# The Distribution of Co-speech Gestures, Information Structure and Prosody: A Corpus Study on Prominence Peak Alignment

Alina Gregori & Frank Kügler
*Goethe Universität Frankfurt*

This study investigates the impact of different degrees of prominence in the prosody-gesture alignment in spontaneous speech. Generally, co-speech gestures and prosodic events align in their occurrence (e.g. [1], [2], [3]). Non-referential hand gestures usually signal structural information on the discourse [1], rather than showing a semantic connection to speech. Following [4], we investigate the smallest gestural constituent, the **stroke**, and its peak, the **apex**. The stroke is the only obligatory component of a gesture, building its core and being a short, prominent movement [4]. The apex is the most prominent point of each stroke, having no or low velocity and often being the end- or turning point of the gesture [5]. Regarding prosodic prominence, we follow [6]'s proposal that different pitch accents are associated with different levels of prominence resulting in a **Pitch Accent Prominence Scale** from lowest to highest prominence: no accent < L* < !H* < H*+L < H* < L+H* (cf. [7], [2]; according to GToBI [8]). Given that prosodic prominence varies as a function of **information structure** (IS), cf. [6], [9] we address the research question whether the alignment of prosody and gesture is sensitive to IS in German spontaneous speech.

The data are taken from the **SaGA corpus** [10]. 18 dialogues (204 min.) of task-oriented spontaneous speech (direction description) were analysed. The corpus provides word and gesture type annotation [10]. Stroke and apex annotation was done following [5], pitch accent annotation with GToBI [8]. Following [11], annotation of IS was done considering information status (Given, Accessible, New) and focus (new-information focus NF and contrastive focus CF). The occurrence of gestures in relation to pitch accents and IS categories was extracted.

In total, 775 non-referential gestures were found in the corpus, from which 39,5% of all apex occurrences aligned with IS referents (Fig 1a). The remaining 60,5% of all apices occurred on words not coded for IS. Still, 51,8% of their strokes (31,4% of all gestures) overlapped with the nearest referent and were included in the analysis. When an apex aligned with an IS referent, the word always carried a pitch accent (Fig 2). Results showed no direct correlation between pitch accent type and IS, contrary to previous findings [6], [9]. When looking at pitch accent distribution on IS referents accompanied by an apex, the occurrence of L+H* is higher on "accessible" referents, while "given" showed less L+H* and most L* suggesting lower prominence (Fig 2). Regarding "accessible" referents accompanied by a gesture, the apex aligned only in 32,4% of the cases (Fig 1a, b, 3c). Comparing pitch accent and gesture distribution under the different information status levels, apices aligned best with L* accents in the "given" condition (Fig 3a), while in the "accessible" (Fig 3b) and "new" conditions (Fig 3c), the apex alignment increased along the Prominence Scale.

Independent of gesture occurrence, most L+H* accents occurred with "discourse-new" material, which is in line with the Prominence Scale [6]. With respect to focus, strikingly more L+H* were observed for NF than for CF. When a focus is accompanied by an apex, downstepped !H* occurrences were more frequent with CF than NF.

Strikingly, most IS material was accented, even though "given" referents are usually prosodically less prominent [9] (and NF is less prominent than CF [8]). Presumably, this result might be task-specific behaviour of the interlocutors to signal that for memorising a route, emphasizing every detail, even already active referents, is relevant. Although more than half of the non-referential gesture apices were not likely to align with IS referents, these preliminary results suggest that pitch accent-gesture alignment is sensitive to Information Structure.



doi:10.11576/pundp2022-1029

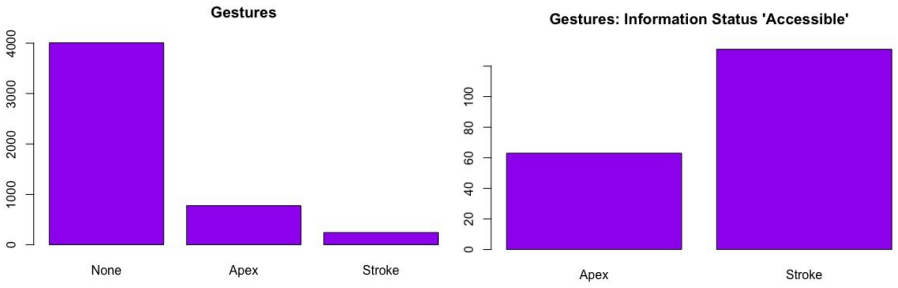


Figure 1a, b. *Distribution of Gesture occurrence a) on all IS-referents (left) and b) on referents with Information Status 'Accessible' (right).*

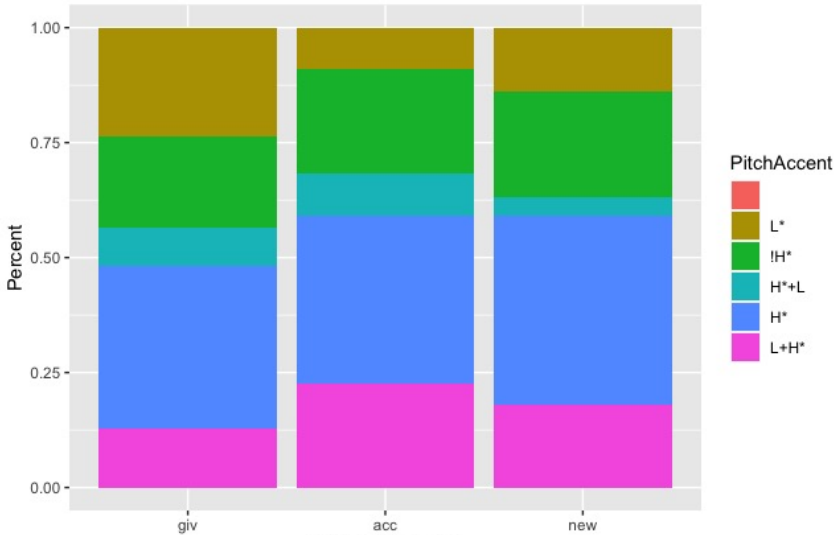


Figure 2. *Distribution of Pitch accents for each Information Status level, when accompanied by a gesture.*

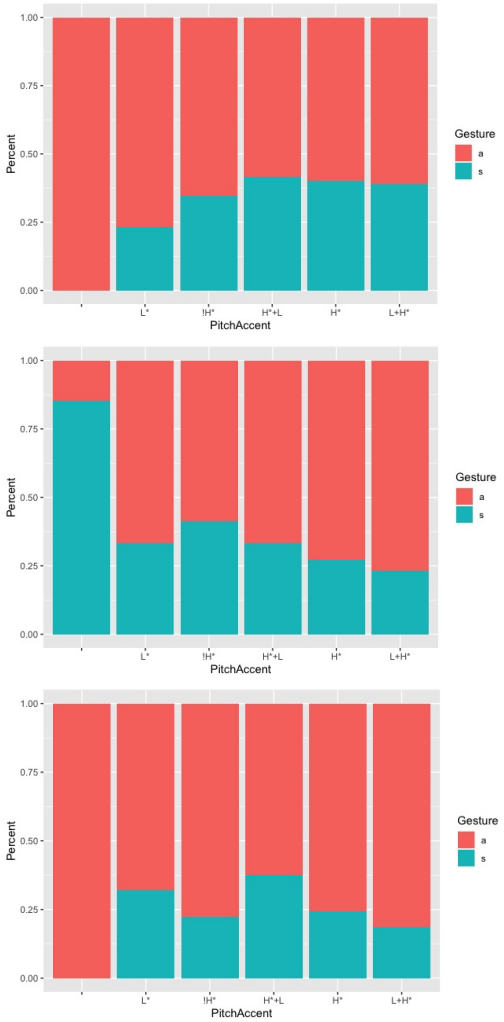


Figure 3a, b, c. *Distribution of Pitch accents and Gesture parts (a = apex, s = stroke) under different Information Status; Top to bottom: a) Given, b) Accessible, c) New.*

## References


[1] McNeill, David. 1992. *Hand and mind: What gestures reveal about thought.* Chicago: UCP.

[2] Im, Suyeon & Stefan Baumann. 2020. Probabilistic relation between co-speech gestures, pitch accents and information status. *Proceedings of the Linguistic Society of America* 5(1). 685–697.

[3] Loehr, Daniel P. 2012. Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 3(1). 71–89.

[4] Kendon, Adam. 1980. Gesticulation and speech: Two aspects of the process of utterance. In Mary R. Key (ed.), *The Relationship of Verbal and Nonverbal Communication*, 207–227. Berlin: Mouton.

[5] Rohrer, Patrick Louis, Ingrid Vilà-Giménez, Júlia Florit-Pons, G. Gurrado, Núria Esteve-Gibert, Ada Ren, Stefanie Shattuck-Hufnagel & Pilar Prieto. 2020. The MultiModal MultiDimensional (M3D) labeling system for the annotation of audiovisual corpora: Gesture Labeling Manual. MS, UPF Barcelona.

[6] Baumann, Stefan & Christine Röhr. 2015. The perceptual prominence of pitch accent types in German. Proc. 18th ICPhS, Glasgow, paper 384.

[7] Baumann, Stefan, Martine Grice & Susanne Steindamm. 2006. Prosodic Marking of Focus Domains - Categorical or Gradient? Proc. Speech Prosody 2006, Dresden, 301–304.

[8] Grice, Martine, Stefan Baumann & Ralf Benzmüller. 2005. German Intonation in Autosegmental-Metrical Phonology. In Sun-Ah Jun (ed.) *Prosody Typology: The Phonology of Intonation and Phrasing*. Oxford: OUP, 55-83.

[9] Kügler, Frank & Sasha Calhoun. 2020. Prosodic Encoding of Information Structure: A typological perspective. In Gussenhoven, C. & A. Chen (eds.): *The Oxford Handbook of Language Prosody*, Oxford: OUP, 454-467.

[10] Lücking, Andy, Kirsten Bergmann, Florian Hahn, Stefan Kopp & Hannes Rieser. 2010. The Bielefeld Speech and Gesture Alignment Corpus (SaGA). *ELRA*. 92–98.

[11] Götze, Michael, Thomas Weskott, Cornelia Endriss, Ines Fiedler, Stefan Hinterwimmer, Svetlana Petrova, Anne Schwarz, Stavros Skopeteas & Ruben Stoel. 2007. Information structure. In Stefanie Dipper, Michael Götze & Stavros Skopeteas (eds.), *Information Structure in Cross-Linguistic Corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics, and Information Structure* (Interdisciplinary Studies on Information Structure (ISIS) 7), 147–187. Potsdam: Universitätsverlag Potsdam.




doi:10.11576/pundp2022-1029