

Impossible boundaries? On broken words and prosodic phrasing

Ludger Paschen
(Leibniz-Institut Allgemeine Sprachwissenschaft)

Keywords: pause, prosodic word, syntax-prosody interface, corpus phonetics

Pauses are an integral part of human speech and function as one cue for boundaries between major higher prosodic units such as Intonational Phrases. Words, the basic building blocks of language, may display a certain degree of internal prosodic complexity (such as layered prosodic words), but are widely considered to be inseparable by major prosodic boundaries. A notable exception was reported by Evans et al. (2008), who noted that certain prefixes in Dalabon may occur as independent units separated from the rest of the word by a silent pause. Mansfield (2023) reports that pauses may also occur between stem elements in careful speech in other Australian languages. In this talk, I will investigate the occurrence and properties of word-internal pauses (WIP) from a broader perspective, building on data from 53 small and endangered languages from the DoReCo corpus (Seifart et al. 2024). The corpus contains transcribed speech data with annotations of various special speech events, including WIP, comprising a total of ca. 500,000 time-aligned word tokens. The survey reveals that WIP are more cross-linguistically common than previously thought, with highly synthetic languages having more WIP than isolating ones. In the corpus data, a total of over 700 WIP are attested, and WIP appear in more diverse contexts compared to the ones described in Evans et al. (2008) and Mansfield (2023), including situations where WIP break up morphologically atomic units (roots, affixes). Some WIP are accompanied by speech disfluencies such as hesitations or false starts. WIP do not necessarily constitute prosodic word boundaries, as some languages show evidence for word-level phonological rules applying across a WIP. However, in cases where the WIP coincides with a grammatical boundary, the chunks separated by WIP often behave as independent prosodic domains. The wide-spread occurrence of WIP across languages sheds light on the prosodic make-up of complex word forms, and highlights the importance of speech planning to account for prosodic phenomena encountered in corpus data representing natural and spontaneous speech.

References

- Evans, Nicholas, Janet Fletcher & Belinda Ross (2008), Big words, small phrases: Mismatches between pause units and the polysynthetic word in Dalabon, *Linguistics* 46(1), 89-129.
- Mansfield, J (2023), The prosodic structure of Australian polysynthetic verbs. In: Ksenia Bogomolets and Harry van der Hulst (eds), *Word Prominence in Languages with Complex Morphologies*. Oxford: OUP.
- Seifart, Frank, Ludger Paschen & Matthew Stave (eds), (2024), *Language Documentation Reference Corpus (DoReCo) 1.3*. Lyon: laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2).