

Goal: Make recordings searchable for phonetic analysis

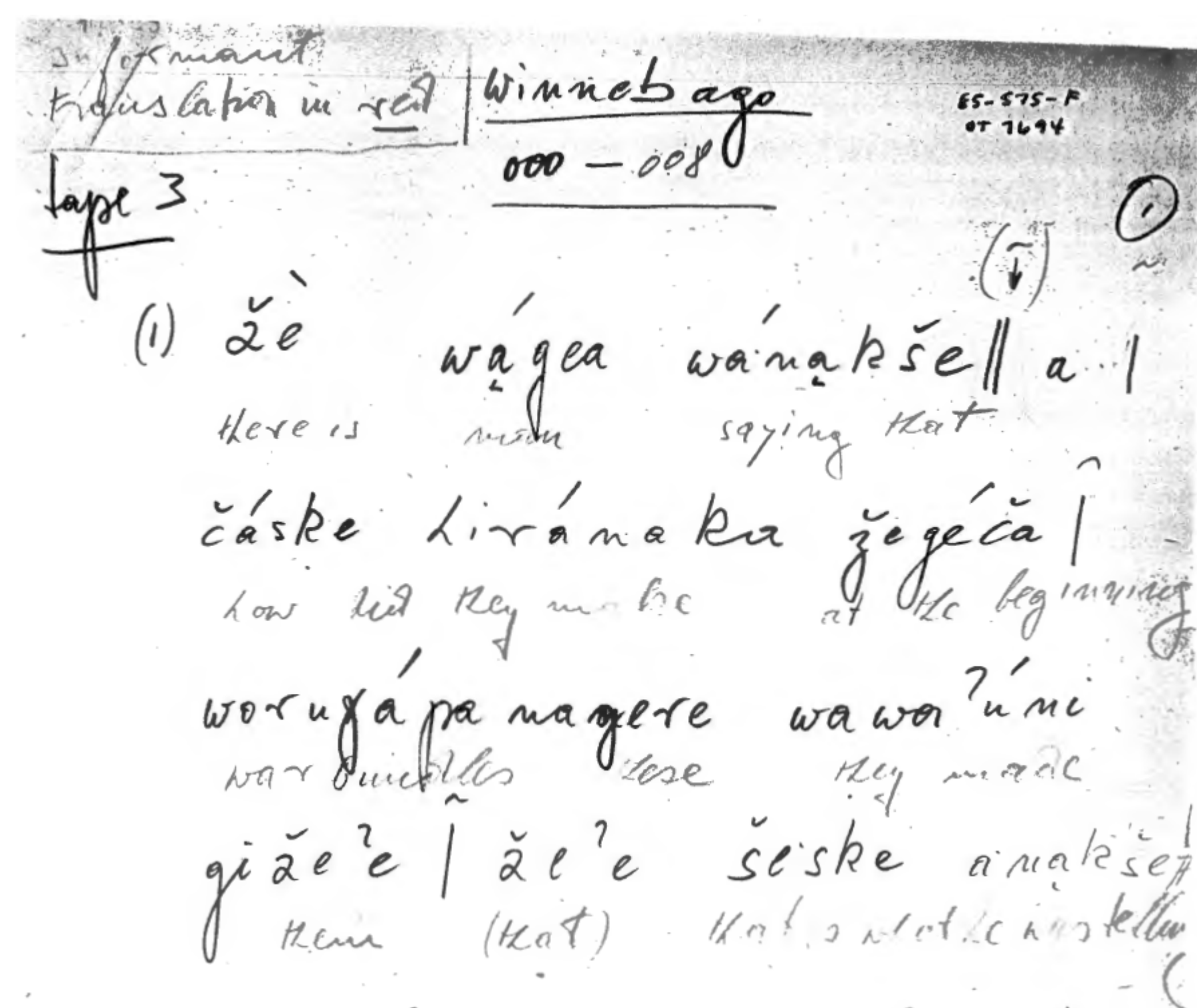
Two Hocank corpora

The Miner Corpus (1974-1975)

- ~28 hours of elicitation sessions with linguist Kenneth Miner
- ~7800 utterances
 - Mostly words in isolation
 - ~120 verb paradigms (often partial)
 - Most speech by Lavina Thorud (Maaxiánažiwiga)
 - ~400 utterances by Robinson Johnson (Xiigugá)
- Miner's handwritten field notebooks
- Wisconsin Native American Languages Project, digitized by University of Wisconsin-Milwaukee Libraries
- Annotation project: summer 2017-present

The Fraenkel Corpus (1959)

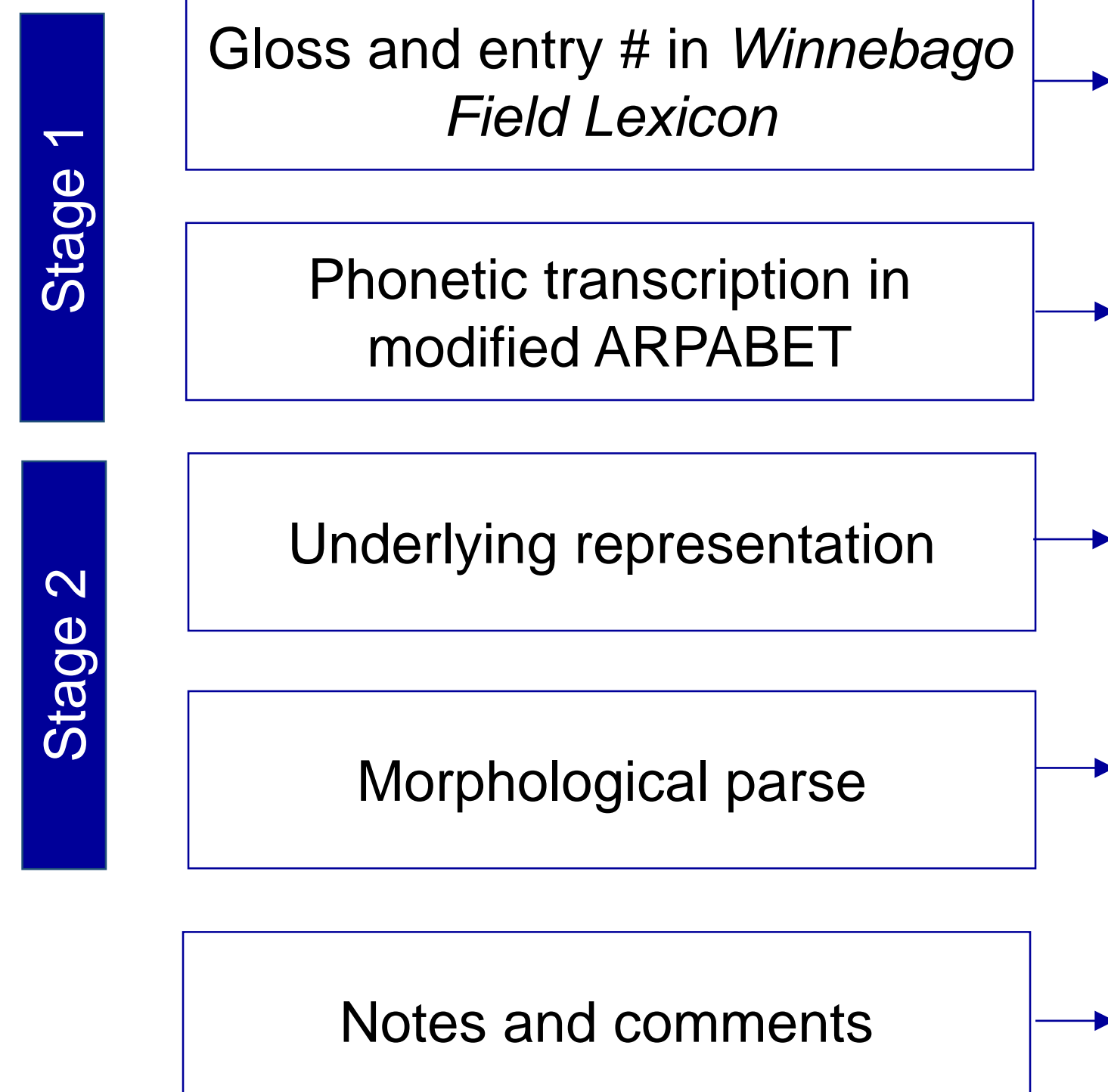
- 24 tapes ~30 minutes each; archived at Indiana U; 15 available digitally
- Recorded in Wisconsin
- Some (not all) accompanied by handwritten field notes
- Little known about linguist Gerd Fraenkel, who apparently never published on Hocank
- Speakers: Chief Daybreak Elias James Smoke, Stella Stacy
- Varied material
 - Retelling of stories previously recorded on wax cylinders
 - Words and phrases
 - Much is apparently a re-elicitation of data from Amelia Susman, Hans Wolff, and Paul Radin
- Annotation project is in early stages



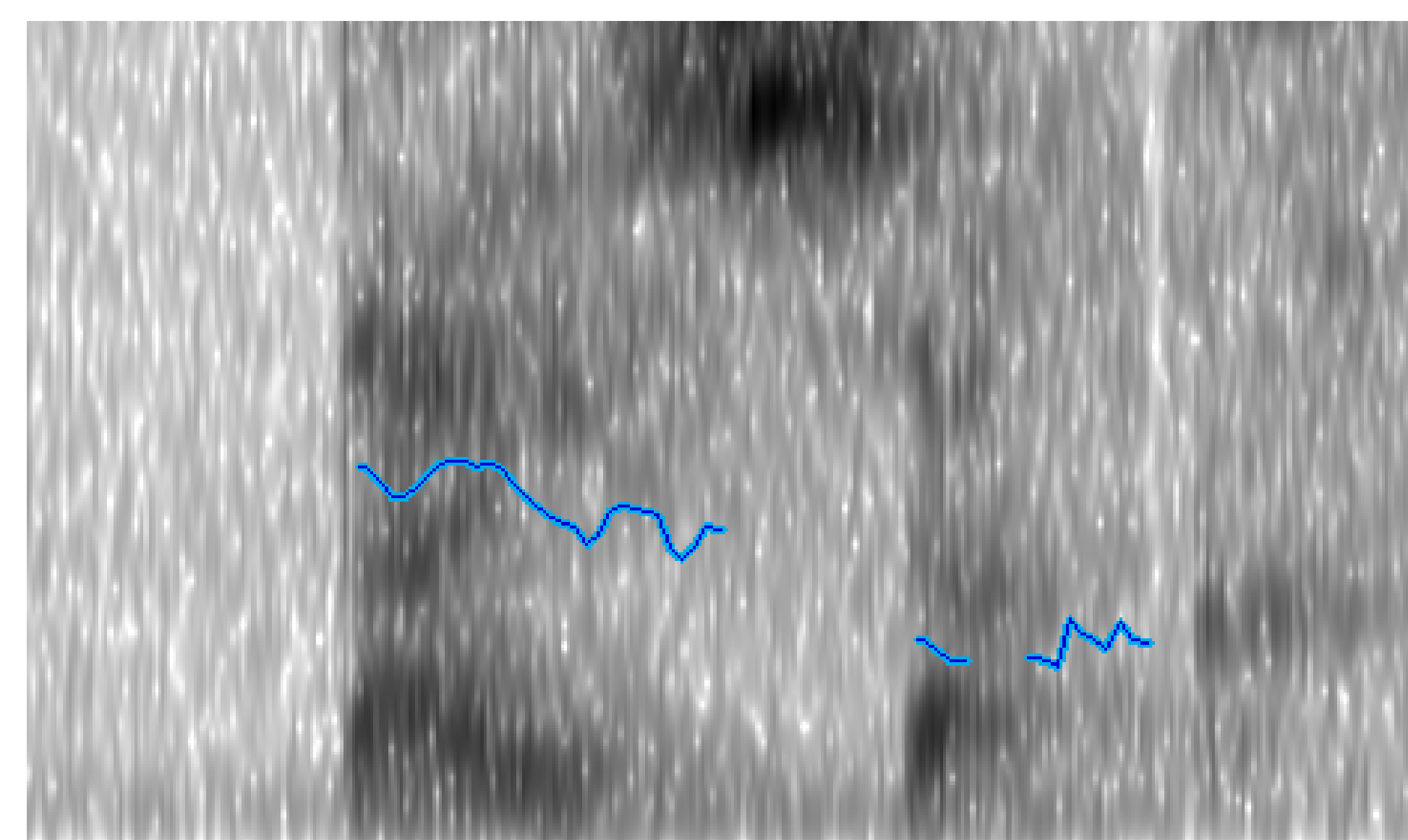
Sample of Fraenkel's field notes

9 annotators have contributed. Annotations are based on:

- Ear
- Field notes
- Miner's *Winnebago Field Lexicon*
- Helmbrecht & Lehmann's 2010 dictionary



Procedure



3231 I am cold
T AIN S A K
T AA - H IN - S AA K
by.heat-1E.U-ROOT
not sure I hear vowel nasalization

Issues in Annotation

Uncertain transcriptions

Field notes and references may disagree on:

- Vowel length
- Vowel nasalization
- Location of accent

Morphological parsing is leading us to revise some transcriptions (like adding nasalization in the example above).

Why ARPABET?

This makes it easy to search by phonological characteristics, e.g.:

- Syllable weight (relevant to accent)
- Nasalized vowels

ARPABET can be automatically transliterated to IPA.

Acknowledgements Thanks to Stacey Jacobson, Miles Haisley, Molly Rosenfeld, Rodolfo Ponce, Andie Niederecker and Joshua Lester for contributing to the annotation of these recordings. This research was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Numbers 8UL1GM118979-02, 8TL4GM118980-02, 8RL5GM118978-02. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References [1] Helmbrecht, J., & Lehmann, C. (Eds.). (2010). *Hocak Teaching Materials, Volume 1: Elements of Grammar/Learner's Dictionary* (Vol. 1). SUNY Press. [2] Miner, K. (1979). Dorsey's law in Winnebago-Chiwere and Winnebago accent. *IJAL* 45(1), 25-33. [3] Susman, A. (1943). *The accentual system of Winnebago*. Columbia University.

Ongoing phonetic studies

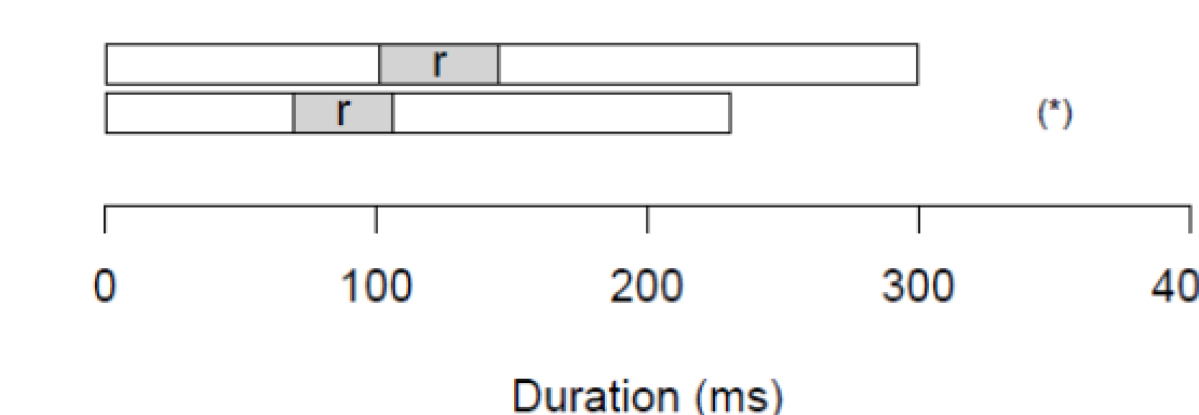
Epenthetic V durations

Phenomenon

Miner 1979 and Susman 1943 report impressionistically that CVCV sequences where V₁ is epenthetic are shorter than other disyllables.

Preliminary findings

This seems to be correct. Lexical VCV sequences (top) tend to be longer than VCV sequences where V₁ is epenthetic (bottom), when matched for context.



Secondary accent on epenthetic Vs?

Phenomenon

Disyllables have final accent. Miner 1979 also transcribes secondary accent on initial epenthetic vowels: /kre/ → kèrè 'depart returning'.

He expresses uncertainty about the secondary accent in a footnote, and omits it in later work.

Preliminary findings

Epenthetic and non-epenthetic initial vowels have similar pitch contours when normalized for duration; we have not found evidence of secondary accent.

