

Analyzing Deixis in Multimodal Genres: extending the FrameNet model to account for invited shifts in joint attention in visual narratives

Natália Sathler Sigiliano¹

¹ Federal University of Juiz de Fora | FrameNet Brasil | PROFLETRAS, natalia.sigiliano@ufff.br

Keywords: Deixis, Multimodality, Joint Attention, Frame Semantics, FrameNet, Analytical Categories

In this presentation, I propose the extension of the FrameNet Brasil model for the annotation of multimodal genres so as to include categories capable of accounting for invited shifts in joint attention in visual narratives.

More than one century ago, de Saussure (1916) defined the object of Linguistics as a system composed of linguistic signs—*langue*—separating the field from Semiotics. Since then, linguists of various theoretical and methodological affiliations have been working on the description and explanation of varied phenomena in language, deixis being one of the most prominent of them. Despite the fact that many of the foundational research on this topic propose correlations between deixis in verbal language and other semiotic modes—such as gesture and image—, the main focus of research lies on the linguistic material in which the expression of deixis is grounded (Bühler, 1934; Fillmore, 1975; Levinson, 2004).

As the discussion on the relation between language use and multimodality advances, the need for a multidisciplinary approach to many phenomena once restricted to language has been emphasized, deeming the analysis of verbal text as the sole system for meaning construction insufficient. In Cognitive Linguistics, research on co-speech gesture (Sweetser, 2007; Steen et al., 2018; Cienki, 2022) has expanded the analytical coverage of models such as Mental Spaces Theory and Construction Grammar. As for Frame Semantics, since the original propositions of the theory, Fillmore (1976, 1982, 1985) has recruited the notion of visual scenes to explain the role of frames in organizing knowledge. Moreover, this theory has highlighted the fact that frames play a role in understanding that goes beyond lexical semantics. Nonetheless, the diverse possibilities for the application of frames to the analysis of meaning construction was restricted—for the sake of feasibility—as the theory was implemented in FrameNet (Fillmore et al., 2003).

Recent research has been extending the FrameNet methodology to the analysis of both textual genres (Dutra & Sigiliano, 2021) and image (Belcavello et al., 2020; Torrent et al., 2022; Viridiano et al., 2022). Those analyses have been uncovering correlations which are fundamental for the development of a Multimodal Frame Semantics. Nonetheless, they still take verbal language as a starting point, associating entities depicted in the images to verbal language superimposed to them. Considering the varied resources that the different communicative modes recruit for meaning construction, in this paper, we propose the extension of FrameNet's analytical framework to include analytical categories not grounded on verbal language. Namely, we propose an initial non-exhaustive set of categories for accounting for invited shifts in the deictic center in films.

By bringing together contributions from human communication studies—especially on the role of pre-linguistic strategies for building joint attention scenes (Tomasello, 2008, 2018)—, from the analysis of deixis in narratives (Segal, 1995; Zubin & Hewitt, 1995) and from Frame Semantics, we claim that semiotic devices of different types can be mobilized in filmic narratives so as to invite viewers to promote shifts in the center of the joint attention scene: whole scene sequences, specific visual elements, sound effects, camera takes, editing choices and, also, verbal language. To illustrate the proposed typology, we provide example analyses of deictic center shifts in films and describe the annotation methodology to be incorporated to the FrameNet model so as to account for them.

References

- Belcavello, Frederico, Marcelo Viridiano, Alexandre Diniz da Costa, Ely Edison da Silva Matos & Tiago Timponi Torrent. 2020. Frame-based annotation of multimodal corpora: Tracking (a)synchronies in meaning construction. In *Proceedings of the international framenet workshop 2020: Towards a global, multilingual framenet*, 23–30. Marseille, France: European Language Resources Association. <https://aclanthology.org/2020.framenet-1.4>.
- Bühler, Karl. 1934. *Sprachtheorie: Die darstellungsfunktion der sprache*. Jena: Gustav Fischer.
- Cienki, Alan. 2022. The study of gesture in cognitive linguistics: How it could inform and inspire other research in cognitive science. *Wiley Interdisciplinary Reviews: Cognitive Science* 13(6). e1623.
- Dutra, Lívia & Natália Sigiliano. 2021. Ferramenta linguístico-computacional como facilitadora para o ensino de gramática na escola. In *Anais do xiii simpósio brasileiro de tecnologia da informação e da*

- linguagem humana*, 432–436. Porto Alegre, RS, Brasil: SBC. doi:10.5753/stil.2021.17824. <https://sol.sbc.org.br/index.php/stil/article/view/17824>.
- Fillmore, Charles. 1985. Frames and the semantics of understanding. *Quaderni di semantica* 6. 222–254.
- Fillmore, Charles J. 1975. *Santa cruz lectures on deixis, 1971*, vol. 28. Indiana University Linguistics Club Bloomington.
- Fillmore, Charles J. 1976. Frame semantics and the nature of language. In *Annals of the new york academy of sciences: Conference on the origin and development of language and speech*, vol. 280 1, 20–32. New York.
- Fillmore, Charles J. 1982. Frame semantics. In Linguistic Society of Korea (ed.), *Linguistics in the morning calm*, 111–137. Seoul: Hanshin Publishing Co.
- Fillmore, Charles J., Miriam R.L. Petruck, Josef Ruppenhofer & Abby Wright. 2003. FrameNet in action: The case of attaching. *International Journal of Lexicography* 16(3). 297–332.
- Levinson, Stephen C. 2004. Deixis. In L. R. Horn & G. Ward (eds.), *The handbook of pragmatics*, 97–121. Malden, MA: Blackwell.
- de Saussure, Ferdinand. 1916. *Cours de linguistique générale*. Paris: Payot.
- Segal, Erwin M. 1995. Narrative comprehension and the role of deictic shift theory. In J. F. Duchan, G. A. Bruder & L. E. Hewitt (eds.), *Deixis in narrative: A cognitive science perspective*, 3–17. London: Psychology Press.
- Steen, Francis F, Anders Hougaard, Jungseock Joo, Inés Olza, Cristóbal Pagán Cánovas, Anna Ple-shakova, Soumya Ray, Peter Uhrig, Javier Valenzuela, Jacek Woźny et al. 2018. Toward an infrastructure for data-driven multimodal communication research. *Linguistics Vanguard* 4(1).
- Sweetser, Eve. 2007. Looking at space to study mental spaces: Co-speech gesture as a crucial data source in cognitive linguistics. In M. Gonzalez-Marquez, M. J. Spivey, S. Coulson & I. Mittelberg (eds.), *Methods in cognitive linguistics*, 201–224. Amsterdam: John Benjamins.
- Tomasello, Michael. 2008. *Origins of human communication*. Cambridge, Mass: MIT press.
- Tomasello, Michael. 2018. *Becoming human: A theory of ontogeny*. Cambridge, Mass: Harvard University Press.
- Torrent, Tiago Timponi, Ely Edison da Silva Matos, Frederico Belcavello, Marcelo Viridiano, Maucha Andrade Gamonal, Alexandre Diniz da Costa & Mateus Coutinho Marim. 2022. Representing context in framenet: A multidimensional, multimodal approach. *Frontiers in Psychology* 13. doi:10.3389/fpsyg.2022.838441. <https://www.frontiersin.org/article/10.3389/fpsyg.2022.838441>.
- Viridiano, Marcelo, Tiago Timponi Torrent, Oliver Czulo, Arthur Lorenzi, Ely Matos & Frederico Belcavello. 2022. The case for perspective in multimodal datasets. In *Proceedings of the 1st workshop on perspectivist approaches to nlp @lrec2022*, 108–116. Marseille, France: European Language Resources Association. <https://aclanthology.org/2022.nlperspectives-1.14>.
- Zubin, David A. & Lynne E. Hewitt. 1995. The deictic center: A theory of deixis in narrative. In J. F. Duchan, G. A. Bruder & L. E. Hewitt (eds.), *Deixis in narrative: A cognitive science perspective*, 129–155. London: Psychology Press.