Speech Prosody 2022
23-26 May 2022, Lisbon, Portugal

Information status and tonal context jointly modulate prosodic prominence relations in German

Janne Lorenzen, Simon Roessig, Stefan Baumann

I*f*L Phonetik, University of Cologne, Germany
{janne.lorenzen, simon.roessig, stefan.baumann}@uni-koeln.de

Abstract

We conducted an interactive online production experiment on German in which participants were asked to read aloud stories for a fellow player who then had to sort picture cards corresponding to single sentences of the stories in the correct order. The target sentences contained two target words, an indirect object followed by a direct object, which were either *new* or *accessible* in their discourse context.

Our aim was to investigate the paradigmatic and syntagmatic effects of information status as well as the (syntagmatic) influence of tonal context on the prosodic prominence relation between the target words. Results show that (i) *new* referents are generally marked by more prominent accent types than *accessible* referents, (ii) the prominence of sentence-final accents increases/decreases when their referents are *newer*/*more accessible* than non-final referents (carrying 'medial' accents), and (iii) the type of medial accent determines the relation between accent type and information status in the final pitch accent.

The results support the idea of a *prominence budget* that is distributed across pitch accents within an utterance. Furthermore, modulations of a specific distribution of prominences may reflect changes in both meaning-related factors (such as information status) and form-related factors (such as accent type) simultaneously.

Index Terms: prosodic prominence, information status, pitch accent type, tonal context, prominence budget

1. Introduction

The production and perception of prominence is affected by a large number of factors, among them semantic-pragmatic, lexical, and prosodic-phonetic parameters (see, e.g., [1], [2], [3]). Exponents of prosodic prominence, in particular pitch accents, are in turn sensitive to their tonal context, i.e., they are influenced by other pitch accents and boundary tones within the same utterance [4], [5]. Although both effects of meaning-related factors and (form-related) prosodic structure on prosodic prominence have been addressed independently, they have rarely been dealt with in conjunction. The present paper attempts to fill this gap to some extent by investigating effects of semantic-pragmatic content *and* tonal context on the realization of prosodic prominence relations of two succeeding pitch accents in German.

Aylett & Turk [6] propose that there is an inverse relationship between 'language redundancy' and 'acoustic redundancy', a principle known as the *Smooth Signal Redundancy Hypothesis*. Referents that are encoded as highly predictable through lexical, pragmatic or semantic factors are prosodically less prominent and *vice versa*. In the present study, *information status* is investigated as a factor contributing to language redundancy. Referents can either be *new*, *accessible*, or *given*, depending on whether they have been explicitly mentioned in the discourse before or have been cued through the context provided. *New* referents have previously been found to be produced with higher prosodic prominence in many languages, *given* referents are often produced in an attenuated way and are thus prosodically less prominent, while *accessible* referents, which are considered to be cognitively 'semi-activated', fall between these two extremes [7], [8], [9].

Furthermore, pitch accents are presumably not processed independently but are influenced by their syntagmatic tonal context and thus need to be viewed as part of a larger, coherent unit. Schettino & Wagner [4], for example, find that pitch accents are perceived as more prominent the further away they occur from preceding pitch accents. Rather than investigating single-accent utterances, we are thus interested in the prominence *relation* of pitch accents produced on two subsequent noun phrases (NPs).

Rump & Collier [5], looking at effects of focus conditions on peak height, find that the information structure of a sentence affects the prominence level of multiple elements in the phrase simultaneously and in a relative way. Their perception results suggest that information structure is not only encoded locally (i.e., by position and type of the nuclear pitch accent) but it is distributed across the sentence in the *prominence relation* of multiple accents. For contrastive focus, the participants of their study judged lower peak accents to be more appropriate when the focus domain was preceded by another contrastive focus NP. When the contrastive focus was preceded by an NP in the background, higher peak accents for contrast were found to be more appropriate.

One way in which prosodic prominence is encoded is via the categorical choice of accent type realized on a metrically strong syllable. Following [10], the ranking of (GToBI [11]) accent types listed below will be assumed, with an increase in perceived prominence from left to right:

L* < H+L* < H* < L*+H < L+H*

In this paper, we ask three main research questions: 1) To what extent and in which way is the information status of two successive referents in an utterance encoded by prosodic prominence, operationalized in terms of accent type choice? 2) How does the prominence *relation* between the accents change depending on the information status of the referents? 3) How does the tonal context, i.e., the form of a preceding pitch accent, modulate the form of a subsequent accent? In order to answer these questions, we collected speech data in an interactive card sorting task.

10.21437/SpeechProsody.2022-2

2. Method

2.1. Speech material

A production experiment was carried out in which the information status on two target words (an indirect and a direct object) were manipulated. The target words were disyllabic with stress on the initial syllable (e.g., *Maler* 'painter'). The carrier sentence was embedded in a story consisting of four sentences. The second sentence provided context towards both, one or neither of the target words in the subsequent sentence in order to render them *accessible* or *new*. *Accessible* referents were preceded by a definite article, *new* referents by an indefinite article.

We created eight such stories, each came in four different versions: *accessible-accessible*, *new-new*, *accessible-new*, *new-accessible* (see Table 1 for an example of the *accessible-new* condition). The stories were presented to the participants in a Latin square design, so that every participant read only one version of each story. This was done to prevent target words from becoming lexically given through repeated mentions.

Table 1: *Example story used in the production task*[1]

	1. Anna hatte einen erfolgreichen Tag. "Anna had a successful day."
	2. Am Mittag sind einige Handwerker von der Baustelle nebenan in ihr Geschäft gekommen. "At noon, some craftsmen from the construction site next door came into her store."
	3. Unter anderem hat sie **dem Maler eine Waage** verkauft. "Among other things, she sold the painter a scale."
	4. Jetzt ist sie zu Hause und entspannt sich bei schöner Musik. "Now she is at home and relaxes by listening to beautiful music."

[1] All visual elements used to create the collages were designed by macrovector / brgfx / pikisuperstar / Freepik (distributed via Freepik https://www.freepik.com/)

2.2. Procedure

The stories were presented in the form of an interactive card sorting task to foster engagement with the task and prevent monotonous speech. Participants were asked to read stories aloud for a fellow player (a confederate of the experimenter) in such a way that the fellow player would be able to remember them and, after a short delay, sort pictures into the correct order. Participants saw the corresponding pictures while reading the stories. After they finished reading, participants were able to simultaneously watch their partner sort the pictures and were asked to provide feedback (see Figure 1 for task set-up). The sorting was pre-programmed and always resulted in the correct order. In case of hesitations, repairs, exaggerated segmental articulation, or continuation rises produced by the participant in the target sentences, they were asked to repeat the reading and sorting task at the end with the experimenter feigning technical difficulties.

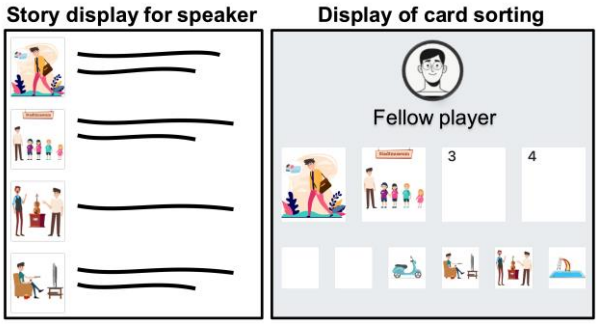

Figure 1: *Schematic depiction of the participant screen during the card sorting task*

The data collection task was conducted online via Zoom [12], and participants were recorded both via Zoom and Ennuicastr [13]. Ennuicastr records participants directly onto their own devices on separate audio tracks, which prevents unstable internet connections from distorting the audio. Zoom's built-in recording function was used as a back-up.

2.3. Hypotheses

In line with previous research (e.g., [9]), we expect to find that *new* referents are produced prosodically more prominently than *accessible* referents. In addition, we hypothesize that there is a balancing between the prominence level of the pitch accents realized on the first and second target word, depending on their information status: the more salient the second accent is produced, the less salient the first one becomes and *vice versa*. The following ranking in terms of prosodic prominence can be expected for the second pitch accent (the reverse ranking is hypothesized to emerge for the first accent):

new-accessible < accessible-accessible = new-new < accessible-new

Finally, the type of accent realized on the first target word may constrain the accent type on the second target word. Specifically, our target words are prone to being produced with *hat patterns*, i.e., sequences of rising and falling accents. In addition, this global contour could be modulated by the

information status of the referents, such that hat patterns may be more frequent when the second referent is *accessible* and less frequent when it is *new* (i.e., we expect more peak accents in the latter case).

2.4. Subjects and analysis

We recorded 32 native speakers of German (eight male, aged 20 to 38 years), each of them reading eight stories. Thus, we collected a total of 256 stories. For the purposes of this paper, we present preliminary results based on a subset of ten speakers, i.e., 80 utterances.

Recordings were automatically segmented via WebMAUS [14], [15] and subsequently manually corrected in Praat [16]. The suprasegmental annotation was conducted independently by two annotators following the DIMA guidelines [17]. DIMA is an annotation system rooted in the autosegmental-metrical (AM) framework of intonation analysis, which aims at being phonetically more transparent than other AM-based labelling systems (e.g., GToBI) but nevertheless reflects the phonological core of a contour. In cases of disagreement, a consensus was reached between the annotators and the authors of the present paper.

Specifically, two levels of phrase boundaries (strong and weak) were annotated. On the tone tier, accent and non-accent tones were annotated, the latter as target points before and after pitch accents. For the accent type analysis, we extracted accent tones in the target words together with any non-accent tones present in the following or preceding word, resulting in monotonal or bitonal accent types. It is important to note that the extracted pitch accent types are not necessarily equivalent to GToBI labels but that their features nevertheless often correspond to each other. An LH* accent in the DIMA-based notation, e.g., is usually more steeply rising than an H*, just as an L+H* accent in the standard GToBI system.

Due to the small sample size and the preliminary character of the analysis, we decided to report only descriptive results in the present paper. A small sample size could cause convergence issues or force us to simplify our model in ways that lead to unreliable results. Our future work will apply inferential methods to the full data set.

3. Results

Our target sentences are typically produced with three pitch accents, which we call *initial*, *medial*, and *final*. The initial pitch accent falls on the utterance-initial adverbial phrase, the medial and final pitch accents are aligned with our target words. We do not use the terms "prenuclear" and "nuclear" pitch accents in this paper, since roughly 40% of the utterances are produced with a phrase boundary between the two target words. Here, we view the utterance as the relevant unit and aim to compare accent strength within this unit. Consider also that potential differences in the presence or absence of phrase boundaries, i.e., the interpretation of a pitch accent as prenuclear or nuclear, may arise due to non- or paralinguistic aspects such as speech tempo or degree of speaker expressivity and not necessarily due to the intention to signal different levels of linguistic importance of a referent.

In terms of paradigmatic effects of information status on both medial and final pitch accents, we find that if the referent is *new*, a larger proportion of LH* accents is produced on both referents (see Figure 2). If the referent is *accessible*, the less prominent accent types L*H and H* are produced more frequently in medial and final position, respectively. Newness thus triggers more prominent accents in both the indirect and direct object in our utterances.

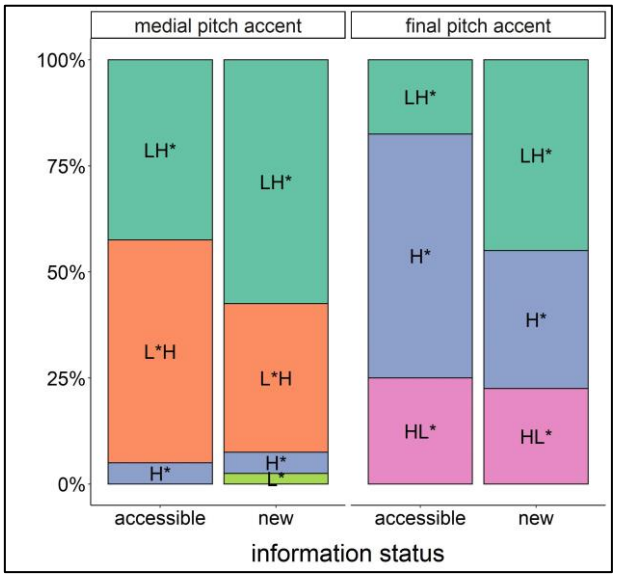

Figure 2: *Distribution of accent types as a function of accent position and information status*

Turning to the syntagmatic effects of information status on the final pitch accent, the proportion of LH* accents is largest on a *new* referent that follows an *accessible* one and smallest on an *accessible* referent that follows a *new* one (see Figure 3). The final pitch accent is thus marked most prominently when it is *newer* than a preceding referent and least prominently when it is *more accessible* than a preceding referent.

A weaker trend is that final pitch accents in the *accessible-accessible* condition appear to be produced with a smaller proportion of LH* accents, i.e., they are less prominent than in the *new-new* condition (see Figure 3). However, the proportion of HL* accents (the least prominent accent type we observe in our final pitch accents) is largest in the final pitch accent in the *accessible-accessible* and *accessible-new* condition.

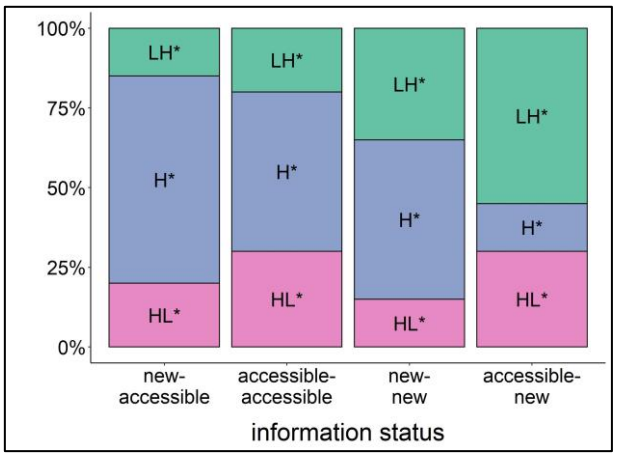

Figure 3: *Distribution of accent types in final position as a function of information status in both referents*

Finally, we find syntagmatic effects of the tonal context. Figure 4 shows the proportions of accent types on the *final* accent grouped by the *medial* accent. The left panel shows the final accent proportions when the preceding medial accent was LH*

and the right panel shows the same proportions after L*H accents. In each panel, we see the proportions of accents differentiated by whether the final referent is *accessible* or *new*.

The effect of information status on the final accent seems to be stronger when the preceding tonal context is an L*H accent: while there is only a very small proportion of LH* accents on *accessible* referents following an L*H accent, the LH* proportion clearly increases when the final referent is *new*. When the referent is preceded by an LH* medial pitch accent, the probability of LH* accents marking *accessible* final referents is considerably larger than following L*H. However, for both tonal contexts, the same trend is obtained: final referents that are *new* are more likely to be marked by LH* accents than final referents that are *accessible*.

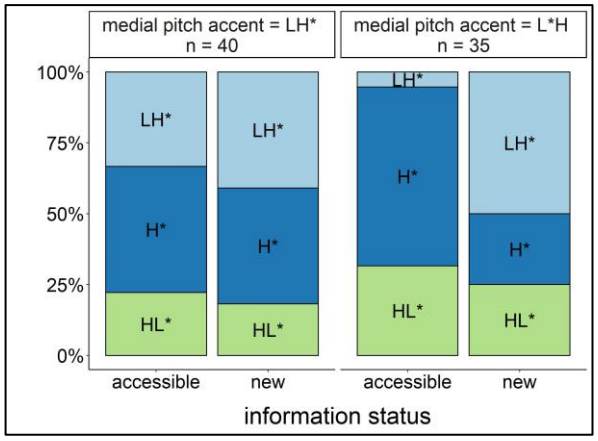

Figure 4: *Distribution of accent types in final position as a function of accent type in medial position*

4. Discussion

Our preliminary results show effects of both information status and tonal context on prosodic prominence. The overall relation between accessibility and accent types reflects previous findings, in that *new* referents are prosodically more prominent than *accessible* ones, irrespective of their position in the utterance. However, this relation is also modulated by the tonal context of an utterance, revealing a more complex picture.

In particular, we observe a stronger effect of information status on prosodic prominence after an L*H accent as opposed to succeeding an LH* accent. A possible reason is that the rise in L*H occurs later than in LH*. Since two referents follow each other directly, it may require more effort to reach an L target before the final accent after a medial L*H accent. The situation may reflect some kind of undershoot resulting in a bias towards H* rather than LH*. The larger proportion of LH* in *new* referents may be indicative of the stronger communicative intent to reach the L target in order to produce a more prominent nuclear accent (LH*). We also observe slightly more HL* accents after L*H, which are typically part of *hat contours*. This supports the above analysis and at the same time underscores the need to consider pitch accents as a part of larger units.

Furthermore, we find support for the idea of a *prominence budget* that is distributed across pitch accents within an utterance. Pitch accents are most prominent when they occur on a referent that is *newer* than a preceding one and least prominent when they occur on a referent that is *more accessible* than a preceding one, suggesting that overall prominence is distributed across the utterance. However, the preliminary results do not allow us to conclusively determine whether the prominence ranking for the final pitch accent outlined in Section 2.3 can be attested. In particular, the realizations of the final accent in *new-new* and *accessible-accessible* seem to be different, indicating that the influence of information status on the realization of the individual referent is strong. That is, if both referents are *new*, the overall prominence of the utterance is slightly higher than if both referents are *accessible*, which is in accordance with [5]'s observation on peak distribution in different focus types. However, we see important traits of the ranking in the distributions of LH*: a *new* referent in the medial position reduces the probability of LH* on the final referent. Hence, looking only at LH*, we find the following ranking that is similar to the hypothesized order:

> new-accessible < accessible-accessible < new-new < accessible-new

However, this – or an equivalent – effect may be blurred by the plateau-like shape of the (frequently occurring) hat pattern. Future research will have to investigate important aspects of the distributions in more detail, such as the proportions of HL*.

The results presented in this paper underline the importance of a distributional approach to intonation that is able to deal with the probabilistic mapping of phonological categories to post-lexical meaning [18]. Attractor landscapes represent one possibility to model the fuzziness of the form-function mapping [19], [20]. In future research, we will explore the potential of probabilistic models for our current data. In particular, it may be interesting to think about how a joint control parameter (or set of control parameters) can modulate attractor landscapes for multiple accents in a single utterance simultaneously, thereby accounting for the coherence of intonational contours.

5. Conclusions

In this investigation we showed both paradigmatic and syntagmatic effects of information status on prosodic prominence as well as a syntagmatic effect of medial pitch accent type on this relation in the final pitch accent. In particular, (i) *new* referents are marked by more prominent accent types than *accessible* referents, (ii) the prominence of final accents increases or decreases when their referents are *newer* or *more accessible*, respectively, than non-final referents (carrying medial accents), and (iii) the type of medial accent determines the relation between accent type and information status in the final pitch accent: it is strong after L*H accents but weak after LH* accents.

While the present paper operationalized prosodic prominence in terms of a categorical variable (pitch accent type), we plan on investigating how prominence relations are reflected in terms of continuous phonetic parameters, e.g., duration or pitch excursion. In addition, the effect of tonal context has to be considered in the light of the complex interaction of meaning-related (e.g., information status of referents) and form-related aspects (both phonological type and phonetic shape of pitch accents) of prominence.

6. Acknowledgements

The research for this paper has been funded by the German Research Foundation (DFG) as part of the SFB 1252 *Prominence in Language* (ID 281511265) in the projects A04 *Dynamic modelling of prosodic prominence* and A07 *Metrical prominence – Scales and Structure*s at the University of Cologne. We thank Tiffany Zukas and Sölve Lorenzen for their invaluable help with picture creation and data collection.

7. References

[1] S. Calhoun, "How does informativeness affect prosodic prominence?," *Language, Cognition and Neuroscience*, vol. 25, no. 7, pp. 1099–1140, 2010.

[2] R. Turnbull, "The role of predictability in intonational variability," *Language and Speech*, vol. 60, no. 1, pp. 123–153, 2017.

[3] S. Baumann, and B. Winter, "What makes a word prominent? Predicting untrained German listeners' perceptual judgments," *Journal of Phonetics*, vol. 70, pp. 20–38, 2018.

[4] V. Schettino, and P. Wagner, "Prosodic prominence and contextual patterns: Evidence from German," Presented at the *ICPL 2015*, Cologne, Jun. 2015.

[5] H. H. Rump, and R. Collier, "Focus conditions and the prominence of pitch-accented syllables," *Language and Speech*, vol. 39, no. 1, pp. 1–17, 1996.

[6] M. Aylett, and A. Turk, "The smooth signal redundancy hypothesis: a functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech," *Language and Speech*, vol. 47, no. 1, pp. 31–56, 2004.

[7] W. Chafe, *Discourse, Consciousness, and Time*. Chicago/London: University of Chicago Press, 1994.

[8] S. Calhoun, "The centrality of metrical structure in signaling information structure: A probabilistic perspective," *Language*, vol. 86, no. 1, pp. 1–42, 2010.

[9] S. Baumann, and A. Riester, "Coreference, lexical givenness and prosody in German," in J. Hartmann, J. Radó, and S. Winkler (eds.), [Special Issue "Information Structure Triggers"], *Lingua*, vol. 136, pp. 16–37, 2013.

[10] S. Baumann, and C. T. Röhr, "The perceptual prominence of pitch accent types in German," *Proceedings of the 18th International Congress of Phonetic Sciences*, vol. 298, Glasgow, UK: University of Glasgow, pp. 1–5, 2015.

[11] M. Grice, S. Baumann, and R. Benzmüller, "German Intonation in Autosegmental-Metrical Phonology," in S.-A. Jun (ed.), *Prosodic Typology: The phonology of intonation and phrasing*, Oxford: Oxford University Press, pp. 55–83, 2005.

[12] Zoom Video Communications, *Zoom Meetings* [Computer program], Version 5.8.3, https://zoom.us/, 2021.

[13] Wereweasel, *Ennuicastr* [Computer program], https://ecastr.com/, 2021.

[14] F. Schiel, "Automatic phonetic transcription of non-prompted Speech," in *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, USA, pp. 607–610, 1999.

[15] T. Kisler, U. D. Reichel, and F. Schiel, "Multilingual processing of speech via web services," *Computer Speech & Language*, vol. 45, pp. 326–347, 2017.

[16] P. Boersma, and D. Weenink, *Praat: doing phonetics by computer* [Computer program], Version 6.1.56, http://www.praat.org/, 2021.

[17] F. Kügler, S. Baumann, and C. T. Röhr, "Deutsche Intonation, Modellierung und Annotation (DIMA) – Richtlinien zur prosodischen Annotation des Deutschen," in C. Schwarze & S. Grawunder (eds.), *Transkription und Annotation gesprochener Sprache und multimodaler Interaktion*, Tübingen: Narr, to appear.

[18] F. Cangemi, and M. Grice, "The importance of a distributional approach to categoriality in autosegmental-metrical accounts of intonation," *Laboratory Phonology*, vol. 7, no. 1:9, pp. 1–20, 2016.

[19] M. Spivey, *The continuity of mind*. Oxford: Oxford University Press, 2007.

[20] S. Roessig, *Categoriality and continuity in prosodic prominence*. Berlin: Language Science Press, 2021.