*Extended version of talk at the Dan Everett Festival*

# *Cohesive Rhythms*

## *Choral Narrative in Ega*

http://wwwhomes.uni-bielefeld.de/gibbon/EGA/

Dafydd Gibbon

Bielefeld University

Göttingen University, 22 June 2023

# *Overview*

1. Methods for analysing rhythms
    1. Inducing rhythm from speech
    2. Deducing rhythm from grammar
2. The ubiquity of rhythms
    1. Rhythms, melodies and the modulation of speech
    2. AM and FM modulation
3. Ega
    1. Geographically and typologically
    2. Phonology, tonology: does Ega have downstep? Maybe!
4. Rhythms in Ega choral narrative
    1. Rhythms in the time domain and the frequency domain
    2. Demodulation of long-term speech events
5. Conclusion

# *Why rhythms of orature?*

Beyond ‘linguistic rhythms’ derived from grammar:

- Partially reversible numerical tree encodings
- Qualitative empical + formal methodology
- Deductive, just one rhythm:
  - Either non-relational binary categories, e.g. [+/- stress]
  - Or *iambic*, *trochaic* metaphors:

    Where are the *dactyls*, *anapaests, spondees, pyrrhics*?

    How about *amphibrachs* and Jassem’s *anacrusis* theory?
- Phonetic correlates of rhythms remain undefined

Induction of ‘realtime rhythms’ from real speech registers:

- Quantitative, physical signal measurement and processing
- Rhythm as *beats/cycles* with *frequencies, magnitudes, ...*

  Plenty of *boom-de-boom-boom!*
- Variability of rhythm patterns, parallel ‘rhythms of rhythm’
- Variability of linguistic correlates
- Depends on contextual style, register, not only grammar:

  e.g. ***reading*** → *formal* → ***orature*** → *informal spontaneous*

# *Deduction of 'linguistic rhythms' from grammar?*

Data:

- normative 'educated registers', not extempore speech
- metalinguistic quoted (written, transcribed) data

Method:

- Individual expert opinion, hermeneutic interpretation
- qualitative judgment and analysis
- data comparison with algebraic and set-theoretic models

Modelling conventions:

- rhythm as relations between nodes in structures:
  - (partly reversible) numerical tree encodings
- binarity constraints on
  - categories and relations: [+/- stress], [s, w] / [w, s]
  - sequences: 'iambic', 'trochaic', 'metrical' metaphors
- physical grounding (correlates) not essential

# *Induction of 'real rhythms' from recorded orature*

Data:

- well-defined speech registers
- reproducible realtime recordings

Method:

- quantitative, physical measurement
- signal processing
- data comparison with hierarchical clustering

Modelling conventions:

- rhythm as *beats/cycles* with *frequencies, magnitudes, …*
- systematic variability of rhythm patterns
- multiple parallel 'rhythms of rhythm'
- physical grounding for linguistic correlates
- dependence on contextual style, register:

  ***reading*** → ***orature*** → *…*

# *Deduction of 'linguistic rhythms' from grammar?*

Very successful 'clear case' methodology, but ...

1. Too much reliance on metaphors rather than models:
   - Where are the dactyls, anapaests, spondees, pyrrhics?
   - How about amphibrachs and Jassem's anacrusis theory?
   - Mutual citation bubbles (*aka* paradigms) prevail.
   - Individual intuitive judgments are opinions rather than facts.
2. Timing, including 'real rhythms', involve many factors, e.g.:
   1. Phonotactics: syllable structure
   2. Morphotactics: word structure
   3. Phrase and sentence structure
   4. Information structure
   5. Dialogue structure
3. And, to misquote a line by Fred Cummins:

   **Where's the 'boom-de-boom-boom' in numerical tree encodings?**

# *Why induce rhythms from orature?*

Data:

- well-defined speech registers
- reproducible realtime recordings

Meth

- 
- 
- 

Mod

- 
- systematic variability of rhythm patterns
- multiple parallel ‘rhythms of rhythm’
- physical grounding for linguistic correlates
- dependence on contextual style, register:

  ***reading*** → ***orature*** → *…*

So let’s take a look at ‘realtime rhythms’.

And where better to start than West Africa?

*After all, that’s where prosodic, autosegmental, metrical phonologies came from –*
*Firth, …, Leben, Goldsmith, Clements, ...*

.

# *Rhythms are ubiquitous...*

**Phonetician**
**Phonologian**

**Radio engineer**

**Musician**

# *Rhythms are ubiquitous...*

**Phonetician**
**Phonologian**

**Radio engineer**

*... Prosody*

*intonation*
*pitch accent*
*tone*

*sonority*
*curve etc.*

**Musician**

# *Rhythms are ubiquitous...*

**Phonetician**
**Phonologian**

**Radio engineer**

*... Prosody*

*intonation*
*pitch accent*
*tone*

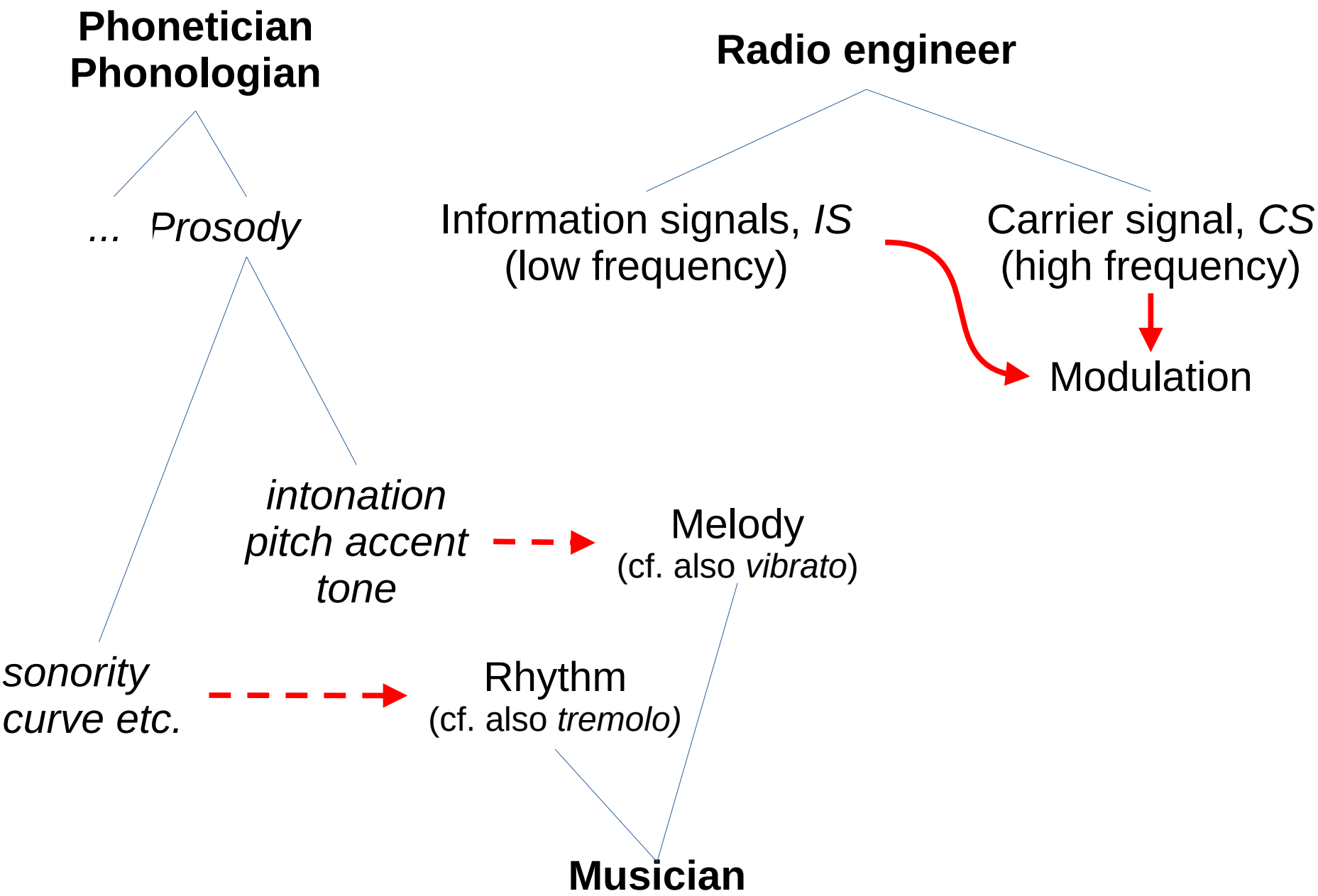
Melody
(cf. also *vibrato*)

*sonority*
*curve etc.*

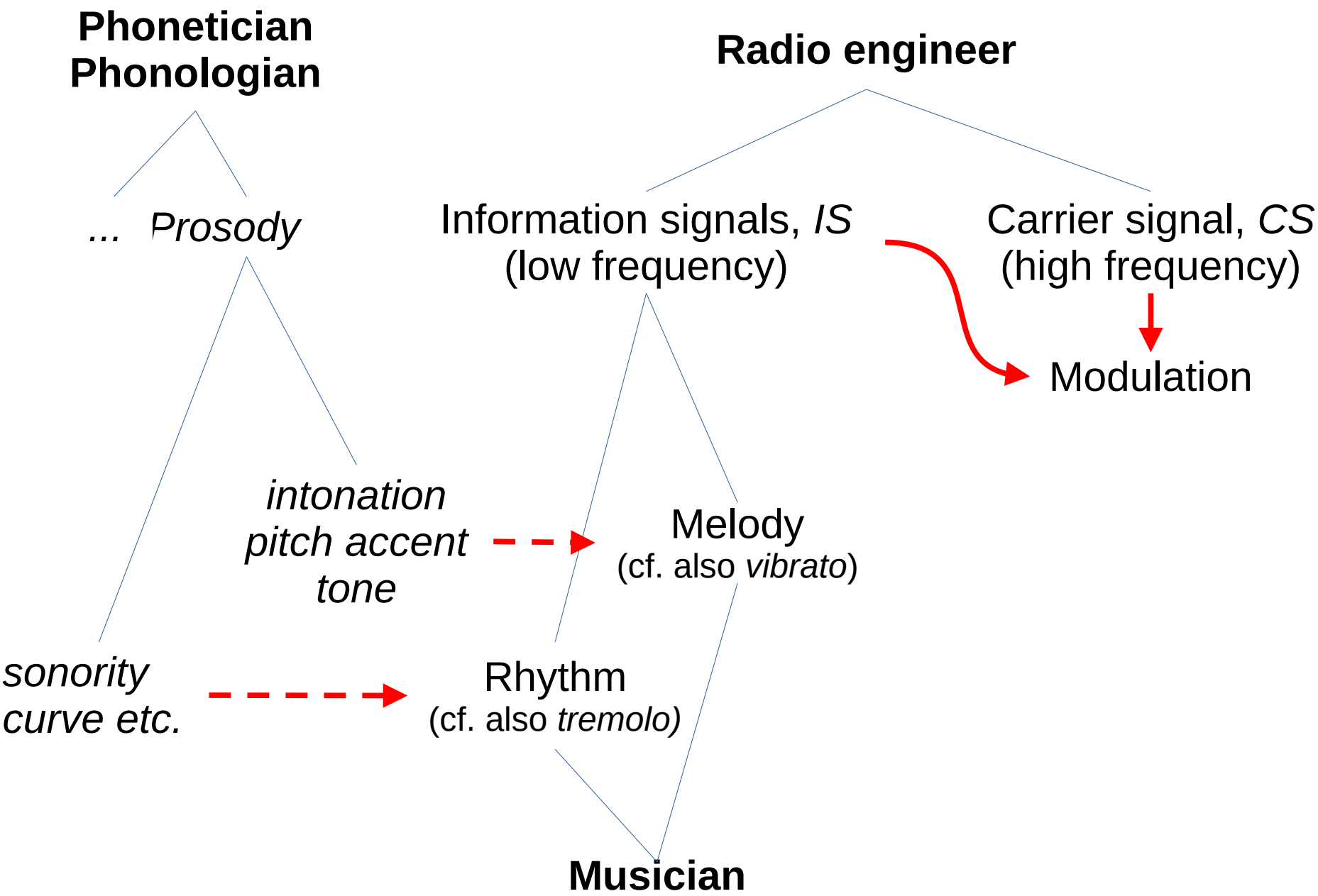
Rhythm
(cf. also *tremolo)*

**Musician**

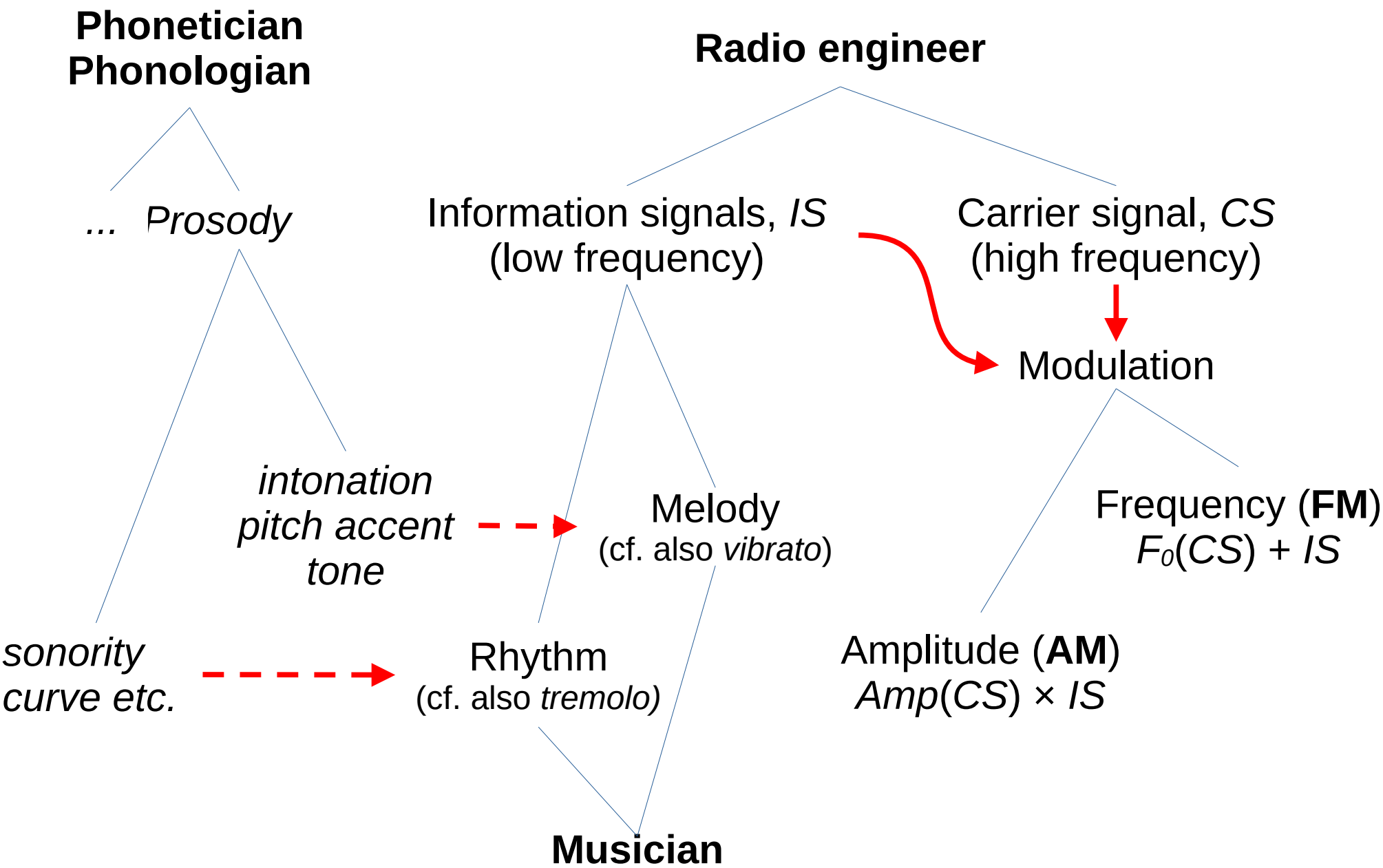
# *Rhythms are ubiquitous...*



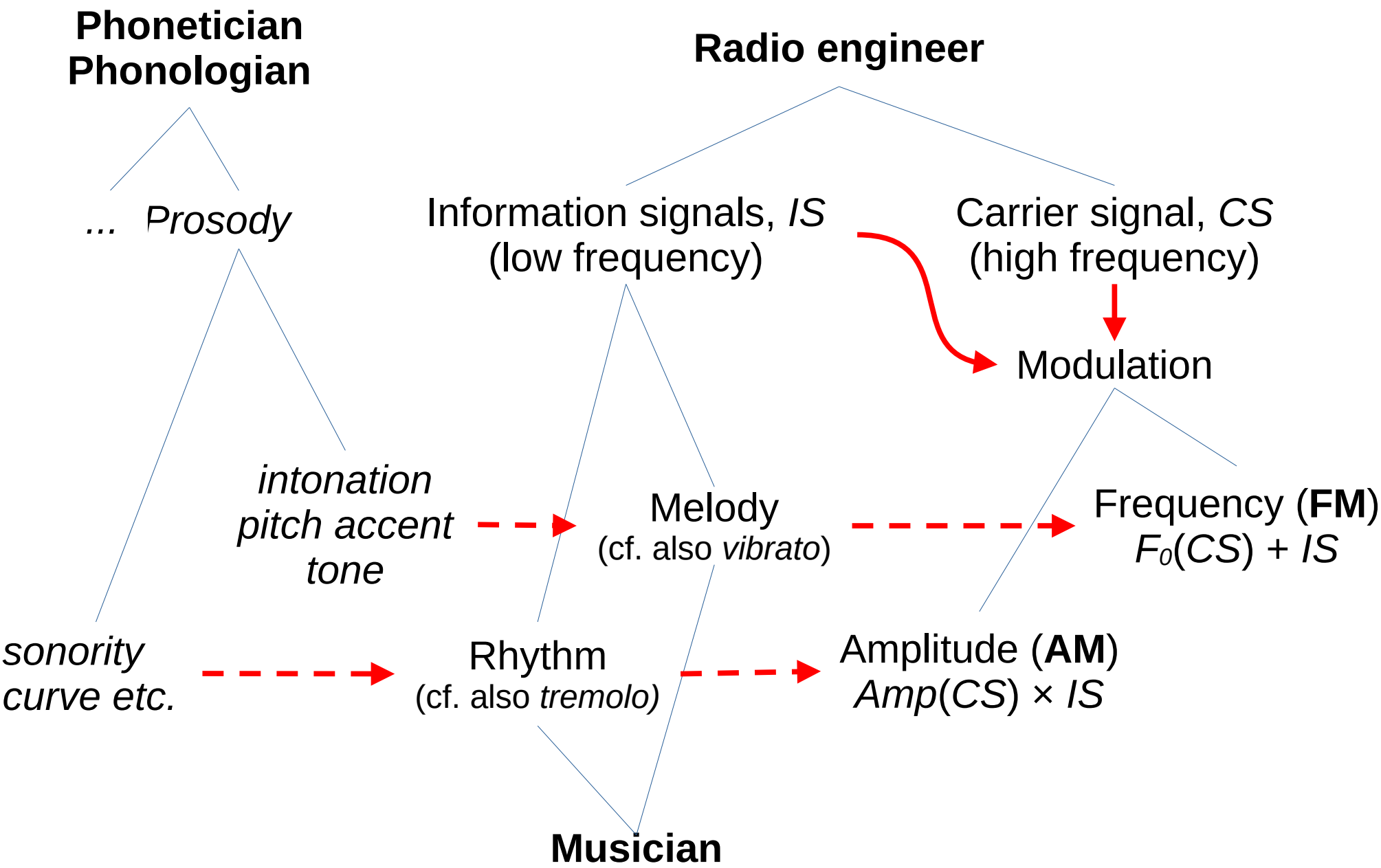
# *Rhythms are ubiquitous...*



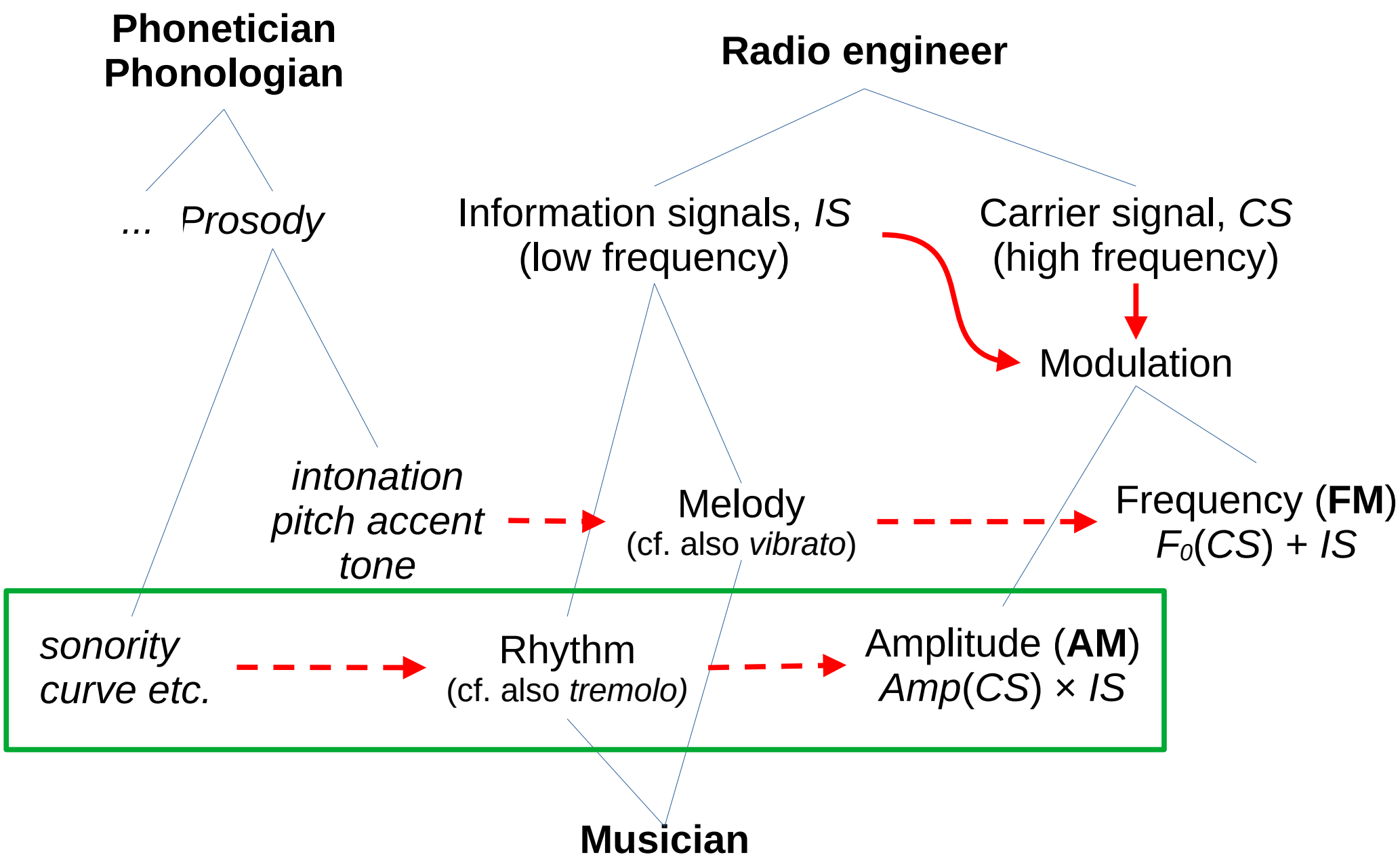
# *Rhythms are ubiquitous...*



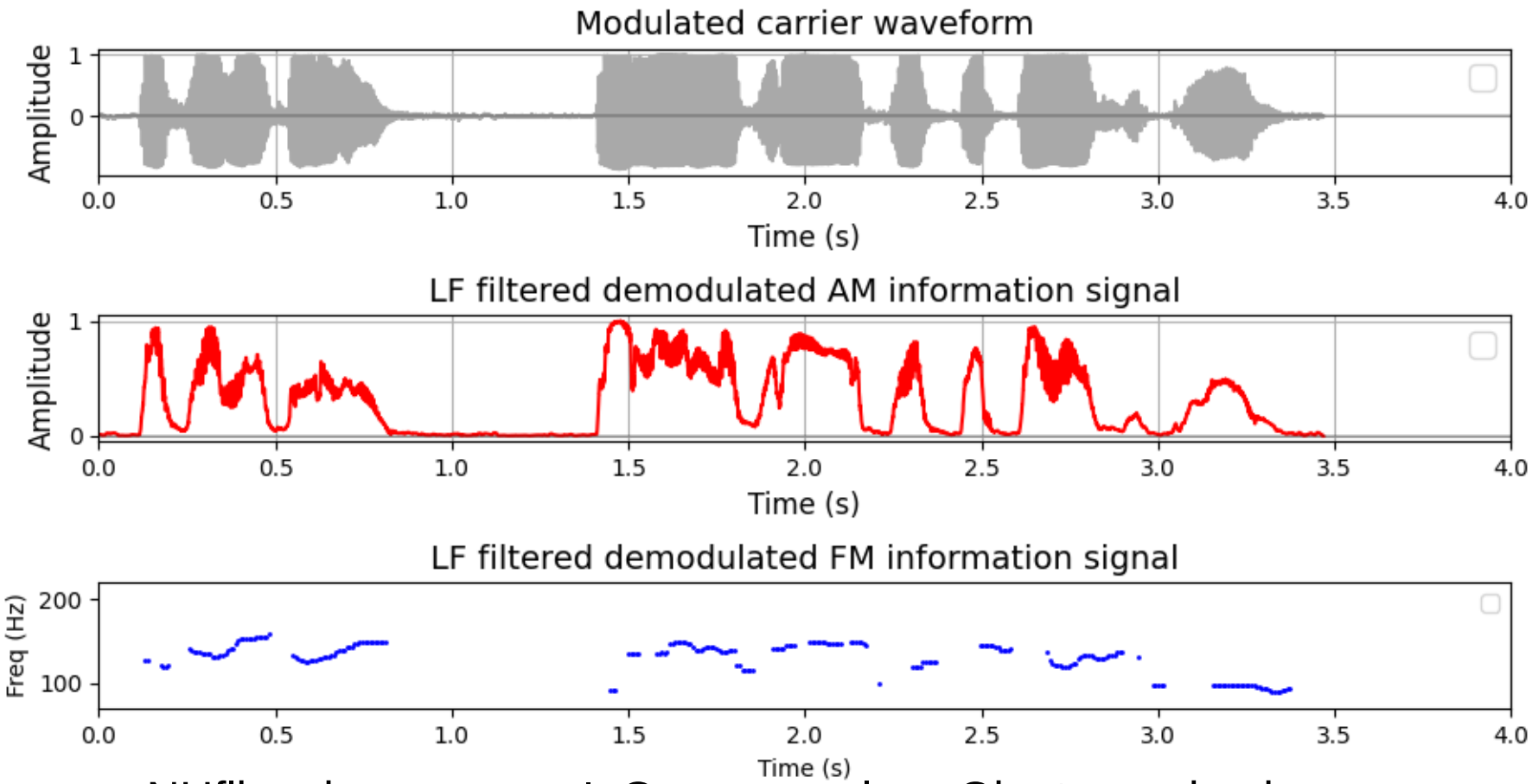
# *Rhythms are ubiquitous...*



# *Rhythms are ubiquitous...*



# *AM and FM demodulation: an Ega example*




NUfli mukanu
*good morning*

InOn ma muglu
*my name is*

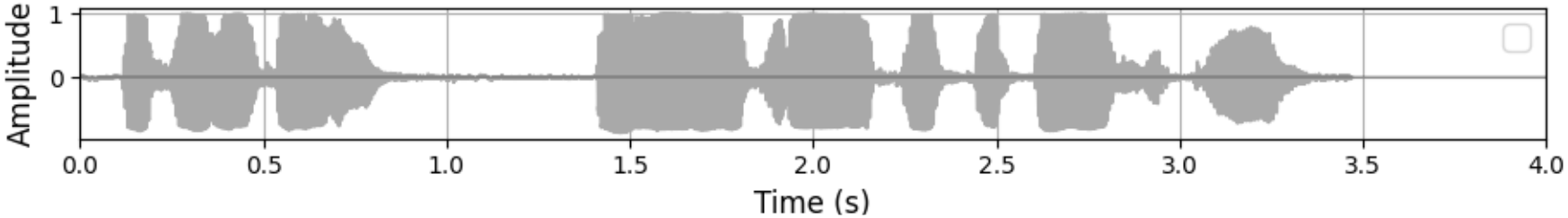
Oku towe siyprie
*Oko Towe Cyprien*

**Note:**

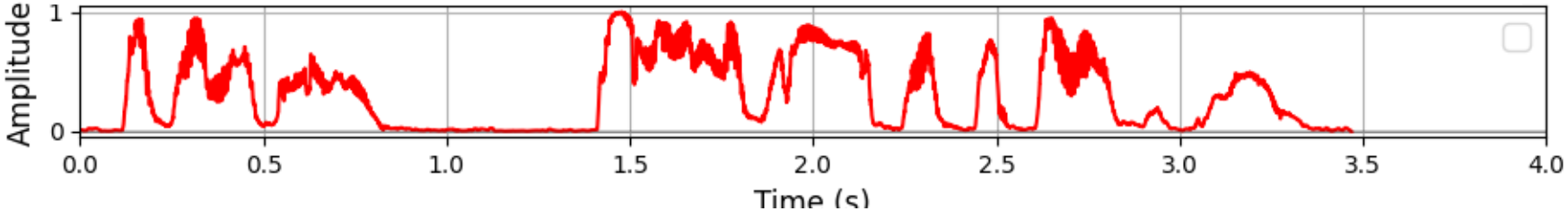
L raising and H lowering on “Òkú tòwé siprìè”

# *AM and FM demodulation: an Ega example*

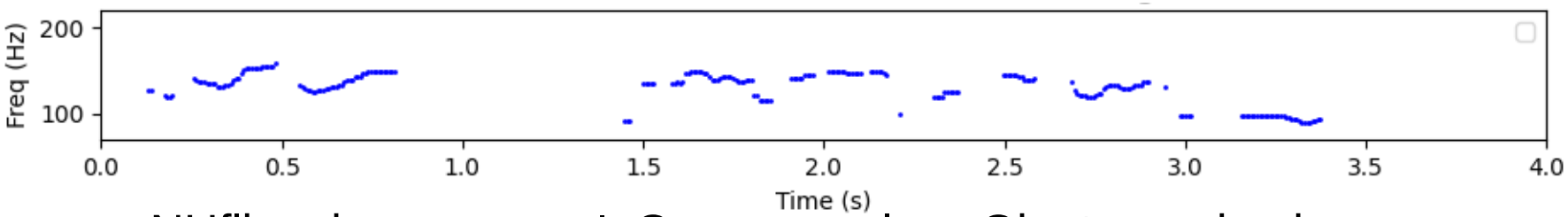
1. Amplitude and frequency modulated carrier waveform




2. AM demodulation (absolute values, low frequency filter)




3. FM demodulation (frequencies of carrier cycles)




| | | |
|---|---|---|
| NUfli mukanu | InOn ma muglu | Oku towe siyprie |
| *good morning* | *my name is* | *Oko Towe Cyprien* |

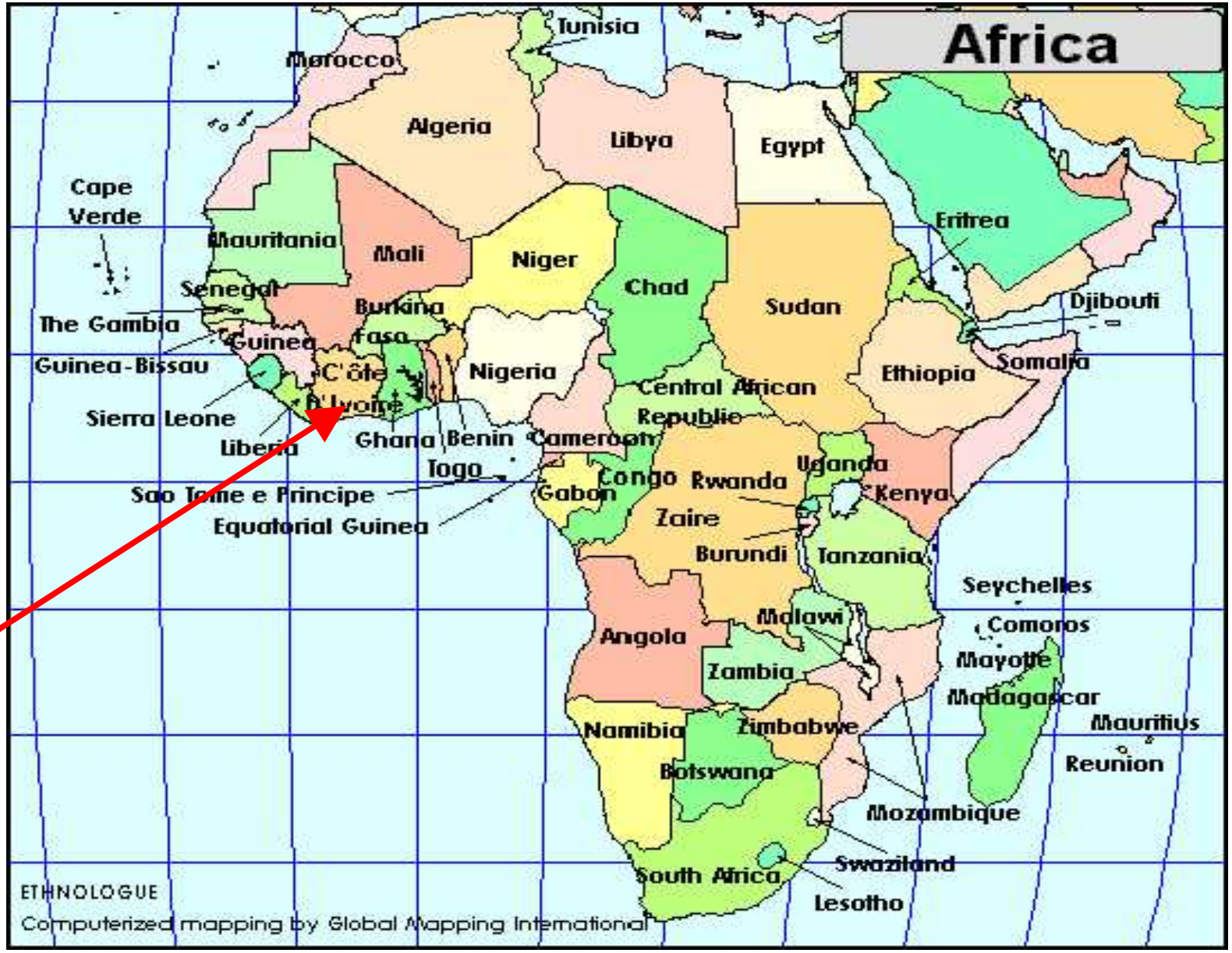
**Note:**

L raising and H lowering on “Òkú tòwé siprìè”

***Where are we geographically?***

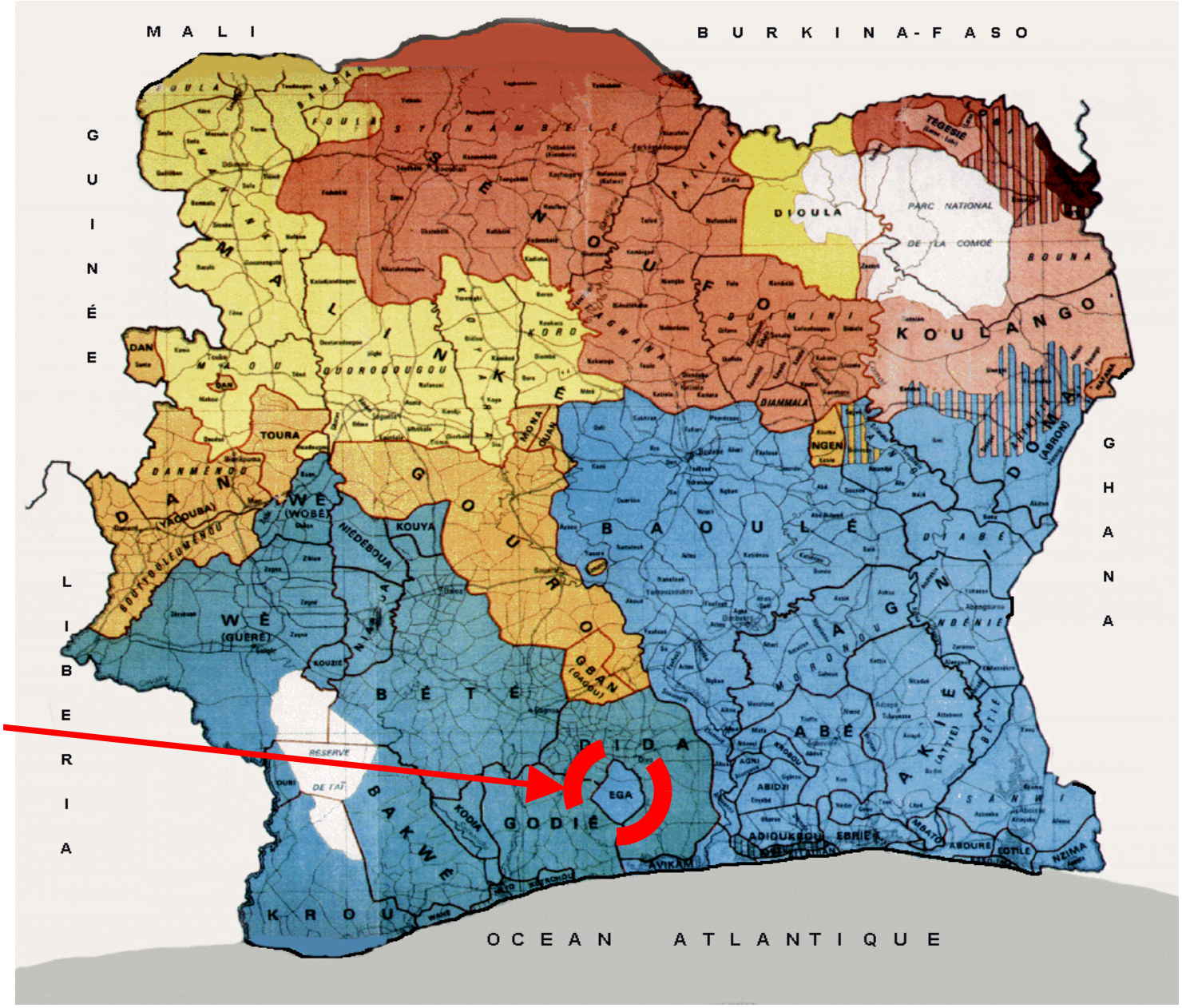
# *Where are we geographically?*

Côte d'Ivoire



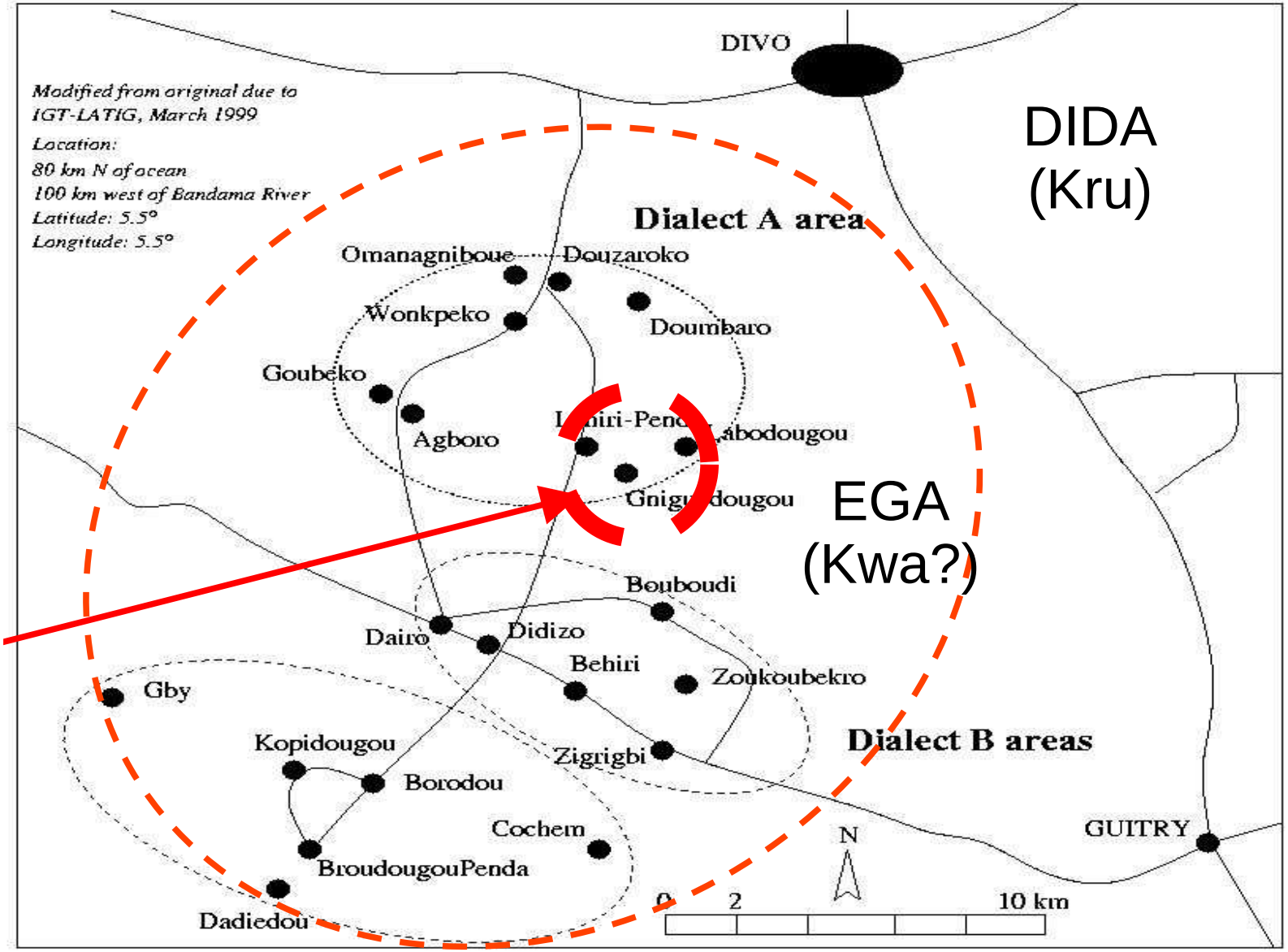
# *Where are we geographically?*

Ega
(ISO 639-3 *ega*)
enclave



# *Where are we geographically?*

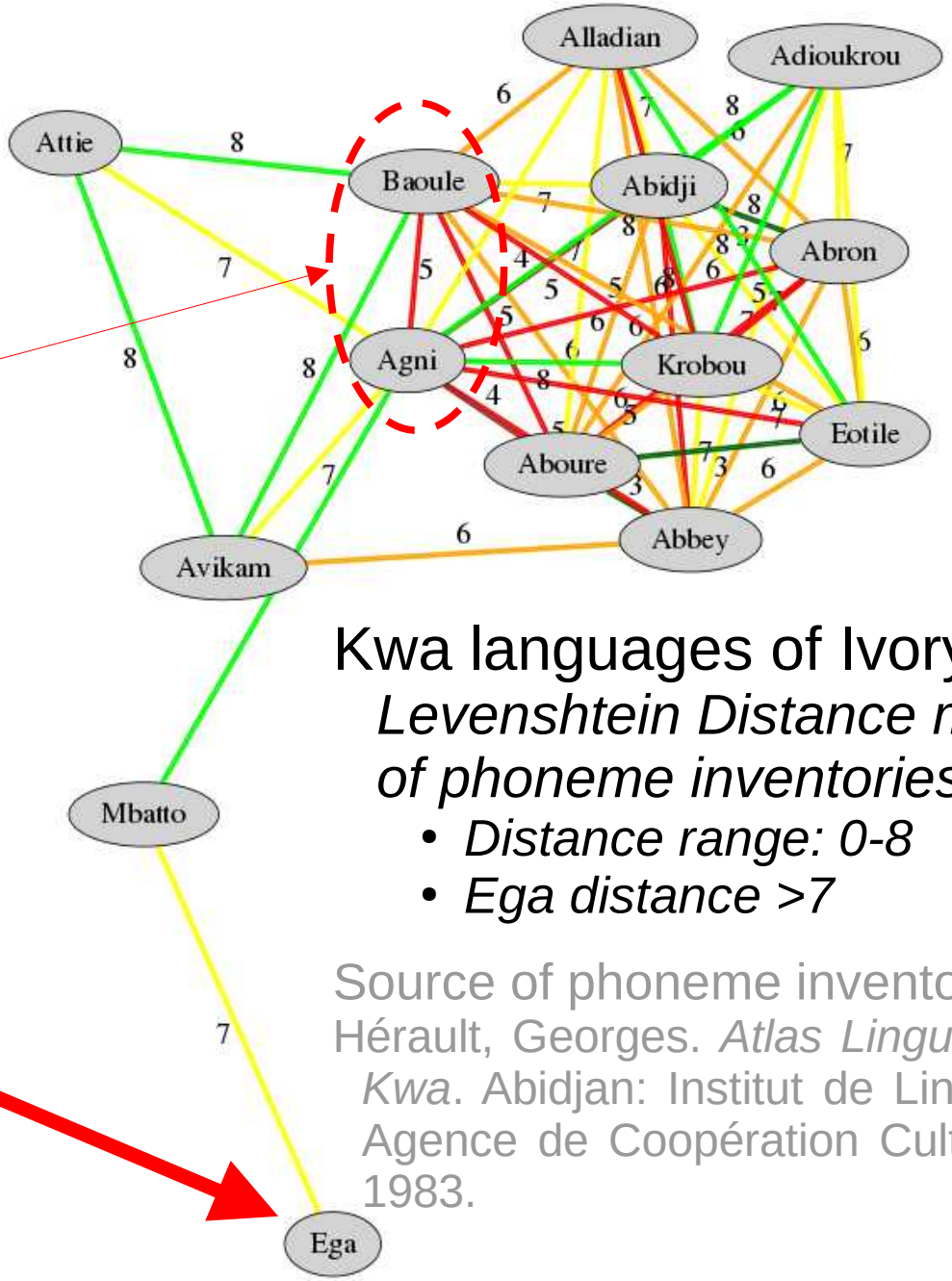
Gniguedougou
*fieldwork*



***Where are we typologically?,***

# *Where are we typologically?*

Better-known Ivory Coast Kwa languages




Kwa languages of Ivory Coast

*Levenshtein Distance mapping of phoneme inventories:*

- *Distance range: 0-8*
- *Ega distance >7*

Ega
ISO 639-3 *ega*

Source of phoneme inventories:
Hérault, Georges. *Atlas Linguistique des Langues Kwa*. Abidjan: Institut de Linguistique Appliquée, Agence de Coopération Culturelle et Technique. 1983.

## *Outline of Ega phonetics, phonology, tonology*

# *Ega Phonetics: Consonants*

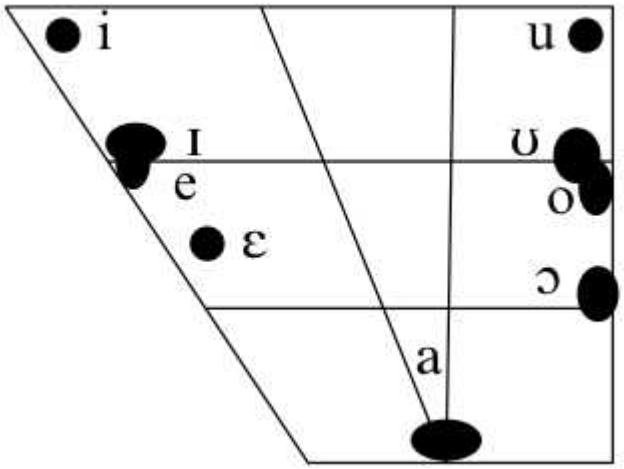
| | Bilabial | | Alveolar | | Palatal | | Velar | | Labial Velar | |
|---|---|---|---|---|---|---|---|---|---|---|
| Plosive | p | b | t | d | c | ɟ | k | g | k͡p | g͡b |
| Implosive | | ɓ | | ɗ | | ʄ | | ɠ | | ɠ͡ɓ |
| Nasal | | m | | n | | ɲ | | ŋ | | |
| Fricative | f | v | s | z | | | x | | | |
| Approximant | | | | | | j | | | | w |
| Lateral Approximant | | | | l | | | | | | |

1. Extensive set of implosives
2. Depressor consonants:
   - The voiced obstruent series /b, d, K, g, gb , v, z/
   - Lowering of following tone (lowering effect on following M and H tones is a salient perceptual cue in distinguishing these from implosives )

# *Ega Phonetics: Consonants*

| | | | |
|---|---|---|---|
| pá | ‘fish’ | īg͡ɓɛ́ | ‘vegetable’ (sp) |
| bá | ‘send away’ | ùg͡bò | ‘cooking pot’ |
| ɓá | ‘complete’ | ùg͡ɓò | ‘mouse’ (sp) |
| tá | ‘speak’ | má | ‘leave nothing’ |
| dá | ‘show’ | nà | ‘walk’ |
| ɗá | ‘hide’ | ɲà | ‘suck’ |
| cá | ‘understand’ | ŋʷà | ‘tie’ |
| ɟé | ‘become numerous’ | fá | ‘consult’ |
| ƒè | ‘swear’ | vá | ‘excuse’ |
| ká | ‘and’ | sá | ‘break in two’ |
| gà | ‘cook’ | sù | ‘break (a piece of wood)’ |
| ɠà | ‘count’ | ùzùzù | ‘shadow’ |
| k͡pá | ‘build a hedge to enclose a field’ | xɛ́ | ‘carry’ |
| g͡bá | ‘finish’ | jé | ‘fill’ |
| k͡pɛ́ | ‘peel’ | wá | ‘kill’ |
| g͡bɛ́ | ‘become thin’ | lá | ‘sleep’ |

# *Ega Phonetics: Vowels*




| | |
|---|---|
| ɓĩ | ‘give birth’ |
| àɓì | ‘beans’ |
| ɓè | ‘swallow’ |
| ɓɛ̀ | ‘plant’ |
| ɓù | ‘become rotten’ |
| ɓʊ̀ | ‘climb up’ |
| ɓó | ‘trap’ |
| ɓɔ́nɛ̀ | ‘begin’ |
| ɓá | ‘complete’ |

ATR vowel harmony within NPs and within VPs

Frequency of vowel hiatus phenomena, regularly resulting in elision, coalescence or assimilation of one of the two vowels

# *Ega Phonetics: Tone*

| | | | |
|---|---|---|---|
| H | ní | tá | 'I speak' |
| M | ní | tā | 'I spoke' |
| L | ní | tà | 'I chew' |

Depressor consonants

The voiced obstruent series /b, d, K, g, gb , v, z/

Lowering of following tone (lowering effect on following M and H tones is a salient perceptual cue in distinguishing these from implosives )

Tone sequences:

Diachronic vowel deletion may leave two tones on one syllable

Downtrend:

Possibly, not investigated.

## *Tones in sequence*

## *A cyclical linear model for downstep*

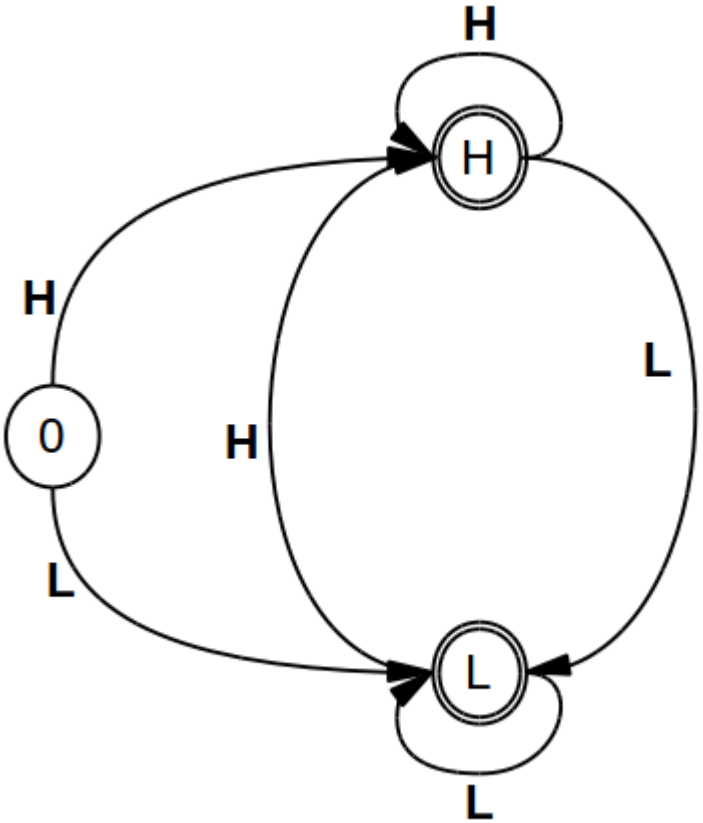
***(Note: there are only borderline cases of downstep in Ega)***

# *Generic downstep model: Finite State Automaton*

- Tone strings of languages with **downstep** are often represented as terminal strings of right-branching trees.
- The set of such terminal strings is a **regular language**, which can be generated by a **regular grammar**.
- The same set of strings can be accepted by a **Finite State Automaton** or **Finite Transition Network**.
- The FSA or FTN for tonal downstep sequences is **cyclical**.

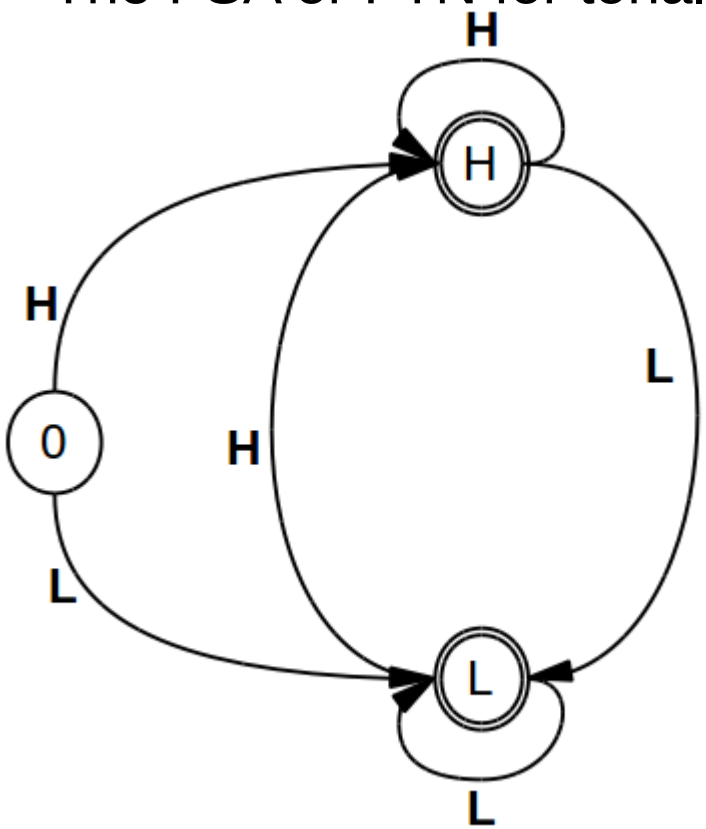
# *Generic downstep model: Finite State Automaton*

- Tone strings of languages with **downstep** are often represented as terminal strings of right-branching trees.
- The set of such terminal strings is a **regular language**, which can be generated by a **regular grammar**.
- The same set of strings can be accepted by a **Finite State Automaton** or **Finite Transition Network**.
- The FSA or FTN for tonal downstep sequences is **cyclical**.



# *Generic downstep model: Finite State Automaton*

- Tone strings of languages with **downstep** are often represented as terminal strings of right-branching trees.
- The set of such terminal strings is a **regular language**, which can be generated by a **regular grammar**.
- The same set of strings can be accepted by a **Finite State Automaton** or **Finite Transition Network**.
- The FSA or FTN for tonal downstep sequences is **cyclical**.



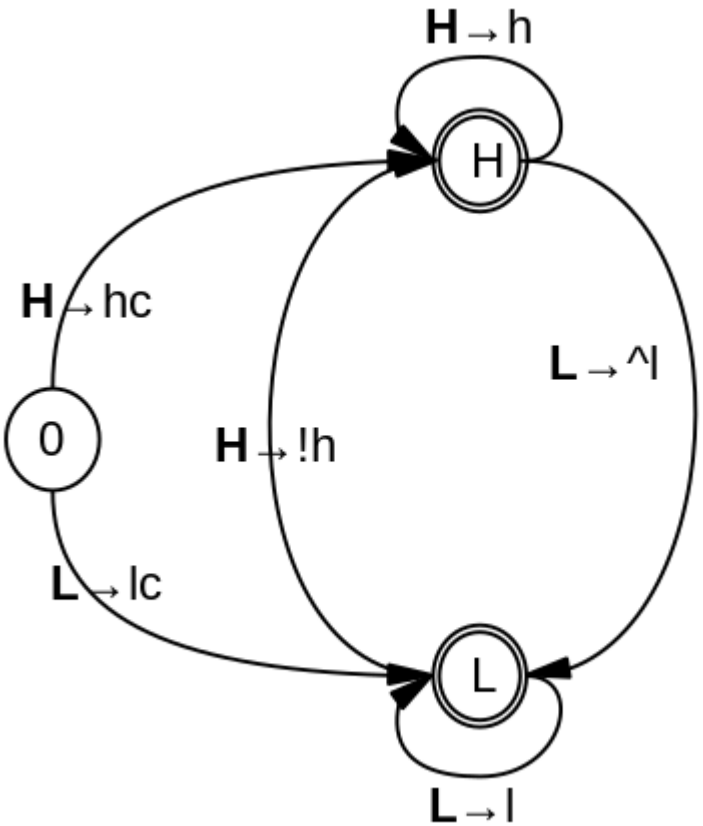

This FTN defines constraints on tone sequences, e.g. tonal terraces, as
- a sequence of H tones
- followed by a sequence of L tones

modelled as cycles around the two main nodes of the FTN.

If terraces are not needed, e.g. there is no downstep, the cyclical tones can be collapsed into a single node.

# *Generic downstep model: Finite State Automaton*

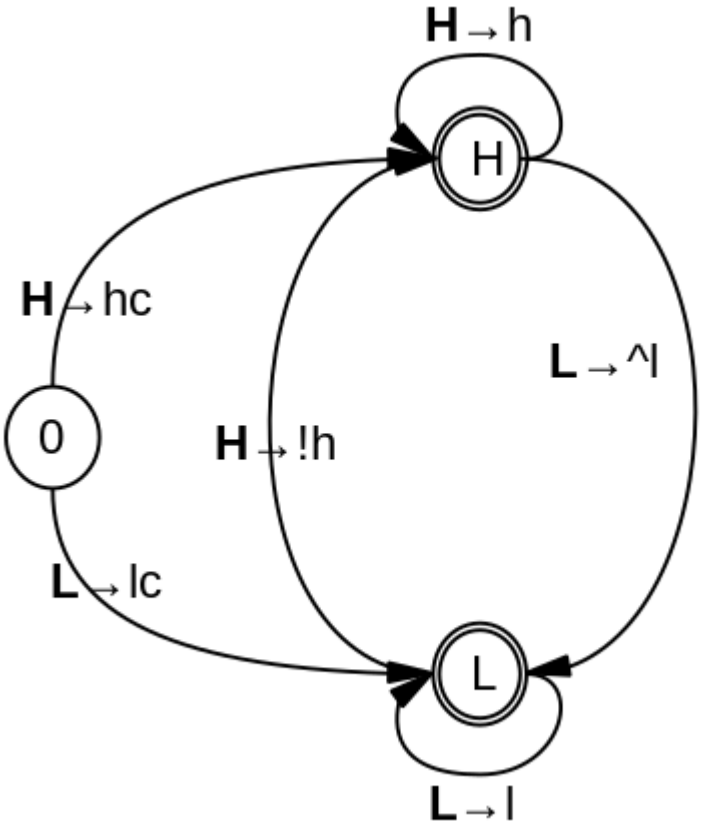
- The FSA or FTN for tonal downstep sequences can be extended into a Finite State Transducer for mapping one set of strings on to a second set of strings, e.g.
  - phonological categories (lexical tones)
    mapped to
  - phonetic categories (phonetic tones)

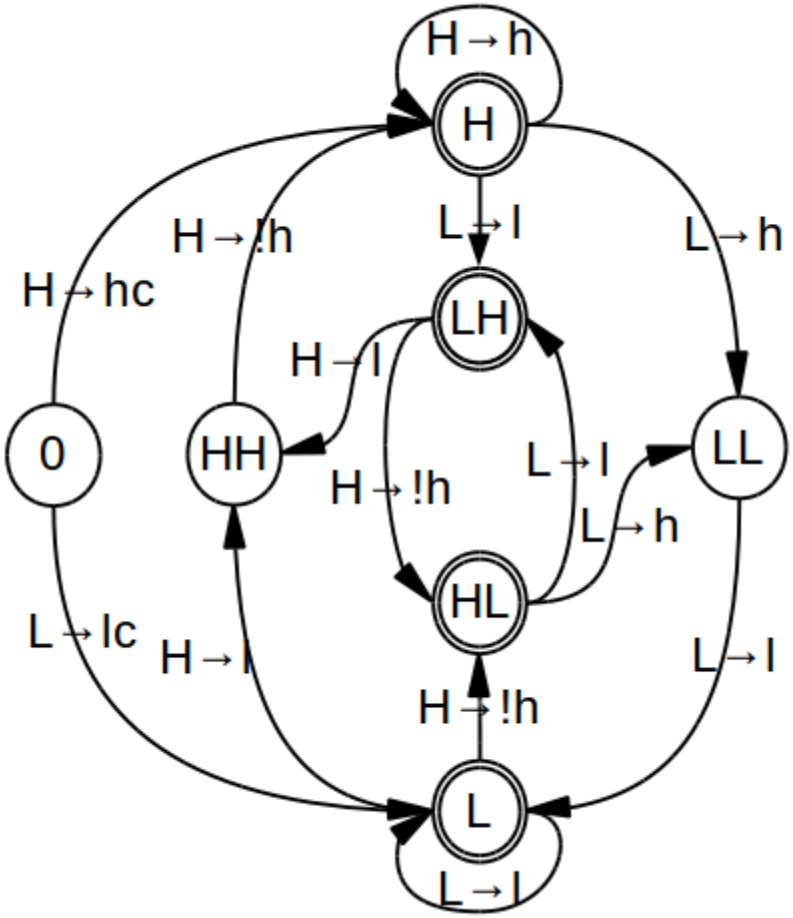



In other words, contexts for phonological rules are provided by a Finite Transition Network.

# *Baule downstep model: Finite State Automaton*

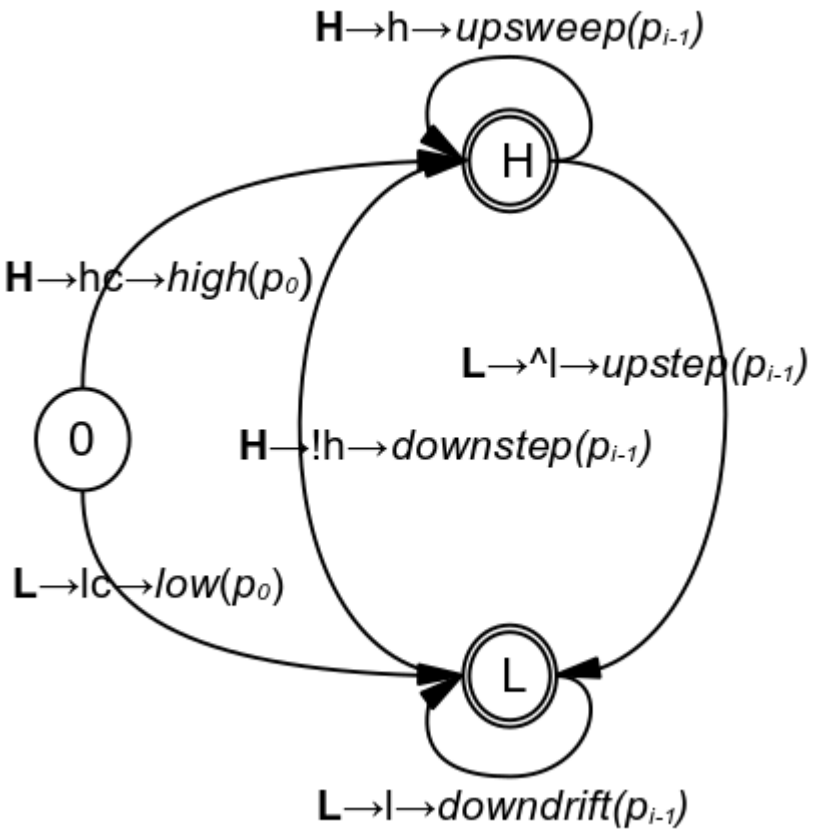
Source:
Ahoua, Firmin. 1996. Prosodic Aspects of Baule.
Cologne: Rüdiger Koppe Verlag.



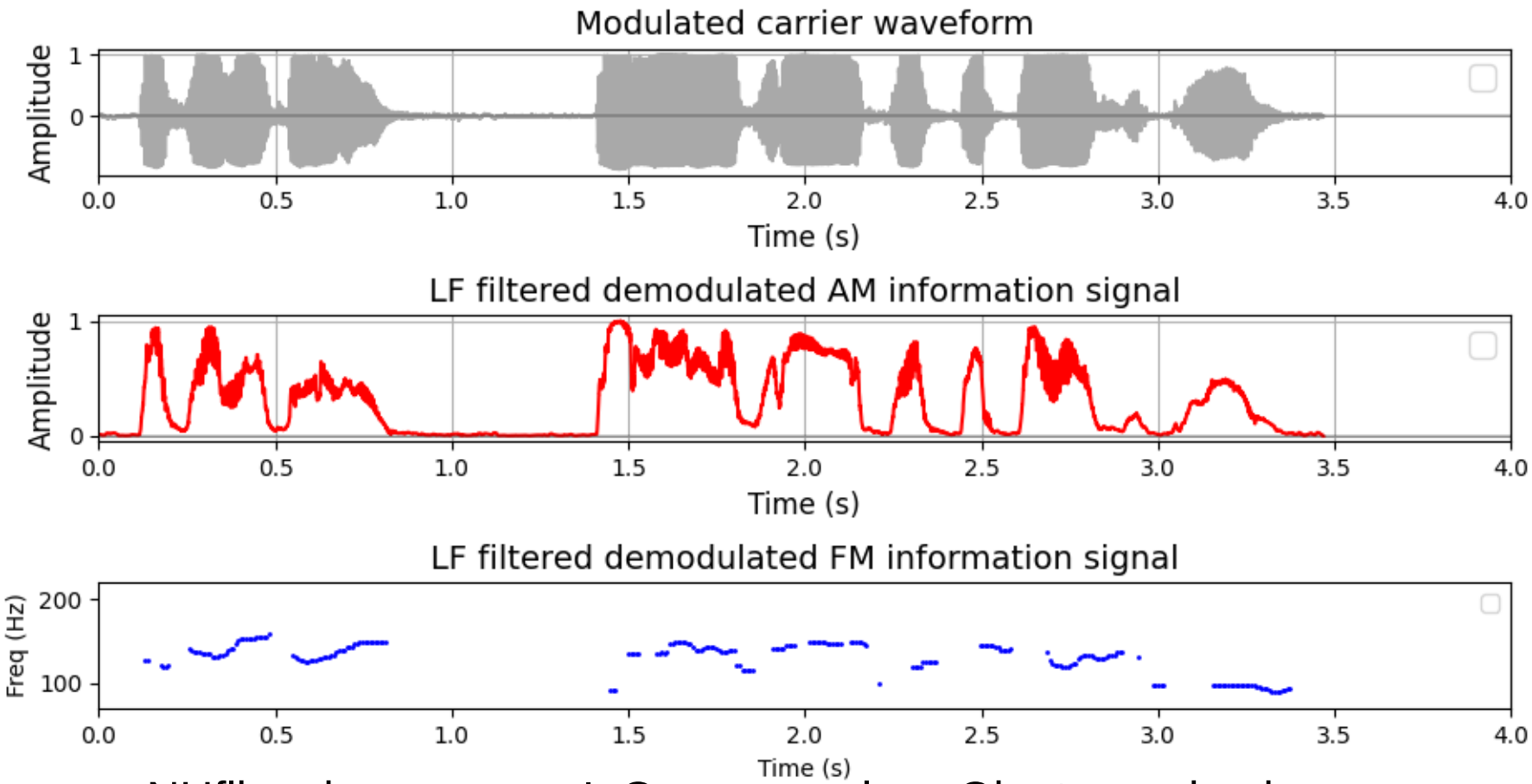



# *Ega Phonetics: Tone*

- The Finite State Transducer for mapping two sets of strings can be extended by a mapping on to a third set of strings, e.g.
  - phonological categories (lexical tones) to
  - phonetic categories (phonetic tones)
  - physical phonetic functions, e.g. *downstep*, *upstep*, *downdrift*, *upsweep*



# *AM and FM demodulation: an Ega example*




NUfli mukanu
*good morning*

InOn ma muglu
*my name is*

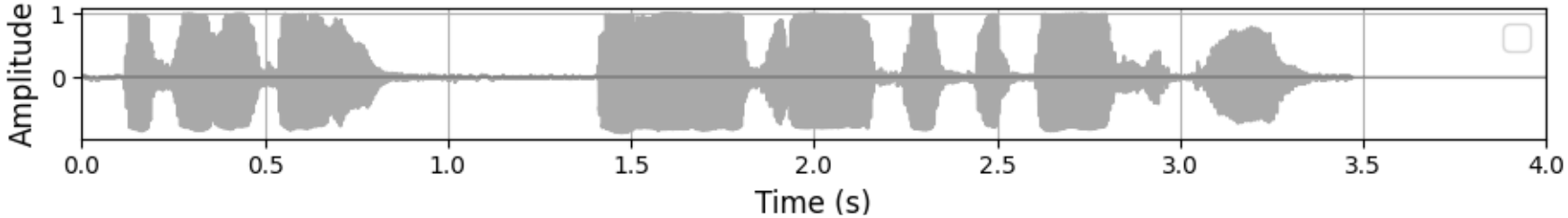
Oku towe siyprie
*Oko Towe Cyprien*

**Note:**

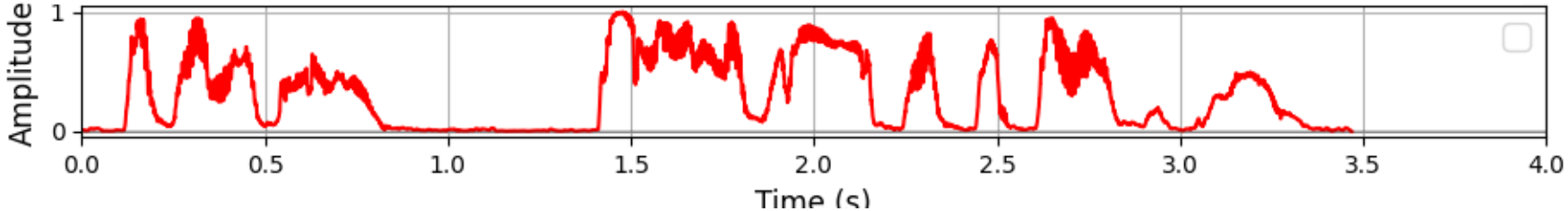
L raising and H lowering on “Òkú tòwé siprìè”

# *AM and FM demodulation: an Ega example*

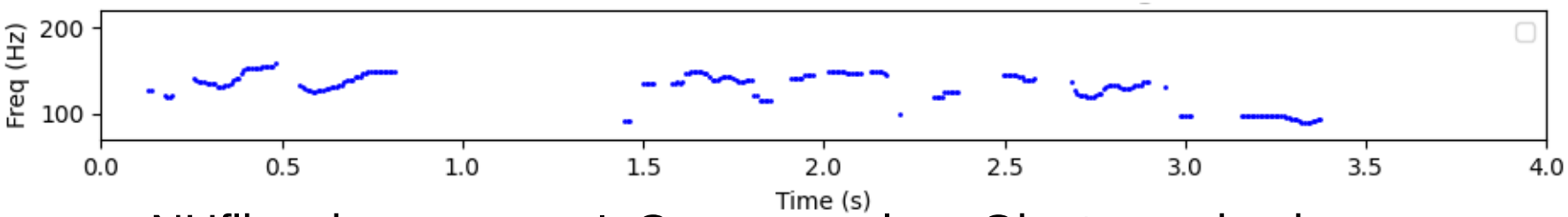
1. Amplitude and frequency modulated carrier waveform




2. AM demodulation (absolute values, low frequency filter)




3. FM demodulation (frequencies of carrier cycles)



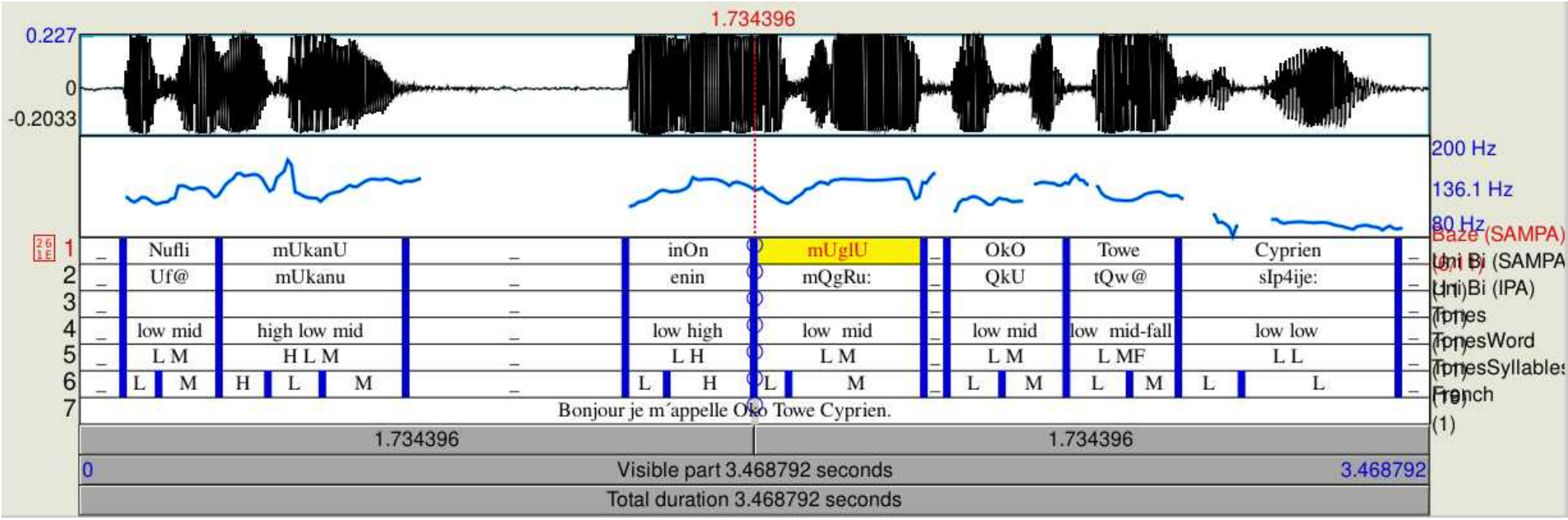

| NUfli mukanu | InOn ma muglu | Oku towe siyprie |
|---|---|---|
| *good morning* | *my name is* | *Oko Towe Cyprien* |

**Note:**

L raising and H lowering on “Òkú tòwé sipríè”

## *Ega rhythms: annotation as a useful starting point*

# *Ega Phonetics: Praat Annotation*




| **Ega:** | Nufli mUkanU | inOn mUglU | Oko Towe Cyprien |
|---|---|---|---|
| **French:** | bonjour | je m´appelle | Oko Towe Cyprien |

**Pause distribution:**

Nufli mUkanU _

inOn mUglU _

OkO Towe Cyprien

**Note:**

L – M/H alternation

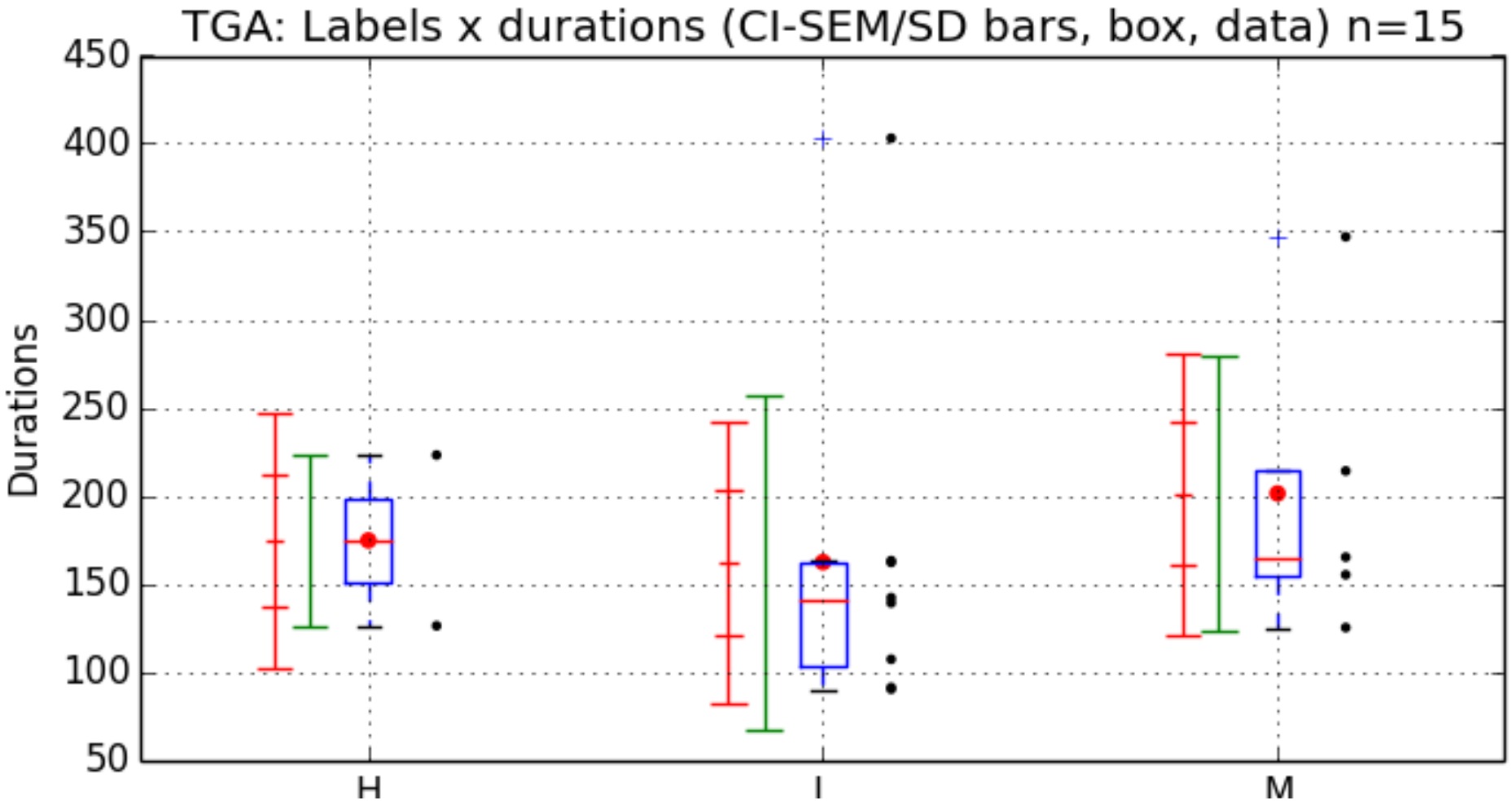
# *Ega Phonetics: Syllable and Word Duration*

| Syllable duration measures | | | | Word duration measures | | | |
|---|---|---|---|---|---|---|---|
| **Attribute** | **Value** | **Attribute** | **Value** | **Attribute** | **Value** | **Attribute** | **Value** |
| *n* | 15 | *intercept* | 117.7 | *n* | 7 | *intercept* | 319.714 |
| *min* (ms) | 90 | *slope* | 8.443 | *min* (ms) | 246 | *slope* | 19.714 |
| *max* (ms) | 403 | *SD* | 86.544 | *max* (ms) | 565 | *SD* | 107.912 |
| *range* (ms) | 313 | *coeff var* | 48.95 | *range* (ms) | 319 | coeff var | 28.484 |
| *mean* (ms) | 176.8 | ***nPVI*** | **50** | *mean* (ms) | 378.86 | *nPVI* | 39 |
| ***mean rate*** | **5.66** | *rPVI* | 96 | ***mean rate*** | **2.64** | *rPVI* | 153 |
| ***total* (ms)** | **2652** | | | ***total* (ms)** | **2652** | | |

**Two points to note:**

1. *nPVI* (next-door-neighbour distance) is higher for syllables, indicating either influence of final lengthening or, perhaps, the presence of stress/accent.
2. The rate values are potential *rhythm frequencies* (but note the influence of final lengthening).

# *Ega Phonetics: Tone X Syllable Duration*




**Note:**

1. Statistically, the diagram is meaningless (data sample too small).
2. However, a sources of hypotheses on dependencies.
3. The syllable type would also need to be taken into consideration.

## *Choral narrative in the village*

# *The parable*

Audience + Choir
Narrator
+
Caller
Backchannel
Responder
Fieldworkers: notes, audio, camera

# *The parable*

# *The parable*

1. A faithless young bride-to-be left her promised bridegroom
   and is on her way to the next village
   to marry her new-found future husband (exogamy)
2. She is fruitlessly warned by a twittering bird
   with a prophecy about her impending end
   at the hands of her formerly betrothed suitor
   who had worked for free for her father to win her hand.
3. Alas, the girl did not understand
4. The bride and the jilted suitor met along the way ...
   he reproached and threatened her, she spat at him
   whereupon the prophecy came true
5. Moral: one should learn other languages like the fieldworkers
6. Punchline: the birdsong is in Dida
   the enclaving Eastern Kru language
   not in endangered Ega.

***A cyclical model of the dialogue***

***<u>turns</u> (e.g. narrative sequences)***
***<u>adjacency pairs</u> (e.g. call-response)***

# *A cyclical dialogue model*

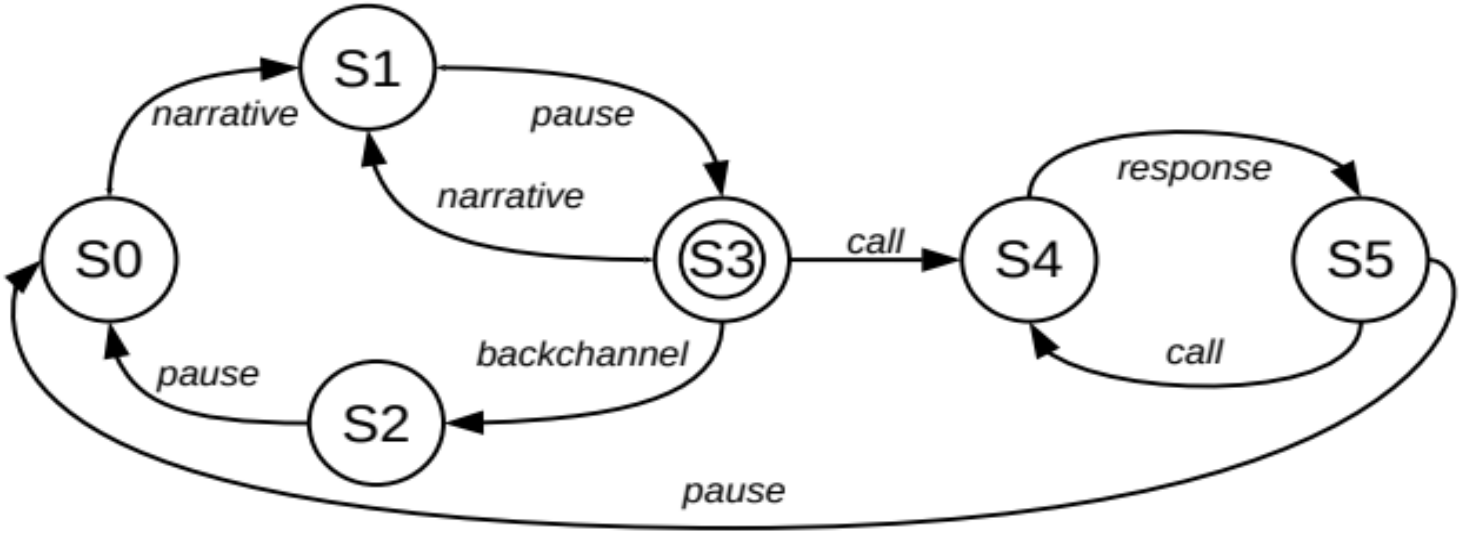
**Form:**

1. There are **multiple parallel speech rhythms**, not a single rhythm
2. Speech rhythms relate to linguistic units in **cycles**
3. Multiple speech rhythm beat levels / cycles can be modelled as regular languages by **cyclical (iterative) FTNs**, **FSTs**, etc.

# *A cyclical dialogue model*

## Form:

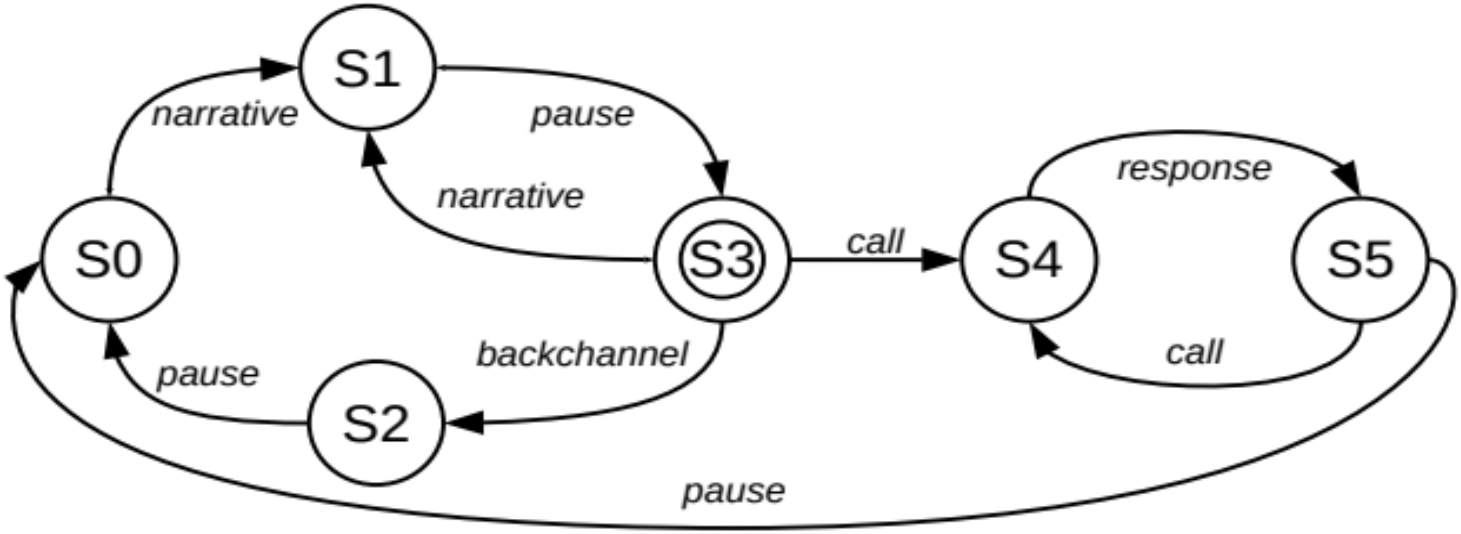
1. There are **multiple parallel speech rhythms**, not a single rhythm
2. Speech rhythms relate to linguistic units in **cycles**
3. Multiple speech rhythm beat levels / cycles can be modelled as regular languages by **cyclical FTNs**, **FSTs**, etc.



# *A cyclical dialogue model*

## Form:

1. There are **multiple parallel speech rhythms**, not a single rhythm
2. Speech rhythms relate to linguistic units in **cycles**
3. Multiple speech rhythm beat levels / cycles can be modelled as regular languages by **cyclical FTNs**, **FSTs**, etc.




## Function

1. *Semantics*: Speech rhythms are **metalocutions** which **denote** units at **temporal points and intervals** in locutions
2. *Pragmatics*: Domains of speech rhythms depend on **register**, **style** and **discourse** (at least as much as lexical and phrasal grammar)

# *A cyclical dialogue model*

**Recursive computation is unnecessary**
whether right, left or centre branching grammar
(unrealistic: requires nonfinite memory)

**Only iteration is required**
through the multiple cycles of
turns, exchange adjacency pairs
(realistic: finite constant working memory)

*Recursion is delegated to function not form.*

# ***Rhythms are ubiquitous ...***

1. An inheritance hierarchy:

*Physiological rhythms*: breathing, heartbeat, brain ...

*Behavioural rhythms*: walking, chewing, ...

*Bonding rhythms*: dance, music, waving, nods, handshakes, …

*Speech rhythms*: syllable, word, phrase, turn, exchange, ...

# *Rhythms are ubiquitous ...*

1. An inheritance hierarchy:

*Physiological rhythms*: breathing, heartbeat, brain ...

*Behavioural rhythms*: walking, chewing, ...

*Bonding rhythms*: dance, music, waving, nods, handshakes, …

*Speech rhythms*: syllable, word, phrase, turn, exchange, ...

**Physiological, behavioural, communication, speech:**
**Low Frequency rhythms in the range**

**0.1 Hz … 1Hz … 10Hz**

# ***Rhythms are ubiquitous ...***

1. An inheritance hierarchy:

*Physiological rhythms*: breathing, heartbeat, brain ...

*Behavioural rhythms*: walking, chewing, ...

*Bonding rhythms*: dance, music, waving, nods, handshakes, …

*Speech rhythms*: syllable, word, phrase, turn, exchange, ...

2. Characterisation:

*A rhythm is perceived, and can be measured, when a series of <u>similarly structured events</u> occurs with a <u>specific frequency</u> at approximately <u>equal intervals in time</u> and motivates a <u>prediction</u> of a further similarly structured event after a similar interval.*

# *Rhythms are ubiquitous ...*

1. An inheritance hierarchy:

*Physiolo*

*Behav*

*Bo* akes, …

ange, ...

2. Charac

*A rhyt* *a series of*

*simila* *quency* at

*appro* *prediction*

*of a fu* *erval.*

In Econometrics, Granger-Causality
(thanks, Alex!)

*A time series X is said to Granger-cause Y if it can be shown [...] that those X values provide statistically significant information about future values of Y.*

Cf. the rhetorical role of timing (X) in ‘predicting’ turn-taking points in locutions (Y).

# ***Rhythms***

1. An inheritance hierarchy:

   *Physiological rhythms*: breathing, heartbeat, brain ...

   *Behavioural rhythms*: walking, chewing, ...

   *Bonding rhythms*: dance, music, waving, nods, handshakes, …

   *Speech rhythms*: syllable, word, phrase, turn, exchange, ...

2. Characterisation:

   *A rhythm is perceived, and can be measured, when a series of <u>similarly structured events</u> occurs with a <u>specific frequency</u> at approximately <u>equal intervals in time</u> and motivates a <u>prediction</u> of a further similarly structured event after a similar interval.*

3. Attributes:

   1. *frequency*, *magnitude*, *bandwidth, persistence, resonance*
   2. *attack, decay; staccato, legato, …*

# *Rhythms*

1. An inheritance hierarchy:

   *Physiological rhythms*: breathing, heartbeat, brain ...

   *Behavioural rhythms*: walking, chewing, ...

   *Bonding rhythms*: dance, music, waving, nods, handshakes, …

   *Speech rhythms*: syllable, word, phrase, turn, exchange, ...

2. Characterisation:

   *A rhythm is perceived, and can be measured, when a series of <u>similarly structured events</u> occurs with a <u>specific frequency</u> at approximately <u>equal intervals in time</u> and motivates a <u>prediction</u> of a further similarly structured event after a similar interval.*

3. Attributes:

   *1. frequency, magnitude, bandwidth, persistence, resonance*

   *2. attack, decay; staccato, legato, ...*

***Ega choral narrative:***

***Rhythms in the time and frequency domains***

# *Phonetic Methods*

**1. Time domain dialogue annotation** (e.g. with Praat, …)

Use case:

Heuristic exploration, validation

Time-stamped labels for linguistically defined units

- Phone, Syllable, Word, Phrase
- Turn, Exchange, Episode, Encounter

**2. Frequency domain LF signal analysis of entire event**

Speech Modulation Theory

→ Rhythm Formant Theory and Rhythm Formant Analysis

→ Demodulation of AM and FM information signals

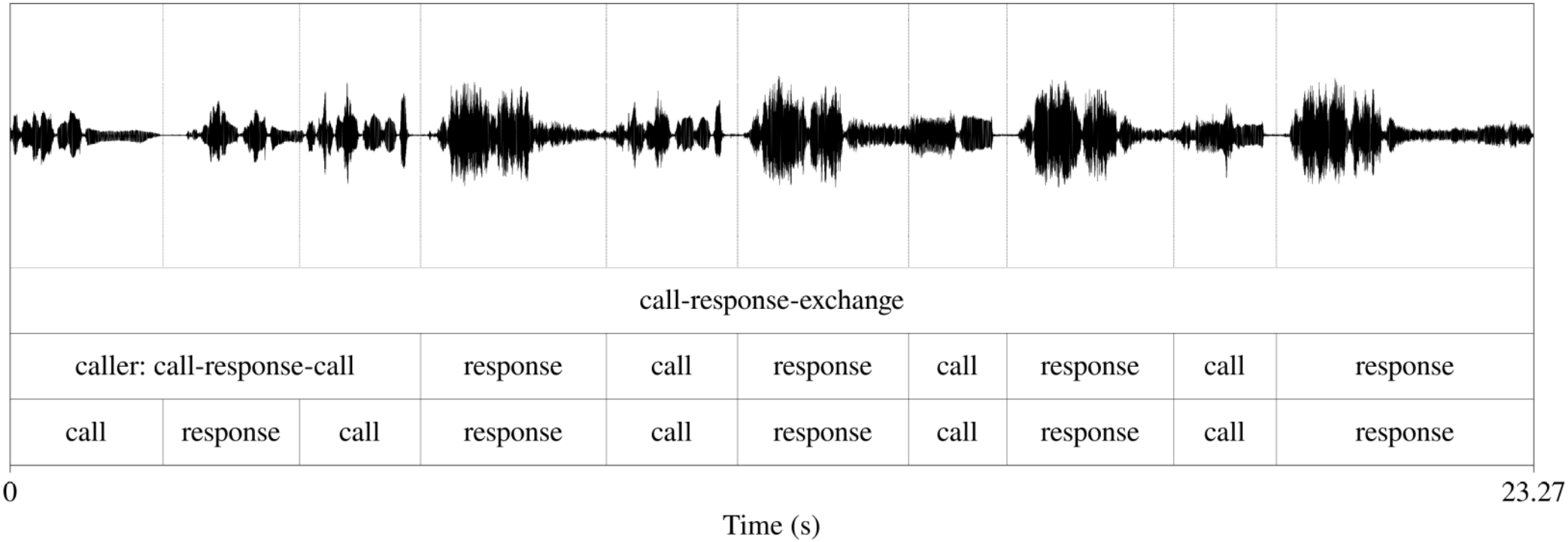
→ Fourier Analysis

*'Rhythm Formants'*

→ Spectral magnitude peaks: vector extraction

→ Distance network calculation, hierarchical clustering

# *Rhythms in the time domain – turn annotation*




Typical use cases for annotation triples (e.g. with Praat):

Use case in speech recognition: train label recognition

Use case in linguistics: illustrate phonetic observations

**Use case in phonetics: calculate duration regularity indices:**

1. **Variance, SD, Coefficient of Variation, other dispersion measures**
2. **Next-door-neighbour distances: *rPVI*, *nPVI***

**Again… Where's the "*Boom-de-boom-boom*" in '65' or '33'?**

# *Speech rhythm in the frequency domain*

We need to

- demodulate the AM to recover the information signal (rhythm)
- analyse the information signal

And speech rhythms have...

1. *frequencies*, so let's apply an **FFT** to get the **LF spectrum**
2. *magnitudes*, so let's apply a **peak detector** to get the salient spectral frequencies (functionally: '**rhythm formants**')
3. *persistence*, so let's make an **LF spectrogram** to see the persistence/variability of frequencies in time
4. *resonance*, so let's check the **stability of persistent frequencies in time**

# *Method 2: Rhythms in the Frequency Domain*

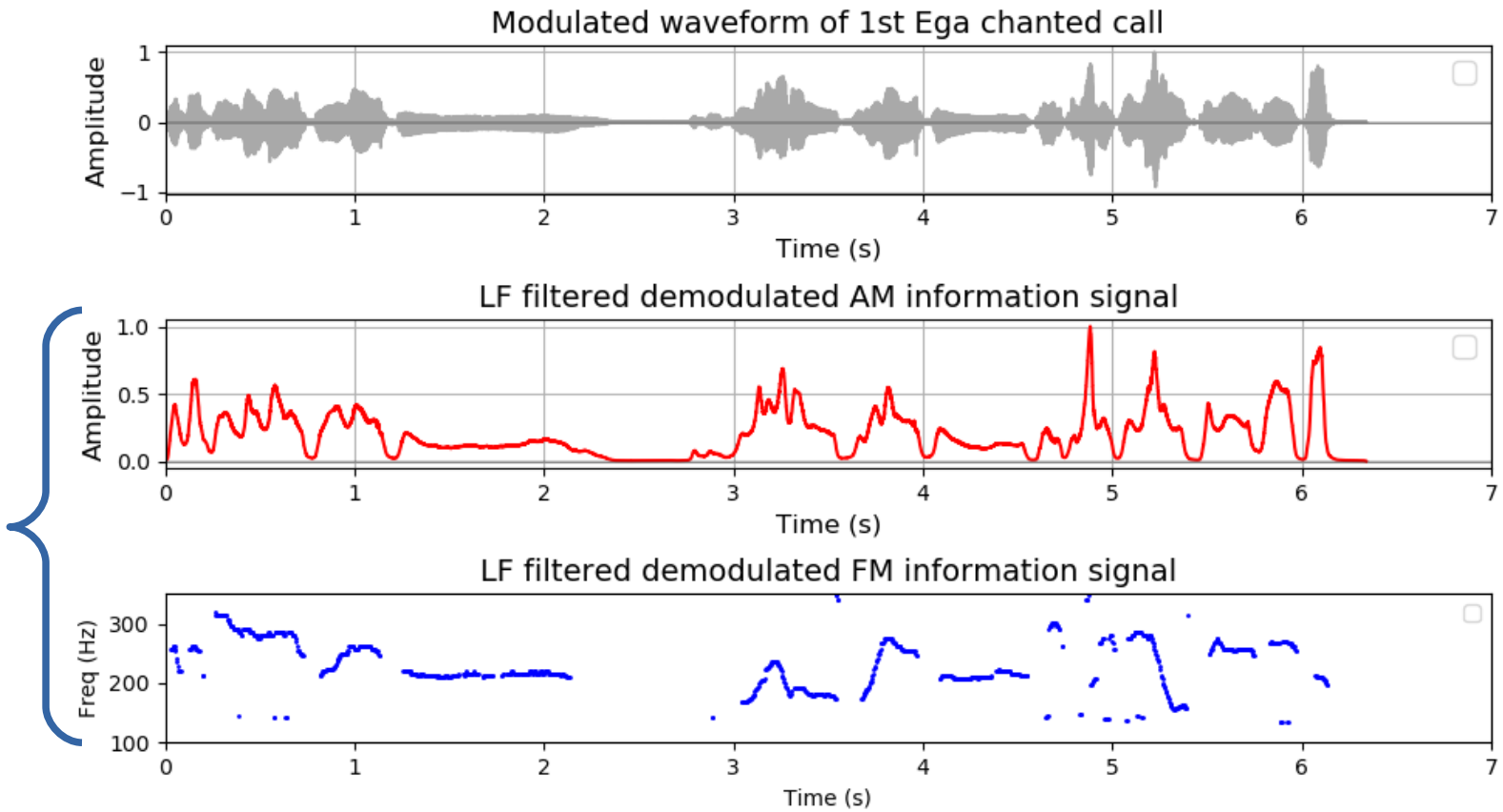
Methods:

1. Extract AM (amplitude demodulation), the ‘sonority curve’
2. Extract FM (frequency demodulation), the ‘pitch curve’

**First call by the narrator-caller**

Carrier signal

Information signals

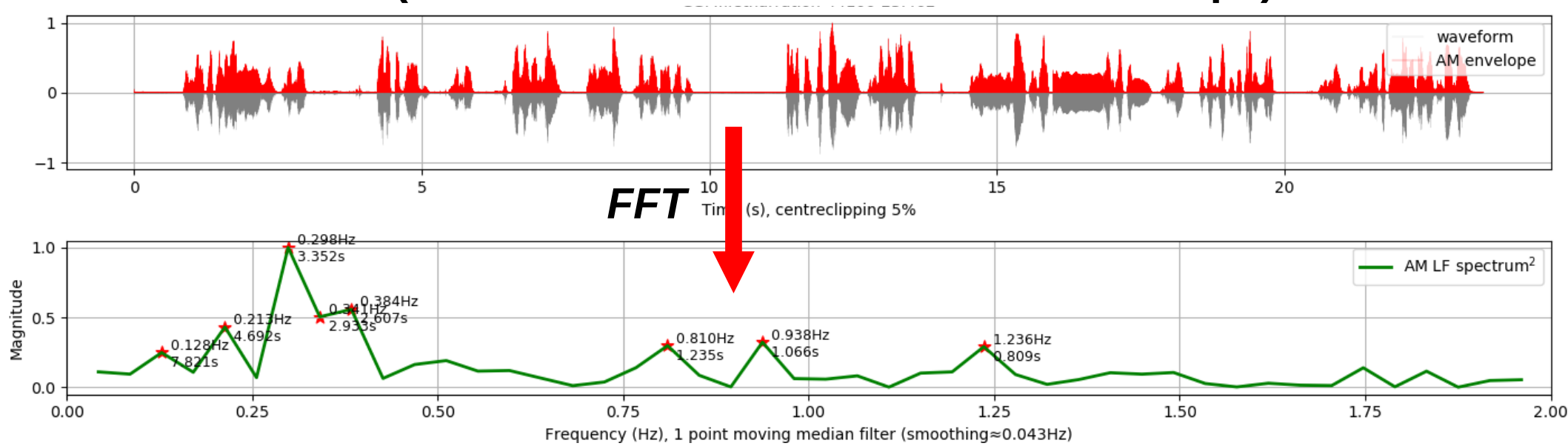


# *Rhythm Formant Analysis: AM demodulation*

# *Rhythm Formant Analysis: global AM spectrum*

## First narrative turn

**Time domain (waveform and demodulated AM envelope)**

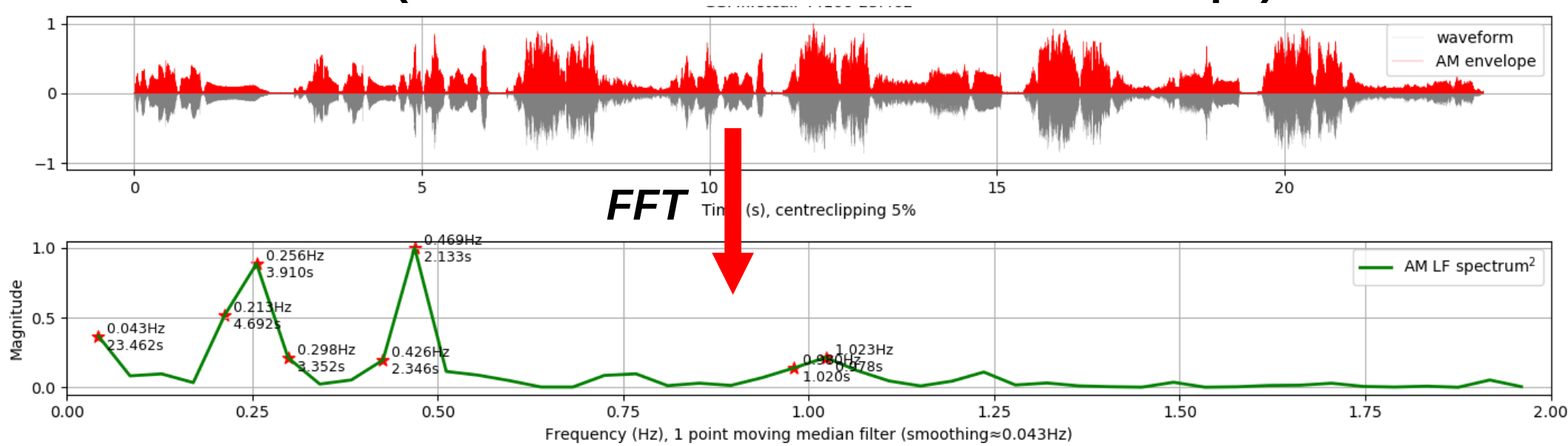



**Frequency domain (Long term Low Frequency Spectrum < 2 Hz)**

# *Rhythm Formant Analysis: global AM spectrum*

## First caller-choir exchange

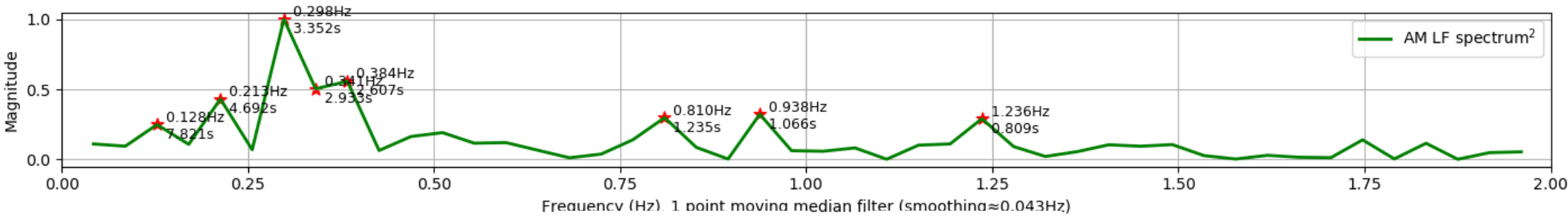
**Time domain (waveform and demodulated AM envelope)**



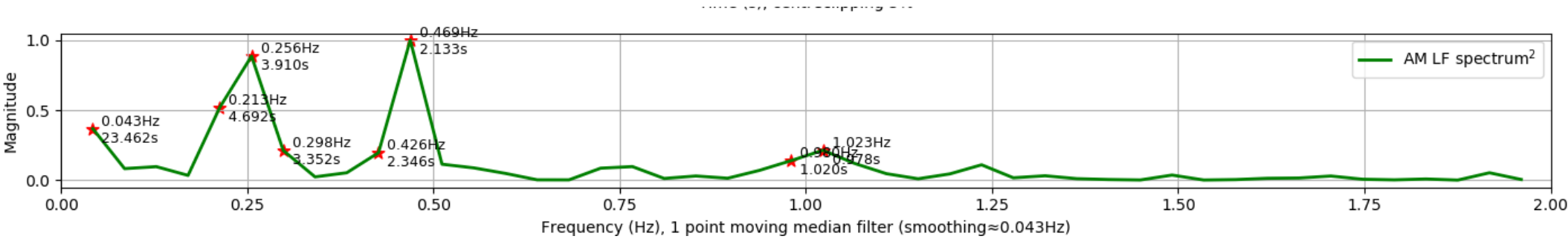

**Frequency domain (Long term low frequency Spectrum < 2 Hz)**

# *Rhythm Formant Analysis: global AM spectrum*

## Comparison of narrative turn cycles and choral turn cycles




**Narrative turns** **Phrases** **Words** **Syllables**

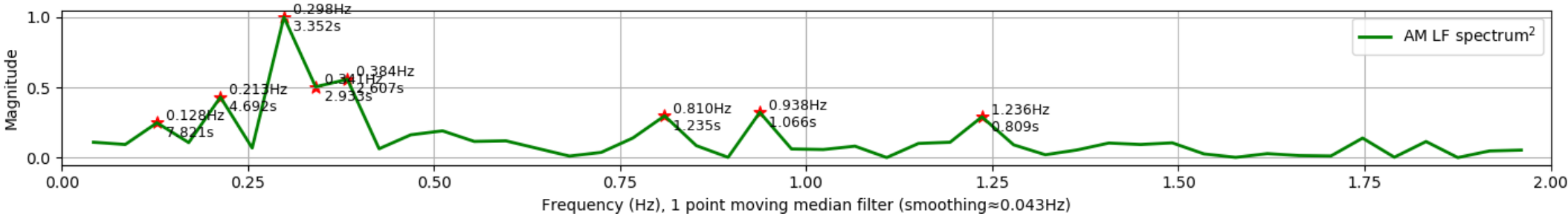



**Caller-Choir exchanges, 'stanzas'** **Phrases, 'lines'** **Words ...**

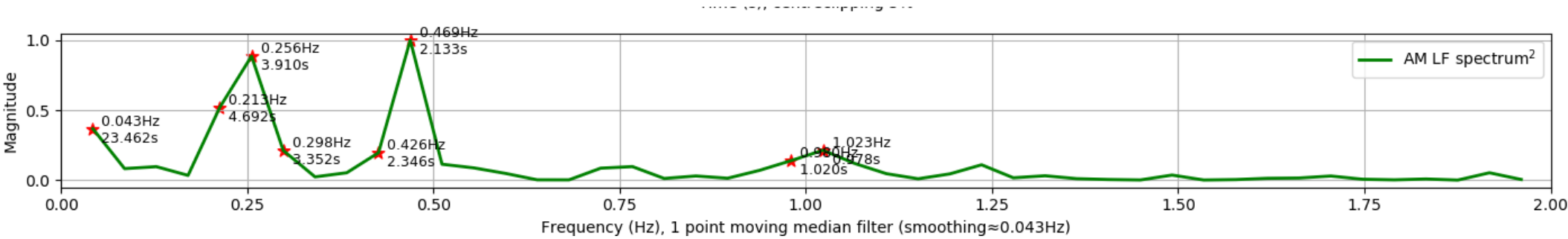
*Cyclical rhythm hierarchy as partial correlate of 'Prosodic Hierarchy'*

# *Different dialogue cycles – different rhythms*

## Narrative turn: single dominant rhythm formant



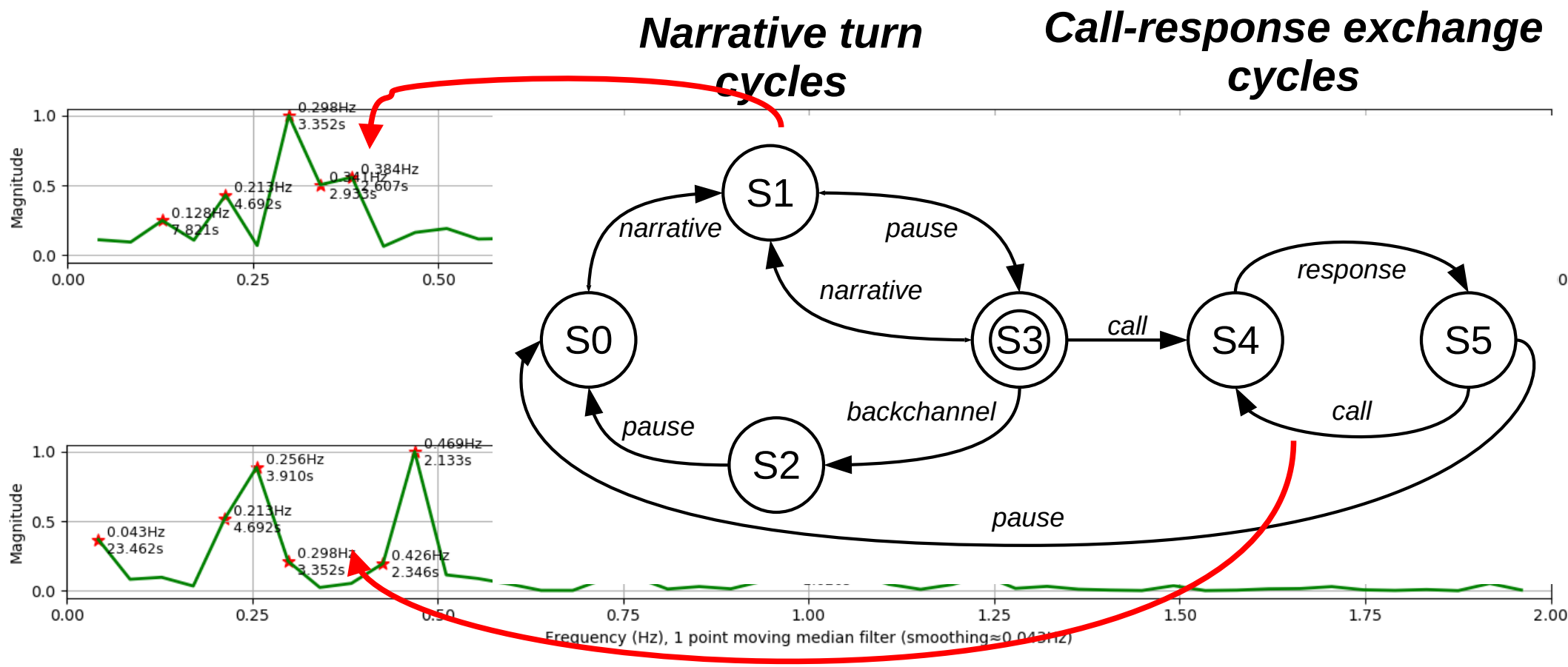

## Call-response exchange: two dominant rhythm formants




**Note the different cyclical structures:**

- Low frequency for narrative cycles
- Lower frequency for exchange (adjacency pair) cycles
- Higher frequency (octave higher) for exchange constituents

# *Different dialogue cycles – different rhythms*



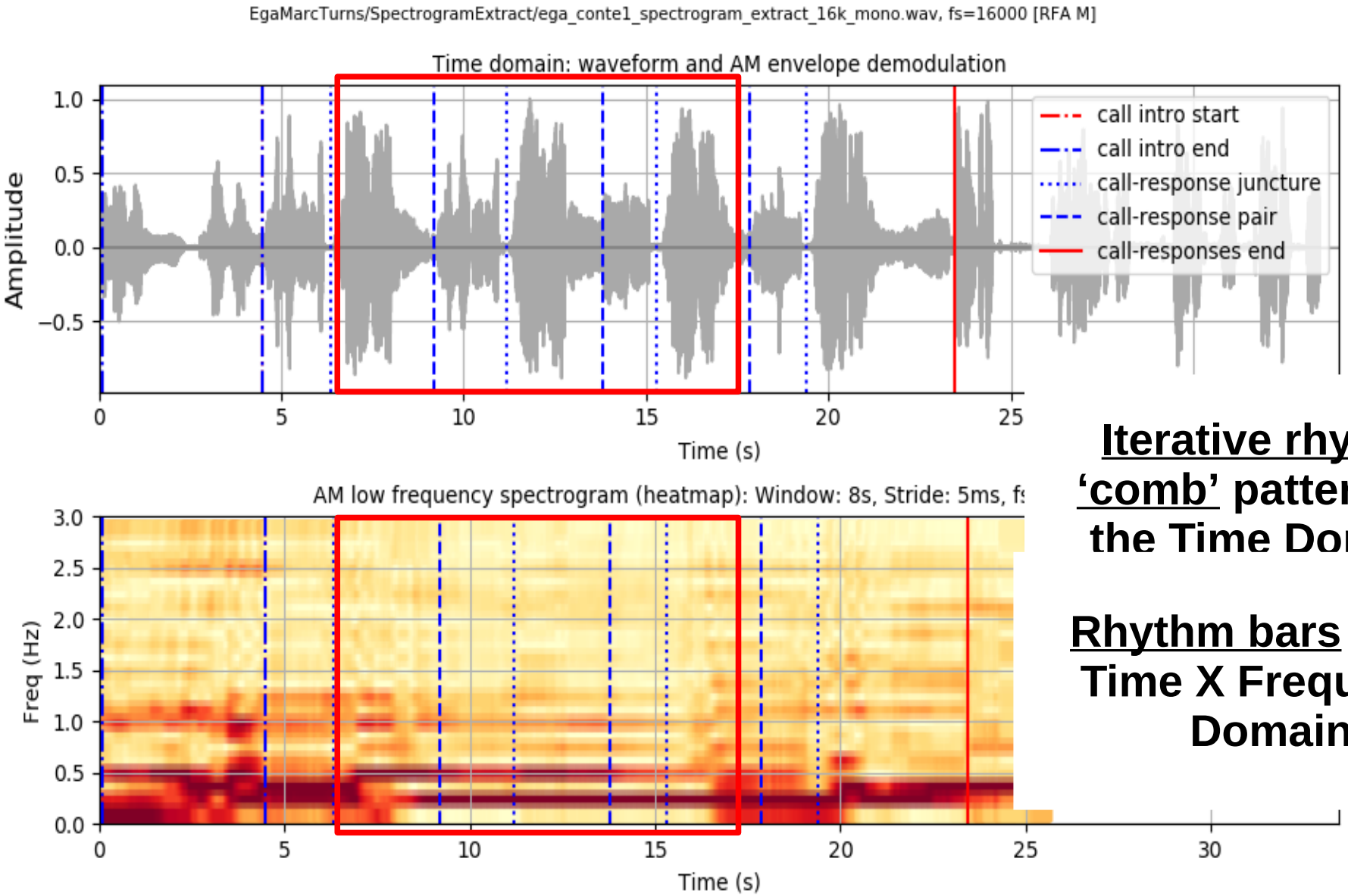

**Note the different cyclical structures:**

- Low frequency for narrative cycles
- Lower frequency for exchange (adjacency pair) cycles
- Higher frequency (octave higher) for exchange constituents

## *Restoring the time domain: LF AM spectrograms*

# *Restoring the time domain: AM spectrograms*

First Caller-Choir exchange

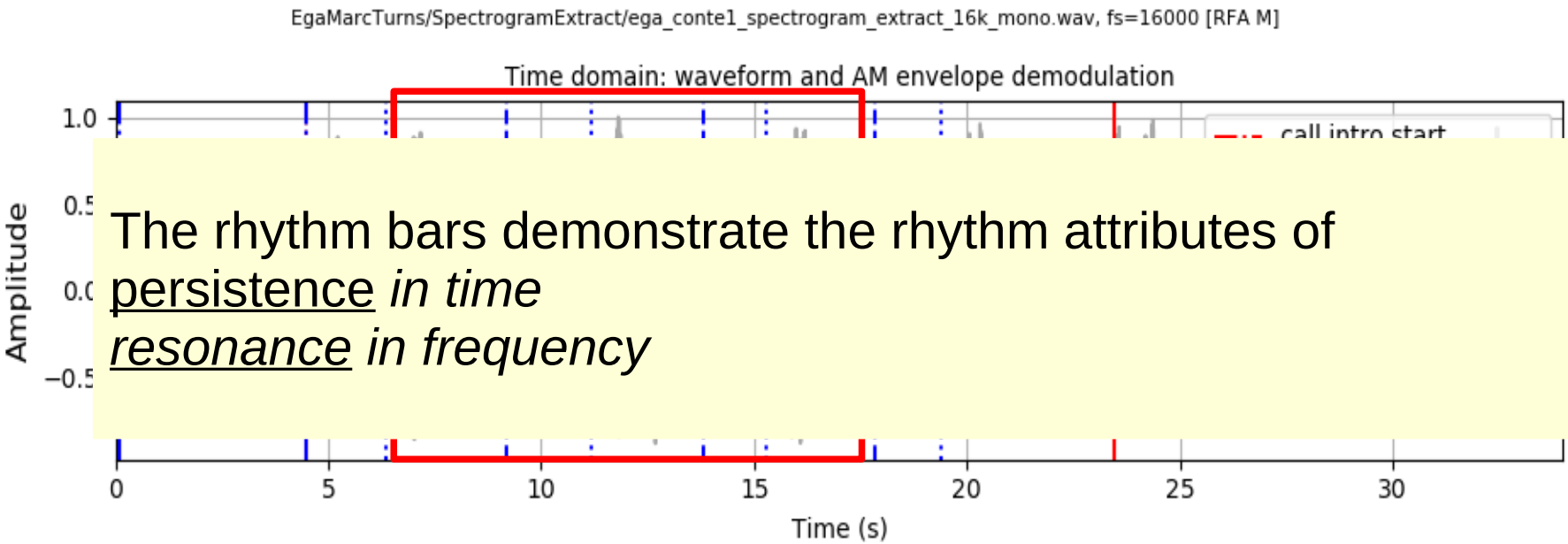



**<u>Iterative rhythm</u> <u>'comb'</u> patterns in the Time Domain**

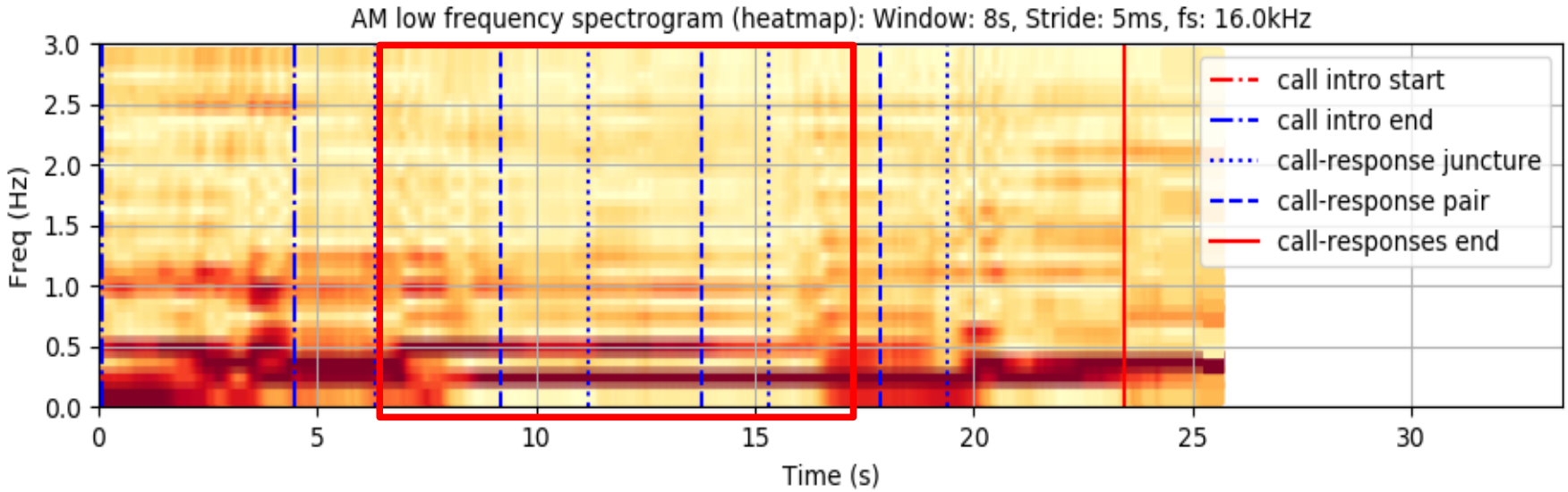
**<u>Rhythm bars</u> in the Time X Frequency Domain**

# *Restoring the time domain: AM spectrograms*

First Caller-Choir exchange




The rhythm bars demonstrate the rhythm attributes of
<u>persistence</u> *in time*
*<u>resonance</u> in frequency*



***How can the insights from demodulation be generalised?***

***From Analysis to Comparison of Rhythms***

***with spectrum and spectrogram differences using distance metrics and clustering***

# *From Analysis to Comparison*

## Extraction of attribute vectors from

**Low Frequency global spectrum**

- LF spectrum as vector
- *N* highest magnitude frequencies
  - Magnitudes of these frequencies
- *N* highest magnitude frequency peaks
  - Magnitudes of these frequencies

**Low Frequency global spectrogram**

- Trajectory of highest magnitude peak frequencin each spectral slice
  - Trajectory of magnitudes of these peaks

# *Method 4: Vectors, distances, clusters*

## Comparisons of different languages and styles

**Orature:**

**Ega** (ISO 639-3 *ega*), putative Kwa, …, Niger-Congo

**Anyi** (ISO 639-3 *ani*), Kwa, …, Niger-Congo

**Readings**:

**Ibibio** (ISO 639-3 *ibb*), Lower Cross, …, Niger-Congo

News in **Ivory Coast French**, Anyi reader

*Opportunistic corpus:*

*Fieldwork was cut short by civil war In the Ivory Coast*

# *Method 3: Vectors, distances, clusters*

## Comparisons of different languages and styles

**Orature:**

**Ega** (ISO 639-3 *ega*), putative Kwa, …, Niger-Congo

**Anyi** (ISO 639-3 *ani*), Kwa, …, Niger-Congo

**Readings**:

**Ibibio** (ISO 639-3 *ibb*), Lower Cross, …, Niger-Congo

News in **Ivory Coast French**, Anyi reader

## Method:

**Input:**

Spectrum and spectrogram vectors

**Procedure:**

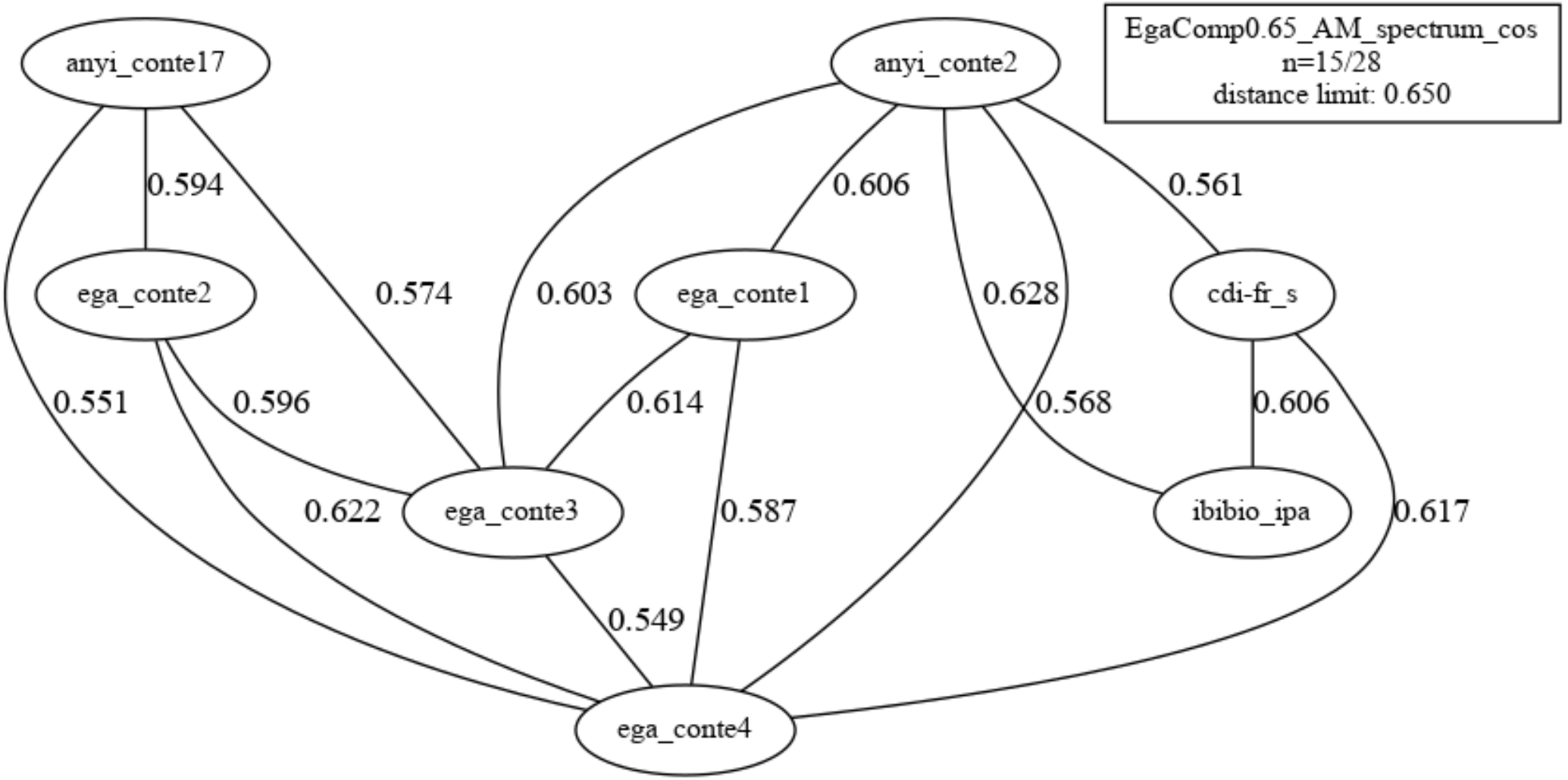
Pairwise distances: Cosine, Manhattan

**Output:**

Distance networks → hierarchical clusters (dendrograms)

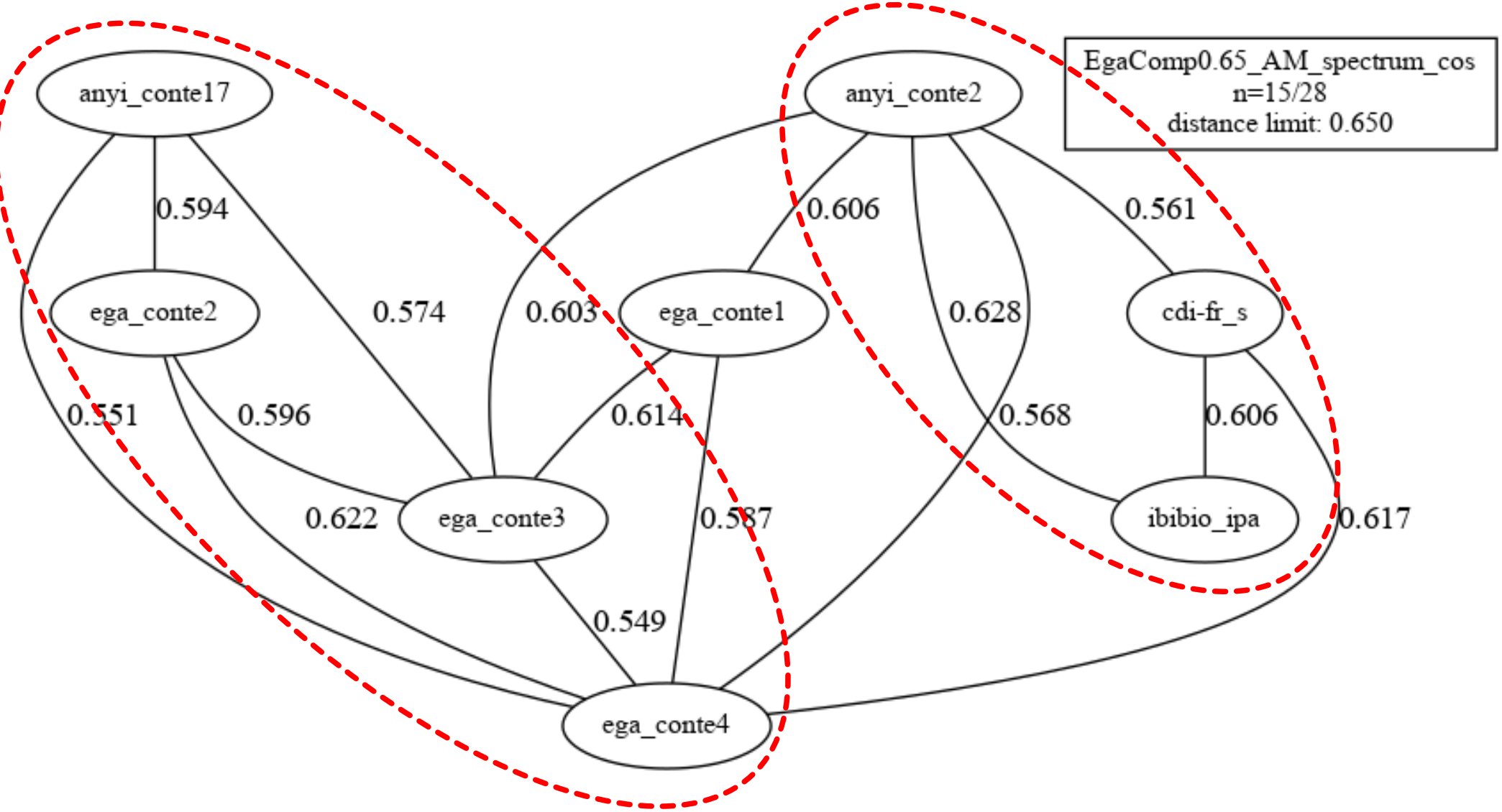
# *Rhythm Formant Analysis of related registers*

## Cosine Distance Network, distances < 0.65 (scale: 0...1)



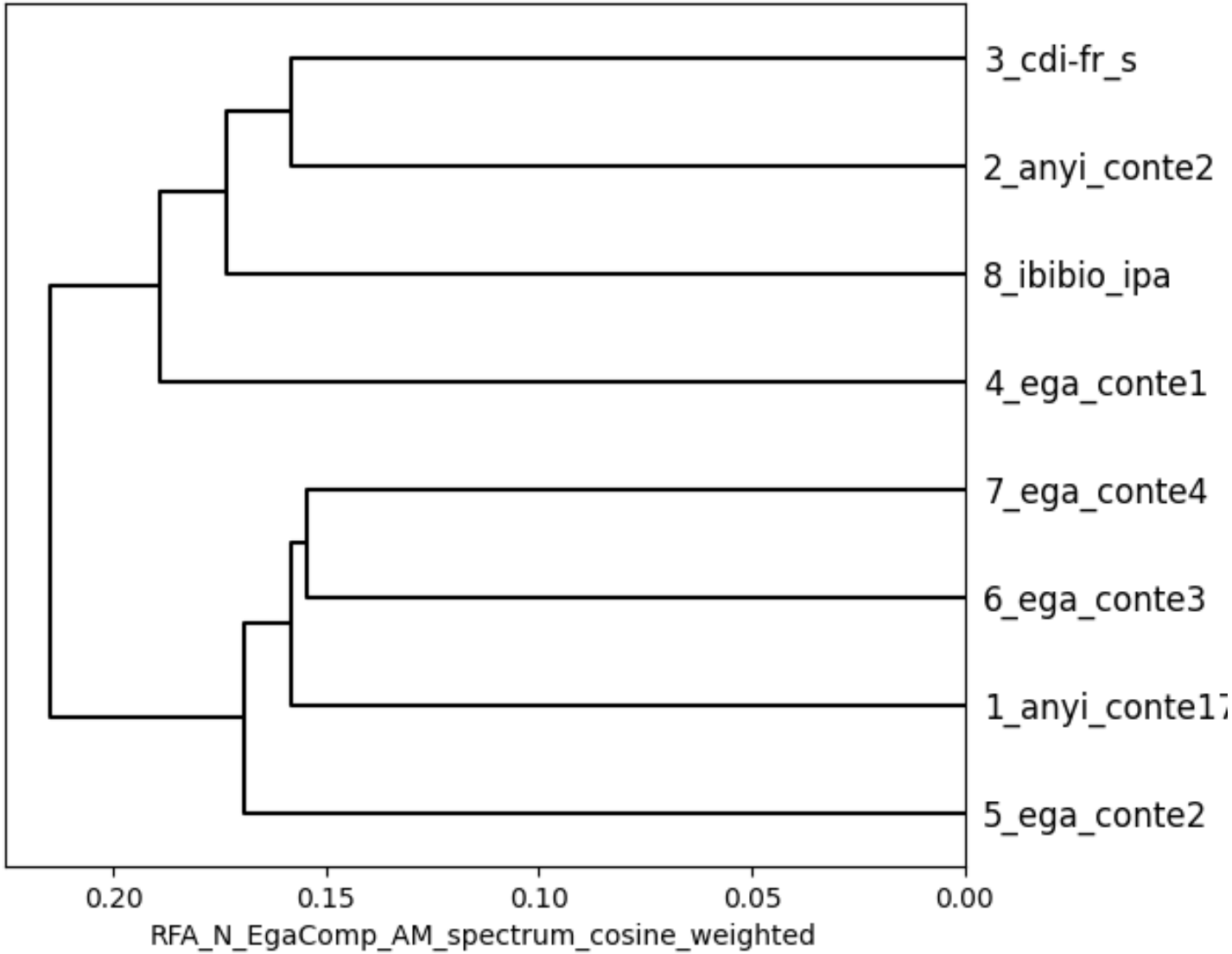
# *Rhythm Formant Analysis of related registers*

## Cosine Distance Network, distances < 0.65 (scale: 0...1)



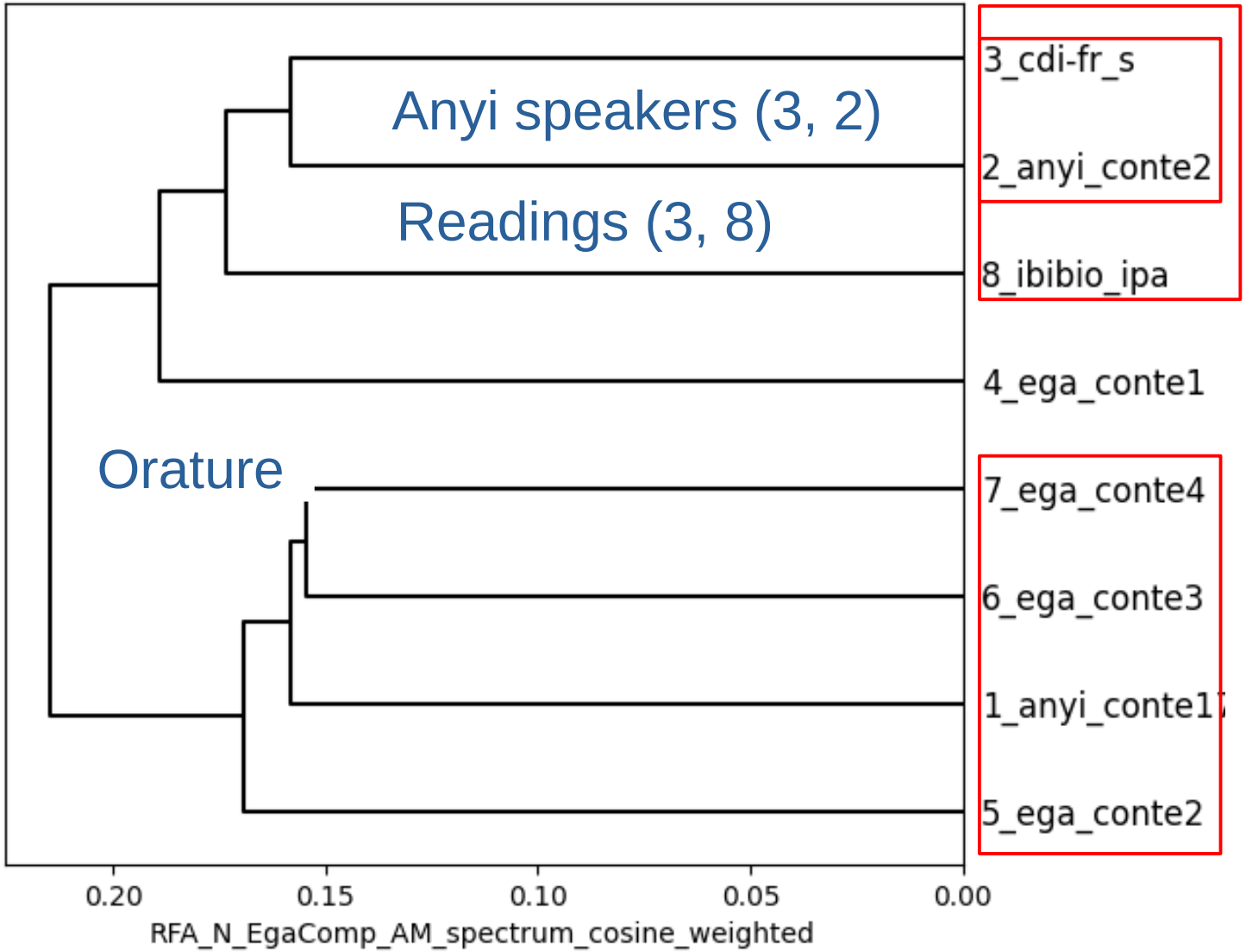
# *Rhythm Formant Analysis of related oral registers*

## Hierarchical clustering
## Cosine Distance & Weighted Mean



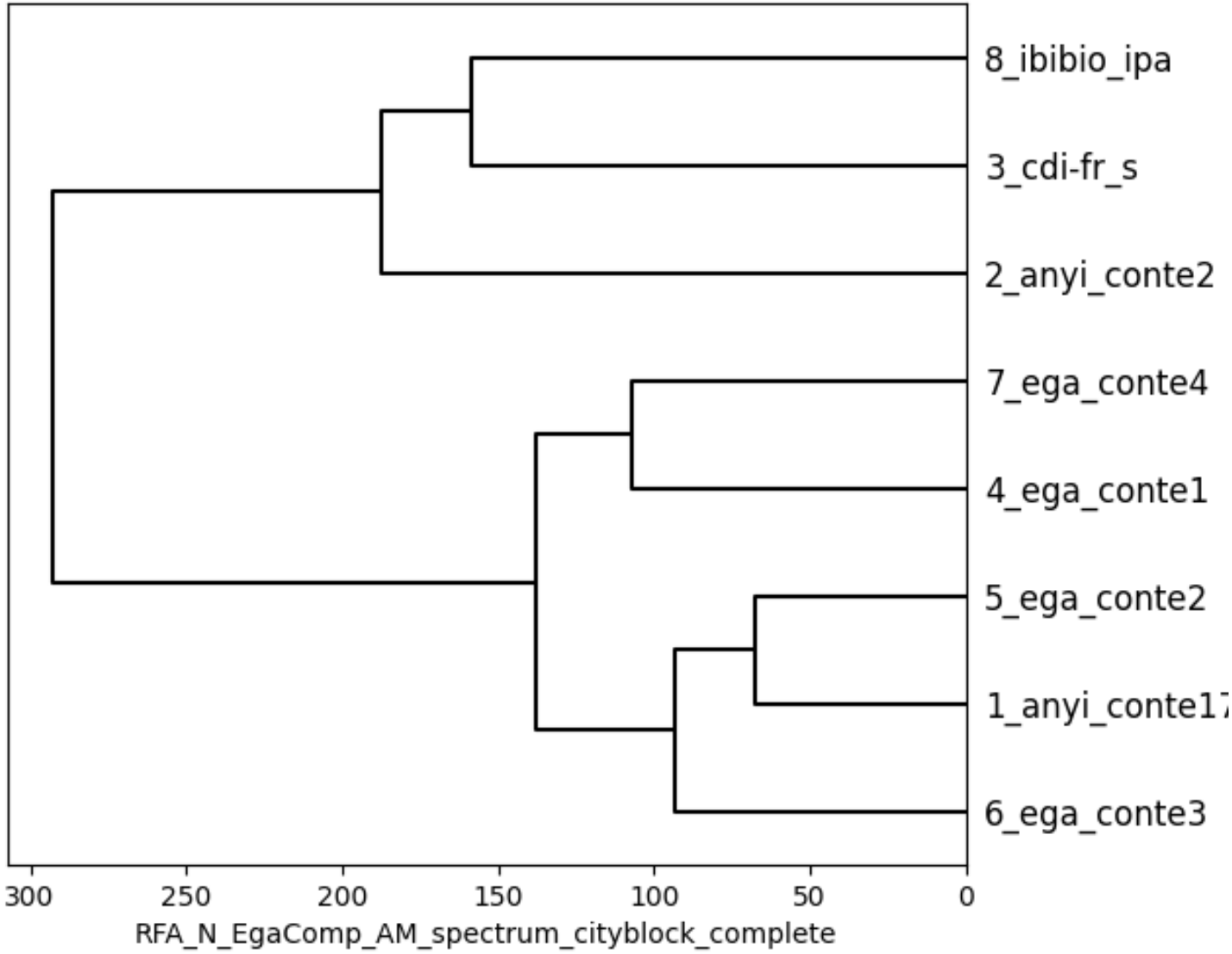
# *Rhythm Formant Analysis of related oral registers*

## Hierarchical clustering
## Cosine Distance & Weighted Mean



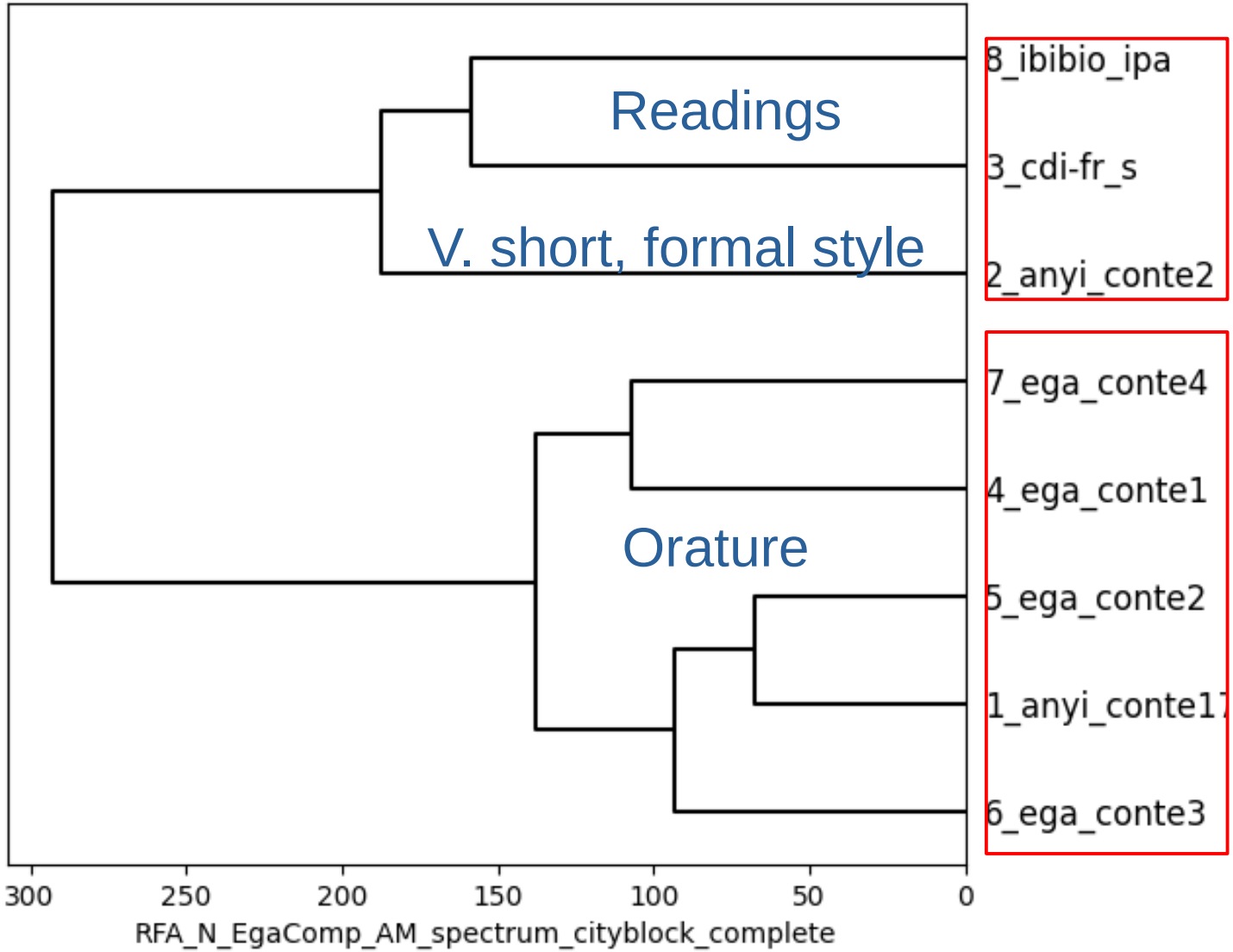
# *Rhythm Formant Analysis of related oral registers*

## Hierarchical clustering
## Manhattan Distance & Farthest Neighbour



# *Rhythm Formant Analysis of related oral registers*

## Hierarchical clustering
## Manhattan Distance & Farthest Neighbour

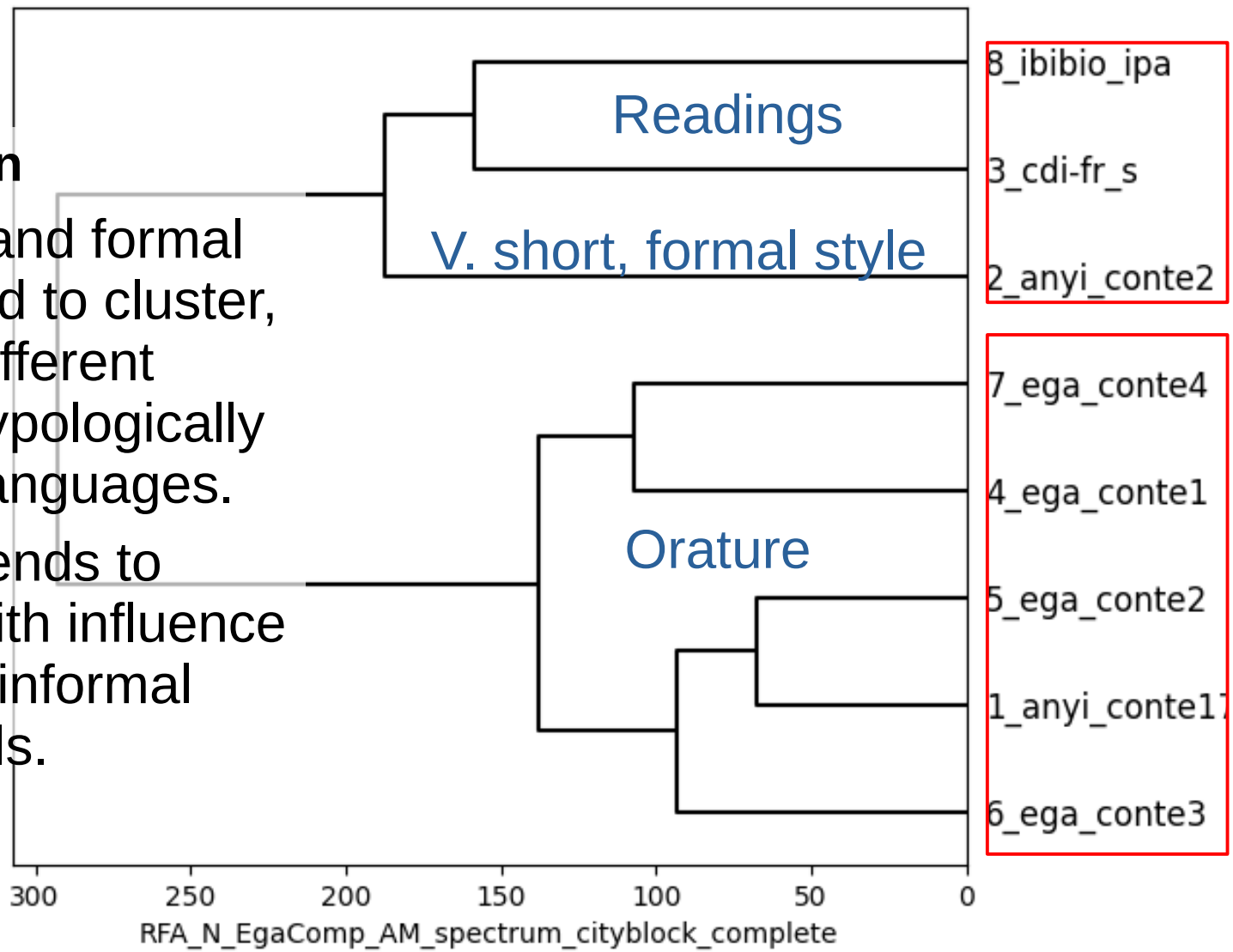


# *Rhythm Formant Analysis of related oral registers*

## Hierarchical clustering
## Manhattan Distance & Farthest Neighbour

**Interpretation**

1. Reading and formal styles tend to cluster, even in different (though typologically related) languages.
2. Orature tends to cluster, with influence of formal-informal style levels.



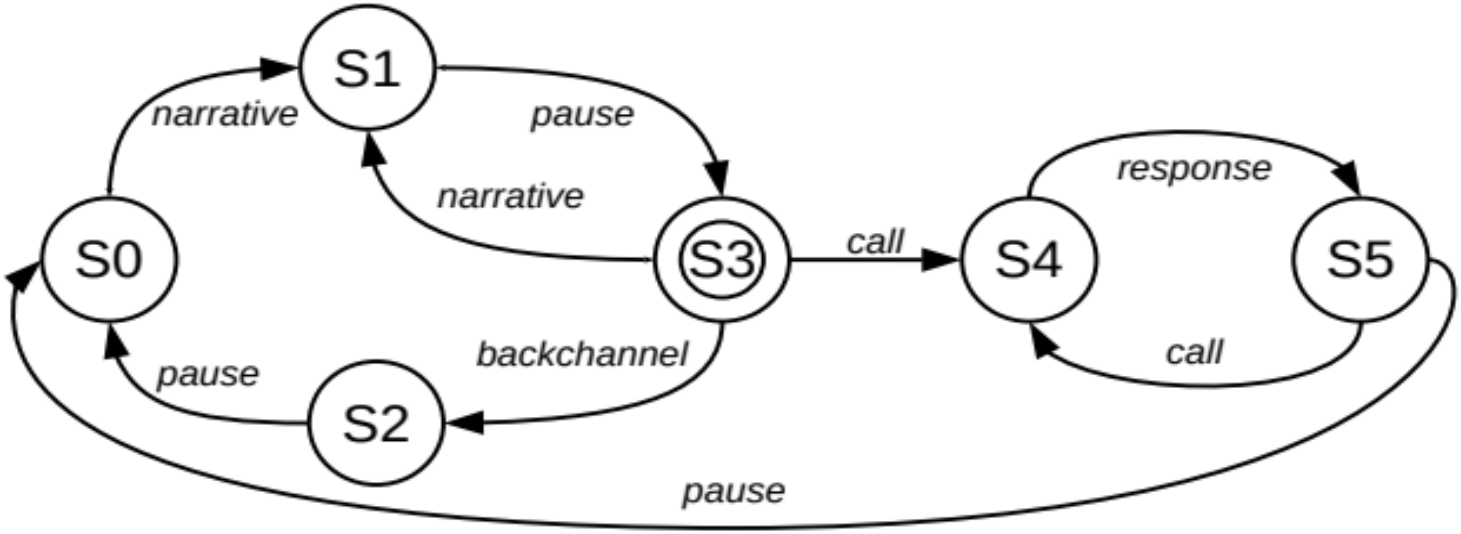
# *Conclusion*

## Form:

1. There are **multiple parallel speech rhythms**, not a single rhythm
2. Speech rhythms relate to linguistic units in **cycles**
3. Multiple speech rhythm beat levels / cycles can be modelled as regular languages by **cyclical FTNs**, **FSTs**, etc.

## Function

1. *Semantics*: Speech rhythms are **metalocutions** which **denote** units in the locution at **temporal points and intervals**
2. *Pragmatics*: Speech rhythms depend on **register**, **style** and **discourse** (as well as the lexical and phrasal grammar)



**The moral of the story:**

***Combine fieldwork with music and engineering!***

# ***Thanks!***