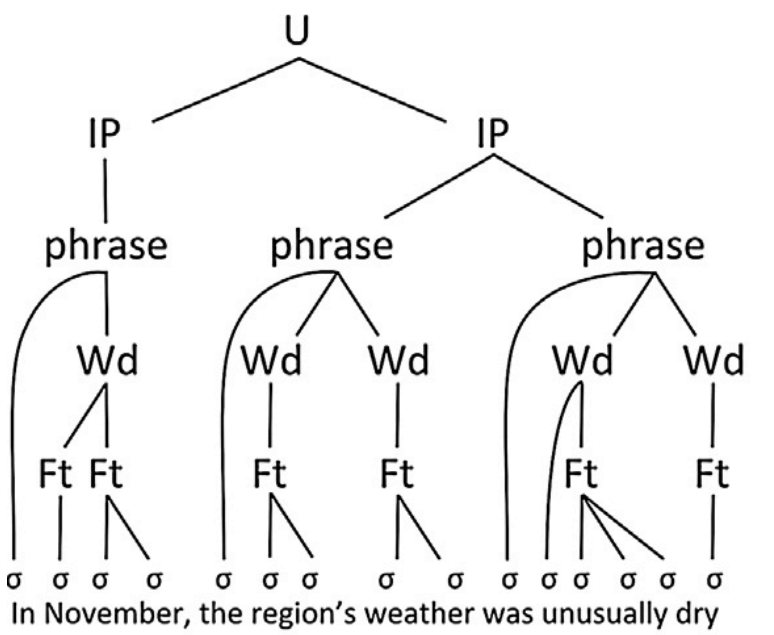
Prosody:
Temporal Structure

Functions of of Prosody

- Phrasing (Grouping)
  - When you make hollandaise slowly, it curdles.
    When you make hollandaise, slowly it curdles.
- Prominence (focus)
  - She didn't earn an A. (she earned a B).
    She didn't earn an A. (but she got one by cheating).
- How are these functions accomplished by speech system?
  - Temporal structure
  - Intonation

Prosodic Hierarchy

- Speech gestures are organized into larger units, beginning with syllables
- No agreement on exactly how many categories (e.g. intermediate phrase (ip).



Utterance

Intonational phrase

Phonological phrase

Prosodic word

Foot

Syllables

Prosodic domains above the word: influenced by syntactic structure as well as rhythm

What are the temporal signatures of phrase boundaries are how are they controlled?

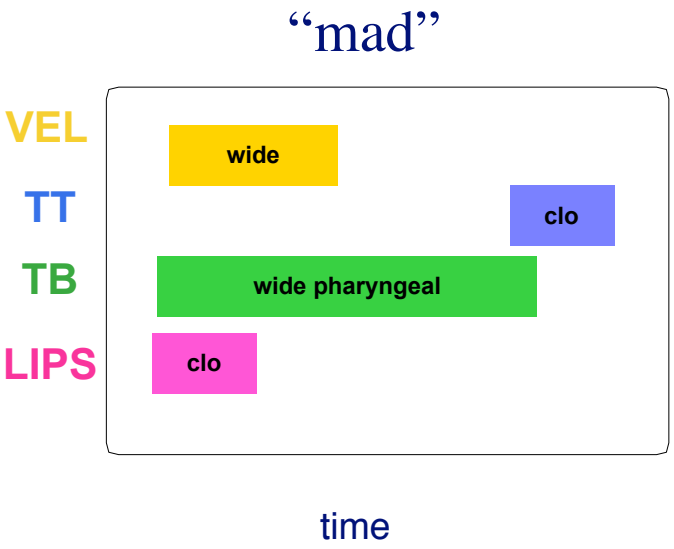
Prosodic domains below the word: influenced by rhythmic/lexical properties

What are the temporal signatures of foot structure are how are they controlled?

Mol, Carien & Chen, Aoju & Kager, Rene & Ter Haar, Sita. (2017). Prosody in birdsong: A review and perspective. Neuroscience and biobehavioral reviews. 81. 10.1016/j.neubiorev.2017.02.016.

To keep in mind…

- Given the gestural score for a given syllable, what changes in the score could result in lengthening or shortening of the syllable?



- Changes in activation intervals
- Changes in relative timing

Boundary-adjacent lengthening

- Vowels (and consonants) are longer at the end of phrase (or at the beginning of a phras)e) than within a phrase (in English and many languages.

- I gave a duck to Doug.
  I gave Doug a duck.

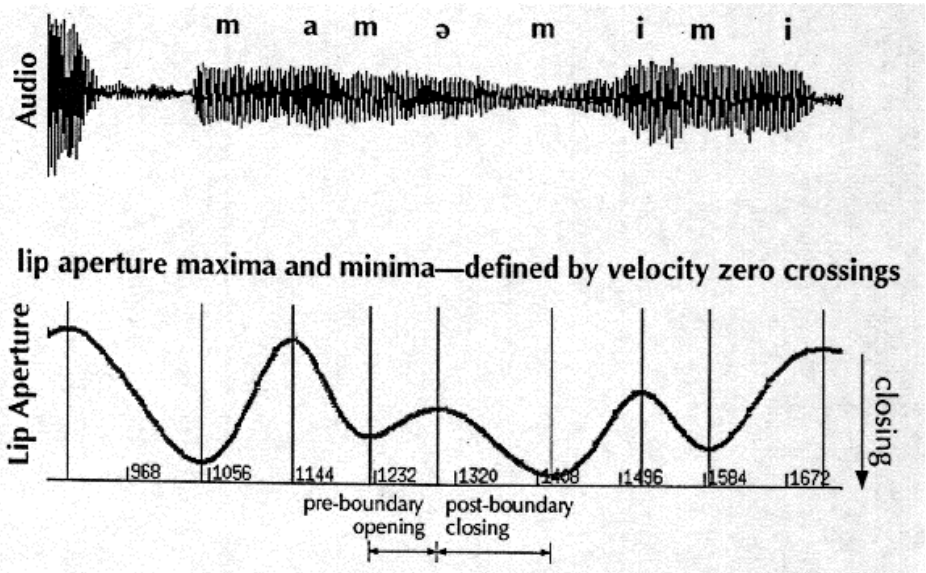
- What kind of boundaries produce lengthening?

- How is the lengthening controlled?

Byrd & Saltzman (1998)

TABLE I. Stimuli sentences for five experimental boundary conditions
(Boldface was not present in the stimuli seen by subjects)

Boundary condition	Sentence
none (word-medial)	Poppa begged "mom**ma**mia" meanly upon coming.
word	Poppa-Pikt and Momma-Mimi tapped Coby.
list	Poppa, Pikt, Momma, Mimi, and Bibi tapped Coby.
vocative	Quick Momma, Mimi tapped Coby.
utterance	Poppa picked Momma. Mimi tapped Coby.



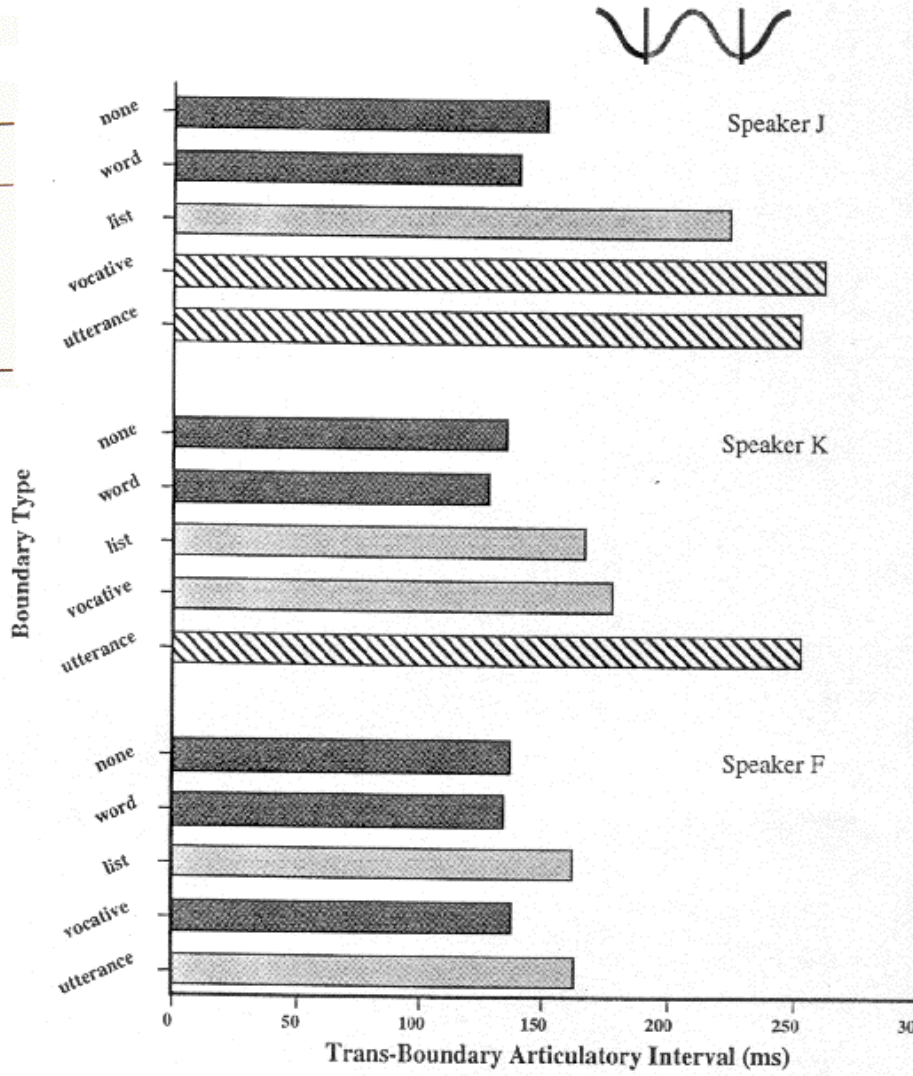


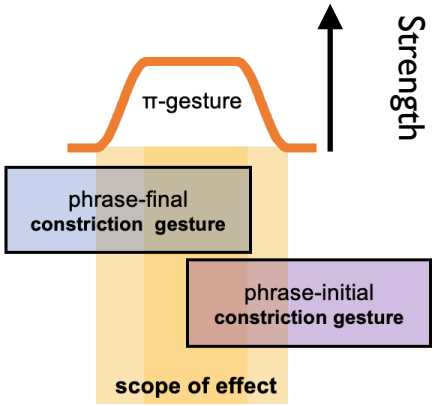
Figure 4. The transboundary articulatory interval for three speakers and five experimental boundary conditions. Bars with like shadings group together in post-hoc tests ($p < 0.005$).

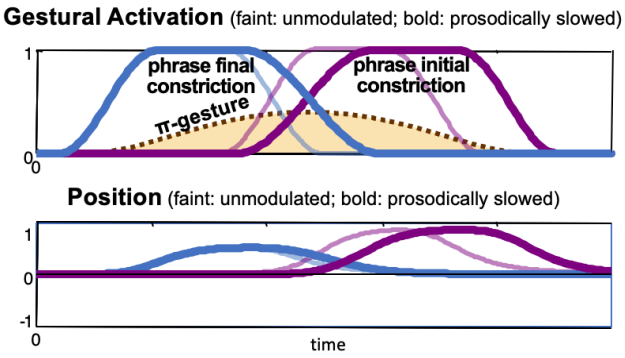
Results

- Subjects exhibit one or two levels of lengthening (compared to within-word).
- Relative lengthening of list, vocative, IP varies across speakers.
- How is it accomplished?
  - Constrictions approach their goals more slowly (appear to have greater stiffness).
  - Activation intervals are longer.

Gestural model of boundaries: π-gestures

- π-gestures are dynamical cognitive units that represent boundaries.
- Like other gesture they have extent in time (activation intervals) and dynamical parameters that govern how their strength waxes and wanes over time.
- Unlike other gestures, they have no articulators, but rather act vicariously to slow the clock that governs the activation dynamics of all the constriction gestures that fall within its scope.
- Slower activation results in:
  Gestures are lengthened
  Gesture onsets are delayed so overlap is reduced.
- Effects are proportional to π-gesture strength.
  - Hierarchically higher boundaries will have stronger activations and consequently there will be more lengthening and less overlap than at lower boundaries, thus accounting for the empirical findings that boundary lengthening is cumulative.





Byrd & Saltzman (2003)

Predictions of π-gesture model

- Effect is local.
- All gestures within scope are affected; no gestures are skipped.
- Effect is greatest at the boundary and decreases at a distance from it.
- Lengthening effects all segments, but could interact with type of segment.
- No difference between "final lengthening" and "initial lengthening"
- Categories: no intrinsic differences between phrases of different types (ip) (IP) (e.g. Beckman & Pierrehumbert, 1986); only strength determines effect.
    - Categories could result from multiple modal values of strength or From Boundary Tone Gestures (for IP, but not ip).

π-gestures as grammatical element

- Tempting to think of slowing as the speaker approaches the end of a phrase as a mechanical consequence of approaching a pause… slow down before stopping.

- But while evidence for boundary-related slowing has been found in several languages (e.g., English, French, Dutch, Greek, German, Spanish, Swedish), it is not universal by any means.

- In the eleven African tone languages from all over Africa in Downing & Rialland (2018):

  - None have final lengthening in declarative IPs

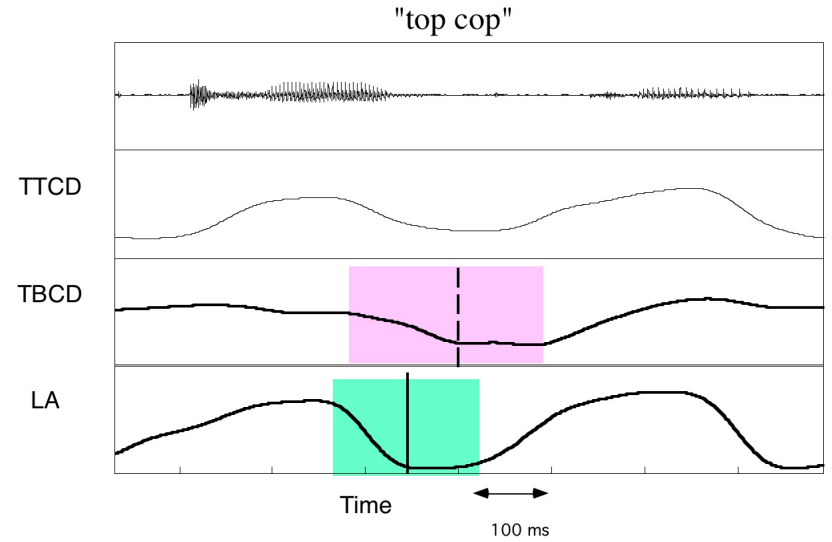
  - Shekgalagari (Botswana) and Tumbuka (Malawi) have penultimate lengthening (common in Bantu).

Reduction in overlap at boundaries

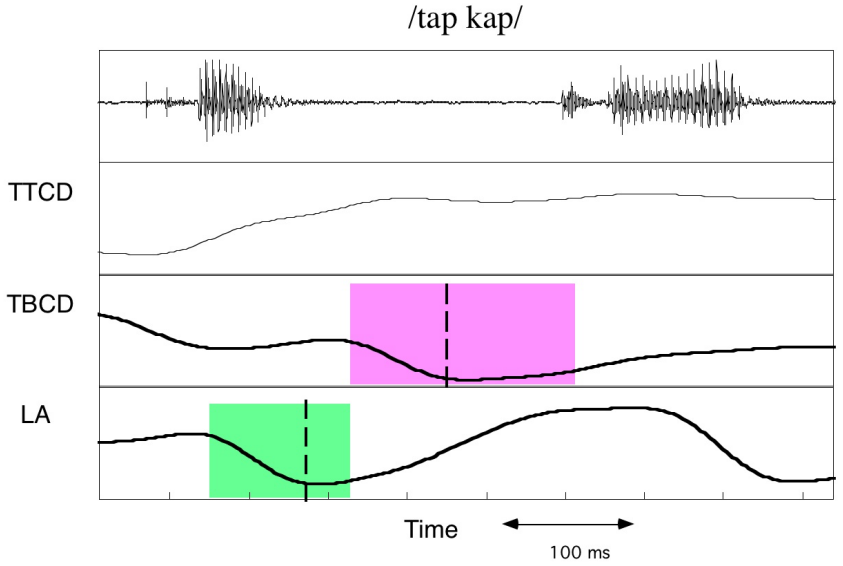
- Much phonetic variation is due to
  - variation in the temporal overlap of invariant gestures
  - Within a phrase, consonant gestures overlap across word boundaries in English.
  - Overlapping gestures jointly determine the output sound, so the acoustic output of a given gesture can vary depending on the gesture that overlaps it.
  - Such effects are less likely to occur across boundaries.
- Reduction of activation intervals of gestures within a phrase can also result in truncation: the gesture doesn't reach its goal when speaking rate is fast.

C#C overlap in English vs. Russian

English



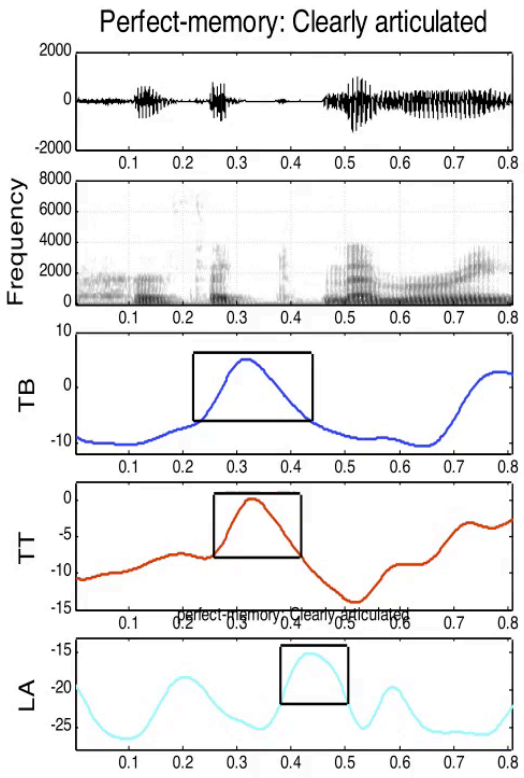
Russian



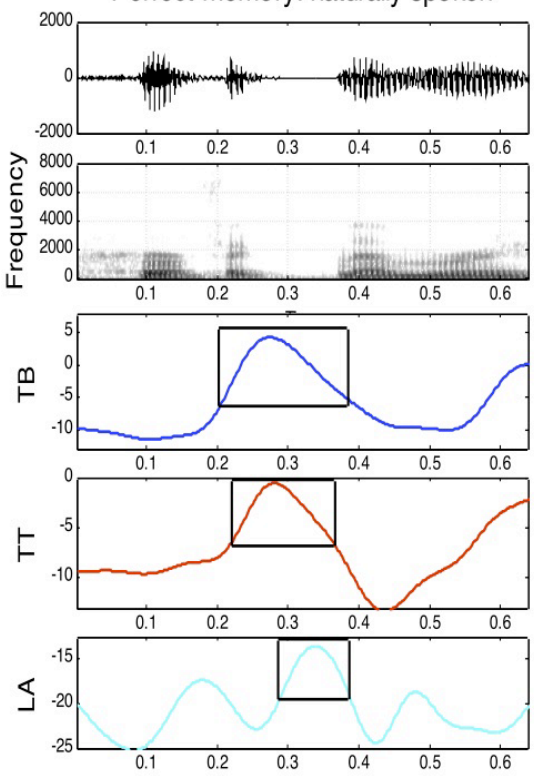
/t/ "deletion": "perfect memory" in three styles

/t/ sounds like it is deleted, but it is not

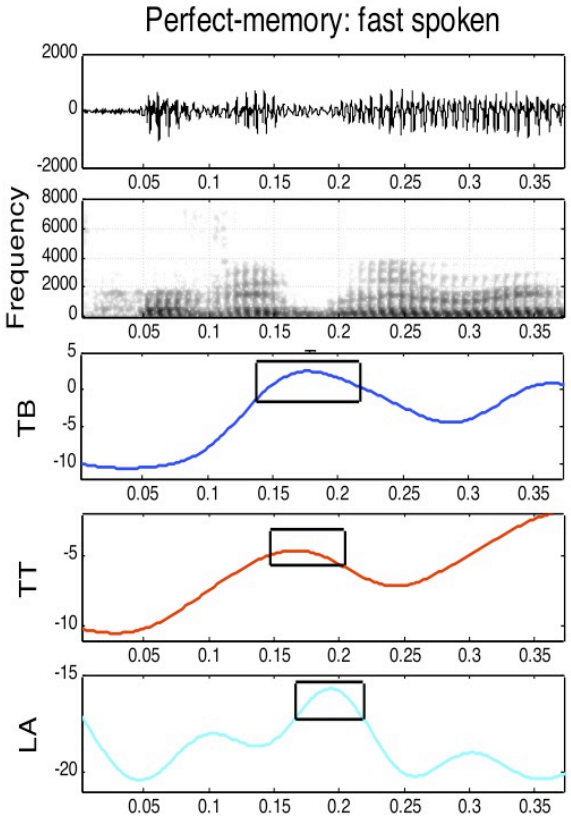
Boundary



Within-phrase



Fast



Release is overlapped by lip gesture and is hidden.

TT gesture is overlapped and truncated.

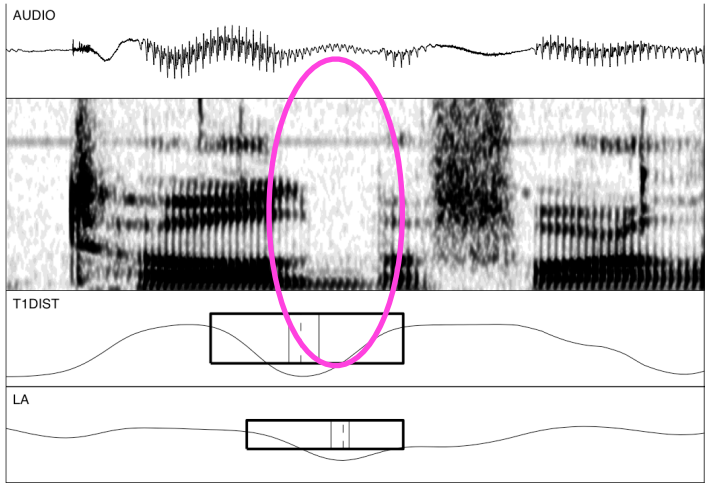
- “Pack my”

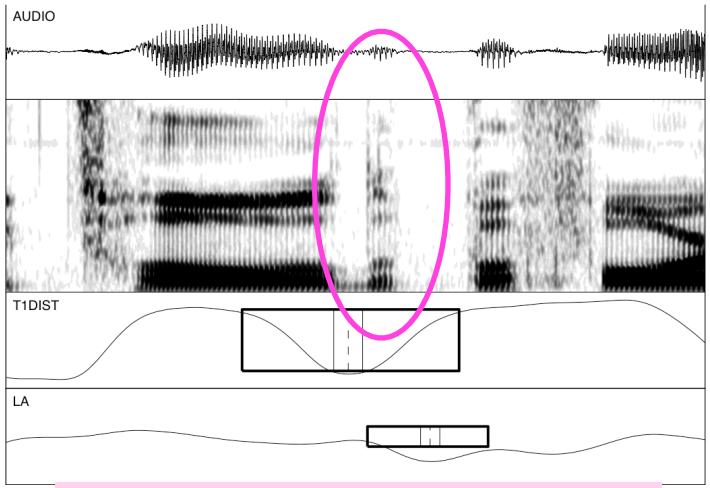
Modeling prosodic variation: "told before"

No [d] audible

[d] release

μbeam Talker 12





μbeam Talker 15

Apertures:

Tongue Tip

Lip

Only 2/38 Transcribers hear boundary

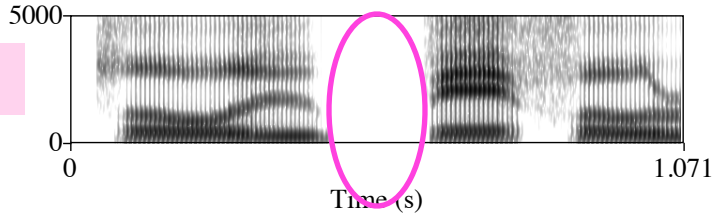
11/38 Transcribers hear boundary

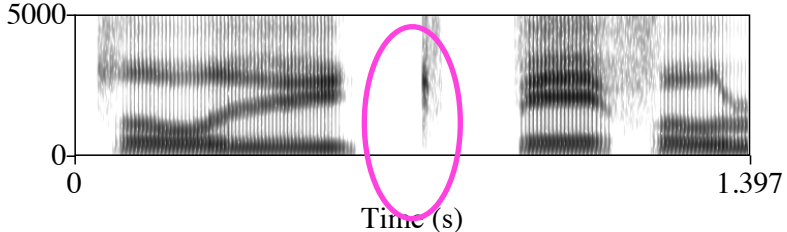
- Phrase was generated automatically by TaDA/HLsyn, then π-gesture slowing effects were added.
- [d] release emerged automatically with slowing, due to decrease in overlap of Tongue Tip and Lip Gestures.

NO π-gesture

π-gesture

TaDA:





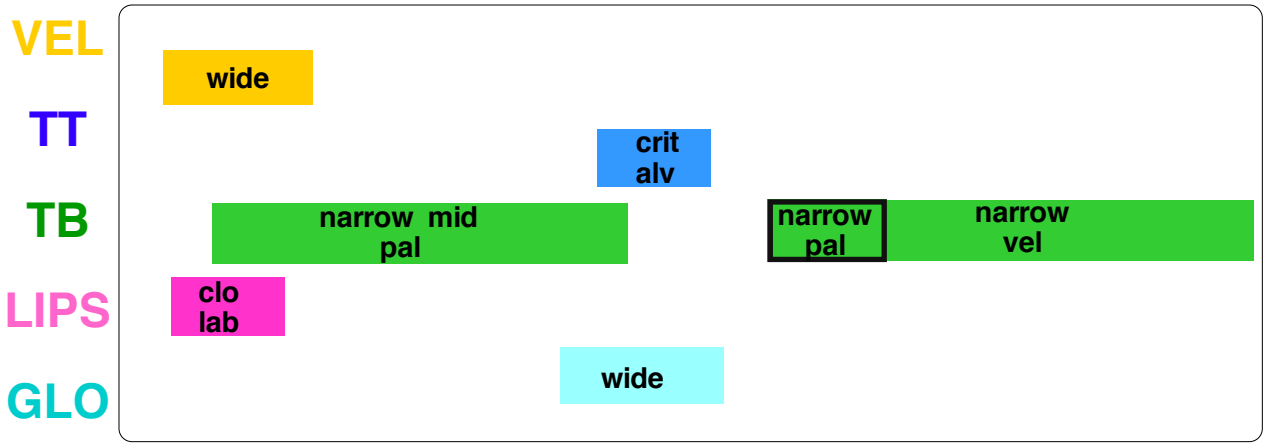
Palatalization of /s/

- /s/ within a phrase exhibits apparent change to [ʃ] before [j]
  - "miss it" [mɪs]
  - "miss you" [mɪʃ]
- Example
  - "I'm sure I'm gonna miss you"

slow fast

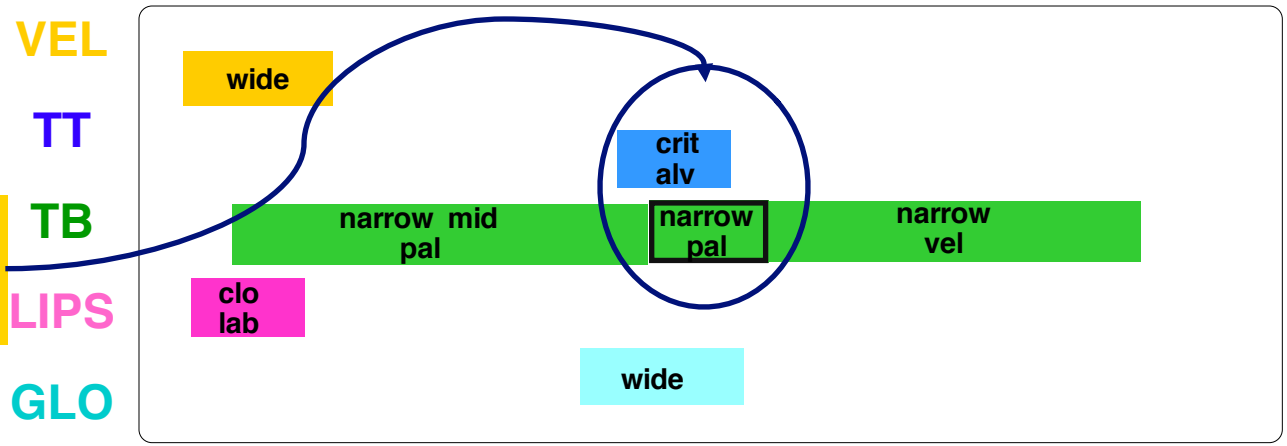
- What is going on here?
  - We change alveolar fricative to palatoalveolar before [j]?
  - More overlap in faster speech (within a phrase).

“miss. you”



“miss you”

Overlap changes fricative acoustics



Zsiga (1995)

Zsiga (1995)

Elizabeth C. Zsiga

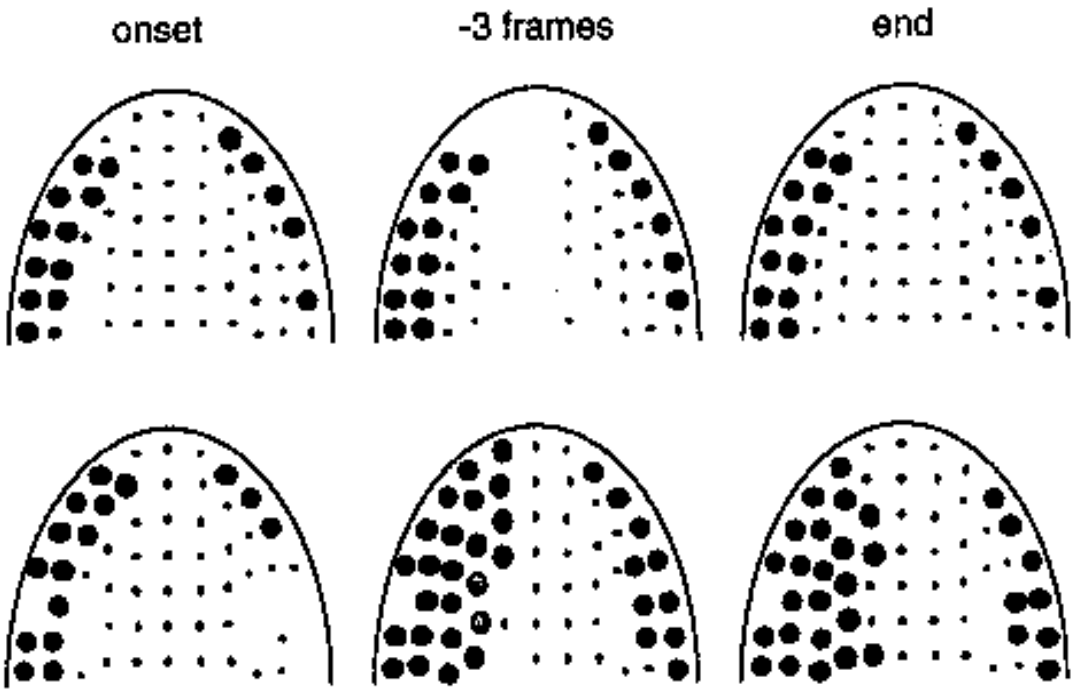


Figure 20.5. Change in contact patterns over time, subject 1. Electrodes shown in black were activated in at least eight of ten repetitions, those in gray, in seven of ten repetitions.

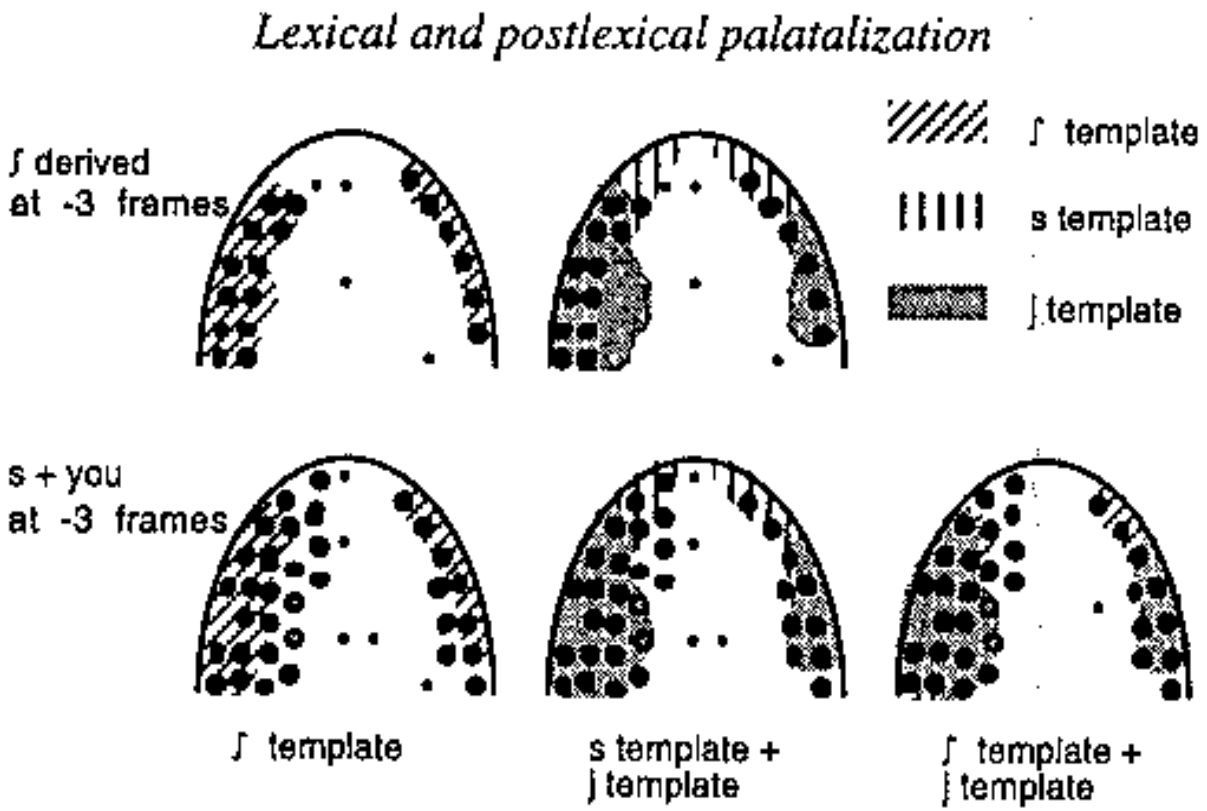
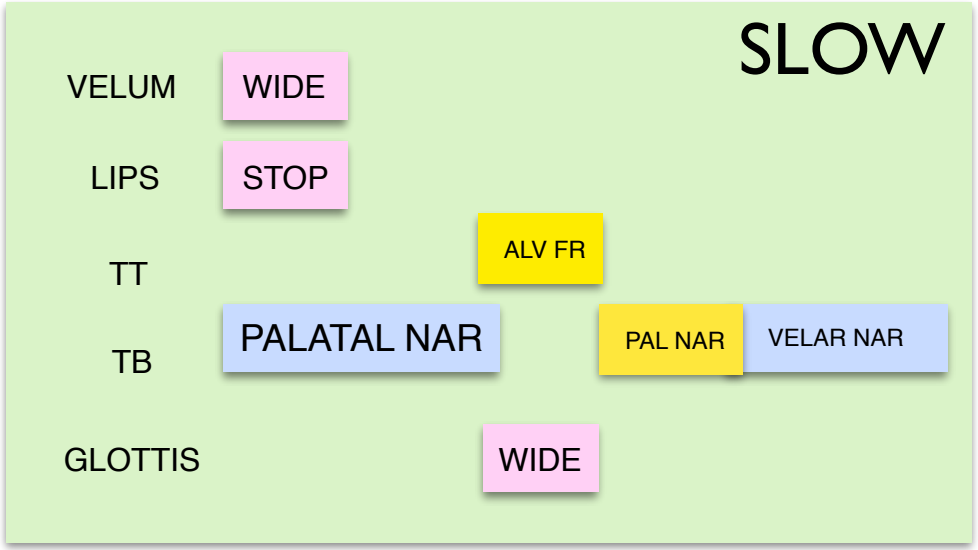
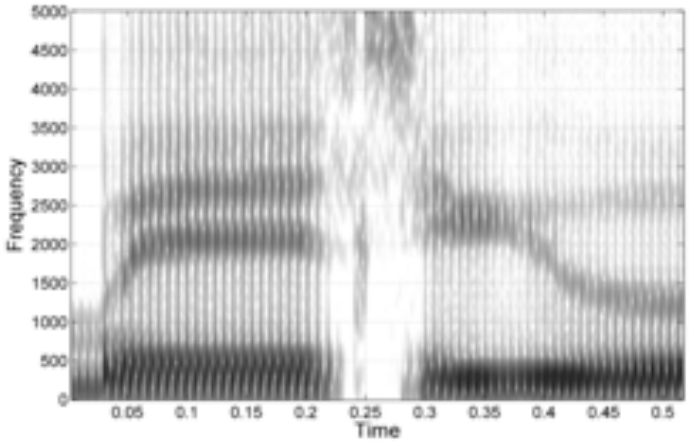
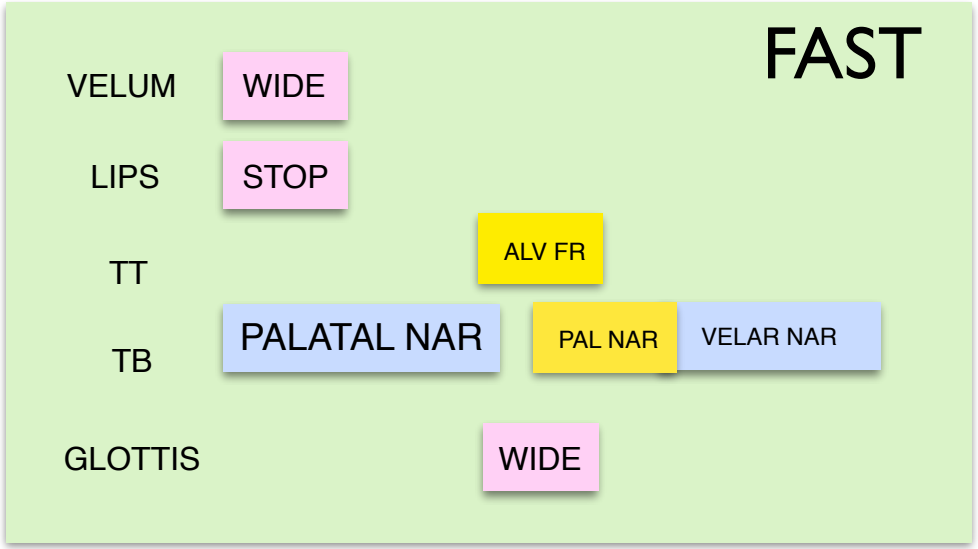


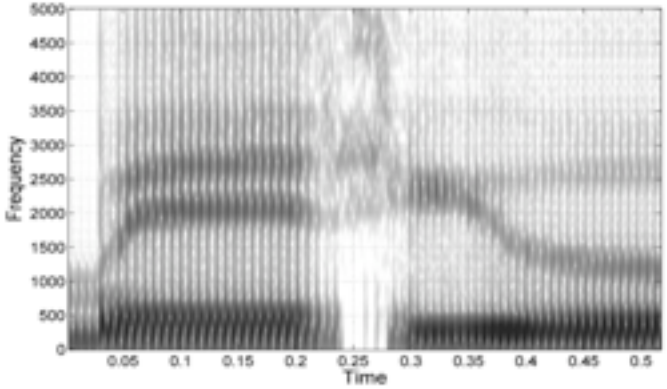
Figure 20.6. Templates from underlying /s/, /ʃ/, and /j/ overlaid on the patterns for *s+you* and derived /ʃ/ at –3 frames.

Effect of Gestural overlap: Synthesis







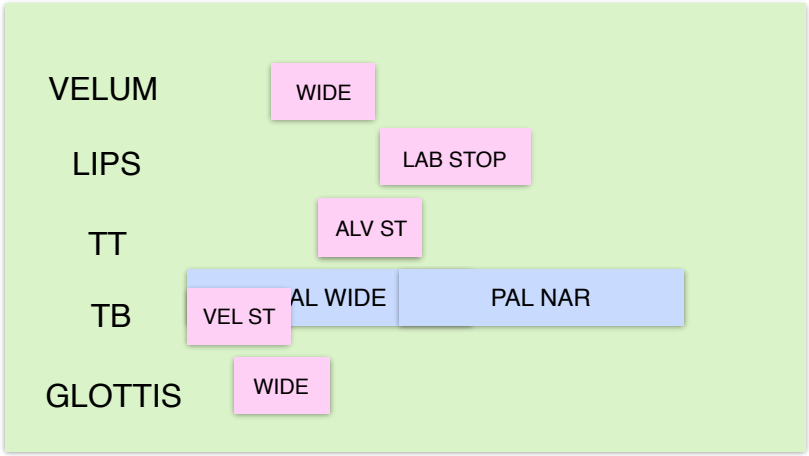


Place Assimilation: nasal

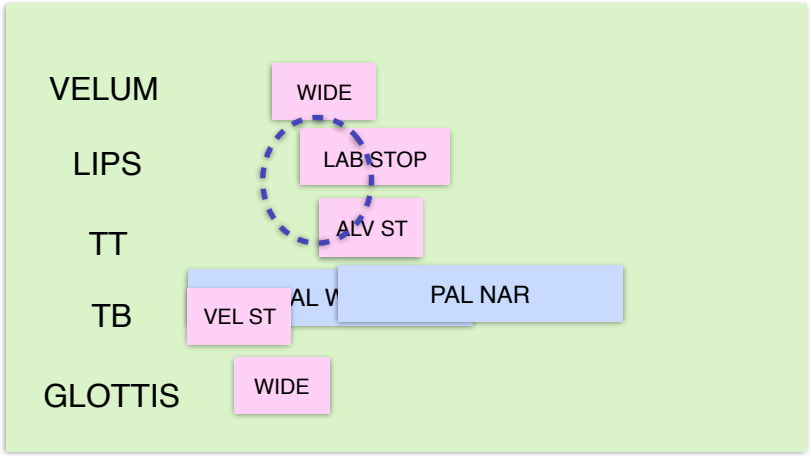
- Final /n/ is sometimes assimilated to the place of a following labial or dorsal stop:
- “can be”
  [kæ̃nbi] slow vs. [kæ̃mbi] fast

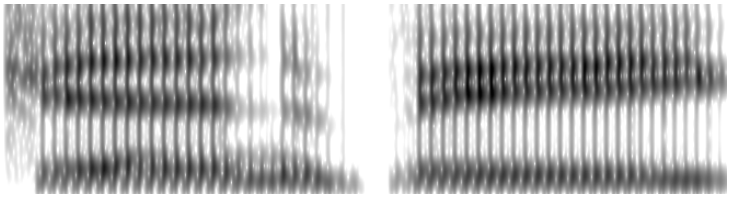
Nasal Assimilation: Synthesis

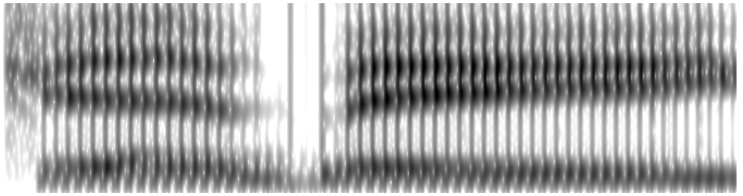
"can be" SLOW



"can be" FAST







- syntax: prosodic structure mediates
  Nespor & Vogel 1986/2007, Selkirk 1984, 2011)
- Directly controlled by syntax Cooper & Paccia-Cooper 1980, Wagner 2005)
- Prominence — also use of clock-slowing?
  - Coordination of PI gesture

Prominence and Focus

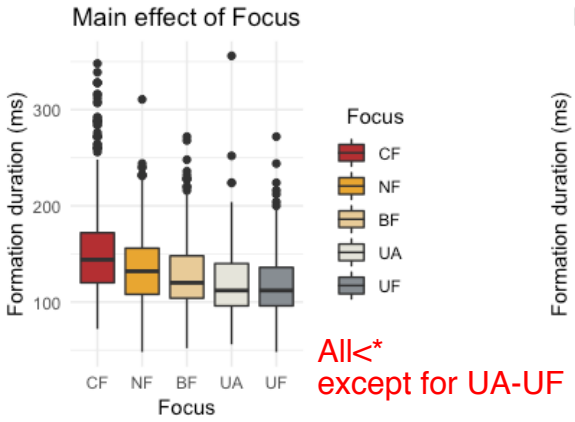
CF	Prompt: Is the botanist going to test the fly with the stripes? Test sentence: No, the botanist will be testing the ***bee*** with the stripes.
NF	Prompt: What is the botanist going to test? Test sentence: Oh, the botanist will be testing the ***bee*** with the stripes.
BF	Prompt: What is the botanist going to do? Test sentence: Oh, the botanist will be testing the bee with the stripes.
UA	Prompt: Is it the zoologist who will be testing the bee with the stripes? Test sentence: Oh, the ***botanist*** will be testing the bee with the stripes.
UF	Prompt: Who is going to test the bee with the stripes? Test sentence: Oh, the ***botanist*** will be testing the bee with the stripes.

Results

bee	Oh, the botanist will be testing the **bee** with the stripes.
baby	Oh, his family will be visiting the **baby** with the long hair.
design	Oh, Ellie's office will be submitting the **design** for the new building.
melody	Oh, Jonathan was praising the **melody** from the movie.
banana	Oh, the baby will be having the **banana** for lunch.
matinee	Oh, the parents were attending the **matinee** by the sea.
military	Oh, her cousin will be joining the **military** by the capital.
humanity	Oh, the historian was applauding the **humanity** of the soldiers.
salmonella	Oh, the teachers were discussing the **salmonella** for hours.

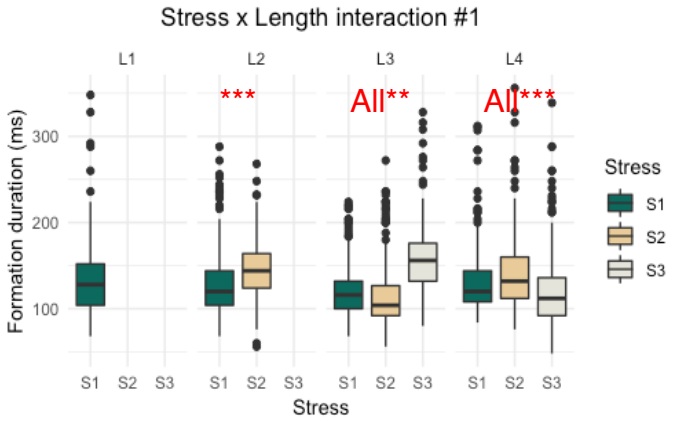
Results

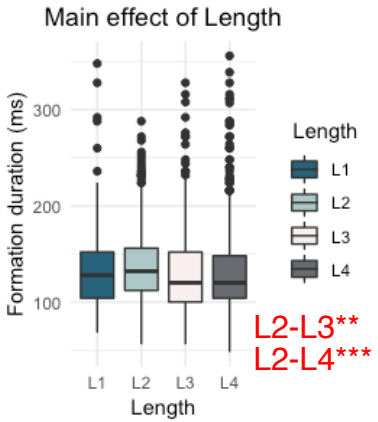
1. Formation duration (pairwise t-test)



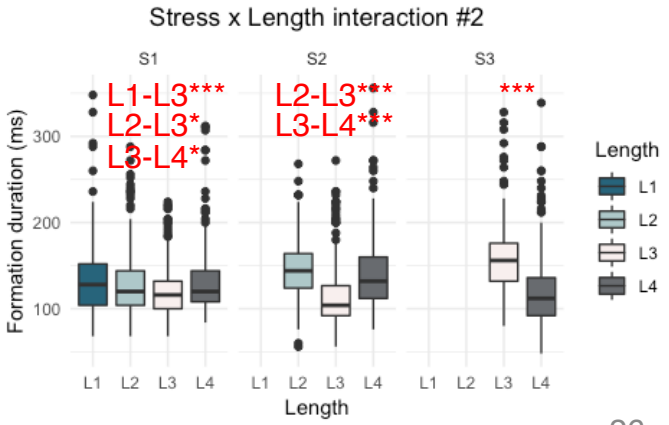
All<*
except for UA-UF

S1-S2*
S1-S3**
S2-S3 tr. .061

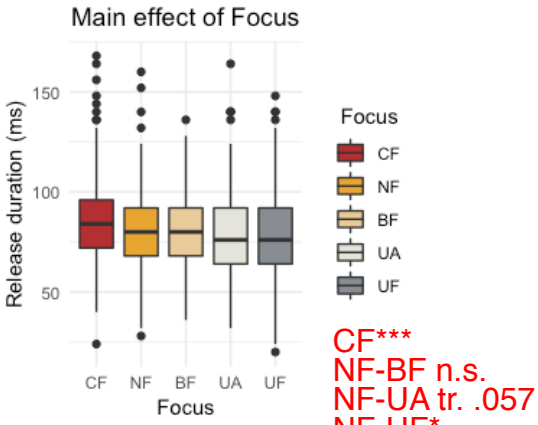




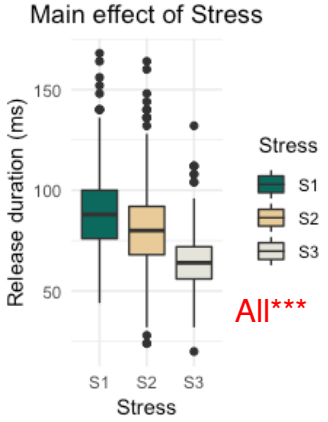
L2-L3**
L2-L4***



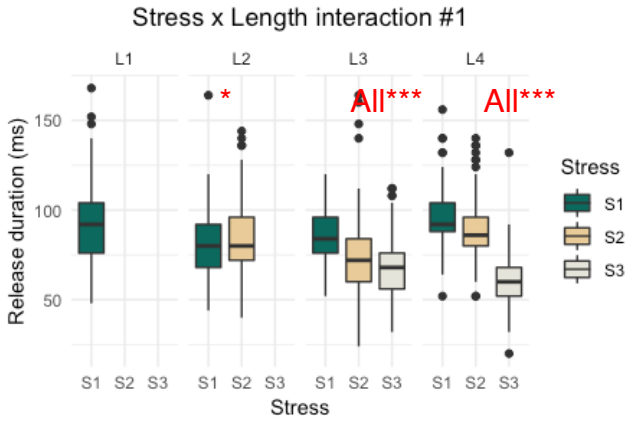
2. Release duration (pairwise t-test)

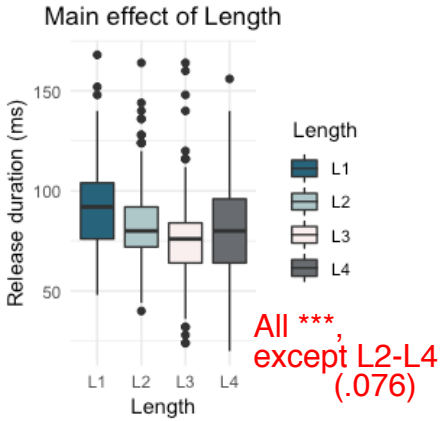


CF***
NF-BF n.s.
NF-UA tr. .057
NF-UF*
n.s. for others

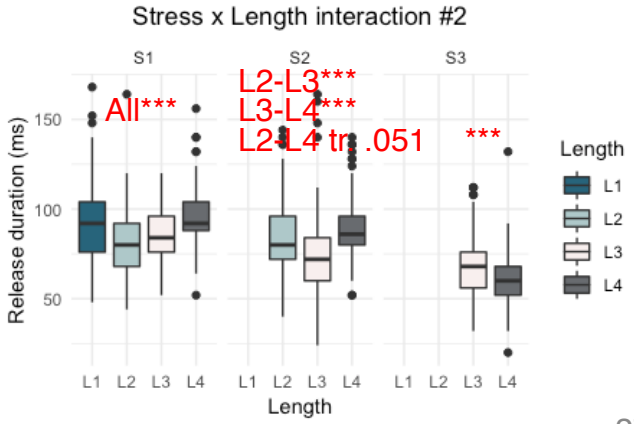


All***





All ***,
except L2-L4
(.076)

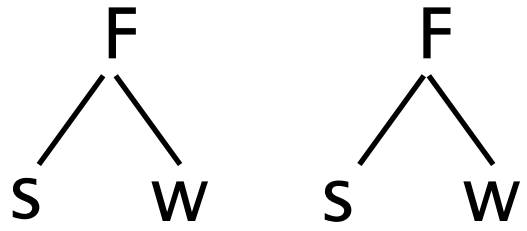


Modeling Prominence

- μ-gesture
  - Temporal modulation gesture
  - Like π-gesture, but centered on stressed vowel

Stress and Foot structure

- Organization of syllables into feet



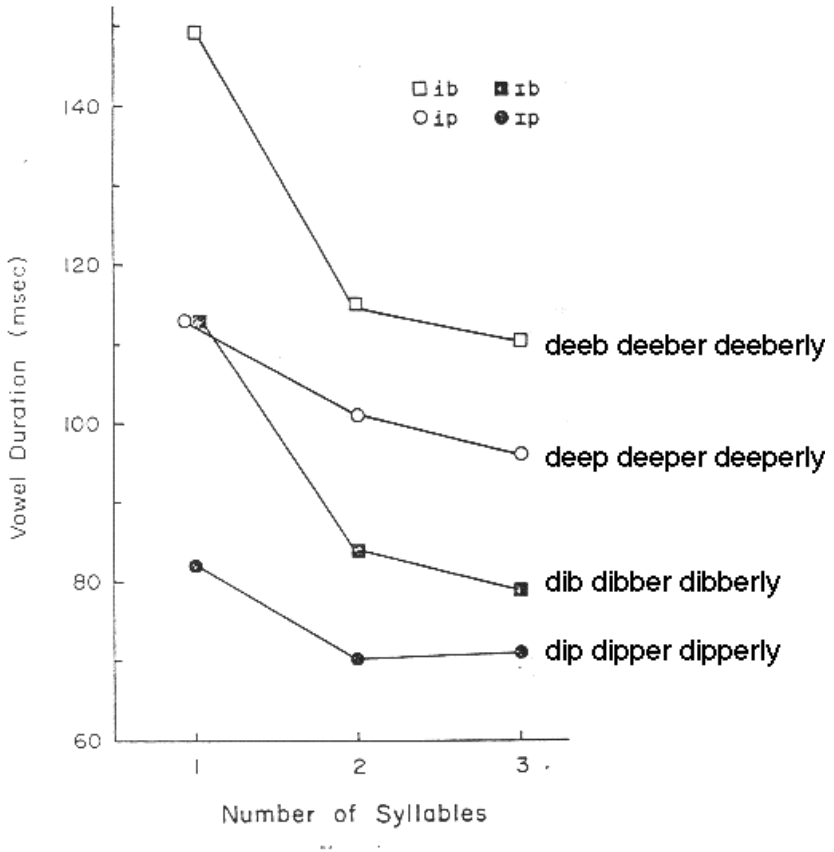
Foot composed of stressed syllable and following unstressed ones

- Possible temporal consequences of foot-structure on rhythm (observed in Germanic languages):
  - Shortening of syllable durations in polysyllabic feet ('stress-timing')
  - Greater length of stressed vs. unstressed syllables (vowel reduction)
- Language differences in rhythm ('stress-timing' vs. 'syllable-timing')

Polysyllabic shortening

- As syllables are added to a foot, the duration of the syllables decreases.
- Tendency to preserve the duration of the foot.
- This tendency is shown in languages that were traditionally called "stress-timed." (e.g., Germanic languages)
- Not shown in "syllable-timed" languages in which syllable durations tend to be preserved (e.g., Romance languages)

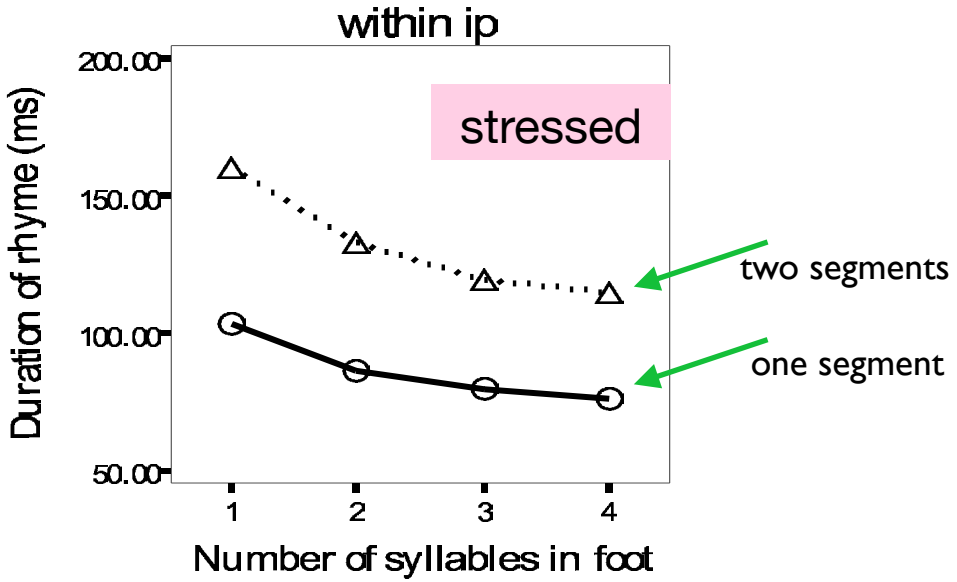
Port (1981)

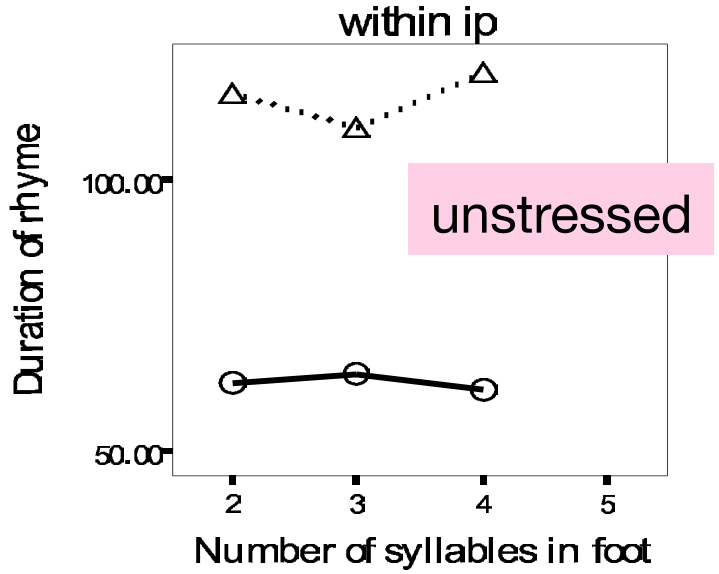


Interaction of:
Number of syllables
Vowel length
Voicing of coda C

Stress and Polysyllabic shortening

- Kim & Cole (2005) study of the BU Radio News corpus shows that polysyllabic shortening affects stressed syllables, but not unstressed ones.
- Stressed syllables are longer than unstressed ones

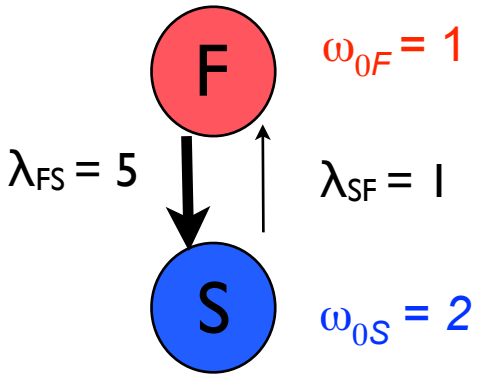




Coupled Oscillator Model of Polysyllabic Shotenimg

- Hierarchical Coupled Oscillators (O'Dell & Nieminen, 1999)
  - Harmonically entrained Foot and Syllable oscillators
  - N:1 entrainment: N syllable cycles per 1 foot cycle
  - Hypothesized inter-level asymmetry of coupling strengths can produce polysyllabic shortening in languages that show it:
  - 'Stress-timed' languages
    - foot-to-syllable coupling (λ_{FS}) >> syllable-to-foot (λ_{SF})
  - 'Syllable-timed' languages
    - foot-to-syllable coupling (λ_{FS}) << syllable-to-foot (λ_{SF})

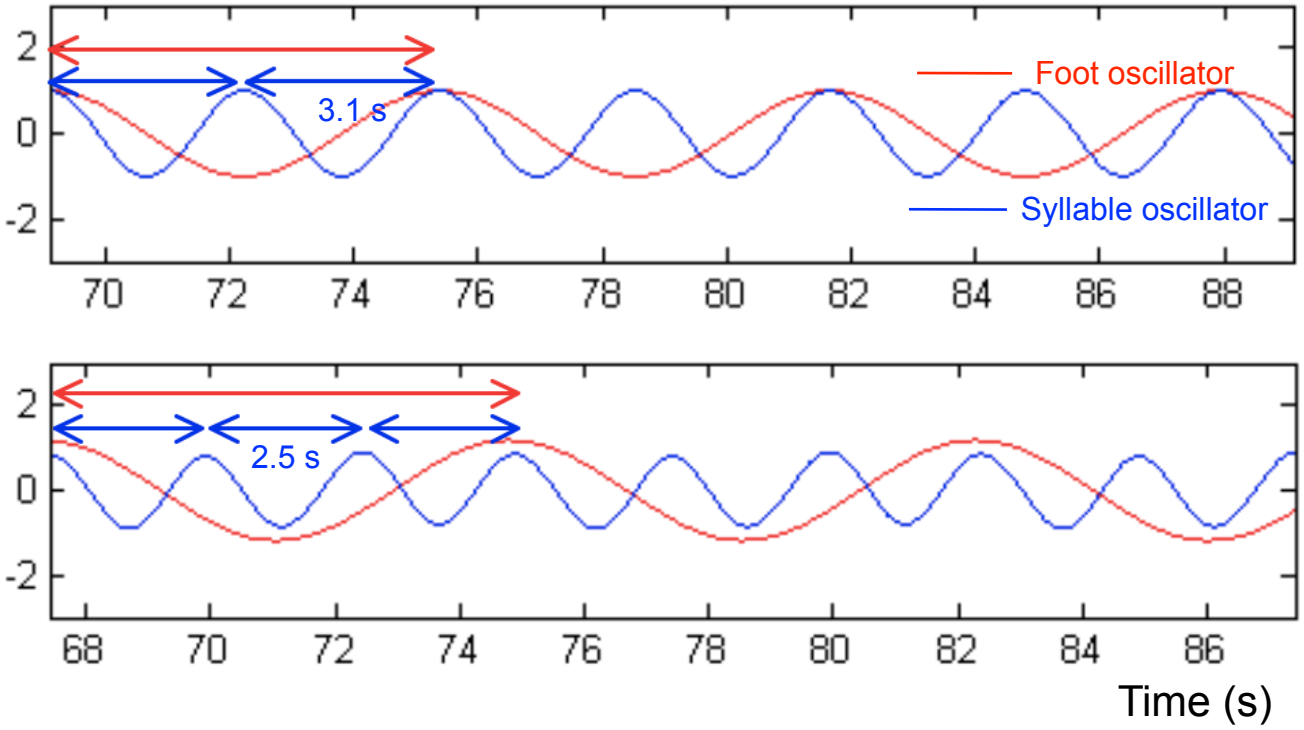
Polysyllabic shortening Simulation



Saltzman et al., 2008

2 syllables per foot

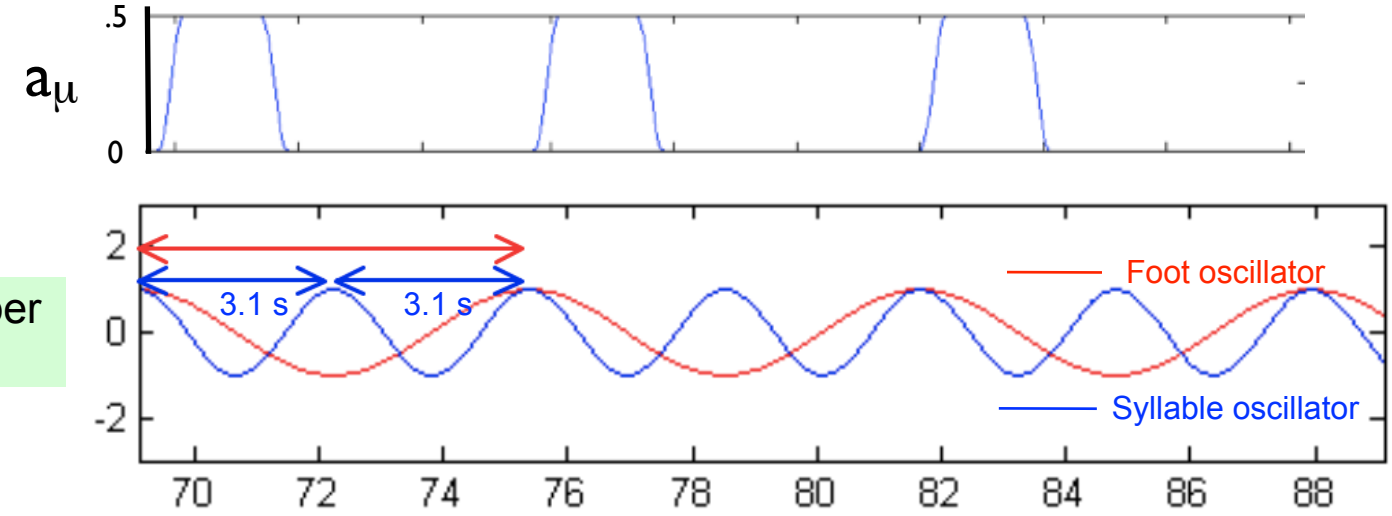
3 syllables per foot

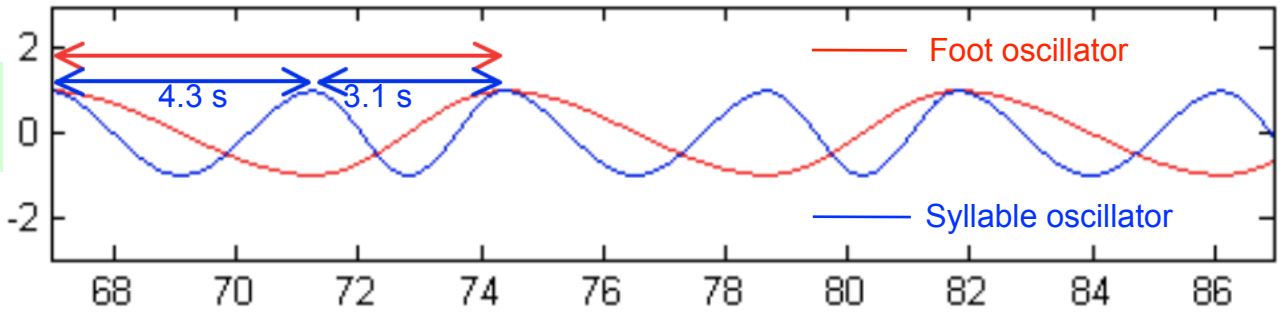


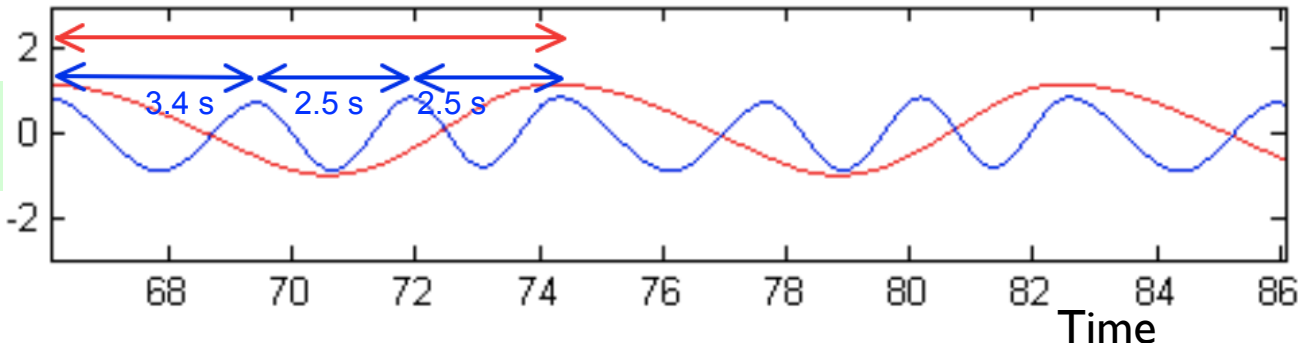
Stress Asymmetry

- How can the differential durations of stressed and unstressed syllables be modeled?
- Hypothesize clock slowing gesture (μ_T) that is active at phases of Foot oscillator corresponding to stressed syllables (similar to π-gesture).
- μ_T slows clock of Foot and Syllable oscillators (and all constriction gesture) in proportion to its activation level (a_μ).
- Maximum strength of μ_T gesture will determine the degree or temporal asymmetry between stressed and unstressed syllables.

Stress Asymmetry Simulation







Asymmetry in Polysyllabic shortening

- Polysyllabic shortening affects stressed syllables more than unstressed ones.
  - modulate coupling strength asymmetry as a function of phase of the foot oscillator
  - STRESS: (λFS) > (λSF)
  - UNSTRESS: (λFS) < (λSF)