experimental condition. OTarget serves as the baseline.

We computed a linear mixed-effect model using the *lme4* package [12] in R [13]. The maximally specified model did not converge and we successively removed the correlation between random intercept and random slopes for participants as well as the random slopes for participants until the model converged (see Table 2 for fixed effects).

	Est.	SE	t	p
Intercept (ø)	6.270837	.043323	144.746	<.0002
LH*	.013495	.048620	.278	.783
L*H	-.008043	.048622	-.165	.870
L*	.035511	.048658	.730	.470

Table 2: Fixed effects from the linear mixed effects model.

We conducted pairwise comparisons using the *emmeans* package [14] adjusting *p*-values for three comparisons using the Sidak method. Pairwise comparisons turned out to be non-significant for all three comparisons that were calculated (see Table 3).

	Est.	SE	t.ratio	p.value
ø vs. LH*	-.01350	.0486	-.278	.9400
ø vs. L*H	.00804	.0486	.165	.8194
ø vs. L*	-.035511	.0487	-.730	.9870

Table 3: Pairwise comparisons for the baseline condition (ø) vs. the accent types. Note that the '-' means that the condition tested against the baseline is slower than the baseline.

4. DISCUSSION

The results do not show the expected effect of prosodic prominence on word identification. We predicted the mean reaction time in the condition *ø* to be the slowest whereas we expected the mean reaction times in the conditions *L**, *L*H* and *LH** to be significantly faster. However, the mean reaction time in the condition *ø* was numerically faster than the mean reaction times in the conditions *LH** and *L**. Our predicted scale of identification times (*LH** < *L*H* < *L** < *ø*) was, thus, not supported by the results. Rather, the results indicate that prosodically prominent accents on the target word did not facilitate word identification. Note that we tried to control for a number of possible confounding variables such as word frequency, the duration in ms and the number of phonemes of the target words as well as for the rating of the written sentences when we created the items and distributed them amongst the experimental groups. Thus we feel confident that these factors were not confounding our results and that the null result we report here is sound.

Our findings contrast with Cutler & Swinney [6] who reported a significant effect of accent in a word-monitoring task. They found that their ten tested adult participants identified open class target words about 36 ms faster when the primary sentence accent was on the target word compared to when it occurred elsewhere. It might be the case that in the experimental sentences of Cutler & Swinney [6: 151] the prosodic structure of the sentence predicted where the accented target would appear (e.g., *The nurse brought a* ***clean*** *towel and took away the dirty one*, bold indicating the target word), while in our experimental sentences hearers might have expected the prenuclear accent to be elsewhere, for example on the subject. Accents are not produced in isolation and expecting the prenuclear accent on a different sentence element than where it actually appeared might have hindered sentence processing and the identification of the target word. Note, however, that having the prenuclear accent on the adverb (= target word) in sentences like *Gerne trinkt Lola* ***morgens*** *ein Glas Milch* is not uncommon in German. Thus, it is unclear why this contour should have a hindering effect on sentence processing. In order to investigate the influence of prosodic prominence on language processing detached from possible expectations based on the syntactic structure or the f0 contour of a sentence, a word monitoring task could be conducted with a serial word list.

Also, Cutler & Swinney [6] presented the same 16 experimental sentences in both conditions, with and without an accented target word, in addition to six filler sentences. Given this relatively small number of sentences, it is likely that word identification times were faster upon the second presentation of the sentence, which might have influenced the results.

Given our failure to find an effect of prosodic prominence on word identification times, it might be the case that the influence of prosodic prominence is interacting with different stages of language processing. While prosodic prominence leads to better word recall in recall tasks, it might be the case that it does not exert a more immediate effect on language processing as measured in our word monitoring task or that its effects during language processing are attenuated. More research is needed to determine whether or not prosodic prominence influences word identification during language processing.

5. ACKNOWLEDGEMENTS

We would like to thank Christine Röhr for recording our items and Max Hörl for support in the statistical analysis. This study was funded by *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation) – Project-ID 281511265 – SFB 1252 "Prominence in Language".

6. REFERENCES

[1] Baumann, S., Grice, M. 2006. The intonation of accessibility. *J. Pragmatics* 38, 1636–1657.

[2] Mücke, D., Grice, M. 2014. The effect of focus marking on supralaryngeal articulation – Is it mediated by accentuation? *J. Phonetics* 44(1), 47–61.

[3] Cutler, A., Pearson, M. 1986. On the analysis of prosodic turn-taking. In: Johns-Lewis, C. (eds), *Intonation in Discourse*. Routledge, 139–155.

[4] Terken, J., Hermes, D. 2000. The perception of prosodic prominence. In: Horne, M. (eds.), *Prosody: Theory and Experiment*. Kluwer, 89–127.

[5] Savino, M., Winter, B., Bosco, A., Grice, M. 2020. Intonation does aid serial recall after all. *Psychonomic Bulletin and Review* 27(2), 366–372.

[6] Cutler, A., Swinney, D. 1987. Prosody and the development of comprehension. *J. Child Language* 14(1), 145–167.

[7] Baumann, S., Röhr, C. 2015. The perceptual prominence of pitch accent types in German. *18th Proc ICPhS 2015*.

[8] Baayen, R., Piepenbrock, R., Gulikers, L. 1995. *The CELEX lexical database (WebCelex)*. Linguistic Data Consortium.

[9] Audacity Team. 2022. Audacity®: Free Audio Editor and Recorder [Computer application]. Version 2.4.2.

[10] Boersma, P., Weenink, D. 2020. Praat: doing phonetics by computer [Computer program]. Version 6.1.16.

[11] Mathôt, S., Schreij, D., Theeuwes, J. 2012. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods* 44(2), 314–423.

[12] Bates, B., Maechler, M., Bolker, B., Walker, S. 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-27.1.

[13] R Core Team. *R: A language and environment for statistical computing* (2021.09.0). R Foundation for Statistical Computing.

[14] Lenth, R., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., Love, J., Migues, F., Riebl H., Singmann, H. 2022. Estimateed Marginal Menas, aka Least-Suqare means. Version 1.8.3.