

Rhythm - a quick look.

1. Isochrony
2. Emergent rhythm
3. Perceptual rhythm

1. Isochrony

The claim: Languages come in three types:

mora timed
syllable timed
stress timed

According to their chosen unit of isochronous timing.

An example of isochrony.

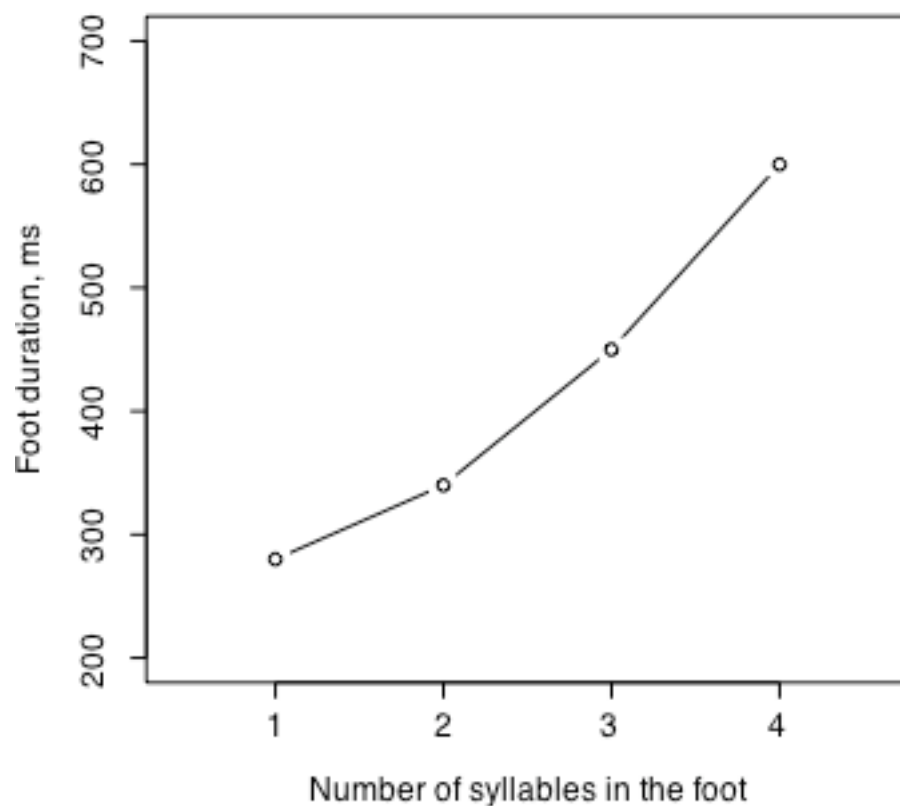
Consider the click of a metronome -

X-----X-----X-----X-----X-----X-----X

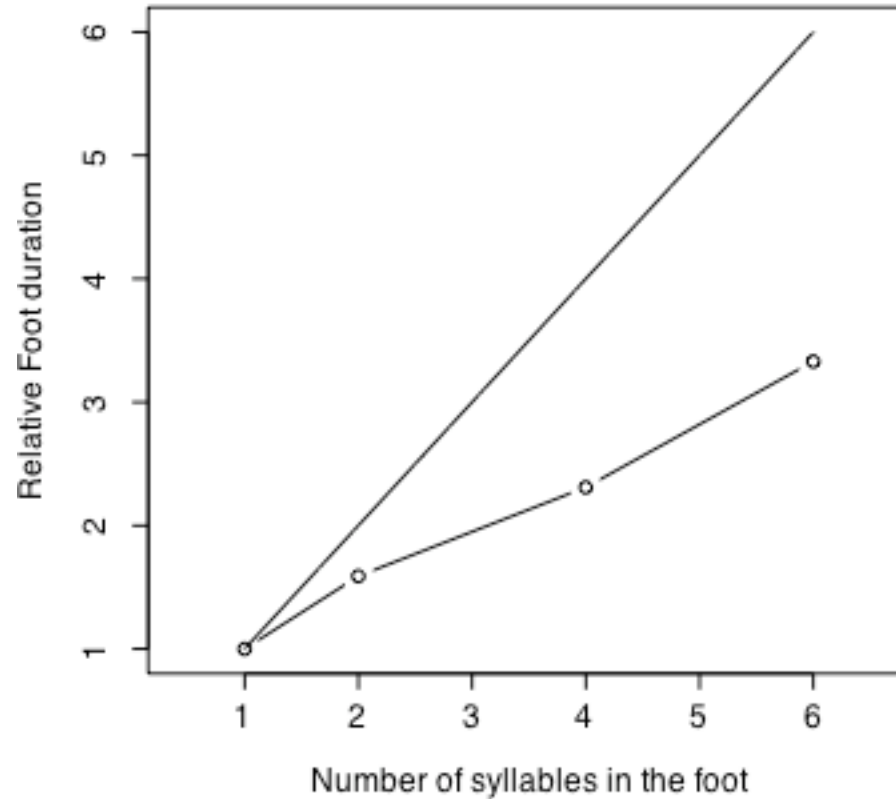
The interval between each click (x) is the same (6 dashes), so this sequence of clicks is isochronous.

Thus, the isochrony claim for English is that the intervals between stressed syllables (foot duration) should be equal.

Here is some data from Nakatani, O'Connor & Aston (1981) that refutes this claim.

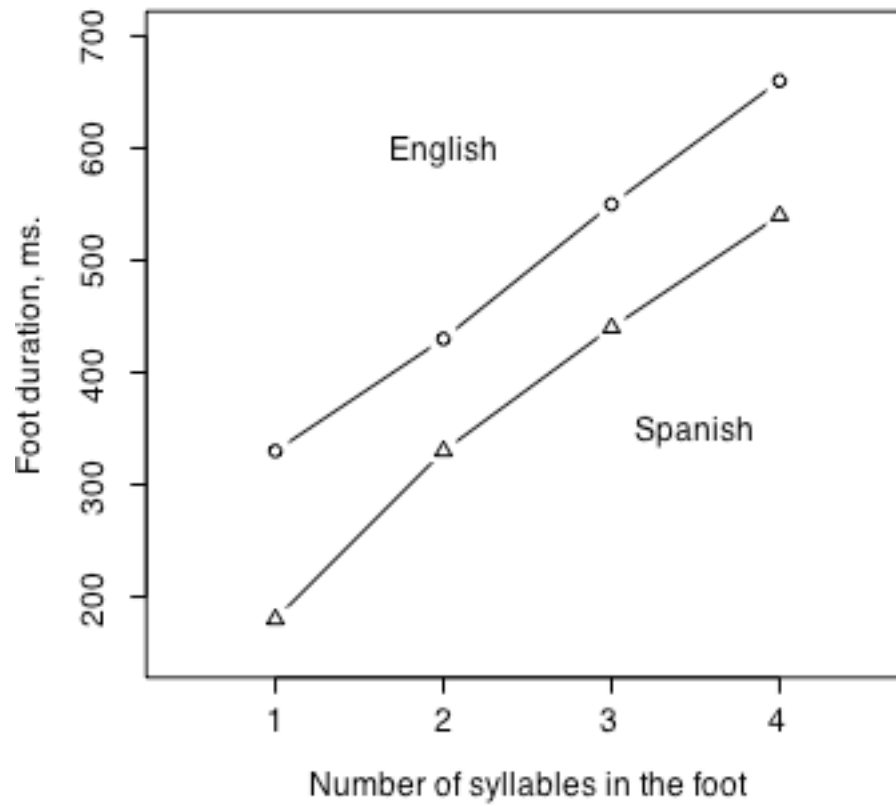


Data on the lack of syllable isochrony in Spanish



PAMIES BERTRÁN, A. (1999) **prosodic Typology: On the Dichotomy between *Stress-Timed* and *Syllable-Timed* Languages** *Language Design* 2 (1999, 103-130)

Direct comparison of interstress intervals by Dauer (1983)



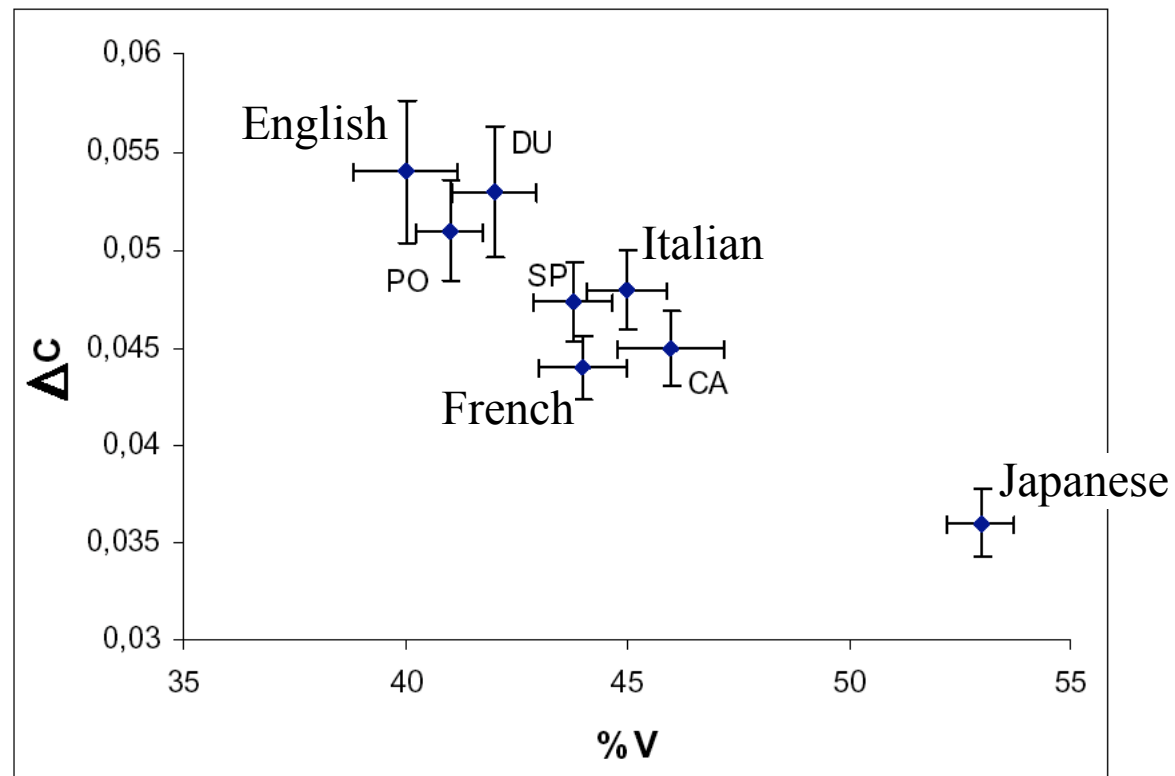
Emergent Rhythm.

Dauer (1983) - linguist's impression of rhythm is due to:

- vowel quality in stressed and unstressed syllables
- consonant clusters
- regularity/predictability of stressed syllables

$\%V$ = the proportion of vocalic intervals in the sentence

ΔC = the standard deviation of consonantal intervals within the sentence



Ramus, F. and Nespors, M. and Mehler, J. (1999) Correlates of linguistic rhythm in the speech signal. *Cognition*, 73, 3, 265-292.

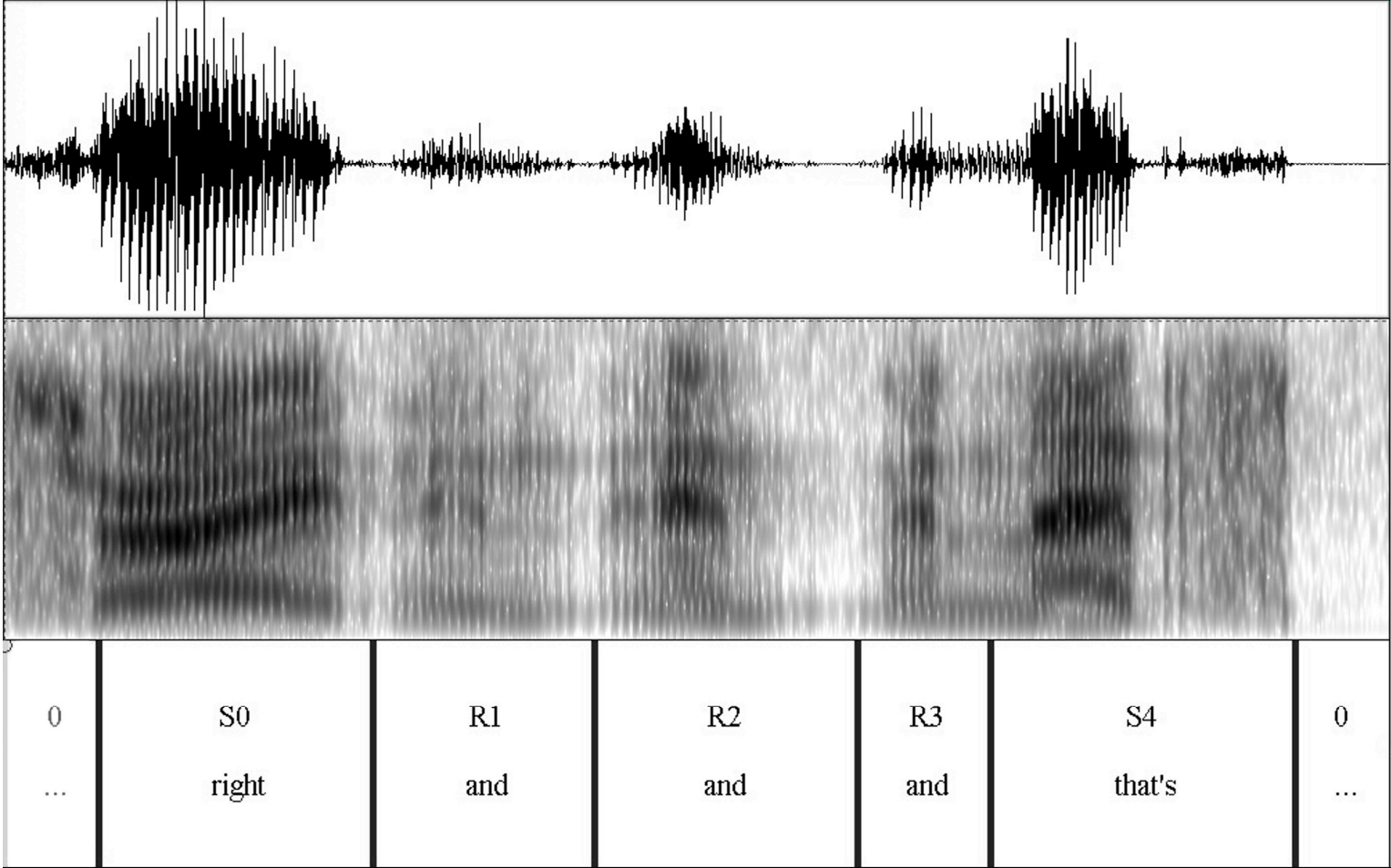
This suggests that Bauer's hypothesis is right - segment sequence properties correlate with "stressed timed" and "syllable timed"

%V - do we tend to have as much time for vowels as for consonants?

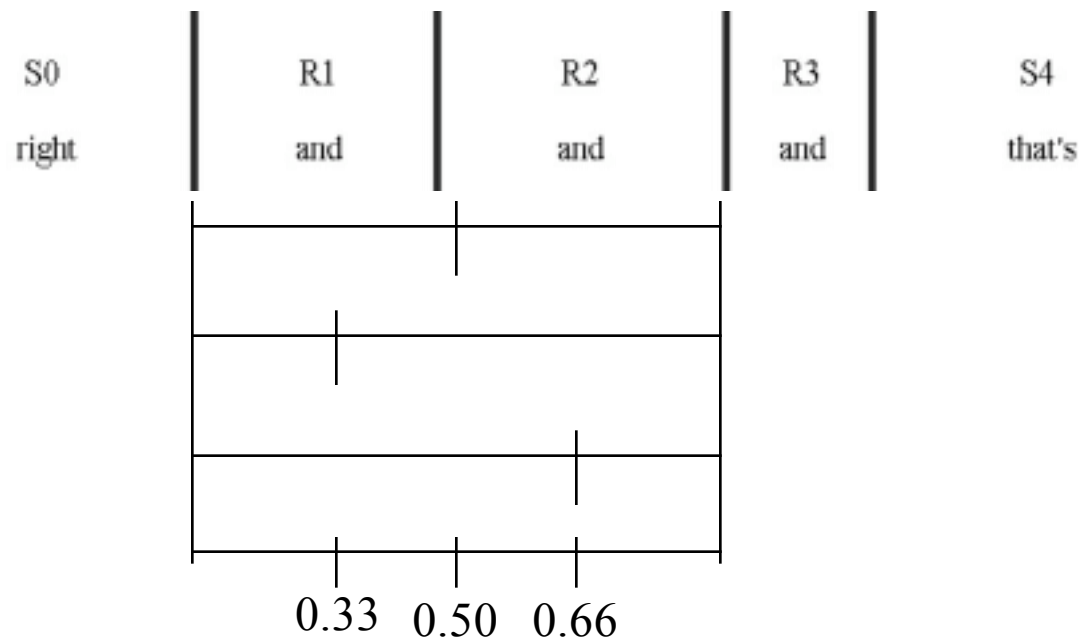
ΔC - do we have a mixture of consonant singletons and clusters?

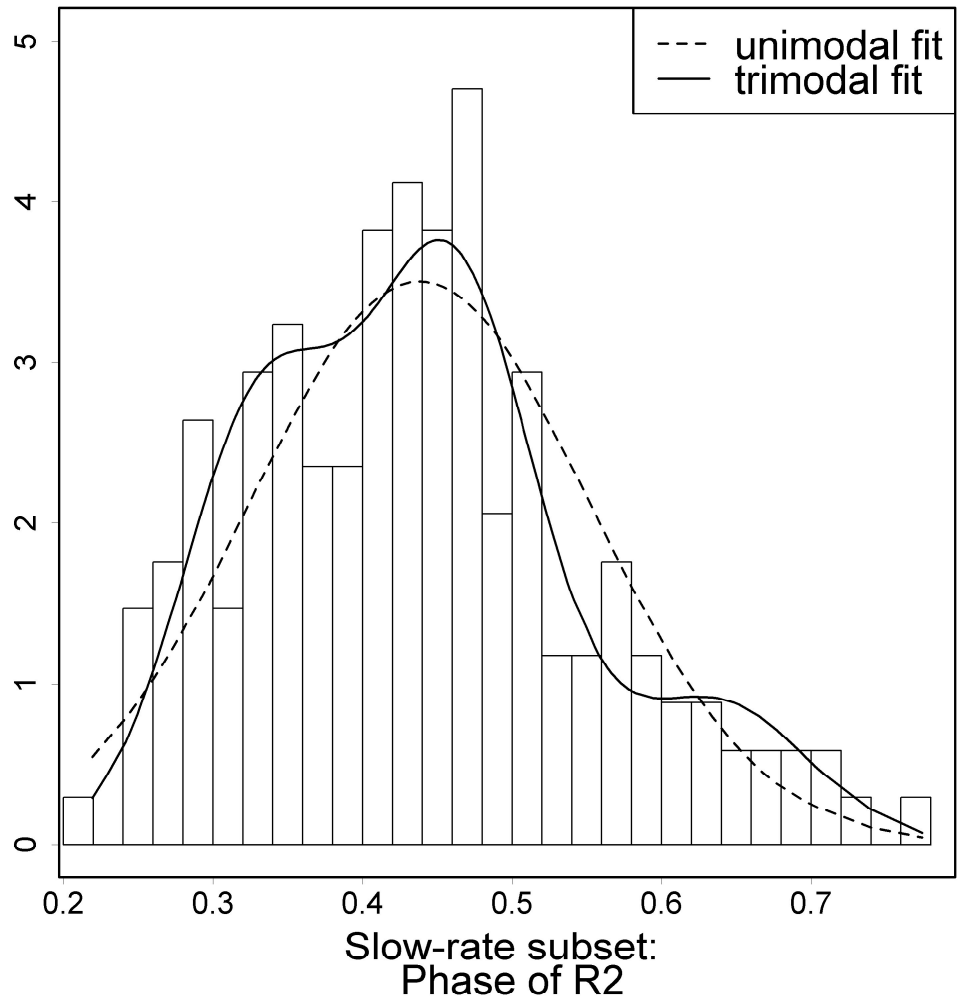
But, Is there rhythm in speech production???

There is room for research here. For example recent research at UCB on the timing of "three-peaks" in conversational speech (Sam Tilsen).

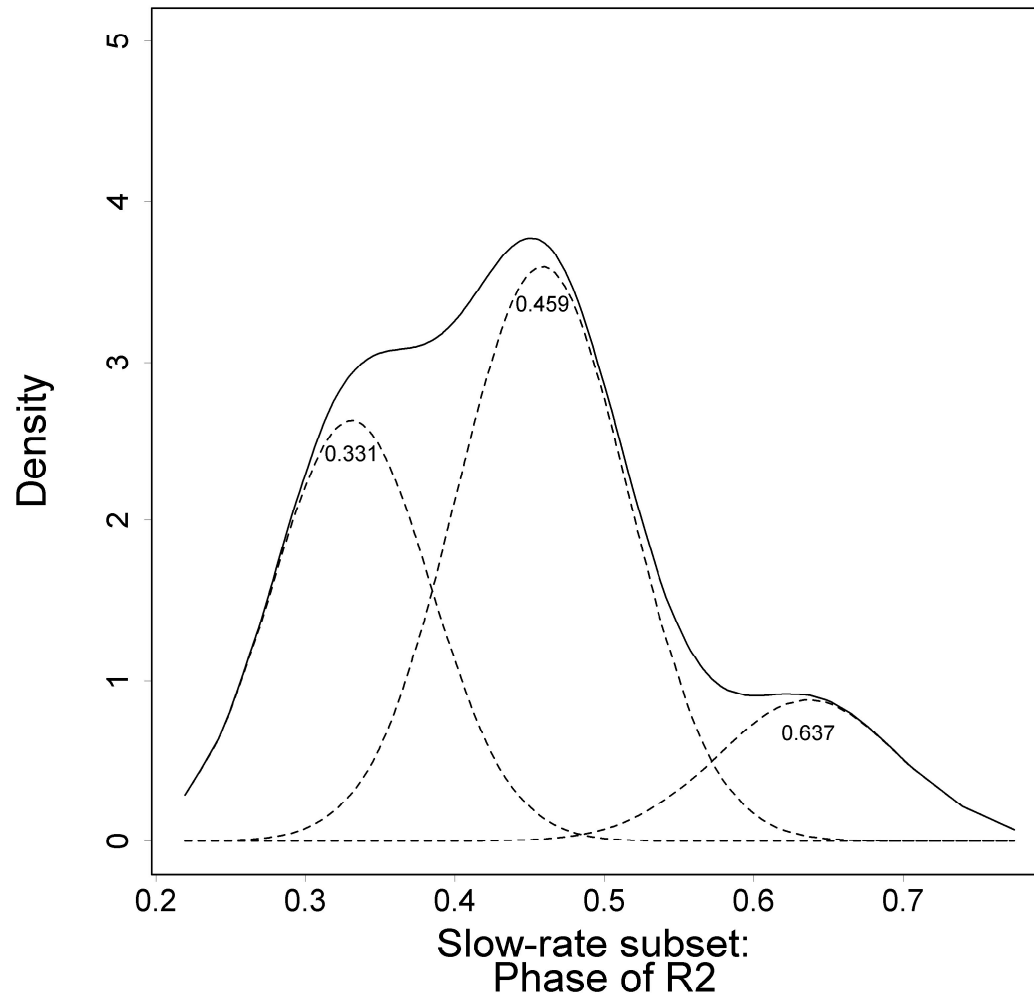


Hypothesis: If three-peaks are on “beats” in the hesitation interval, then the onset of R2 should fall in a non-Gaussian multi-modal distribution.





N1= 59 mean= 0.331 sd= 0.053
N2= 87 mean= 0.459 sd= 0.057
N3= 23 mean= 0.637 sd= 0.061



Perhaps speech is rhythmic after all, but measuring foot duration or syllable duration doesn't capture speech rhythm.