

# БОЛЬШАЯ ЛИНГВИСТИКА

В малом языке:

## локальная структура верхнекускоквимского дискурса

**А.А. Кибрик**

(ИЯз РАН и МГУ имени М.В. Ломоносова)

[aakibrik@gmail.com](mailto:aakibrik@gmail.com)



Российский  
научный  
фонд

МЯБЛ-2

Project #17-18-01649

МГУ, 18 апреля 2020

# Language description / documentation



- A “collection of texts” is a part of the gold standard
- What is “text”?
- Note specialized studies of discourse transcription in field work, e.g. Jung and Himmelmann 2011
- However, in many descriptive and documentary studies the view of “text” is rather naïve

# Some of the naïve assumptions



- Speakers create texts, that are subsequently registered by linguists
  - In fact: text  $\neq$  communication
- There is no crucial difference between spoken and written language
  - In fact: minor languages are mostly unwritten, and there is profound difference between talk and writing
- Identification of sentences in texts is self-evident
  - In fact: no punctuation marks in talk
- Pausing, prosody, disfluencies, etc., can be ignored or registered in an ad hoc way
  - In fact: all this abounds in talk and interacts strongly with the segmental material

# As a result



- Published texts in less studied languages appear as poorly organized chains of words and expressions
- in which much of the original signal is lost
- while various artifacts are introduced

# In this paper I propose that



- Knowledge from discourse theory can be applied to less studied languages
- That enhances the **ecological validity** of transcribed data radically,
- particularly regarding the local discourse structure
- and basic prosodic phenomena.

# Plan of talk

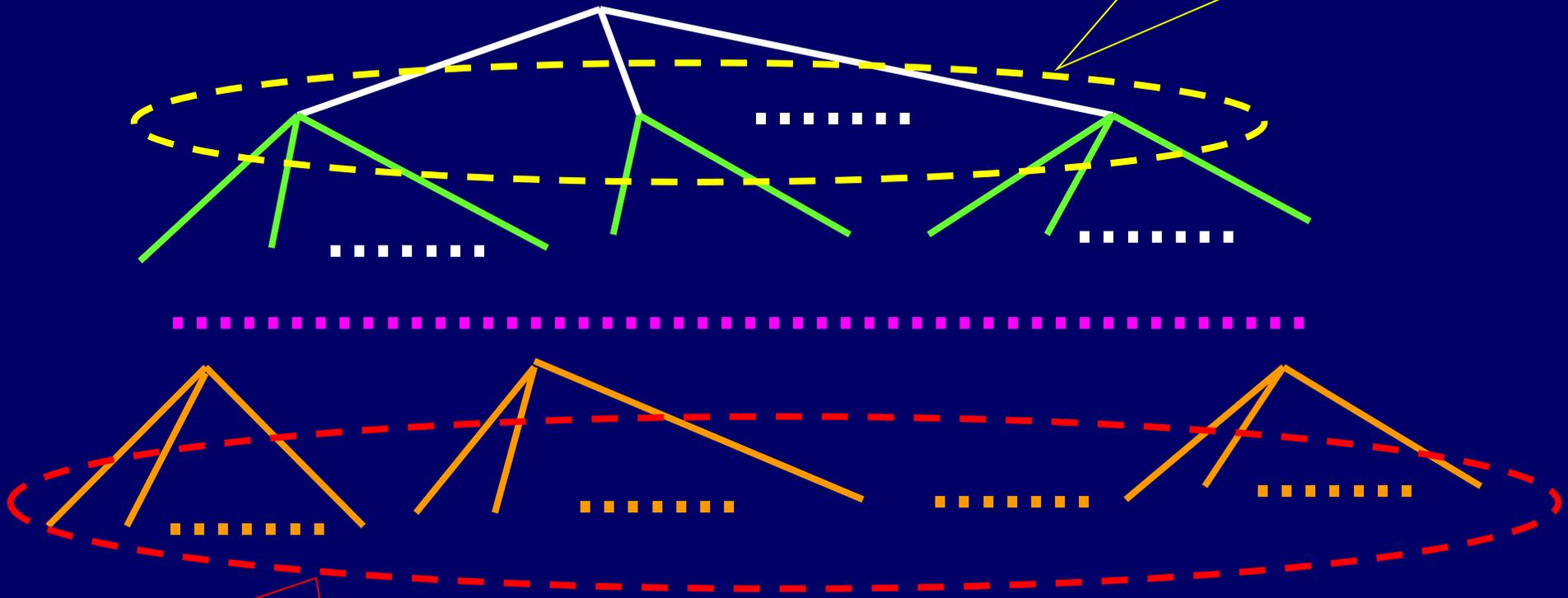


- Elements of discourse theory
- Application to documentary work on Upper Kuskokwim (Athabaskan, Alaska)
- Conclusions

# Hierarchical discourse structure

Discourse

global structure



local structure



# Crucial role of prosody

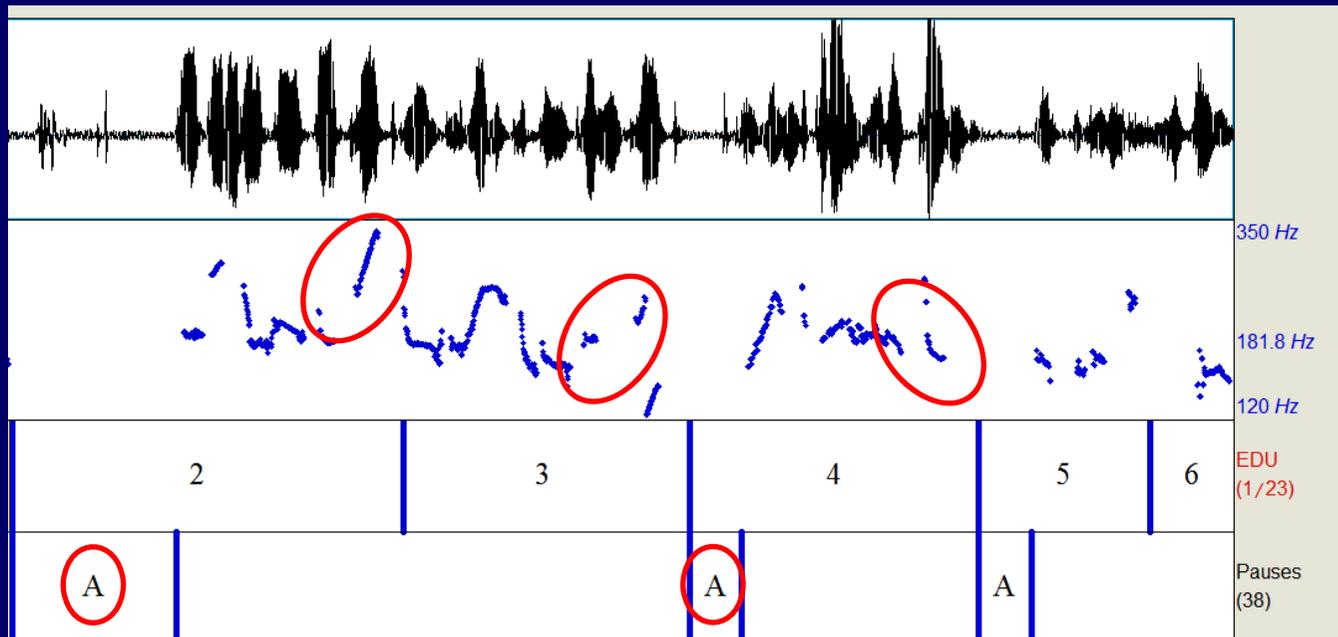


- EDUs are identified on the basis of prosodic criteria:
  - Pausing (inhalation)
  - Integral tonal contour
  - Primary accent (accentual center)
  - Tempo pattern
  - Loudness pattern



# Example: EDUs #2—4

- Two out of three have boundary pauses
- Each EDU has a typical intonation contour
  - which is primary
  - placed on the informational center (rheme)
  - located at the end



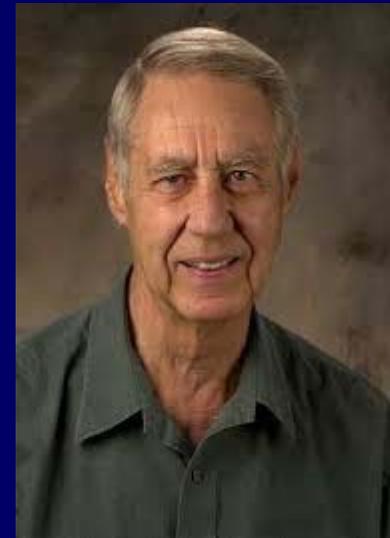
# Example: EDUs #2—4

- Tempo pattern

EDU #	Initial part		Final part	
	Total	Per syllable	Total	Per syllable
2	0.47	0.09	0.73	0.18
3	0.66	0.09	0.85	0.17
4	0.58	0.12	0.66	0.17

# Content of EDUs: general

- EDUs appear to not only be prosodic units, but display unity in terms of their content (Chafe 1994)
  - Cognitively: represent one focus of consciousness
  - Semantically: represent one event/state
  - Grammatically: correlate with a clause



Wallace Chafe  
1927 – 2019

# EDUs and grammatical constituents in various languages

- Prosodically identified EDUs: correlation with clauses in a Russian corpus
- EDU = clause 68%
- EDU < clause (subclausal) 26%
- EDU > clause (superclausal) 6%

Language	Percentage of clausal EDUs
English (Chafe 1994)	60%
Mandarin (Iwasaki and Tao 1993)	39.8%
Sasak (Wouk 2008)	51.7%
Japanese (Matsumoto 2000)	68%

# Well studied and less studied languages



- Expert judgment
- High convergence between trained experts
- Identified purely on prosodic basis or analysts sneak in semantico-syntactic considerations?
- How much prosodic knowledge about a language one must have to tackle this goal?
- To what extent various transcribers converge in how they identify EDUs on the basis of prosody alone?
- A quasi-experiment with an unfamiliar language
  - Q'eqchi' (Mayan, Guatemala)
  - EDU identification is accessible even to minimally trained analysts
  - EDU identification is in principle independent from propositional content
  - There are near-universal prosodic markers of EDU segmentation

# Upper Kuskokwim language

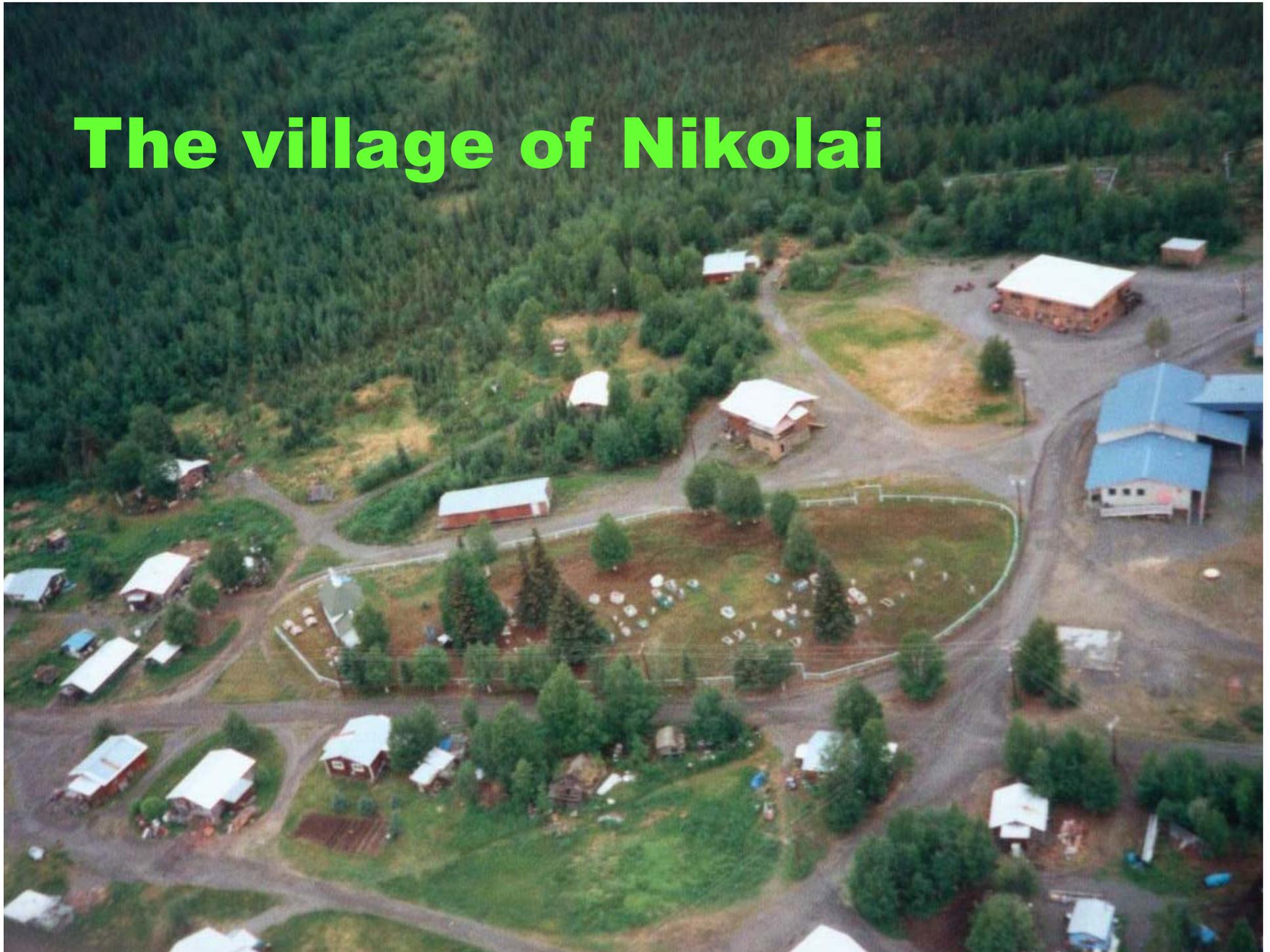


- Athabaskan, central Alaska
- Moribund language
  - Rapid language shift in the 1960s
  - No actual use of UKA (up to four households till 2014)
  - The youngest fully-fledged speaker born in 1952
  - Less than 15 speakers left out of the population of about 450
- Highly complex verb morphology
  - polysynthesis
  - qualitative complexity
- Prior work – Collins and Petruska 1979
- Eight field trips between 1997 and 2019

# Native languages of Alaska (M. Krauss)



# The village of Nikolai



# Available discourse data



## ■ Spoken

- Several dozen recordings by Ray Collins and his aides in 1960s – 1990s
- Over 70 recordings by myself in 1997 – 2019
- By genre
  - Folklore stories
  - Ethnic history
  - Personal stories
  - Comments on a film
  - Conversation
  - Interview
  - Sermon

## ■ Written

- Several dozen school books published in 1970s—1990s
- A couple of personal letters
- One piece of fiction

# The Icon story by Bobby Esai (1919-2014)



# The Icon story

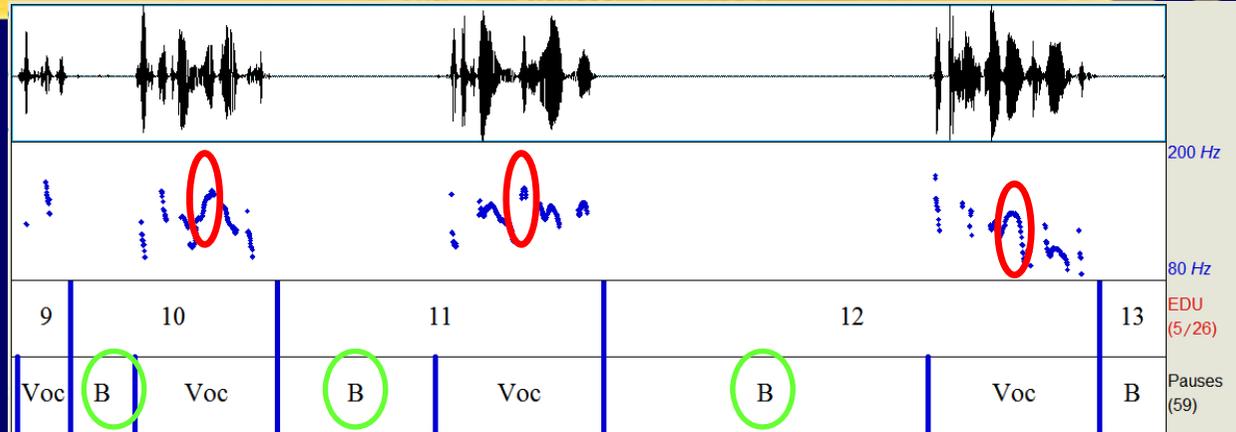


- 23 minutes long
- Summary (beginning)
  - I was hunting in the fall in the Medfra area and I saw many fur animal tracks. I went back home and told my parents I was going to set traps. They told me not to do that (because it did not freeze up yet and the furs were no good). But I still went there with a dog team. As I was riding, there was little snow and very tall grass, so **I took my dogs to the ice of the river. Suddenly my sled broke through ice. I started sinking too. I took an icon from my bosom and prayed.**

# EDU identification (local discourse structure)



- Normal procedure
  - boundary pauses
  - primary accents
  - tone contours



## Peculiarities

- longer boundary pauses
- deceleration effect is unstable

## EXAMPLE 1

EDU #	Initial part		Final part	
	Total	Per syllable	Total	Per syllable
10	0.56	0.28	1.25	0.25
11	0.74	0.37	1.24	0.31
12	0.66	0.22	1.5	0.3

# EDU types in Upper Kuskokwim



- Majority: clausal (71% in a corpus study)
- Subclausal (15%)
- Superclausal (14%)
  - Twice as frequent as in Russian (14% vs. 6.6%)
  - Possible explanation: meaning is more tightly packed in words, and as a result more words fit into an EDU, including predicative words
- Cf. Markus 2009

# Spoken sentences

- Sentences as groups of EDUs are identifiable in spoken discourse (Kibrik 2008)
- Sentence is based on *[interclausal] phase* (trans. continuity)
  - Abstract discourse-semantic category (Sandro V. Kodzasov)
  - conclusion: final falling in accent (period intonation)
  - expectation of continuation: rising or non-final falling in accent (comma intonation)
- But sentence identification is only possible through a rather sophisticated analytic procedure
  - based on **prosodic criteria**
  - involving a **prosodic portrait** of a specific speaker
- The notion of sentence, taken for granted by some linguistic theories (restricted to the verbal component alone), is ultimately prosody-dependent



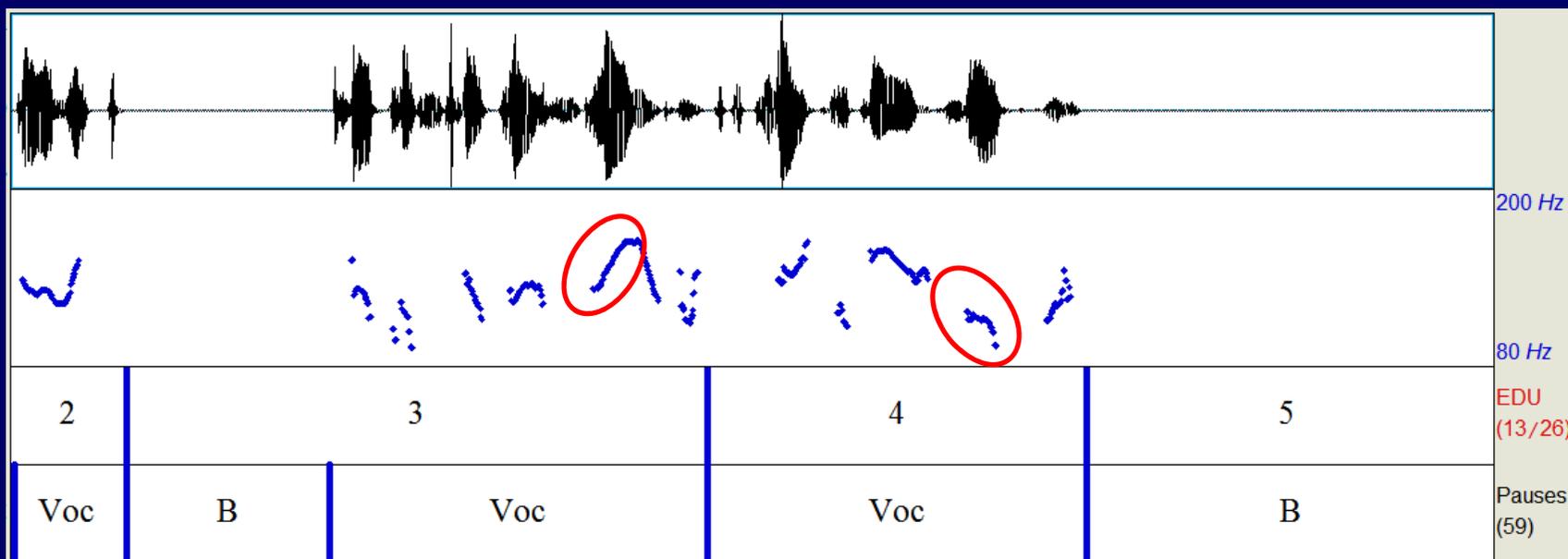
Sandro V. Kodzasov  
1938 – 2014

# Phase and pitch accents

- Main pattern

- non-final: rising #3
- final: falling #4

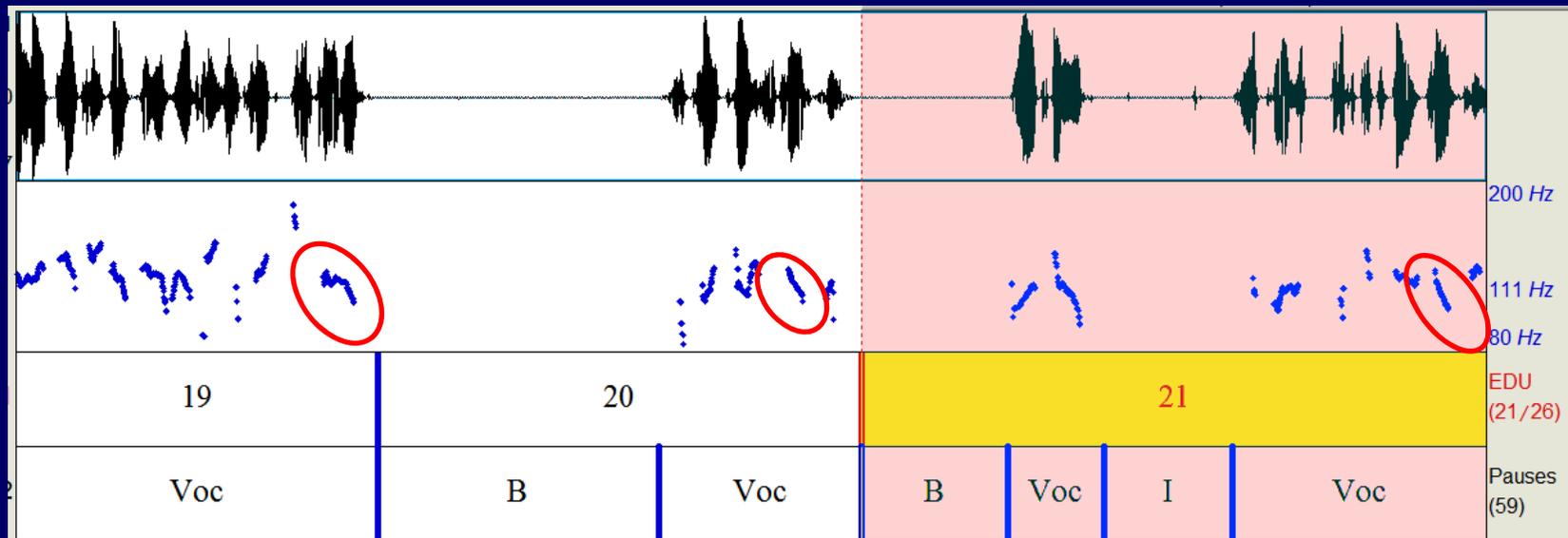
## EXAMPLE 2



# Final and non-final falling

- Final falling: to about 100 Hz or lower
  - E.g. #21
- Frequent non-final falling to about 110 Hz or higher
  - #19, 20, 22, 23

**EXAMPLE 3**



- Consistent contrast allowing to posit clear sentence boundaries

# Conclusions



- EDUs as universal building blocks of local discourse structure are perfectly well identifiable in a moribund and polysynthetic language
- Prosody serving local discourse structure is comparable across very different languages
- and can be represented with the help of punctuation marks
- Spoken sentences are also identifiable
- For more detail, see Kibrik 2020, “Anthropological linguistics”

# More general conclusions



- Transcribing discourse of less-studied languages in a theoretically informed way is perfectly possible
- That takes little training
- and brings radically better and more empirically and ecologically valid results
- Theory of discourse is a very practical tool for descriptive / documentary linguists
- In turn, theoretically informed work on various languages leads to a typology of discourse structures and thus enriches theory



*To the memory of*  
Bobby Esai Sr.  
Philip Esai  
Nick Dennis  
Willie and Betty  
Petruska  
Jim Nikolai

