Is silence a linguistic sign?
Basic Word Order, Silent Pauses, Speech Verbs and Interjecting in the production of Direct Speech

Jocelyn Aznar, `contact@jocelynaznar.eu`

October 2023

Direct speech is discourse figure that has been observed in all documented languages so far, whether it is from past speech communities (for old European narratives, see for instance Louviot 2016) or, among contemporary oral speech communities. The figure, if recurrent across narrative practices, can also be used in various other speech situations, such as conversations (Holt 1996). In this talk, I will argue that silent pauses in certain positions behave grammatically, here not only in the sense that silent pauses can be described according to certain systematically patterns, but also that the production of silent pauses is dependent on the linguistic structures associated with the languages, that is here, basic word order.

Building upon previous research on silent pauses as prosodic cues for marking the beginning or the end of direct speech (Genetti 2011; Malibert and Vanhove 2015) and on interjection (Norrick 2015;Aznar 2019, pp. 175–80;Nikitina, Aplonova, and Roa 2023), this presentation presents findings from two cross-linguistic studies conducted using language documentation corpora. The first study, in collaboration with Seifart[1] (in prep.), focuses on establishing correlations between the production of silent pauses and speech verbs and the basic word order. The second study, conducted with Aplonova[2] (in prep.), delves into interjections and interjection phrases.

Both investigations involve an examination of direct speech sequences in monologic narratives. The first study spans twelve languages from the DoReCo dataset (Paschen et al. 2020) to explore silent pauses and speech verbs, while the second study extends across twenty language corpora, focusing on what with I term with Aplonova "interjecting" (additional corpora from the Discourse Reporting database were incorporated; Nikitina, Aplonova, Jordanoska, et al. 2022)[3]. These studies heavily rely on automated analysis, utilizing tools such as Multitool[4] for compiling annotation files and Jupyter notebooks (Kluyver et al.

---

[1]See `https://frankseifart.info/`.

[2]See `https://isidore.science/a/aplonova_katya`.

[3]The language documentation corpora retained for the study were selected for their variety in terms of genetic, areal and structural properties.

[4]`https://github.com/DoReCo/multitool`

2016) to organize and structure the automated processing into a shareable set of documents.

The presentation will commence by providing context for this research, detailing the initiatives QUEST and DoReCo. This will be followed by an overview of the distinct methodologies employed in the two investigations. Concrete examples will be presented to explain how the direct speech sequence was conceptualized and annotated. I will then present the results of the study with Seifart, on the patterns of silent pauses production and speech verb positions within the direct speech sequence and according to each language and their associated basic word order. For the second study, made in collaboration with Aplonova, I will describe the data that led us to propose a cross-linguistic discourse device that we labelled "interjecting". The term "interjecting" is used deliberately to emphasize that the use of this discourse device is not limited to interjections as a grammatical category, but corresponds to the coordination by a speaker of a silent pause and an interjection phrase to produce a mismatch within the expected coordination between the syntactical structure and the prosodical one (Rochet-Capellan and Fuchs 2013; Spronck and Nikitina 2019) to mark out the beginning of direct speech. To conclude, drawing on the results of both cross-linguistic investigations, I will argue that silent pauses can be, in certain contexts, regarded as linguistic signs.

1 Acknowledgment

References

Aznar, Jocelyn (2019). "Narrer une nabol : La production des textes nisvais en fonction de l'âge et de la situation d'énonciation, Malekula, Vanuatu". Marseille: EHESS.

Genetti, Carol (2011). "Direct Speech Reports and the Cline of Prosodic Integration in Dolakha Newar". In: *Himalayan Linguistics* 10.1. ISSN: 1544-7502. DOI: 10/ghfbmz. URL: https://escholarship.org/uc/item/5h96w08q (visited on 10/11/2020).

Holt, Elizabeth (July 1996). "Reporting on Talk: The Use of Direct Reported Speech in Conversation". In: *Research on Language & Social Interaction* 29.3, pp. 219–245. ISSN: 0835-1813, 1532-7973. DOI: 10/bphbd3. URL: 10.1207/s15327973rlsi2903_2 (visited on 02/26/2021).

Kluyver, Thomas et al. (2016). "Jupyter Notebooks—a Publishing Format for Reproducible Computational Workflows". In: p. 4. DOI: 10.3233/978-1-61499-649-1-87.

Louviot, Élise (2016). *Direct Speech in Beowulf and Other Old English Narrative Poems.* Anglo-Saxon Studies 30. Cambridge: D.S. Brewer. 285 pp. ISBN: 978-1-84384-434-1.

Malibert, Il-Il and Martine Vanhove (2015). "Quotative constructions and prosody in some Afroasiatic languages". In: *Corpus-based Studies of Lesser-described Languages: The CorpAfroAs corpus of spoken AfroAsiatic languages* 68, p. 117.

Nikitina, Tatiana, Ekaterina Aplonova, Abbie Jordanoska Izabela andHantgan-Sonko, et al. (2022). *The SpeechReporting Corpus: Discourse Reporting in Storytelling.* URL: http://discoursereporting.huma-num.fr/index.html.

Nikitina, Tatiana, Ekaterina Aplonova, and Leonardo Contreras Roa (2023). "The use of interjections as a discourse phenomenon". In: *Discourse Phenomena in Typological Perspective* 227, p. 65.

Norrick, Neal R. (2015). "Interjections". In: *Corpus Pragmatics.* Cambridge: Cambridge University Press. URL: www.cambridge.org/9781107015043.

Paschen, Ludger et al. (2020). "Building a Time-Aligned Cross-Linguistic Reference Corpus from Language Documentation Data (DoReCo)". In: *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)* 12, pp. 2657–2666.

Rochet-Capellan, Amélie and Susanne Fuchs (2013). "The Interplay of Linguistic Structure and Breathing in German Spontaneous Speech". In: *14th Annual Conference of the International Speech Communication Association (Interspeech 2013)*, p. 1228. URL: https://hal.archives-ouvertes.fr/hal-00871600/ (visited on 01/04/2017).

Spronck, Stef and Tatiana Nikitina (2019). "Reported speech forms a dedicated syntactic domain". In: *Linguistic Typology* 23, pp. 119–159.