

## Frames with a Vision

Anna Wilson<sup>1</sup>, Irina Pavlova<sup>2</sup> & Peter Uhrig<sup>3</sup>

<sup>1</sup>University of Oxford, [anna.wilson@area.ox.ac.uk](mailto:anna.wilson@area.ox.ac.uk), <sup>2</sup>University of Oxford, <sup>3</sup>University of Dresden

**Keywords:** conceptualisation of the future, frame blending, viewpoint

Much of the frame analysis in linguistics nowadays is based on the English-language Berkeley FrameNet. Although Berkeley FrameNet has grown over years, references to the future need further attention and development. In order to do this successfully we need to address a theoretical question: Is future a temporal perspective / viewpoint, or a temporal frame in its own right, or both? In an attempt to answer this question empirically we first engaged in data-driven analysis of 20 episodes of the RT show SophieCo Visionaries in English from 2020-21. We analysed 47 'future depictions' snippets, each 20 seconds long. These were selected based on corpus searches for the 'auxiliary verb 'will' + hands visible' and the subsequent manual verification of the results returned. We also ran Roth's (2016) PathLSTM on the video transcripts and compared the results of computational processing and manual analysis. Our 'frame activation' analysis of 336 elements – 235 linguistic and 101 gestural 'frame activation' triggers - revealed that both types of units contribute to multimodal communication on semantic and pragmatic levels. From the functions that hand gesture contributes, we selected those directly relevant to the activation of a 'future' frame or the construction of 'future' as a perspective or viewpoint (64 linguistic and 68 gestural). We examined the respective speech-gesture co-occurrences in a case study, while accounting for both future and the future-present-past relation with the structures *going to + back*; *will + back*; *will + return*; *will + was/were*. In analysing those structures, we considered a further 227 snippets from our large RT dataset (RT's SophieCo Visionaries '2019-February 2022' and SophieCo '2013-2022' broadcasts). Our focus on the gestural side was on direction and orientation (e.g. Núñez and Cooperrider 2013; Cooperrider et al. 2014; Valenzuela et al. 2020). We used frame-blending (Fauconnier and Turner 2008) and viewpoint blending (Nikiforidou 2012; Turner 2014) as our core analytical framework. We assessed 'future' as both a distinctive unity and as part of a unified 'future-present-past relation'. We engaged with Minsky 1980, Fauconnier and Turner 2008, Nikiforidou 2012, Turner 2014, and Timponi Torrent et al 2022) to discuss our findings.

Our study revealed the need for the 'future' as a part of FrameNet to be developed in close relation to the development of corresponding parts for past and present and with an appreciation of time conceptualisations as being multimodal frames and viewpoints at the same time. Our findings based on our manual analysis and the computational analysis of the 'linguistic' elements of videos enabled us to understand what aspects of frame analysis can and should be automated, especially for multimodal data. We map possible future directions for doing that.

We developed a dataset composed of video snippets for future, past, and present depictions, which we annotated for respective speech units and gesture, and which will be made available with our paper, once it is published.

This study was made possible thanks to AHRC-DFG research funding.

## References

- Cooperrider, Kensy, Núñez, Rafael and Eve Sweetser 2014. The conceptualization of time in gesture. Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill, and Jana Bressemer (eds.), *Body-language-communication*. Vol. 2, Ch. 134. pp. 1781-1788. Berlin/Boston: De Gruyter Mouton.
- Fauconnier, Gilles & Mark Turner (2008). "Rethinking metaphor." Ray Gibbs (ed.) *Cambridge Handbook of Metaphor and Thought*, 53-66. New York: Cambridge University Press.
- Minsky, Marvin 1980. A Framework for Representing Knowledge. Dieter Metzger (ed.) *Frame Conceptions and Text Understanding*, 1-25. Berlin, New York: Walter de Gruyter.
- Nikiforidou, Kiki 2012. The constructional underpinnings of viewpoint blends: the *Past + now* in language and literature. Barbara Dancygier & Eve Sweetser (eds.) *Viewpoint in language: A multimodal perspective*, 177-197. New York: Cambridge University Press.
- Núñez, Rafael & Kensy Cooperrider 2013. The tangle of space and time in human cognition. *Trends in cognitive sciences* 17.5, 220-229.
- Roth, Michael 2016. Improving frame semantic parsing via dependency path embeddings. *Book of Abstracts of the 9th International Conference on Construction Grammar*. Juiz de Fora, Brazil, 165-167.
- Timponi Torrent, Tiago, 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*, 573.
- Turner, Mark 2014. *The Origin of Ideas: Blending, Creativity, and the Human Spark*. New York: Oxford University Press.
- Valenzuela, Javier, Cristóbal Pagán Cánovas, Inés Olza & Daniel Alcaraz Carrión 2020. Gesturing in the wild: Evidence for a flexible mental timeline. *Review of Cognitive Linguistics. Published under the auspices of the Spanish Cognitive Linguistics Association* 18.2, 289-315.