

## Variation in native speaker transcription

Discourse transcription is often treated as an objective part of the research process, yet it requires interpretive and analytical decisions by transcribers (Himmelman 2018: 35). Choices such as the level of detail or the inclusion of multimodal features depend heavily on the research question or documentation goal at hand (Du Bois et al. 1992: 4; Edwards 1993: 12). Scholars have long called for more methodological reflections (e.g. Ochs 1979; Bucholtz 2000).

Previous work (e.g. Dobrin 2021; Jung & Himmelman 2011; Urban 1996) shows that native speakers often prioritise meaning over form. Rather than transcribing exactly what was said in a recording, they may paraphrase, omit code-switching and repetitions, or add content they deem relevant.

This talk focuses on morphosyntactic variation in native speaker transcripts. It will present the first results from an ongoing project conducting transcription experiments. Eight Bachelor linguistics students and (near-)native speakers of Catalan transcribed a pear story (monologue), which was already segmented into intonation units for better comparability. Despite the controlled conditions, substantial variation emerged, especially in word segmentation (e.g. articles):

(1)

- |     |     |                 |             |          |      |      |
|-----|-----|-----------------|-------------|----------|------|------|
| (a) | veu | la              | oportunitat | d'agafar | una  | pera |
| (b) | veu | la              | oportunitat | de gafar | una  | pera |
| (c) | veu | l'oportunitat   | d'agafar    | una      | pera |      |
| (d) | veu | l'oportunitat   | de gafar    | una      | pera |      |
|     | see | ART-opportunity | to-take     | a        | pear |      |

‘(he) sees the opportunity to take a pear’

Similar patterns emerge in preliminary German data, suggesting that even in WEIRD and LOL languages with established writing systems and traditions that may influence transcribers, substantial morphosyntactic variation can be found. The potential influence of this variation on subsequent linguistic analyses are explored in this talk.

This contribution thus argues that transcript variation is not just noise in the data but a valuable source of metalinguistic knowledge of native speakers. It calls for greater transparency in transcription practices, including documenting transcription processes and incorporating detailed accounts of transcription choices into typological research.

## References

- Bucholtz, Mary. 2000. The politics of transcription. *Journal of Pragmatics* 32(10). 1439–1465.  
[https://doi.org/10.1016/S0378-2166\(99\)00094-6](https://doi.org/10.1016/S0378-2166(99)00094-6).
- Dobrin, Lise M. 2021. The Arapesh “suitcase miracle”: The interpretive value of reproducible research. *Language Documentation and Description* 21. 37–69.  
<https://doi.org/10.25894/ldd16>.
- Du Bois, John W., Susanna Cumming, Stephan Schuetze-Coburn & Danae Paolino (eds.). 1992. *Discourse transcription* (Santa Barbara Papers in Linguistics 4). University of California, Santa Barbara.
- Edwards, Jane A. 1993. Principles and contrasting systems of discourse transcription. In Jane A. Edwards & Martin D. Lampert (eds.), *Talking data: Transcription and coding in discourse research*, 3–31. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Himmelman, Nikolaus P. 2018. Meeting the transcription challenge. In Bradley McDonnell, Andrea L. Berez-Kroeker & Gary Holton (eds.), *Reflections on Language Documentation 20 Years after Himmelman 1998* (Language Documentation & Conservation Special Publication 15), 33–40. Honolulu: University of Hawaii Press.  
<http://hdl.handle.net/10125/24806>.
- Jung, Dagmar & Nikolaus P. Himmelman. 2011. Retelling data: Working on transcription. In Geoffrey L. J. Haig, Nicole Nau, Stefan Schnell & Claudia Wegener (eds.), *Documenting Endangered Languages: Achievements and Perspectives*. Berlin/Boston: De Gruyter Mouton.
- Ochs, Elinor. 1979. Transcription as theory. In Elinor Ochs & Bambi B. Schieffelin (eds.), *Developmental Pragmatics*, 43–72. New York: Academic Press.
- Urban, Greg. 1996. Entextualization, replication, and power. In Michael Silverstein & Greg Urban (eds.), *Natural Histories of Discourse*, 21–44. Chicago: University of Chicago Press.