

# Refining the pragmatic synchrony rule: Interactions between prosody and gesture in the marking of the information status of discourse referents

Patrick Louis Rohrer<sup>1,2</sup>, Pilar Prieto<sup>1,3</sup>

<sup>1</sup>Grup d'Estudis de Prosòdia, Universitat Pompeu Fabra, Barcelona, Catalonia

<sup>2</sup>LLING-UMR 6310, Nantes Université, Nantes, France

<sup>3</sup>Institució Catalana de Recerca i Estudis Avançats, Barcelona, Catalonia

patrick.rohrer@upf.edu, pilar.prieto@upf.edu

A multimodal view of language holds that speakers make use of multiple modes (or strategies) in order to communicate meaning (e.g., Mondada, 2016; Perniss, 2018). Taking this perspective, both gesture and speech prosody are important modes in communicating meaning. According to McNeill's synchrony rules (1992), these two modes of communication are closely temporally coordinated (the phonological synchrony rule) and function together to convey the same pragmatic meaning (the pragmatic synchrony rule). Although much literature has highlighted the close temporal connection between prosody and gesture (e.g., Loehr, 2012, among many others) few studies have investigated how these two modes work together and interact as cues to pragmatic meaning. The information status of referents (ISR) is one aspect of Information Structure which indicates whether entities in the discourse are new to the discourse and to the listener, accessible via context or shared knowledge, or given by having already been introduced into the discourse (see Götze et al., 2007; Krifka, 2008 for a review).

As previously mentioned, most previous studies on this issue have only investigated one mode at a time by focusing either on the *prosodic* marking of ISR or the *gestural* marking of ISR individually. Studies on prosodic marking of ISR have found that given referents tend to be deaccented, while novel or accessible referents are more likely to receive prominence via pitch accentuation (see Kügler & Calhoun, 2020 for a review). However, this relationship remains unclear particularly for *prenuclear* pitch accents (i.e., any non-final pitch accents within a prosodic phrase, see e.g., Ladd, 2008), where some researchers argue that they are ornamental or rhythmically motivated in nature (e.g., Büring, 2009; Calhoun, 2010). Similar to the prosodic marking of ISR, gesture is also more likely to be produced with newer information that advances the discourse than given information (e.g., Debreslioska et al., 2013; Debreslioska & Gullberg, 2019; Marslen-Wilson et al., 1982; McNeill, 1992; Yoshioka, 2008, among others). However, most previous studies have either focused on referential gestures, or have not taken gesture types into account at all. However, previous literature has described non-referential gestures (specifically McNeill's "beat" gesture) as special markers of focus and other discourse-pragmatic meanings (e.g., Loehr, 2012; McNeill, 1992). Interestingly, McNeill (1992) also describes how referential iconic and metaphoric gestures are more likely to be produced with speech that is more communicatively dynamic (i.e., new information which pushes discourse forward). Thus, the role of gesture referentiality in marking the information status of referents remains unclear.

The current study has three main objectives. The first is to assess how gesture and prosody function *together* in the multimodal marking of the information status of referents. Given the contradictory nature of the role of gesture referentiality in marking information structure, the second objective is to assess if there is an interaction between gesture referentiality and the marking of different ISR types. The final objective is to assess whether there is an interaction between gesture and speech prosody, particularly in pre-nuclear positions. In order to reach these objectives, a corpus analysis was carried out on the English M3D-TED corpus consisting of TED Talks carried out by 5 speakers. The corpus contains over 23 minutes of multimodal speech annotated for manual co-speech gesture according to M3D standards (Rohrer et al., 2023), and prosody was manually annotated according to ToBI standards (Silverman et al., 1992). For the annotation of ISR, the simplified LISA labeling system (Götze et al., 2007) was applied to the corpus, in which the information status is applied to the referential phrase (i.e., the entire NP or PP). The analysis of gesture as a marker if ISR is mainly based on a non-strict temporal association (e.g., Rohrer et al., 2019, Debreslioska et al, 2013) and/or on the semantic meaning conveyed by the gesture.

Regarding the first research question, the results of the study show that pitch accentuation and gesture work in a parallel fashion to mark information structure, with novel and accessible referents receiving primarily a double marking by prosody and gesture (followed by a pitch accent alone), while given referents receive significantly less multimodal marking. That is, gesture and pitch accents go together in the marking of ISR. Regarding the second research question, the results of the current study did not find an effect of gesture type for the marking of ISR. That is, both referential and non-referential gestures equally function to mark ISR. Regarding the last research question, when referents are marked by a pre-nuclear pitch accent, only accessible referents were significantly more likely to also be accompanied by a gesture (Figure 1). That is, when pitch accents do not stably mark ISR, only accessible referents were significantly more likely to receive a double marking by both a pitch accents and a gesture. This study is the first to find that gesture are not parasitic on pitch accentuation, and that they may act to disambiguate ISR when the prosodic cues are more ambiguous (for similar results regarding ambiguity in morphosyntax, see Debreslioska & Gullberg, 2020). These results contribute to our understanding of the pragmatic synchrony rule, in terms of multimodal ISR marking, by being the first to shed light on the complex relationship between gestures and speech prosody in pre-nuclear contexts.

**Index Terms:** Pragmatic synchrony, Gesture, Prosody, Information Structure

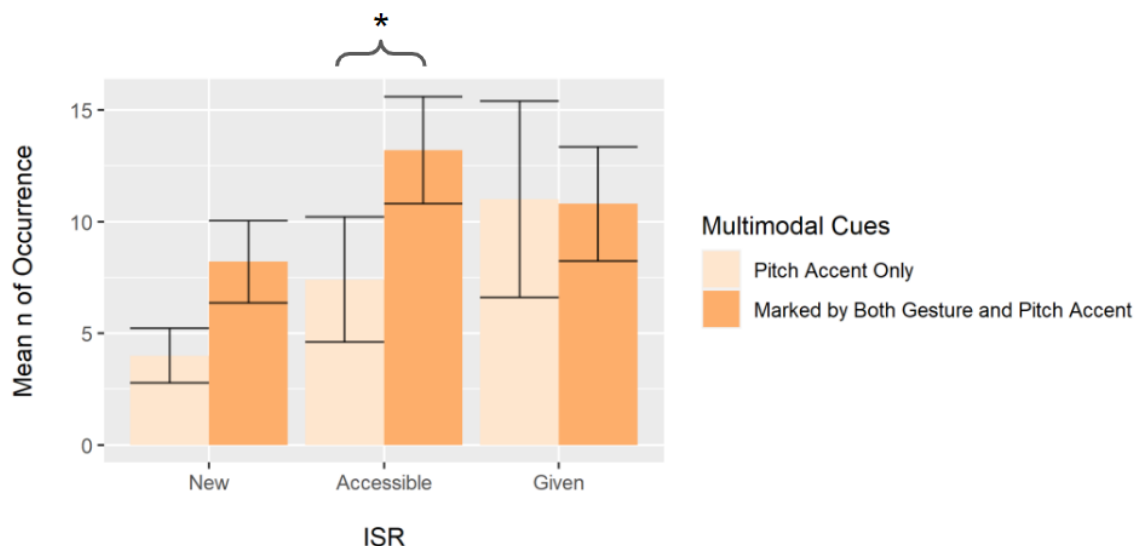


Figure 1: The average number of occurrences of referents with prenuclear pitch accents by speaker as a function of whether they co-occurred with gesture or not (error bars represent standard error).

## References

- Büring, D. 2009. Towards a typology of focus realization. In M. Zimmermann & C. Féry (Eds.), *Information Structure* (1st ed., pp. 177–205). Oxford University Press/Oxford. <https://doi.org/10.1093/acprof:oso/9780199570959.003.0008>
- Calhoun, S. 2010. The centrality of metrical structure in signaling information structure: A probabilistic perspective. *Language*, 86(1), 1–42. <https://doi.org/10.1353/lan.0.0197>
- Debreslioska, S., & Gullberg, M. 2019. Discourse Reference Is Bimodal: How Information Status in Speech Interacts with Presence and Viewpoint of Gestures. *Discourse Processes*, 56(1), 41–60. <https://doi.org/10.1080/0163853X.2017.1351909>
- Debreslioska, S., & Gullberg, M. 2020. What's New? Gestures Accompany Inferable Rather Than Brand-New Referents in Discourse. *Frontiers in Psychology*, 11(1935), 1–11. <https://doi.org/10.3389/fpsyg.2020.01935>
- Debreslioska, S., Özyürek, A., Gullberg, M., & Perniss, P. 2013. Gestural Viewpoint Signals Referent Accessibility. *Discourse Processes*, 50(7), 431–456. <https://doi.org/10.1080/0163853X.2013.824286>
- Götze, M., Weskott, T., Endriss, C., Fiedler, I., Hinterwimmer, S., Petrova, S., Schwarz, A., Skopeteas, S., & Stoel, R. 2007. Information structure. In S. Dipper, M. Götze, & S. Skopeteas (Eds.), *Information Structure in Cross-Linguistic Corpora* (pp. 147–187).
- Krifka, M. 2008. Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3–4), 243–276. <https://doi.org/10.1556/ALing.55.2008.3-4.2>
- Kügler, F., & Calhoun, S. 2020. Prosodic Encoding of Information Structure: A typological perspective. In C. Gussenhoven & A. Chen (Eds.), *The Oxford Handbook of Language Prosody* (p. 0). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198832232.013.30>
- Ladd, D. R. 2008. *Intonational Phonology* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511808814>
- Loehr, D. P. 2012. Temporal, structural, and pragmatic synchrony between intonation and gesture. *Laboratory Phonology*, 3(1). <https://doi.org/10.1515/lp-2012-0006>
- Marslen-Wilson, W., Levy, E. T., & Komisarjevsky Tyler, L. 1982. Producing Interpretable Discourse The Establishment and Maintenance of Reference. In R. J. Jarvella & W. Klein (Eds.), *Speech, Place, and Action* (pp. 339–378). John Wiley & Sons, Ltd.
- McNeill, D. 1992. Gesture and discourse. In *Hand and mind: What gesture reveals about thought*.
- Mondada, L. 2016. Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, 20(3), 336–366. [https://doi.org/10.1111/josl.1\\_12177](https://doi.org/10.1111/josl.1_12177)
- Perniss, P. 2018. Why We Should Study Multimodal Language. *Frontiers in Psychology*, 9, 1109. <https://doi.org/10.3389/fpsyg.2018.01109>
- Rohrer, P. L., Vilà-Giménez, I., Florit-Pons, J., Gurrado, G., Esteve Gibert, N., Ren, A., Shattuck-Hufnagel, S., & Prieto, P. 2023. *The MultiModal MultiDimensional (M3D) labeling system*. <https://osf.io/ankdx/>
- Silverman, K., Beckman, M. E., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. B., & Hirschberg, J. 1992. TOBI: A standard for labeling English prosody. *ICSLP-1992*, 867–870. [https://www.isca-speech.org/archive\\_v0/icslp\\_1992/i92\\_0867.html](https://www.isca-speech.org/archive_v0/icslp_1992/i92_0867.html)
- Yoshioka, K. 2008. Gesture and information structure in first and second language. *Gesture*, 8(2), 236–255. <https://doi.org/10.1075/gest.8.2.07yos>