

Multimodal prominence relations: post-focal de-accentuation and de-gesticulation in German?

Alina Gregori¹, Pilar Prieto^{2,3} and Frank Kügler¹

¹Goethe University Frankfurt, ²ICREA, ³Universitat Pompeu Fabra

gregori@lingua.uni-frankfurt.de, pilar.prieto@upf.edu, kuegler@em.uni-frankfurt.de

For German, it is known that focus [1], [2] is marked prosodically by enhancing cues in focused constituents compared to background constituents and that different focus types can be distinguished by increasing prosodic strength [3], [4], [5]. Prosodic focus marking has been investigated in terms of accentuation, where it was found that more prominent pitch accents are associated with more prominent pragmatic contexts [6]. Thus, higher prosodic prominence is associated with pragmatic prominence. [7] proposed that prominence is a relational concept that evaluates elements in comparison to their surroundings. In accordance with this, not only were focused elements found to be highlighted prosodically in German, by a nuclear pitch accent for instance, but post-focal elements were de-accented, and therefore reduced in prosodic prominence [3]. However, potential decreases of prosodic prominence beyond pitch accentuation have not been explored yet. Since gestures employ similar functions to prosody and the systems have often been described as similar, investigating gestural prominence in focus marking is important. Gestures have been found to occur more frequently on focus than on background constituents [8], [5], [9] and are temporally more tightly aligned with pitch accents in focus [10], [11]. Both measures are instances of gestures marking focus and employing different levels of prominence. However, no study has explored the elements surrounding focus constituents for gestures and whether they show signs of reduced gestural prominence to further enhance the focused constituent. To investigate prominence trade-offs in different focus contexts in the prosodic and gestural modality, the present study examines whether prosodic and gestural prominence varies on post-focal constituents to enhance the focused constituent. We choose holistic and categorical measures for prosodic and gestural prominence to receive a general prominence impression that can be compared between modalities and mapped to focus types with different degrees of contrastivity.

We use German semi-spontaneous speech and gestures from the MultiFocus corpus [12] which are elicited in different focus conditions following [2], that increase in degrees of contrastivity – background, information focus, contrastive focus and corrective focus. Sentences contain target NPs with a narrowly focused adjective and a following post-focal noun. Data were annotated for gestures, as well as prosodic prominence [13] and gestural prominence [14], both on a scale from 0 (low prominence) to 3 (very strong prominence). Based on these annotations, prominence ratings and gesture occurrences for adjectives and nouns were extracted per focus condition. Inter-annotator agreement between three annotators was high for all variables. Results from 895 sentences show that almost all adjectives and nouns are accented prosodically, while only 50% (nouns) to 66% (adjectives) of words are accompanied by a gesture. Regarding prosodic and gestural prominence, ratings for both are higher on the focused adjective than on the post-focal noun, indicating that focused elements are more enhanced (see Figure 1 and Figure 2). On the adjective, prominence in both modalities increases along focus types. On the post-focal noun, prosodic and gestural prominence show different patterns: prosodic prominence shows a stepwise lowering of prosodic prominence on more prominent focus types, with fully deaccented occurrences in contrastive and corrective focus; gestural prominence does not show a distinct pattern across focus types, with a stable reduced prominence across conditions.

These results indicate that a holistic measure of prosodic prominence reflects what has been found for pitch accentuation, namely that post-focal information is reduced. It shows a fine-grained distinction between focus types, gradually increasing prosodic prominence on the focused word, while decreasing prominence on the post-focal noun. The post-focal pattern shows a clear distinction between contrastive and corrective focus, which cannot be distinguished on the focused adjective itself. Gestural prominence shows a weaker increase on the focused adjective per focus type, and a general pattern of gestural prominence lowering on the post-focal noun, but no distinction between focus types. This may lead to the conclusion that gestural marking cannot capture focus type differences in the post-focal domain, or that at least this reduction is not captured by a holistic prominence measure. The difference in post-focal behavior between modalities could emerge from differences in fine-grained modulation that is possible in the acoustic and the visual modality. Prosodic cues may be easier to modulate on two adjacent constituents as they involve less bodily effort than gestures. Ongoing follow-up analyses investigate temporal prosody-gesture alignment to address potential gesture positioning in the post-focal domain.

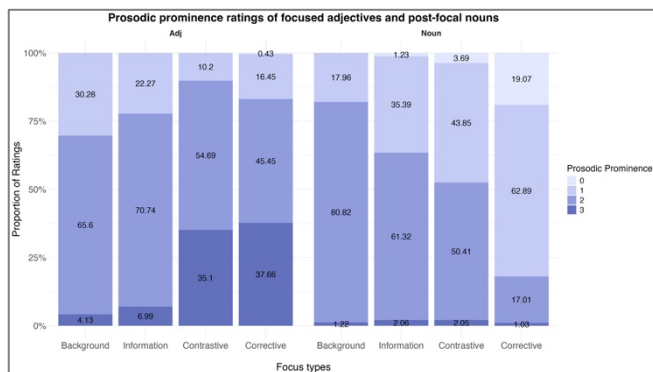


Figure 1: Prosodic prominence ratings per focus condition (background/information focus/contrastive focus/corrective focus) and word (adjective/noun). Lighter color indicates lower prominence; darker color indicates higher prominence.

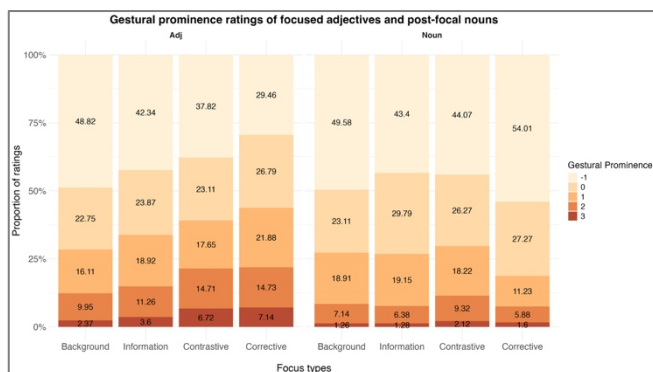


Figure 2: Gestural prominence ratings per focus condition (background/information focus/contrastive focus/corrective focus) and word (adjective/noun). Lighter color indicates lower prominence; darker color indicates higher prominence.

References:

- [1] M. Krifka, "Basic notions of information structure," *Acta Linguist. Hung.*, vol. 55, no. 3–4, pp. 243–276, Dec. 2008, doi: 10.1556/ALing.55.2008.3-4.2.
- [2] S. Cruschina, "The greater the contrast, the greater the potential: On the effects of focus in syntax," *Glossa J. Gen. Linguist.*, vol. 6, no. 1, Jan. 2021, doi: 10.5334/gjgl.1100.
- [3] C. Féry and F. Kügler, "Pitch accent scaling on given, new and focused constituents in German," *J. Phon.*, vol. 36, no. 4, pp. 680–703, Oct. 2008, doi: 10.1016/j.wocn.2008.05.001.
- [4] M. Zimmermann, "Contrastive focus and emphasis," *Acta Linguist. Hung.*, vol. 55, no. 3–4, pp. 347–360, Dec. 2008, doi: 10.1556/ALing.55.2008.3-4.9.
- [5] A. Gregori, P. G. Sánchez-Ramón, P. Prieto, and F. Kügler, "Prosodic and gestural marking of focus types in Catalan and German," in *Speech Prosody 2024*, ISCA, Jul. 2024, pp. 891–895. doi: 10.21437/SpeechProsody.2024-180.
- [6] S. Baumann and C. T. Röhr, "The perceptual prominence of pitch accent types in German.," in *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, Glasgow, Scotland, 2015.
- [7] M. Grice and F. Kügler, "Prosodic Prominence – A Cross-Linguistic Perspective," *Lang. Speech*, vol. 64, no. 2, pp. 253–260, Jun. 2021, doi: 10.1177/00238309211015768.
- [8] C. Ebert, S. Evert, and K. Wilmes, "Focus marking via gestures," in *Proceedings of Sinn und Bedeutung 15*, Saarbrücken, Germany, 2011, pp. 193–208.
- [9] S. Coego, N. Esteve-Gibert, and P. Prieto, "Preschoolers Mark Focus Types Through Multimodal Prominence: Further Evidence for the Precursor Role of Gestures," *Languages*, vol. 10, no. 5, p. 92, Apr. 2025, doi: 10.3390/languages10050092.
- [10] A. Gregori and F. Kügler, "Temporal alignment of prosody and gesture in focus: Unimodal and Multimodal Hyperarticulation," *J. Phon.*, submitted.
- [11] O. Türk and S. Calhoun, "Phrasal Synchronization of Gesture With Prosody and Information Structure," *Lang. Speech*, vol. 67, no. 3, pp. 702–743, Sep. 2024, doi: 10.1177/00238309231185308.
- [12] A. Gregori, P. G. Sánchez-Ramón, P. Prieto, and F. Kügler, "The MultiFocus corpus: An audiovisual database for the elicitation of focus types and semi-spontaneous speech and gestures in Catalan and German." OSF, osf.io/4m53q, 2025.
- [13] F. Kügler, S. Baumann, and C. T. Röhr, "Deutsche Intonation, Modellierung und Annotation (DIMA)," in *Transkription und Annotation gesprochener Sprache und multimodaler Interaktion: Konzepte, Probleme, Lösungen*, C. Schwarze and S. Grawunder, Eds., Narr Francke Attempto Verlag, 2022, pp. 23–54.
- [14] P. L. Rohrer *et al.*, "Multidimensional Labeling of Gesture in Communication: the M3D Proposal," *Corpus Pragmat.*, Jun. 2025, doi: 10.1007/s41701-025-00197-2.