

Individual variability in the prosodic encoding of information status in German

Janne Lorenzen, Simon Roessig, Stefan Baumann
IfL Phonetik, University of Cologne, Germany

Information status in German is encoded via prosodic prominence, such that new referents are realized prosodically most prominently, given referents least prominently and accessible referents lie in between these extremes [1]. Prosodic prominence is multifaceted, encompassing a variety of cues related to timing, spectral properties and the F0 contour. It would seem reasonable to assume that not every speaker makes use of every single cue in order to encode a given (pragmatic) contrast. Previous studies have found, for example, considerable inter-individual variability in focus type marking [2].

In order to investigate different strategies in the prosodic encoding of information status, we collected data from 32 participants in an interactive reading task via Zoom. Participants were presented with eight stories that they read aloud for a partner, who then had to sort corresponding picture cards into the correct order. Stories consisted of four sentences, the third of which was the target sentence. It contained two target words, an indirect object followed by a direct object, which were either accessible through the preceding context or new (e.g., *Unter anderem hat sie dem Maler [accessible] eine Waage [new] verkauft*. ‘Among other things, she sold the painter a scale.’).

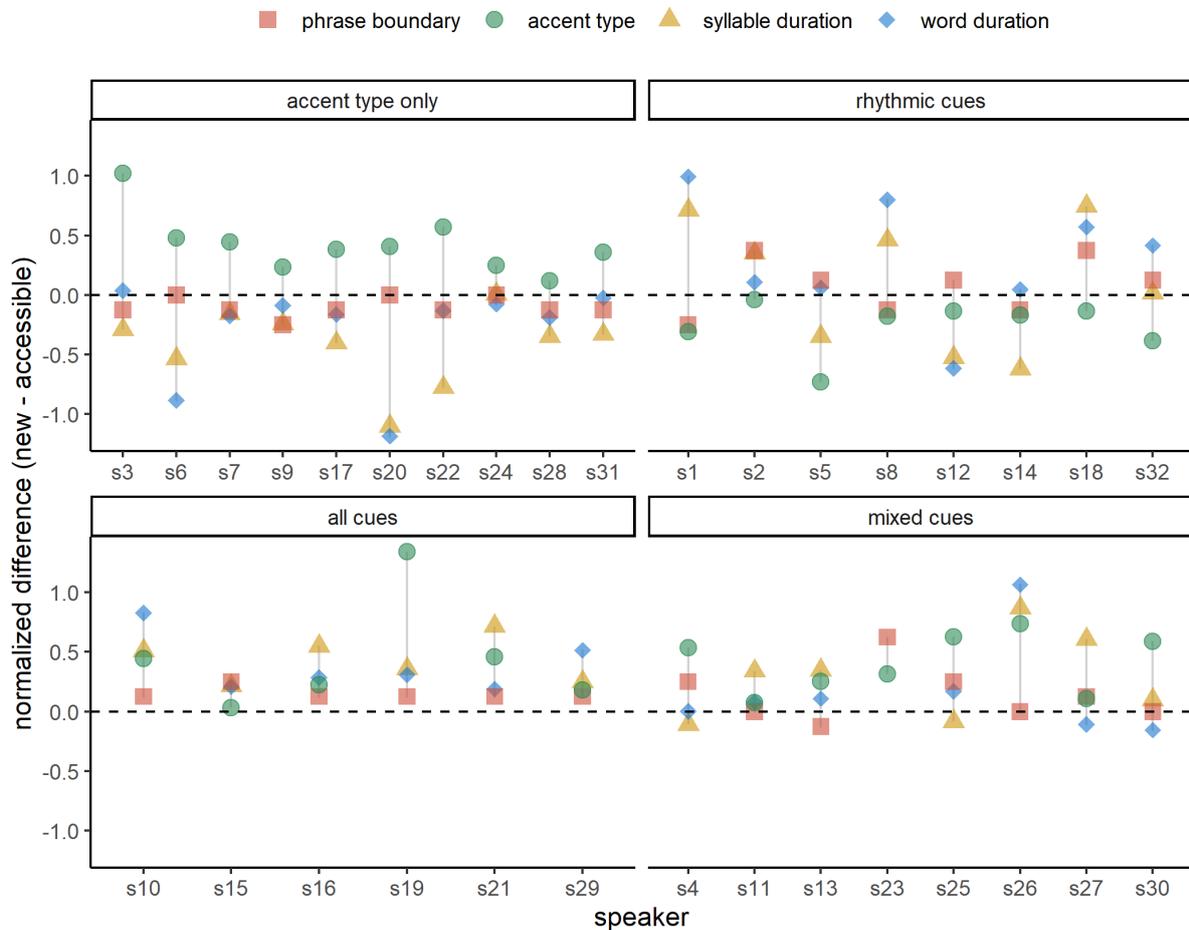
The data were segmented via WebMAUS [3,4] and prosodically annotated following the DIMA guidelines [5]. Specifically, two levels of phrase boundaries (strong vs. weak) were labelled. For the purposes of this paper, these two levels were collapsed to create a variable describing either the presence or absence of a boundary after a target word. The tone annotations were further translated into GToBI [6] accent types, which were then assigned a prominence score based on perceptual judgements of different accent types collected by [7]. In addition, word and (accented) syllable duration of the target words were measured in milliseconds. Cases where the audio was distorted in the target word were discarded ($n = 6$). Target words immediately preceding a phrase boundary were also excluded from duration measures to avoid effects of final lengthening.

Figure 1 shows the differences in the normalized averaged parameter values between new and accessible target words for every speaker. Values above the horizontal line thus indicate that a speaker uses the parameter in a way to make new target words more prominent than accessible ones. Each vertical line represents one speaker.

Overall, accent type appears to be the most robust cue as the majority of speakers ($n = 24$) realize new referents with more prominent accent types than accessible ones. In contrast, less than half of the speakers ($n = 15$) use phrase boundaries as a cue to make the preceding new referent more prominent. Nineteen speakers produce longer referents when they are new, and 17 speakers lengthen the accented syllables of new in comparison with accessible referents. Grouped by type of cues employed, the largest group of speakers ($n = 10$, upper left panel) marks prominence exclusively via accent type. On the other hand, the speakers that do not modulate accent type ($n = 8$, upper right panel) preserve the accessible/new contrast via durational cues and/or phrase boundaries. Only six speakers encode information status maximally redundantly via all four parameters considered here.

To conclude, while prosodic prominence can be encoded by many parameters, only few speakers exploit the full range of these cues, which is in line with the finding by [2]. Speakers tend to reduce this redundancy by following individual strategies in the encoding of prosodic prominence. A perception study will need to verify how successful these strategies are on the side of the listener.

Figure 1. Normalized differences between new and accessible target words averaged within speakers



- [1] Baumann, S., & Riester, A. Coreference, lexical givenness and prosody in German. In Hartmann, J., Radó, J., & Winkler, S. (Eds.), [Special Issue “Information Structure Triggers”], *Lingua*, 136, 16–37, 2013.
- [2] Cangemi, F., Krüger, M., & Grice, M. Listener-specific perception of speaker-specific productions in intonation. In Fuchs, S., Pape, D., Petrone, C., & Perrier, P. (Eds.), *Individual differences in speech production and perception*. Frankfurt: Peter Lang, 123–145, 2015.
- [3] Schiel, F. Automatic phonetic transcription of non-prompted speech. In *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, USA, 607–610, 1999.
- [4] Kisler, T., Reichel, U. D., & Schiel, F. Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326–347, 2017.
- [5] Kügler, F., Baumann, S., & Röhr, C. T. Deutsche Intonation, Modellierung und Annotation (DIMA) – Richtlinien zur prosodischen Annotation des Deutschen. In Schwarze, C., & Grawunder, S. (Eds.), *Transkription und Annotation gesprochener Sprache und multimodaler Interaktion*. Tübingen: Narr, 23–54, 2022.
- [6] Grice, M., Baumann, S., & Benz Müller, R. German intonation in Autosegmental-Metrical Phonology. In Jun, S.-A., (Ed.), *Prosodic Typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press, 55–83, 2005.
- [7] Baumann, S., & Röhr, C. T. The perceptual prominence of pitch accent types in German. *Proceedings of the 18th ICPHS*, vol. 298, Glasgow: Univ. of Glasgow, 1–5, 2015.