ARTICULATORY PHONETICS AND PROSODY

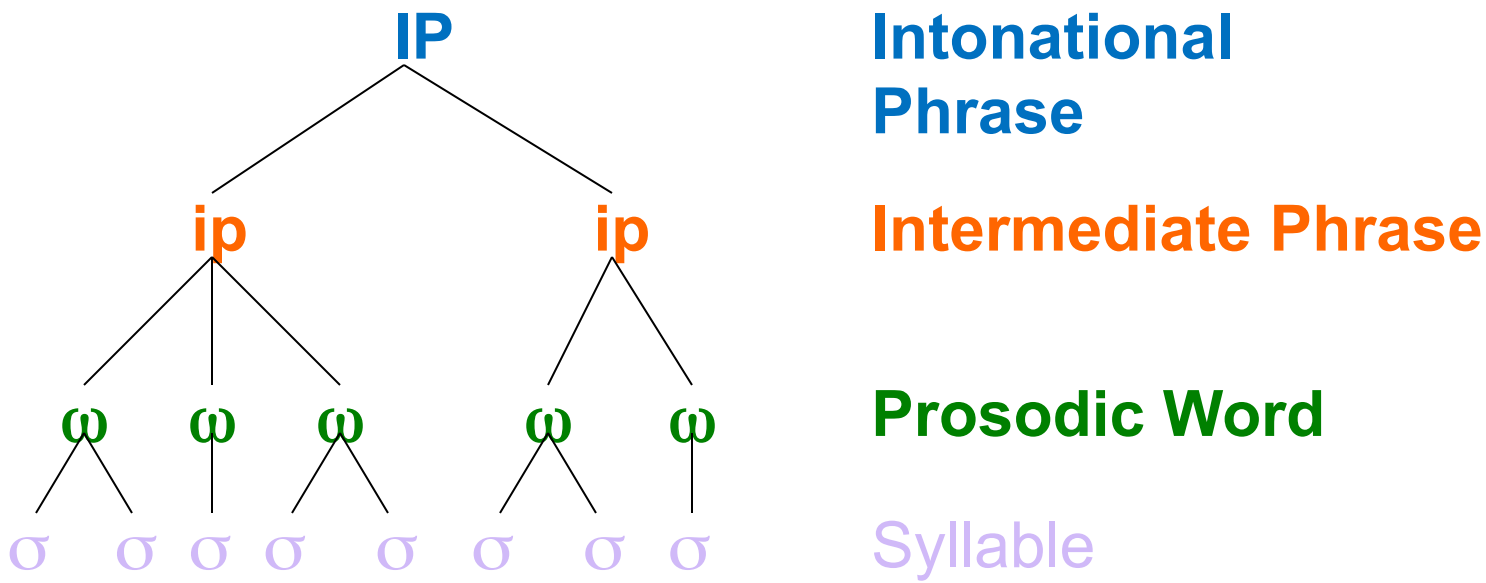
Jelena Krivokapić
University of Michigan, Haskins Laboratories

jelenak@umich.edu

INTRODUCTION

Prosodic hierarchy

- Prosodic structure is hierarchically organized.
  - Speech is grouped into prosodic units, with higher units dominating lower units.



Beckman & Pierrehumbert 1986 model

Articulatory properties of boundaries

- In addition to tonal properties, prosodic boundaries have characteristic **spatial** and **temporal properties**.
  - *Acoustic domain*: Phrase final and initial lengthening
    e.g. Oller 1973, Wightman et al. 1992, Shattuck-Hufnagel & Turk 1998
  - *Articulatory domain*: Domain-edge gestures become larger, longer, and less overlapped in production
    e.g. Beckman, Edwards, & Fletcher 1992; Byrd & Saltzman 1998; Byrd 2000; Tabain 2003; Cho 2004, 2005;

1. Prosodic structure as manifested in articulation

- Overview of studies on the properties of **prosodic boundaries** in articulation
  - Articulatory strengthening at boundaries (Fougeron & Keating 1997, Keating, Cho, Fougeron, & Hsu 2003)
  - Temporal properties of prosodic boundaries (Byrd & Saltzman 1998, Byrd 2000, Byrd, Krivokapić & Lee 2006)

2. Prosodic structure within Articulatory Phonology

- How is prosodic structure accounted for within the framework of Articulatory Phonology (e.g., Browman & Goldstein 1992, 1995, Goldstein & Fowler 2003)
  - The π-gesture model of prosodic boundaries (Byrd & Saltzman 2003)
  - Tone gestures within this model (Gao 2008, Mücke et al. 2012).

3. Recent developments

- Interaction of prosodic boundaries and prominence
  - E.g., Byrd & Riggs 2008, Katsika et al. 2014
- The articulatory properties of pauses and their role in prosodic boundaries
  - Ramanarayanan et al. 2010, 2013, Katsika et al. 2014
- Prosodic structure in manual gestures
  - E.g., Rochet-Capelan et al. 2008, Krivokapić et al. 2015

Outline

- Prosodic structure as manifested in articulation
- Prosodic structure within Articulatory Phonology
- Boundaries, prominence, & pauses

PROSODIC STRUCTURE AS MANIFESTED IN ARTICULATION

Strengthening: Fougeron & Keating 1997, Keating, Cho, Fougeron, & Hsu 2006, electropalatography

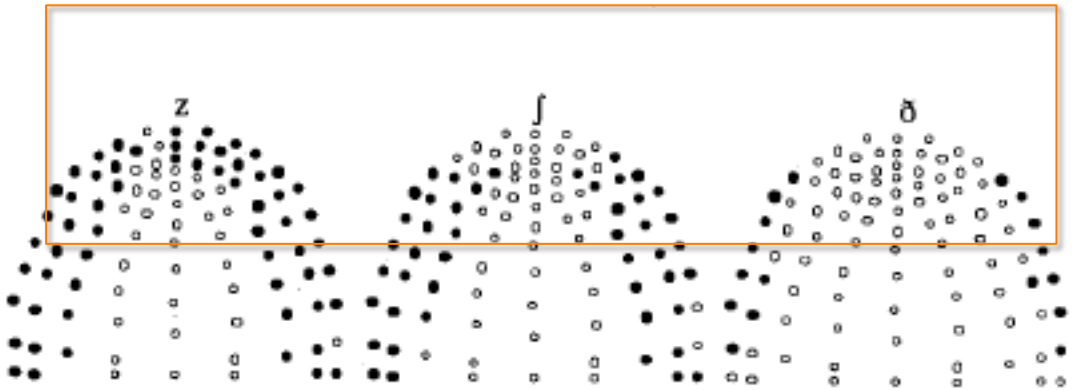
Dynamic electropalatography (Dynamic Palatography, EPG)

- Using an artificial palate, tongue-palate contact over time is recorded
  - On EPG, see D. Byrd. (1994) Palatogram reading as a phonetic skill: A short tutorial. Journal of the International Phonetic Association, 24, 21-34.

FIGURES REMOVED FOR COPYRIGHT REASONS. FOR IMAGES SEE:

http://www.linguistics.ucla.edu/faciliti/facilities/physiology/epg.html

https://www.eda.kent.ac.uk/medical/epg.aspx



Contact in constrictions in [z, ʃ, ð]

Adapted from Figure 2, Byrd, JIPA 1994

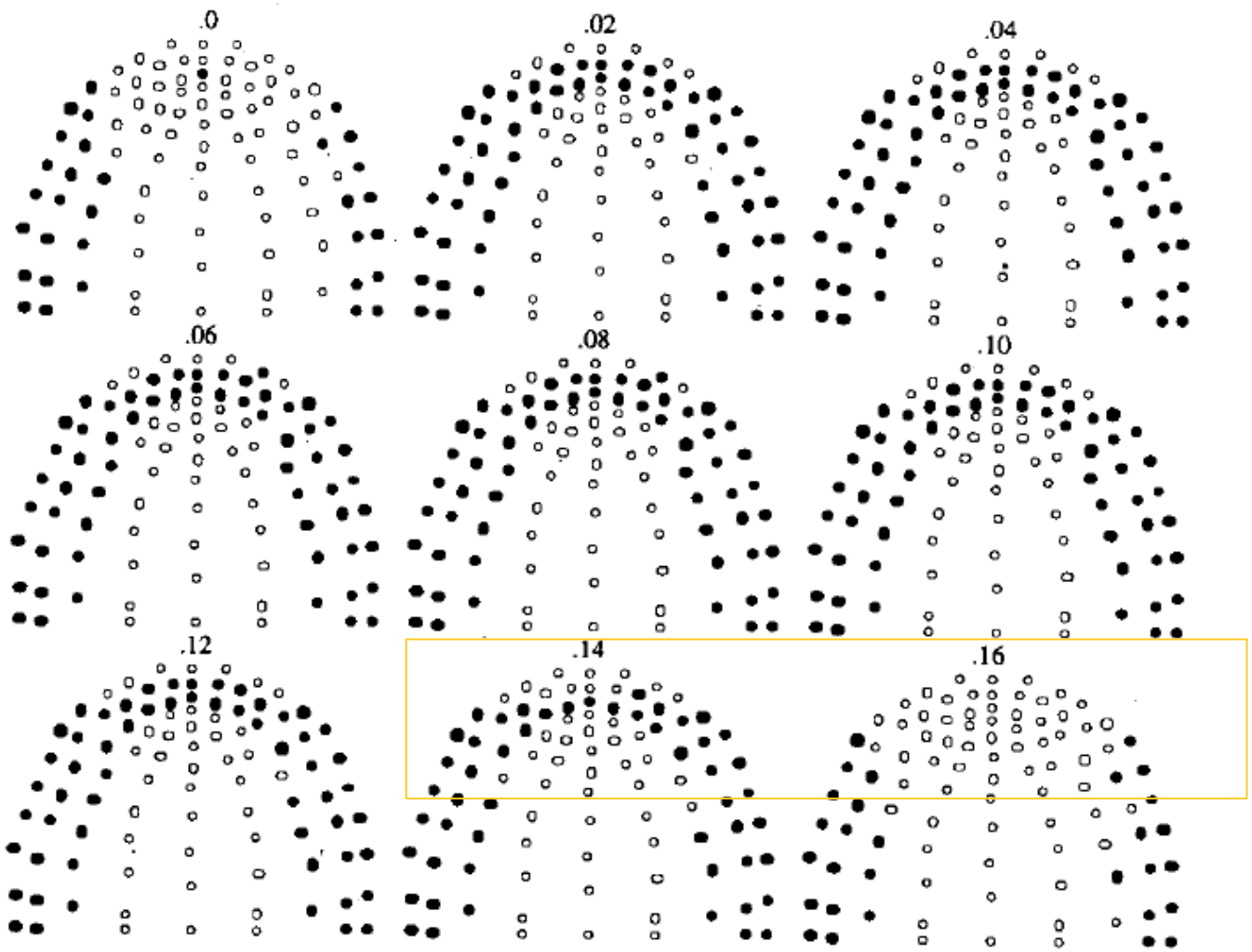
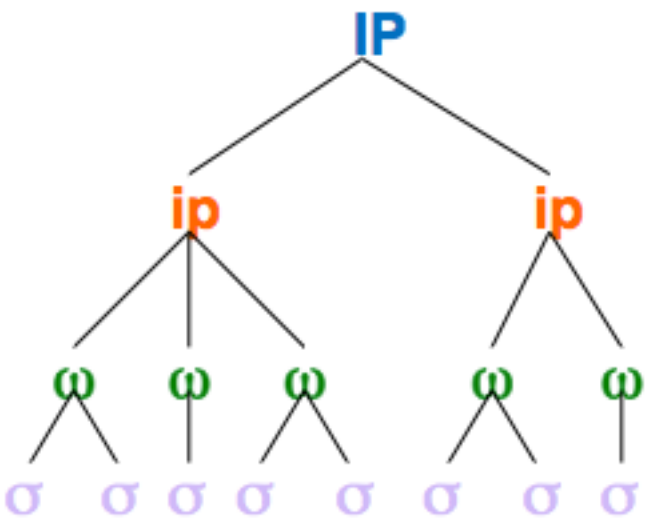


Figure 4. Sequence of frames showing [n] in [ana].

Adapted from Figure 4, Byrd, JIPA 1994

Fougeron and Keating 1997

- Seminal study on the articulatory characterization of the prosodic structure
- Goal: to characterize spatial and temporal properties at domain edges.
- Examine linguopalatal contact (EPG)



Methods

- *Stimuli*: reiterant speech, *no*
  - allows the study of different prosodic positions

Model sentence
1. 89 + 89 + 89 + 89 = a lot.
2. (89 + 89) * (89 + 89) = a lot.
3. 89 * (89 + 89 + 89) = a lot.
4. (89 + 89 + 89) * 89 = a lot.

- E.g., second line:
  (nonono no nonono) no (nonono no nonono) = a lot
- **/n/** phrase initially, **/o/** phrase finally

Table adapted from
Keating & Fougeron 1997

Resulting prosodic domains (example)

Domain:	no	no	no	+ no	no	no	no	+ no	no	no	no	+ no	no	no	no
Utterance (U)	i	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Intonation Phrase (IP)	i	m	m	m	m	m	m	m	m	m	m	m	m	m	m
Prosodic Phrase (PP)	i	m	f	i	m	m	f	i	m	m	f	i	m	m	f
Word	i	m	f		i	m	f		i	m	f		i	m	f

i- phrase initial; m – phrase medial; f – phrase final

Domains: w, ip, IP, Utterance

Measurements: a) contact:
the frame with
- maximum contact for /n/
- minimum contact for /o/

b) acoustic duration

Figure adapted from Fougeron & Keating 1997

Results: Distinguished prosodic levels

		Speaker 1	Speaker 2	Speaker 3
Contact:				
• Initial /n/	(a)	IPi>PPi>Wi=Si (Ui=PPi and Ui=IPi)	Ui>IPi=PPi>Wi>Si	Ui=IPi>PPi=Wi>Si
• Final V	(b)	IPf<PPf<Wf=Sf	IPf=PPf<Wf<Sf	IPf,PPf,Sf<Wf
• Initial V (in *no*	(c)	IPi<Ui<PPi<Wi<Si	Ui,IPi,Si<PPi	Ui<Wi<PPi=IPi<Si
Duration (acoustic):				
• Initial /n/	(f)	IPi>PPi>Wi>Si=Ui	IPi=PPi>Wi>Ui>Si	Ui=IPi>PPi>Wi>Si
• Final V	(g)	IPf=PPf>Wf=Sf	IPf=PPf>Wf=Sf	IPf>PPf>Sf>Wf

Fougeron and Keating 1997, adapted from Table III

Results

- Overall, the higher the prosodic boundary the (acoustically) longer and the more linguopalatal contact for /n/ ***phrase initially.***
- Overall, the higher the prosodic boundary the less contact for /o/, and the acoustically longer its duration, ***phrase finally.***

  => Articulatory strengthening at boundaries
- Speaker dependent, but at least 2 domains above the word level are always distinguished phrase-initially.
- No speaker showed a distinction for all categories, and there was no category that was distinguished by all speakers.
- There are ***effects further away*** from the boundary as well: phrase-initially, there were effects for /n/ and on the following /o/ as well.
- The effect is ***cumulative***, increasing with boundary strength.

Keating, Cho, Fougeron & Hsu (2003)

- Extend the findings from Fougeron & Keating 1997 to three other prosodically different languages
  - French, Korean, Taiwanese

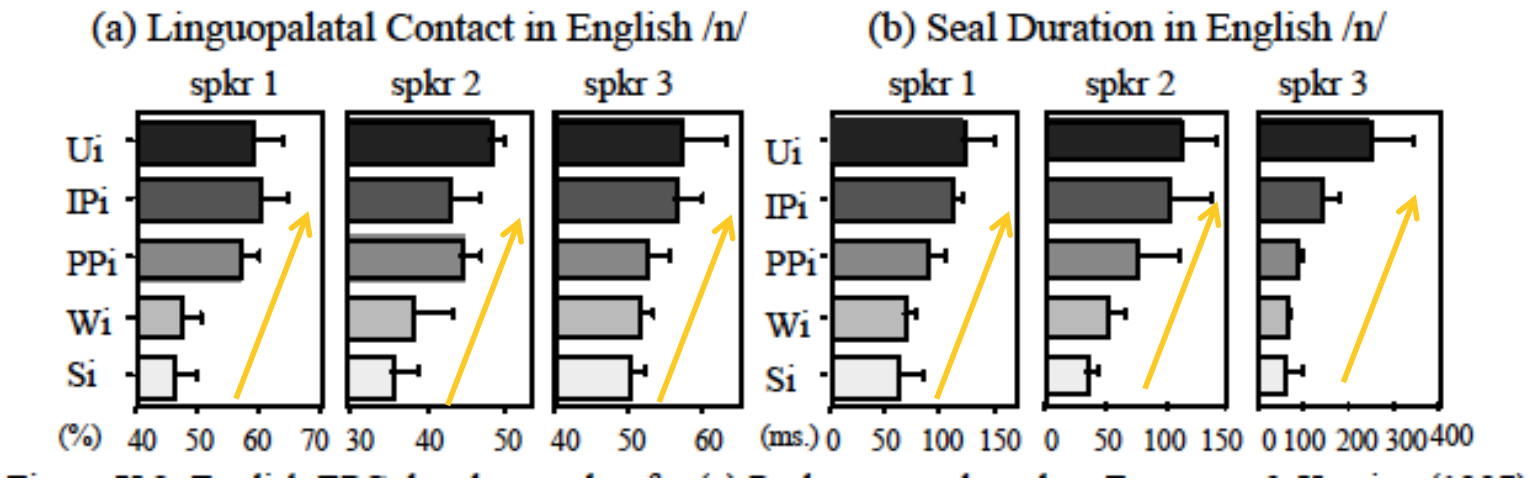
Cross-linguistic differences

Use of different languages in order to further examine strengthening. Some of the differences:

- English has prominent tonal boundary marking, leading to the possibility that temporal properties are not pronounced.
- Taiwanese uses lexical tone and therefore might signal prosodic boundaries with temporal characteristics.
- French has phrase final boundary marking (final lengthening and phrasal accent at the end of an AP).
- Korean has phrase initial marking. There is no AP phrasal accent and no final lengthening but there is AP initial phrasal accent and lengthening.

- Each of the examined languages has a defined prosodic hierarchy, and the prosodic domains are marked phonetically.
- The identity of the categories differs across the languages, but for the purposes of the experiment, that is not relevant.

Results: English



Contact measure as in Fougeron & Keating 1997
Seal duration: the number of frames which have a complete stop closure

Figure adapted from Keating et al. 2003

Results: French

Above the word:
Contact: Utt, IP > AP > W or S

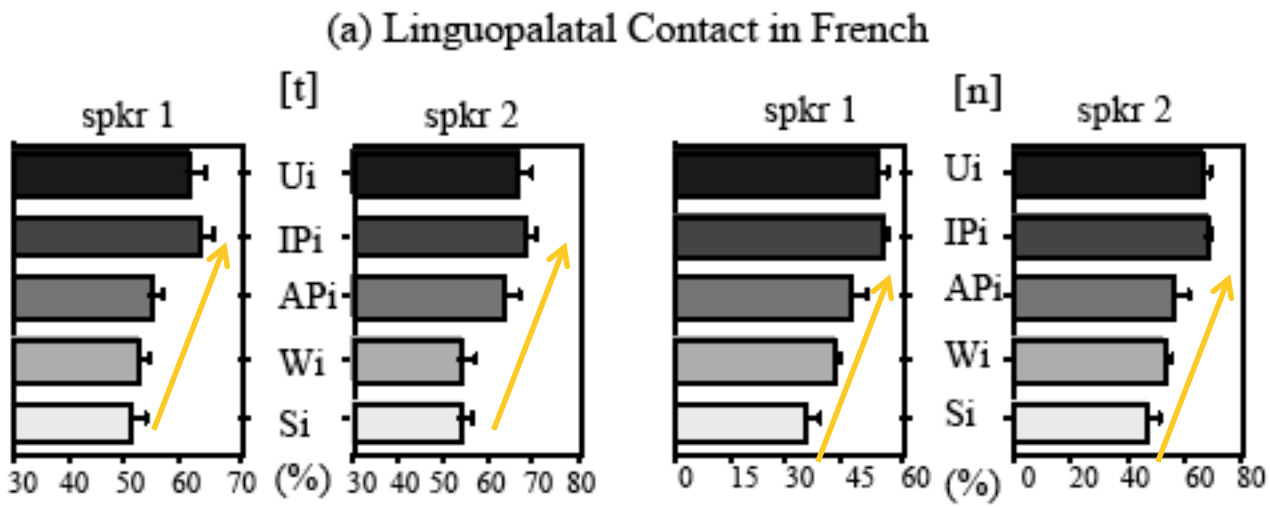


Figure adapted from Keating et al. 2003

Results: Korean

Contact above the word level:
U > IP > AP > W
Except spkr3: AP = W

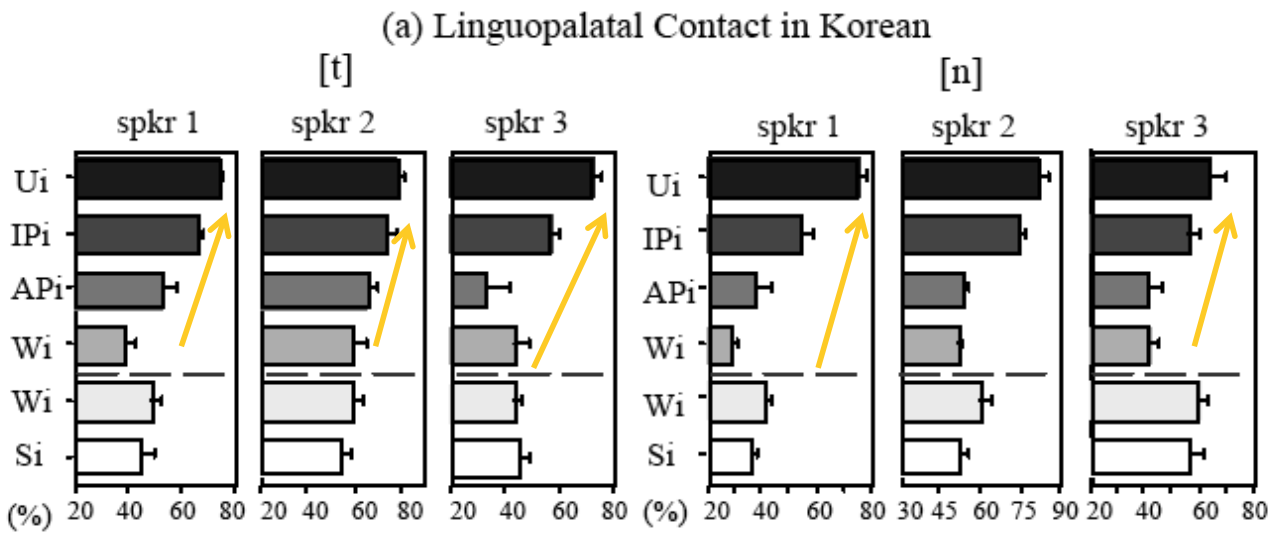


Figure adapted from Keating et al. 2003

Results: Taiwanese

Contact above the word level:
Spkr 1:
/t/: IP > SP > W; /n/: IP>SP
Spkr2:
/t/: U>IP>SP>W; /n/: U>IP>SP

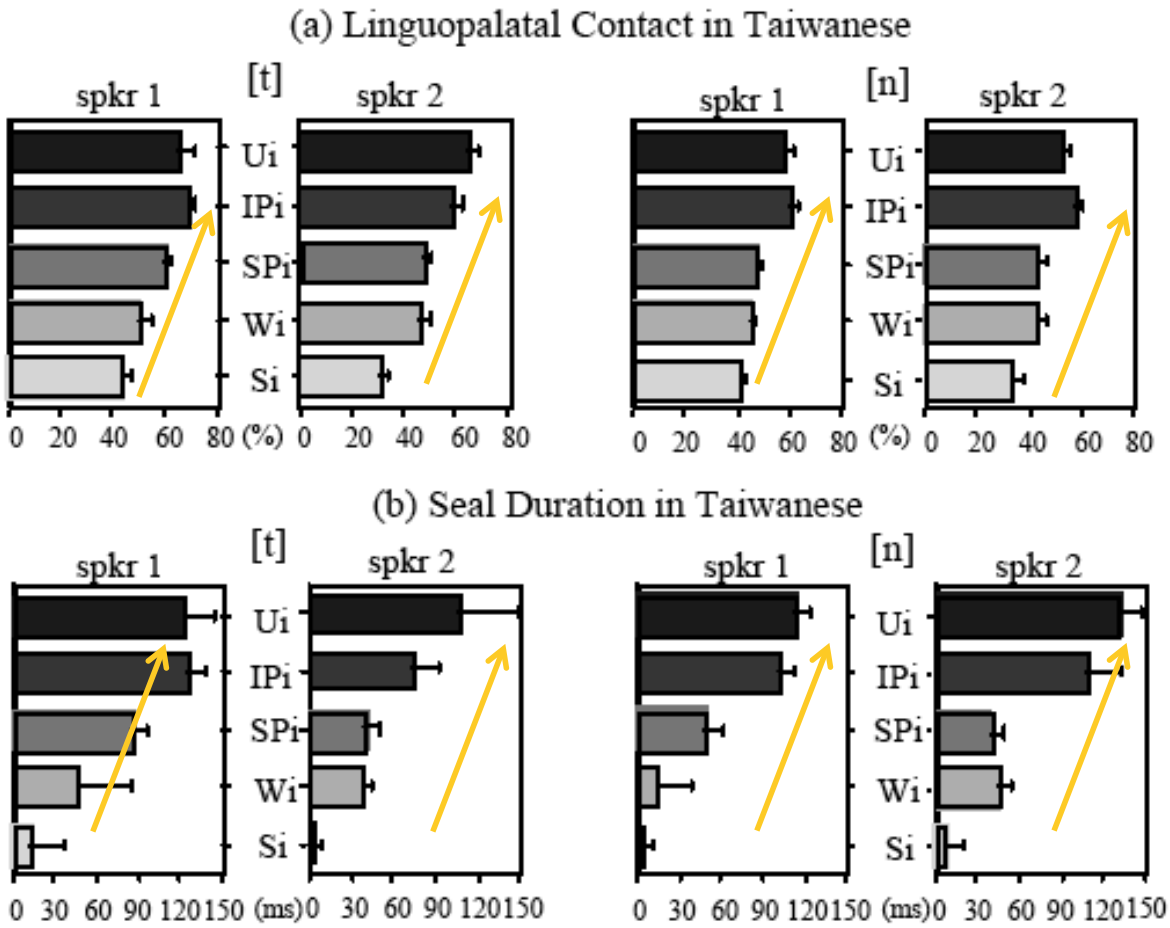


Figure X.10. Data for Taiwanese, displayed as in Figure X.2. (a) Peak EPG contact for /t, n/; (b) articulatory duration for /t, n/.

Figure adapted from Keating et al. 2003

Results

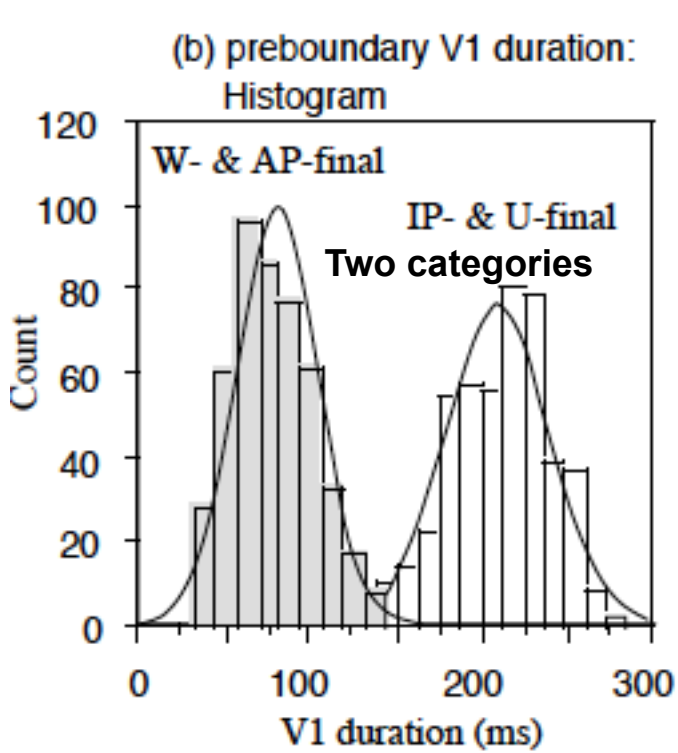
- Cumulative initial strengthening, variable across speakers, similar as for English
- Overall similarity across languages

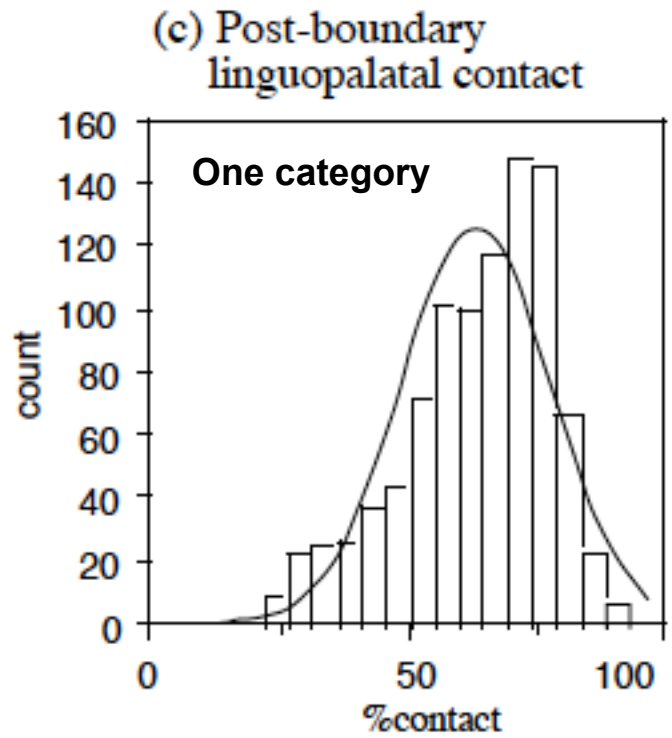
Summary

- These experiments show clearly phrasal/prosodic conditioning of articulation across languages
- Effects are speaker dependent and depend on consonant studied
- *No prosodic hierarchy in a language is exhibited by all speakers.*

Question

- Do these data (and others like it) provide evidence of an instantiation of the prosodic boundaries (in the spatial domain)?

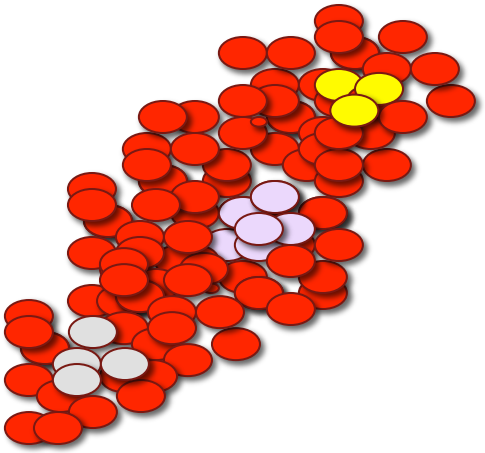




Figures adapted from Cho & Keating 2001, see also Keating 2006

Question

Categoricity



Gradiency

See work by Swerts 1997, Ladd 1998, Cho & Keating 2001, Wagner 2005, Keating 2006, Krivokapic & Byrd 2012

PROSODIC STRUCTURE AS MANIFESTED IN ARTICULATION

Temporal properties: Byrd & Saltzman 1998, Byrd 2000, Byrd, Krivokapić & Lee 2006, magnetometry

Byrd & Saltzman 1998

- Goals:
  - To characterize temporal properties at prosodic boundaries
  - Lowered gestural stiffness account

Byrd & Saltzman 1998

- Goals:
  - To characterize temporal properties at prosodic boundaries
  - Lowered gestural stiffness account

Point-Tracking Systems for Studying Articulatory Movement: Magnetometer

- The subject is seated in a magnetic field generated by three to six coils.
- Small receivers are adhered to the subject's articulators.
- As the receivers move through the magnetic field, the generated voltage fluctuations are recorded.
- This allows the movements of these points on the articulators to be tracked.

Figures removed due to copyright. See

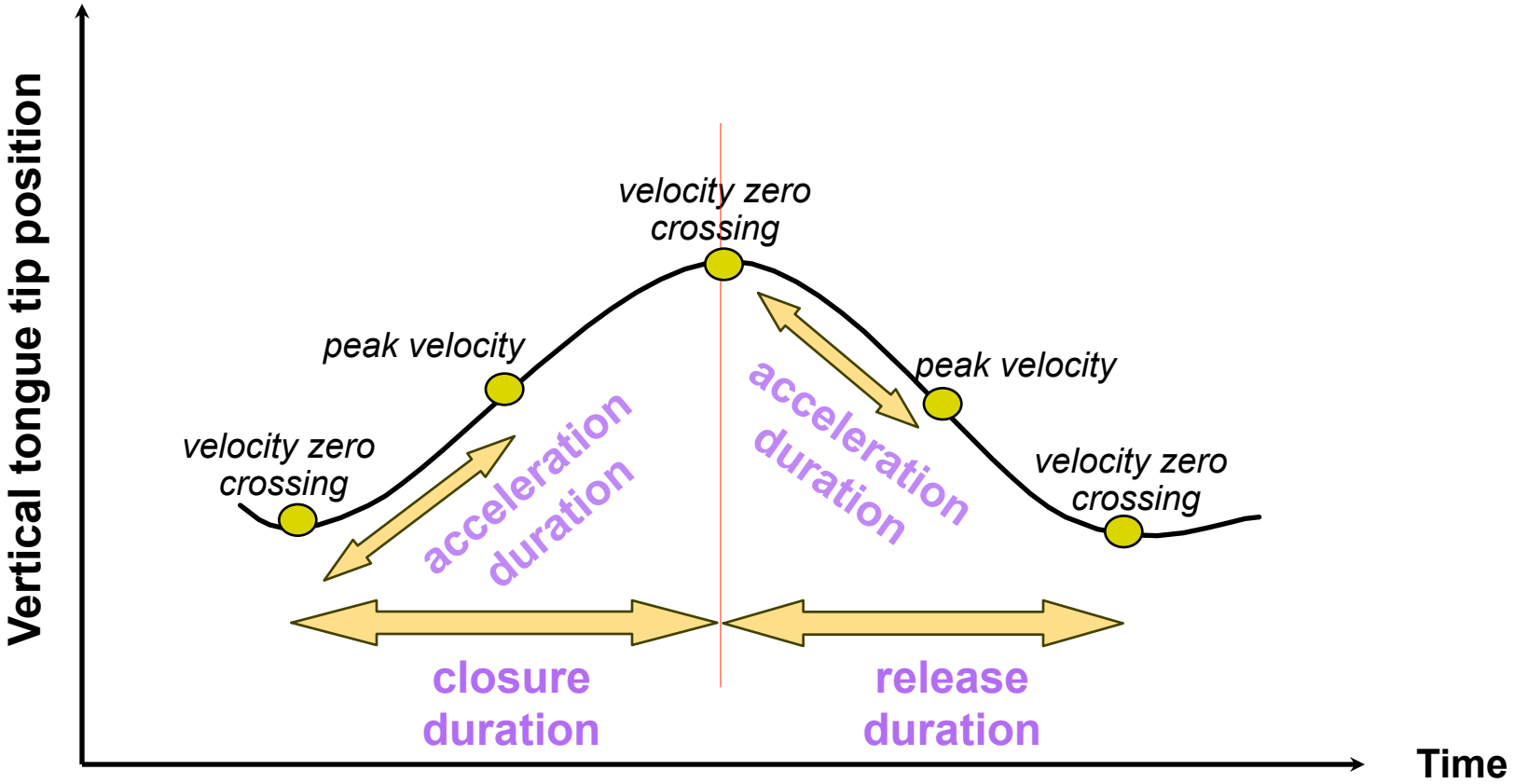
www.phon.ox.ac.uk/~paula/ images/edema.jpg

http://www.articulograph.de/?page_id=711

EMA movie, see:

http://sail.usc.edu/peterladefoged80.html

Measurements of tongue tip movement



Byrd & Saltzman 1998: Boundary-Adjacent Lengthening Across Different Boundaries

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

Table adapted from Byrd & Saltzman 1998

Target word: [mə#mi]

Boundaries of different strength

Examined:

- pre-boundary lip opening duration
- post-boundary lip closing duration

Measured variables: [mə#mi]

- Lip aperture for the pre-boundary opening movement and the post-boundary closing movement:
  - Opening displacement
  - Closing displacement
  - Time to peak velocity openingmovement
  - Time to peak velocity closing movement
  - Opening movement duration
  - Closing movement duration
  - Transboundary interval (opening duration + closing duraiton)

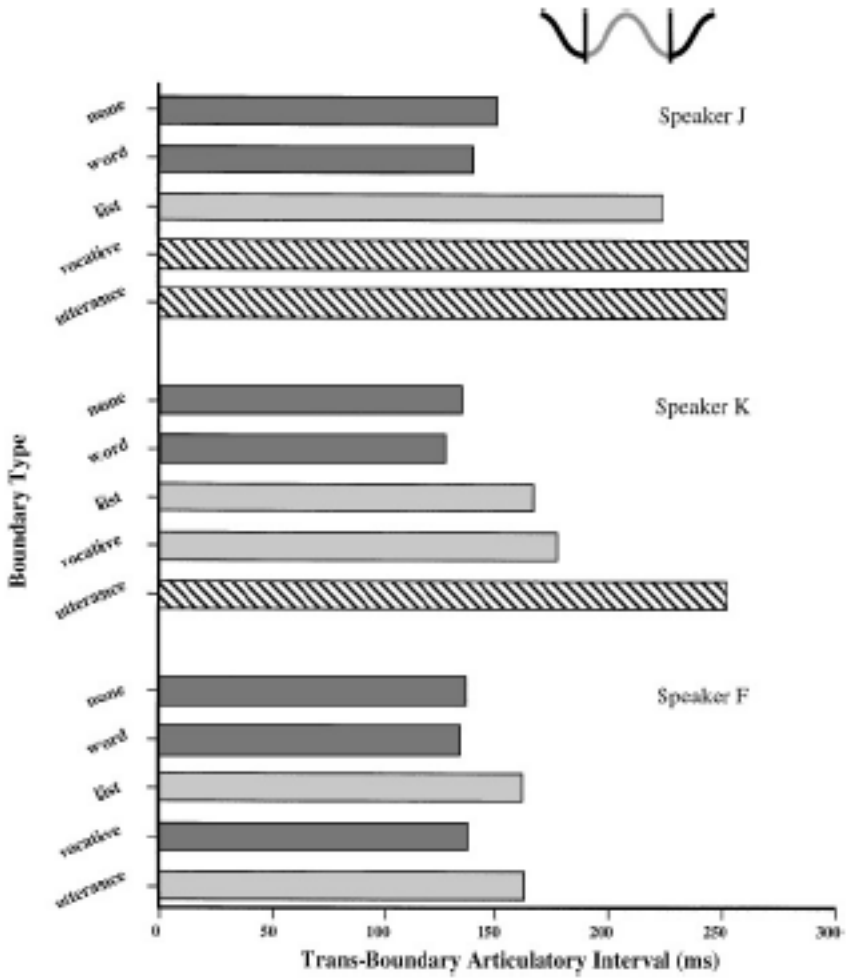


Figure 4. The transboundary articulatory interval for three speakers and five experimental boundary conditions. Bars with like shadings group together in post-hoc tests ($p < 0.005$).

Boundary-Adjacent Lengthening, lip aperture

[mə#mi]

Byrd & Saltzman 1998

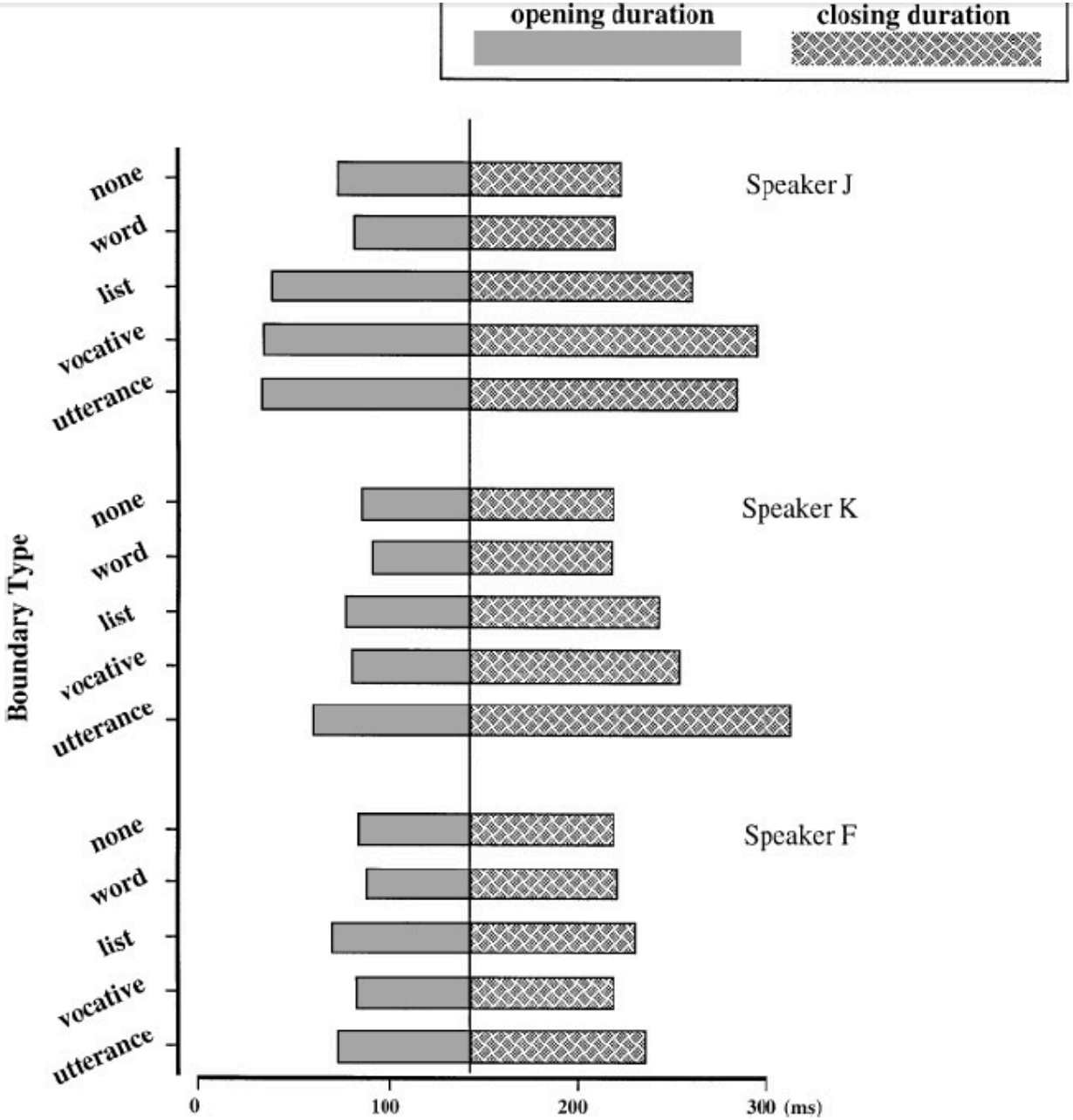


Figure 5. Preboundary opening and postboundary closing durations for three speakers and five experimental boundary conditions. Vertical line marks onset of post-boundary closing (and end of preboundary opening).

Boundary-Adjacent Lengthening

[mə#mi]

Byrd & Saltzman 1998

Reduced gestural overlap

- Using the same set of data, Byrd (2000) examines gestural overlap

[mə#mi]

c1v1#c2v2

Interval:

- C1targ -> V2pkvel
- All participants show an effect of boundary, such that the post-boundary (V2) vowel occurs later with respect to the pre-boundary consonant

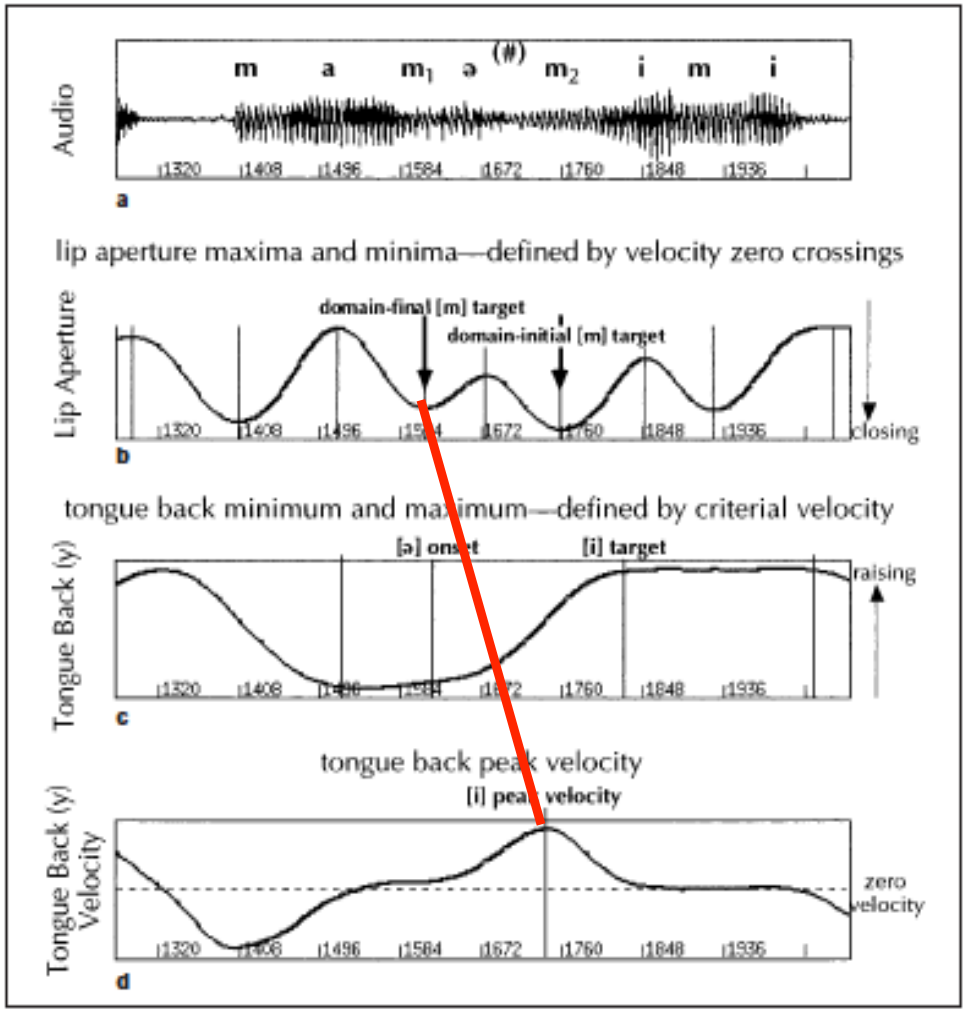


Figure adapted from Byrd 2000

Byrd, Krivokapić, & Lee (2006): Scope of Effect of the Boundary

- Byrd, Krivokapić, & Lee (2006) examine how far from the boundary the effect of the boundary extends.

<u>Design</u>

- **C3** ∨ **C2** ∨ **C1** ∨ **(#$_{IP}$)** **C1** ∨ **C2** ∨ **C3**
  Control [no boundary] versus IP boundary
- **How far left and right does lengthening extend?**

EMA Experiment: Stimuli

- Consonant sequence: /n/ /d/ /d/ /n/

- **Control**: C C C C
  Birdhunting, we were shocked to see a new dodo knocking on wooden posts.

- **Leftward test:** C C C # C
  Birdhunting, we were shocked to see a new dodo. # Knock on wood that there are more.

- **Rightward test:** C # C C C
  At the zoo, we were shocked to see a Gnu. Dodo knocking about, however, would have been more surprising.

Measurements of tongue tip movement

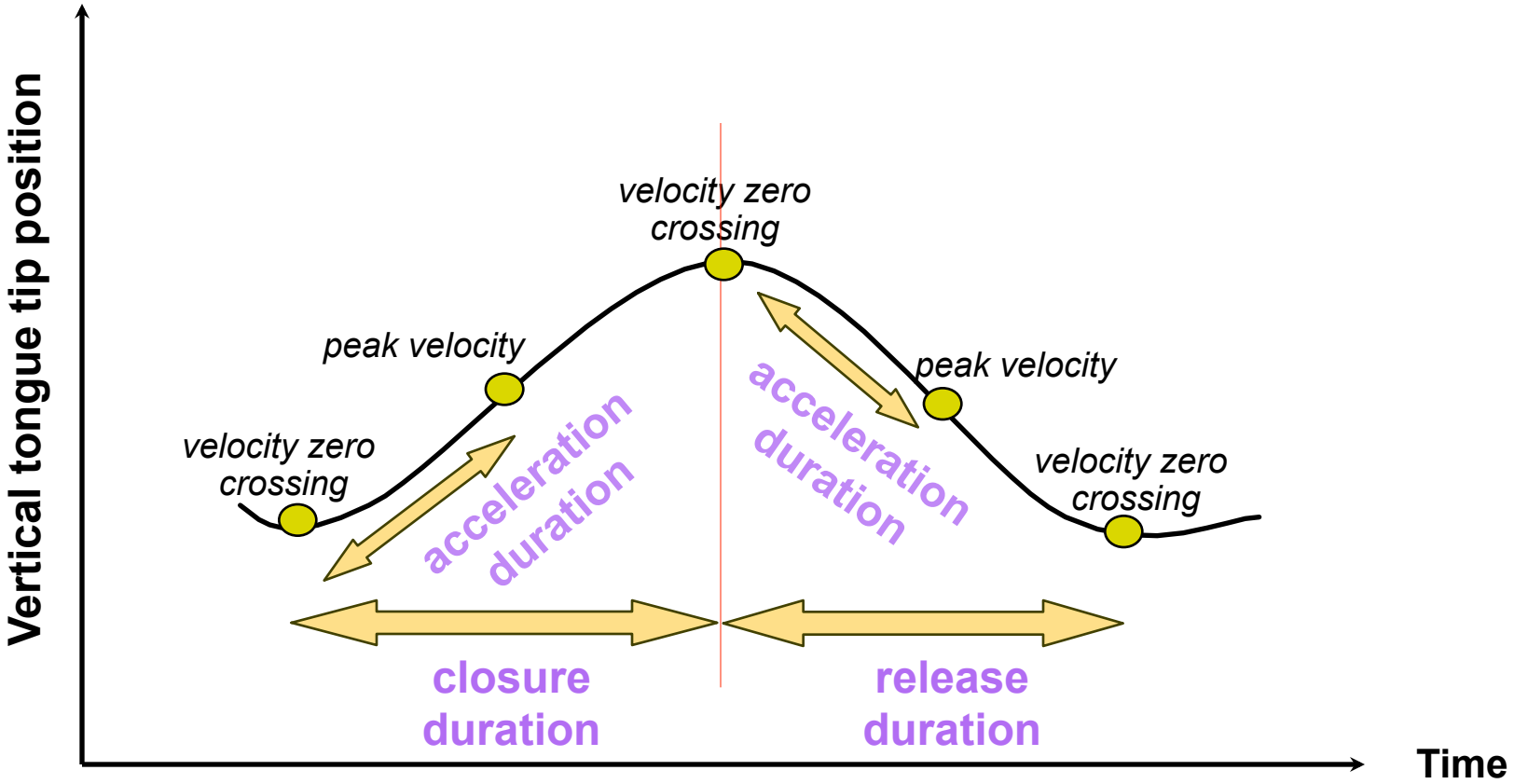
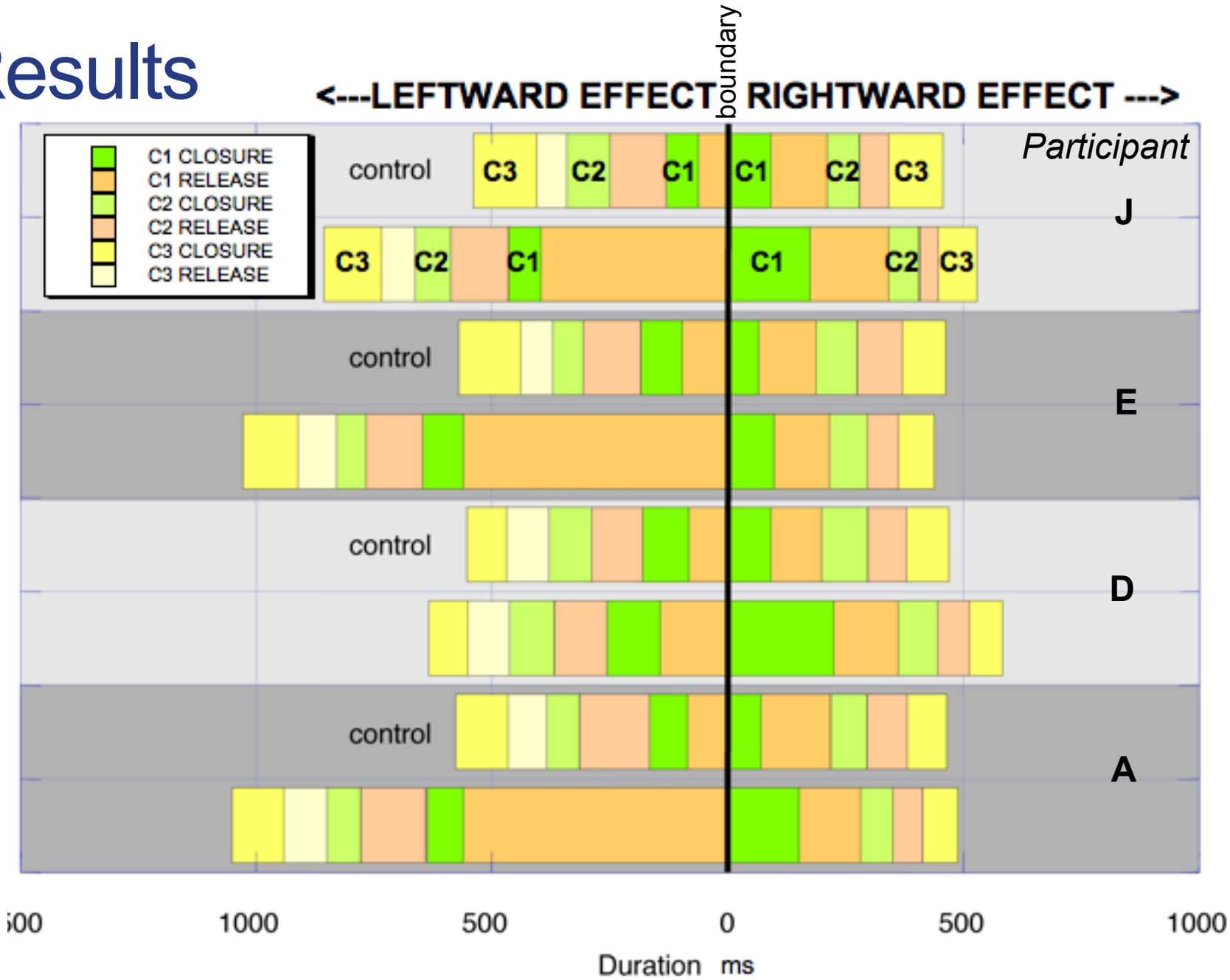


Figure adapted from Byrd, Krivokapić & Lee 2006

Results



Results

- Lengthening extends to the left and to the right of the IP boundary over a fairly stable temporal interval of two consonants for all participants.

Summary of temporal properties

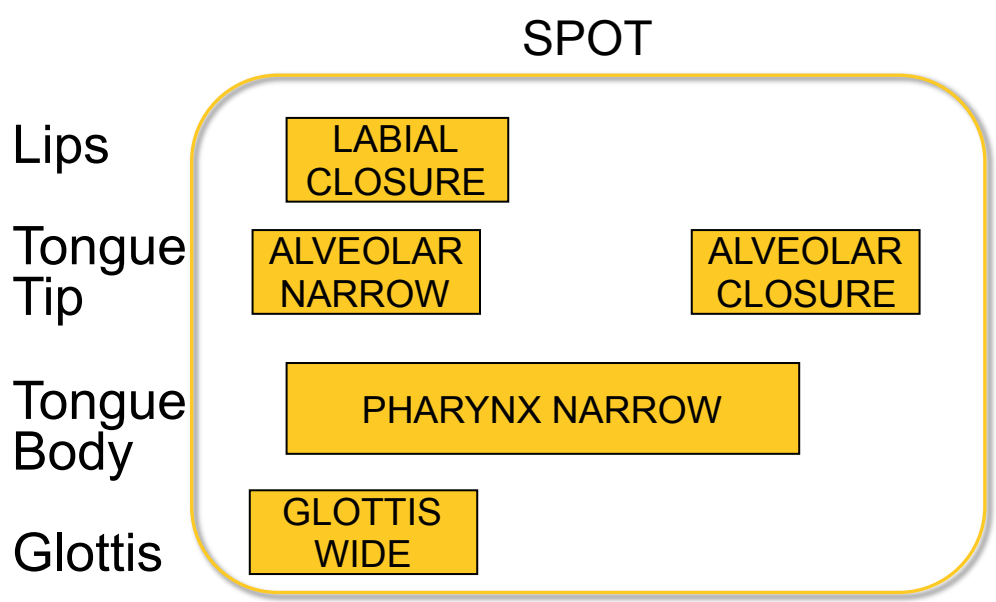
- Gestures lengthen at prosodic boundaries
- Gestures are less overlapped (co-articulated) at prosodic boundaries
- Lengthening is cumulative, increasing with boundary strength
- Lengthening extends over an interval
- There is a lot of variation between speakers
  - While all show lengthening, the categories distinguished and the amount of lengthening differs

PROSODY IN ARTICULATORY PHONOLOGY

Articulatory Phonology framework

Browman & Goldstein 1992 and elsewhere

- The gesture is the basic phonological unit.
  - A linguistically relevant, goal oriented movement of the vocal tract —for example, the lip closure for [b] in “bake.”
- Gestures are “units of action” and “units of representation”
- Gestures are **active over a temporal interval** and **overlap** in time.
- Gestures are **temporally coordinated.**



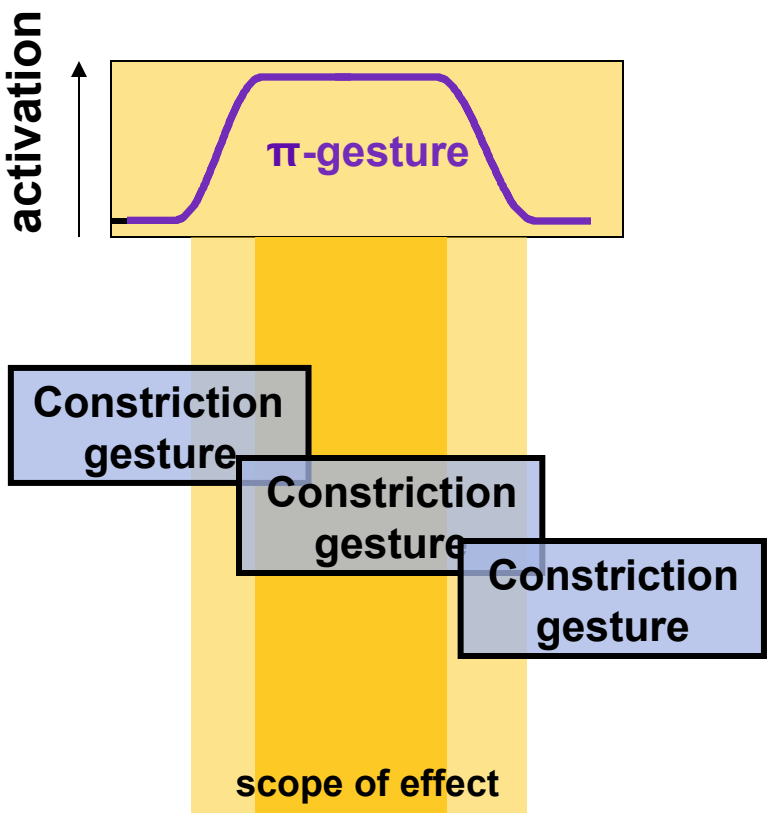
Adapted from
Goldstein, Byrd, & Saltzman, 2006

What about prosodic structure?

- **Prosodic events are understood to be temporal in nature**

  Byrd et al. 1998, Byrd & Saltzman 2003

Prosodic-gestural model Byrd & Saltzman 2003
(as integrated in the Articulatory Phonology framework)



Byrd & Saltzman (2003). Figure adapted from Byrd, Krivokapic, & Lee (2006)

- Within the π-gesture model. boundaries *are* prosodic gestures.
- Like constriction gestures, π-gestures **extend over a temporal interval**.
- They **slow the timecourse** of the constriction gestures co-active during the interval of the π-gesture.
  - Bigger boundaries⇒greater activation⇒more slowing
- Effect: gestures become bigger, larger, and less overlapped
  - The effect increases with boundary strength

Modeled effect (Byrd & Saltzman 2003)

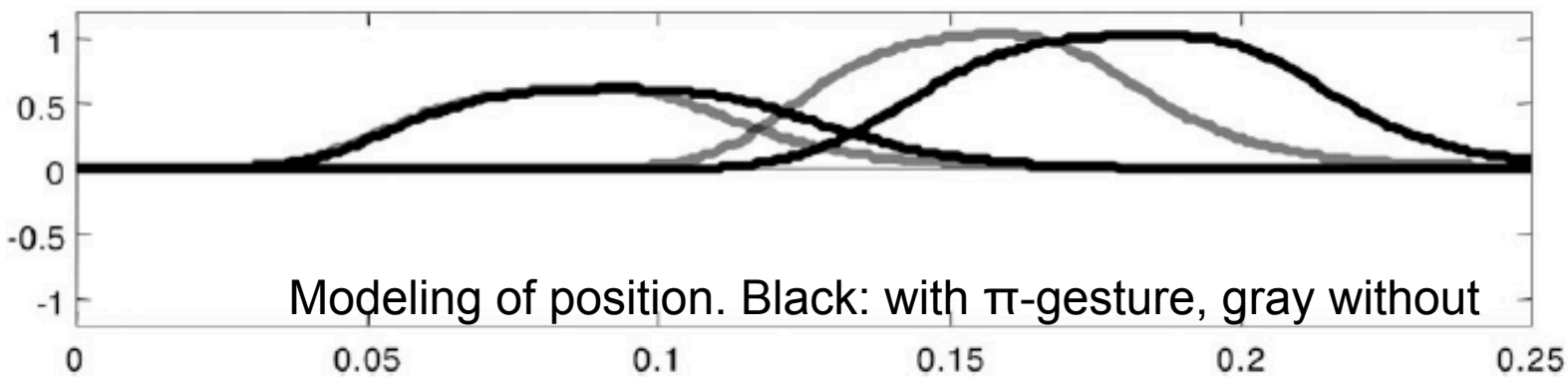


Figure adapted from Byrd & Saltzman 2003

What about other prosodic elements?

- Stress (Saltzman, Nam, Krivokapić, & Goldstein 2008)
- Tone gesture
  - Lexical tone (Gao 2008, 2009)
  - Pitch accent (Mücke et al. 2012),
  - Boundary tone (Katsika 2012, Katsika et al. 2014)
- Prosodic hierarchy
- Foot (Saltzman, Nam, Krivokapić, & Goldstein 2008)

Lexical tones in AP

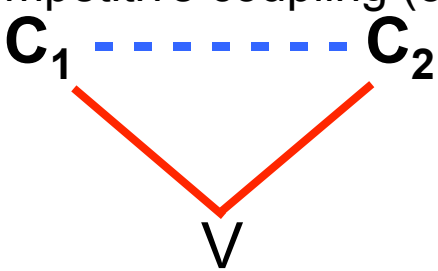
- Modeled as gestures (Gao 2008)
- Task: Linguistically relevant variations in fundamental frequency (e.g., F0 targets H and L). So, the targets are specified in the F0 space
- Articulators: Lungs, trachea, larynx, various muscles.
  - Use F0 as proxy
- Dynamic: Tone gesture defined as dynamic unit (with parameters and activation interval).
- Coordinated with other gestures and interact with them, behaving like a consonant gesture
  - E.g., with onset consonants and with the vowel in Gao (2008)

Oscillator Coupling & Gestural Timing Relations

(adapted from Marin & Pouplier 2010)

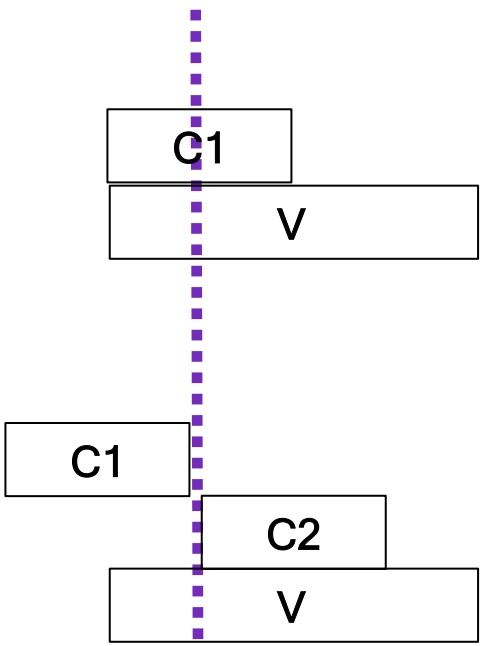
Oscillator coupling

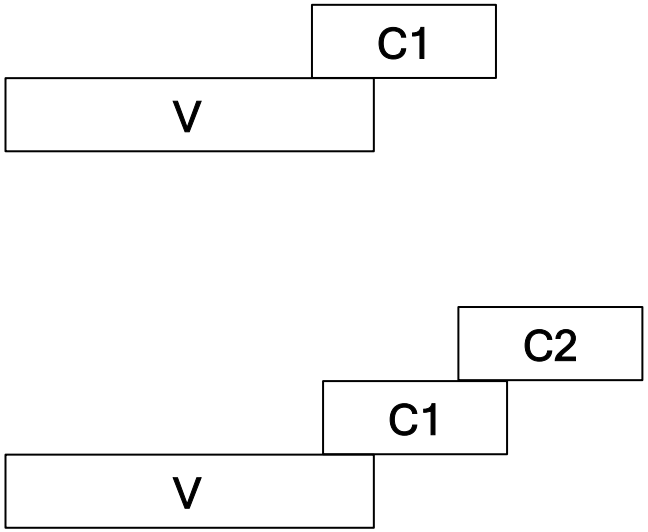
Competitive coupling (onset)



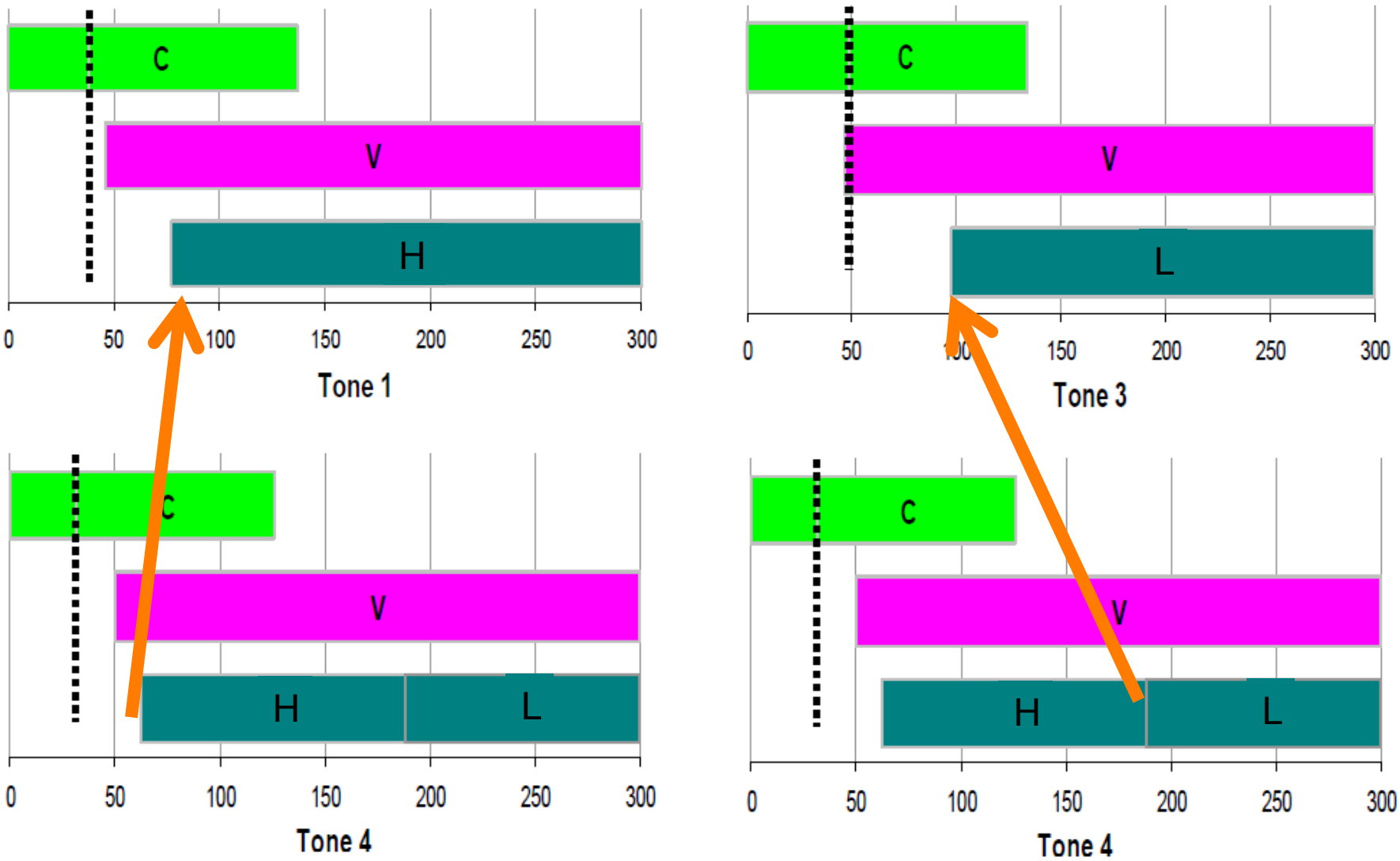
No competition (coda)

Gestural timing





Dotted line: mid points between the onsets of C and T gestures



T1, T2, T3: V onset and the mid point between the onsets of C and T gestures: nearly synchronous

T4: V onset nearly synchronous T1 (H) onset

Figure adapted from Gao (2009)

Tone gestures in pitch accents

- F0 in pitch accents can also be understood as tone gestures
  - e.g., study on Catalan and Standard German (Mücke, Nam, Hermes, & Goldstein 2012)
- Difference between tone gestures in lexical tones and pitch accents:
- In lexical tones the tone gesture affects the lexically specified CV coordination, in pitch accents it does not.
  - Lexical tone gestures are part of the lexical representation (so it can interact) while pitch accent tone gestures are not, so they don't affect the coupling.

BOUNDARY AND PROMINENCE INTERACTIONS

How is the π-gesture coordinated with other gestures

- π-gesture
- Prominence
- Boundary tone gesture

The coordination of prosodic gestures in Greek (Katsika, Ph.D. dissertation)

- The coordination of the boundary tone gesture
- The scope of boundary lengthening, and the interaction with stress and pitch accent
  - (cf., Turk & Shattuck-Hufnagel 2007, Byrd & Riggs 2008, Riggs & Byrd 2010)

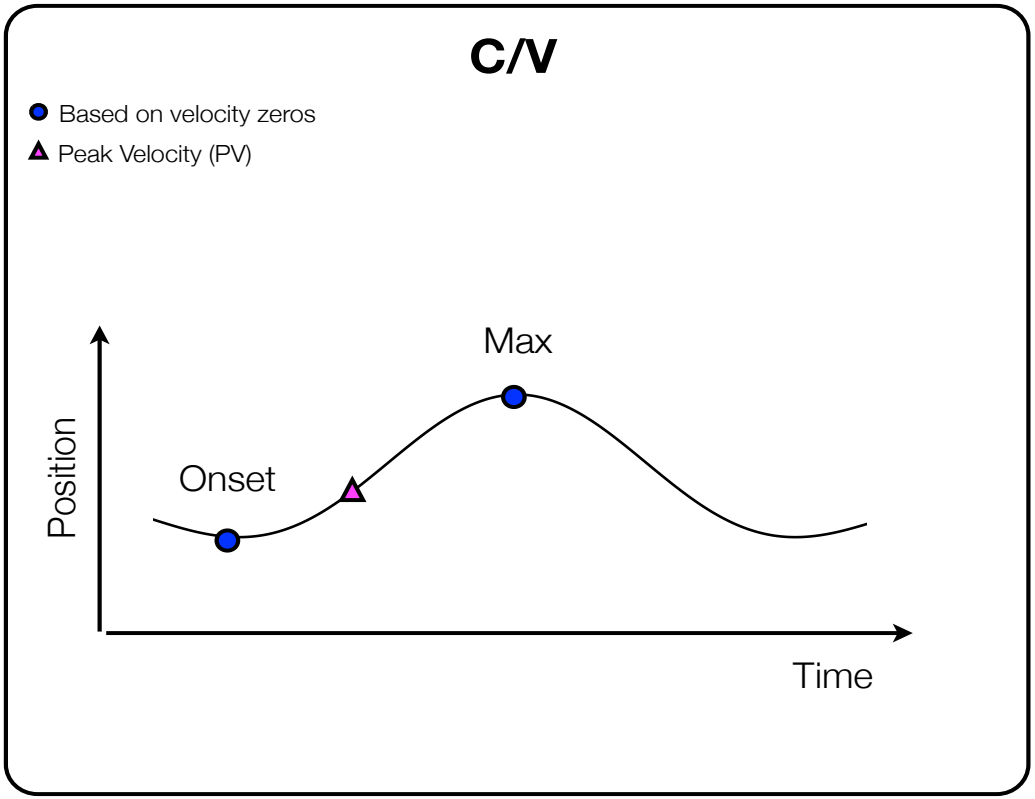
Final lengthening and alignment of the boundary tone
(Katsika, 2012, Katsika et al. 2014)

	Accented	De-accented
S1 (stress on 1st syllable)	**MAmima** # metaKSI	MAmima # metaKSI
S2 (stress on 2nd syllable)	**maMIma** # metaKSI	maMIma # metaKSI
S3 (stress on 3rd syllable)	**mamiMA** # metaKSI	mamiMA # metaKSI

Articulatory magnetometer study
8 participants
10 syntactic constructions *x* 3 words *x* 9 repetitions

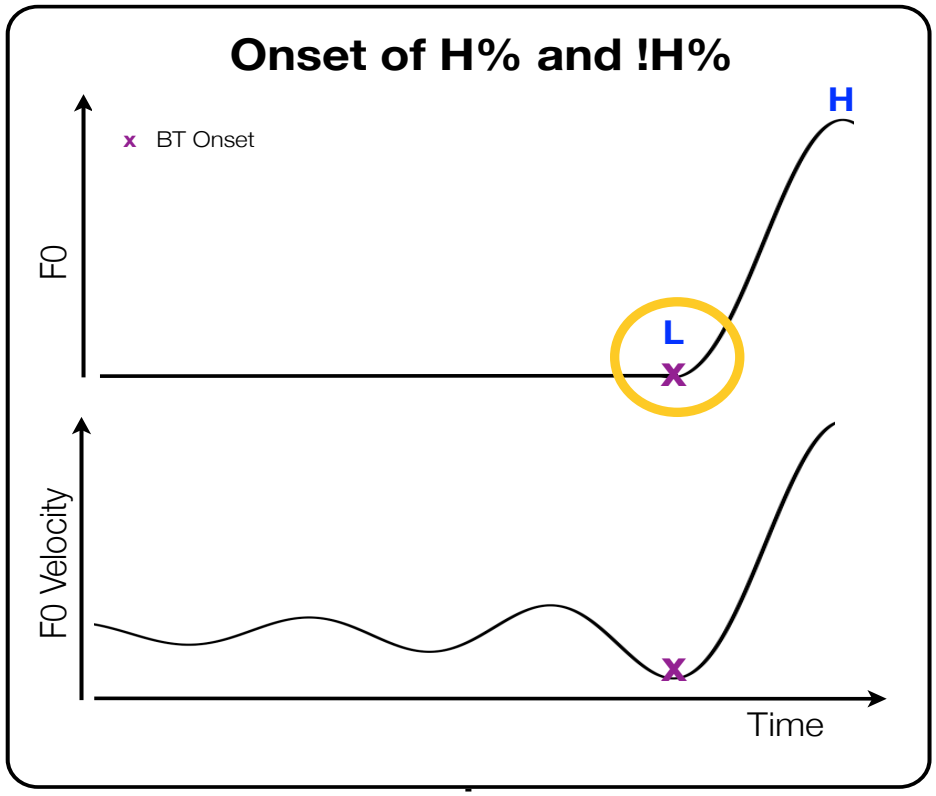
Labelling

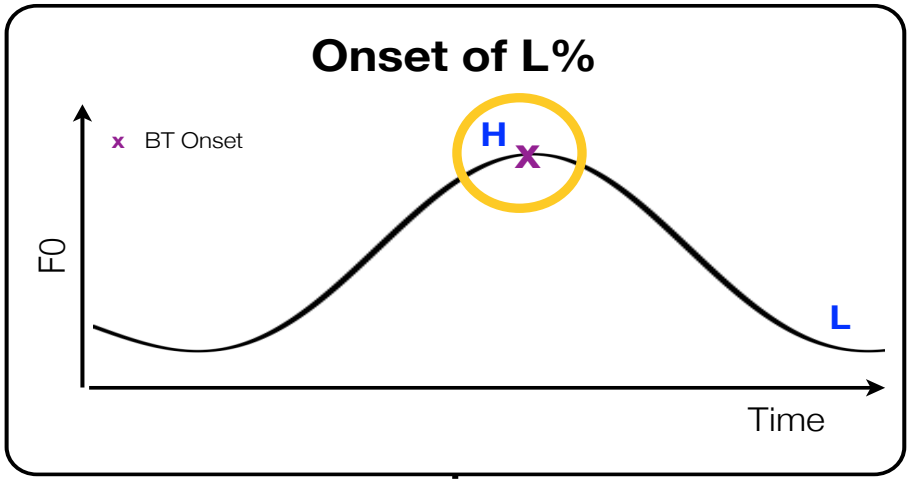
- Lip aperture for the /m/ consonants (C)
- Tongue dorsum for the vowels (V)



From Katsika 2012

F0 data labelling

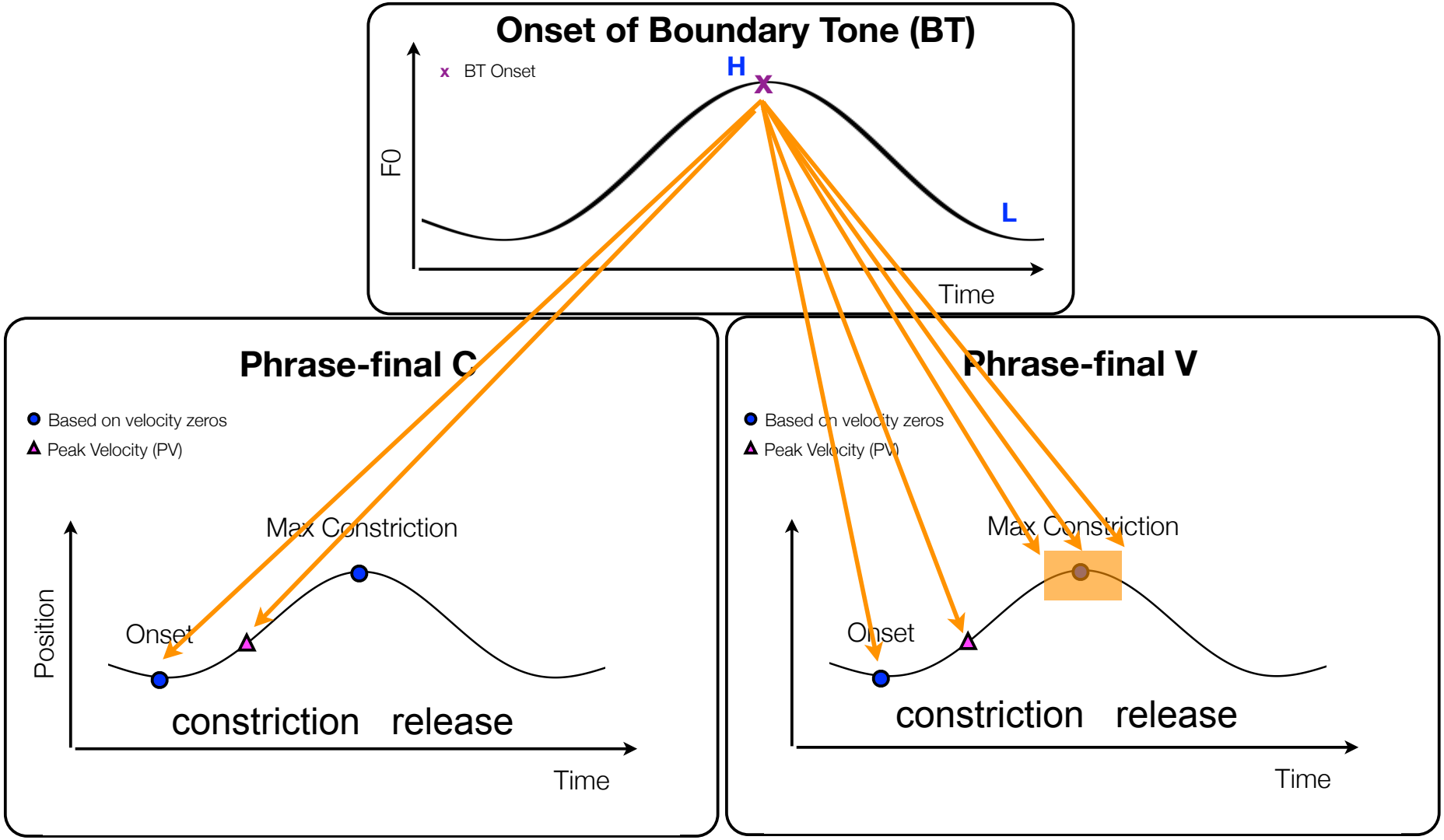




Adapted from Katsika 2012

Examined lags

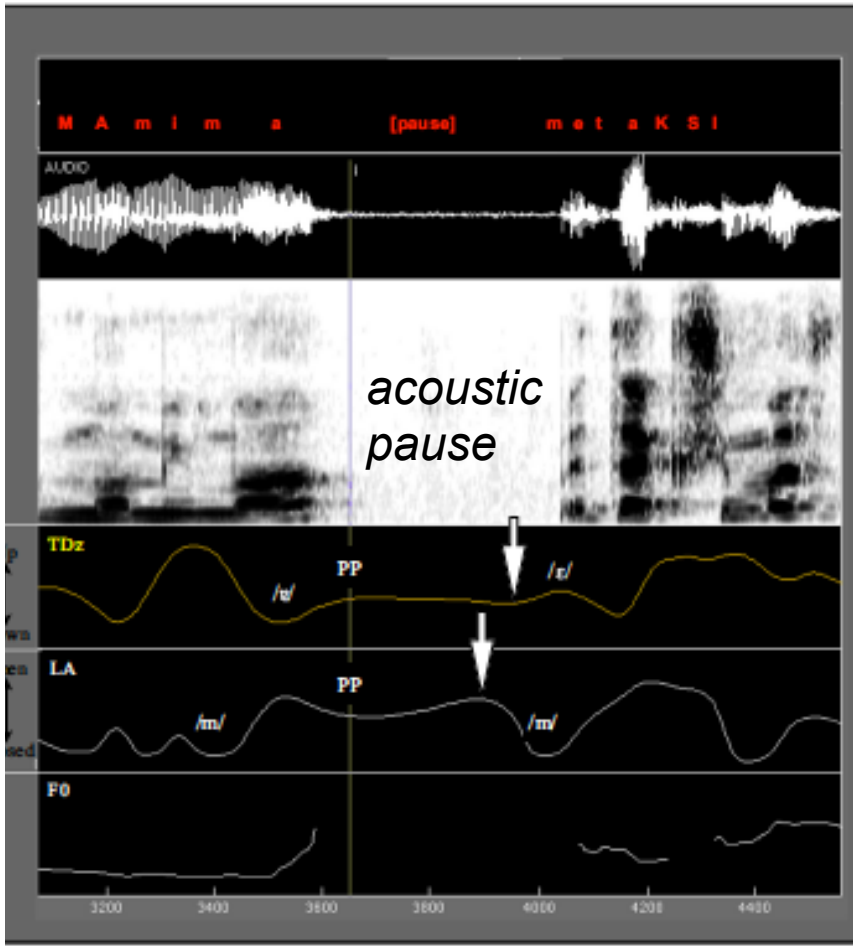
Adapted from Katsika 2012



Results

- The boundary tone onset occurs at the target of the final vowel, but, as the stress comes earlier, shifts leftwards towards the stressed syllable
- Final lengthening extends over the last syllable, but scope also shifts leftward towards the stressed syllable
  - See also Byrd & Riggs 2008, Riggs & Byrd 2010

Pause posture



- A pause posture occurs after the phrase-final gestures
  - Starting at a fixed time from the onset of the boundary tone (independently of stress condition).

PP: pause posture target achievement
Arrow: articulatory highest point before preceding the post-boundary consonant and vowel target.

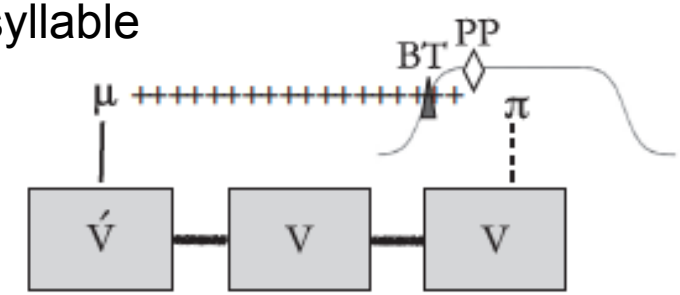
Figure adapted from Katsika et al. 2014

Gestural coordination:

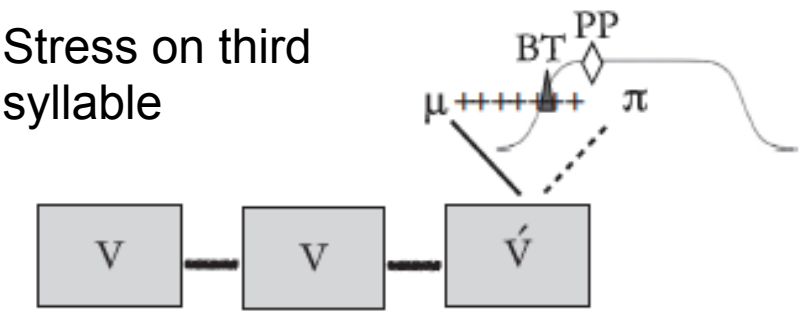
(Katsika 2012, Katsika et al. 2014)

- Coordination of the π-gesture with final syllable.
- Weaker coordination of the π-gesture with μ-gesture (on the stressed syllable)
  - => π-gesture shifts towards stressed syllable (earlier lengthening)
- Boundary tone (BT) gesture is activated when the π-gesture reaches a certain activation threshold
  - => BT onset earlier in words with earlier stress
- Pause posture (PP) is triggered when the π-gesture reaches a certain activation threshold (higher than for the BT).

Stress on first syllable



Stress on third syllable



Adapted from Katsika et al. 2014

Summary

- A complex interaction of boundaries, syllables, and prominence
- π-gesture shifts towards prominence
- π-gesture serves as a trigger for tonal events

Katsika 2012

Summary

- At boundaries, there is spatial strengthening
- Gestures lengthen at prosodic boundaries
- Lengthening is cumulative, increasing with boundary strength
- Lengthening extends over an interval
- Gestures are less overlapped (co-articulated) at prosodic boundaries
- There is a lot of variation between speakers
  - While all show lengthening, the categories distinguished and the amount of lengthening differs
- Within AP, prosodic structure can be understood to arise through the coordination of constriction and prosodic gestures.
  - Not covered here, but a thought for further research: how do body gestures fit in?

REFERENCES

- Beckman, M.E., & Pierrehumbert, J.B. (1986). Intonational structure in Japanese and English. Phonology Yearbook, 3, 255–309.
- Beckman, M. E., Edwards, J. & Fletcher, J. (1992) Prosodic structure and tempo in a sonority model of articulatory dynamics. In Papers in laboratory phonology II: gesture, segment, prosody (G. J. Docherty & D. R. Ladd, editors) pp. 68Ð86. Cambridge, U.K: Cambridge University Press
- Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. Phonetica, 49, 155-180.
  - See also: Haskins Laboratories: Introduction to Articulatory Phonology and the Gestural Computational Model at: http://www.haskins.yale.edu/research/gestural.html
- Byrd, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. Phonetica, 57 (1), 3-16.
- Byrd, D., Kaun, A., Narayanan, S., & Saltzman., E. (2000). Phrasal signatures in articulation. In M. B. Broe, & J. P. Pierrehumbert, (Eds.). Papers in Laboratory Phonology V. New York: Cambridge University Press. Pp. 70-87.
- Byrd, D., Krivokapic, J., & Lee, S. (2006). How far, how long: On the temporal scope of phrase boundary effects. Journal of the Acoustical Society of America, 120, 1589-1599.

- Byrd, D., & Saltzman, E. (1998). Intragestural dynamics of multiple phrasal boundaries. Journal of Phonetics, 26, 173-199.
- Byrd, D., & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. Journal of Phonetics, 31, 149-180.
- Cho, T., & Keating, P. (2001) Articulatory and acoustic studies of domain-initial strengthening in Korean. Journal of Phonetics , 29, 155-190.
- Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. Journal of the Acoustical Society of America, 101, 3728-3740.
- Gao, M. (2008). Mandarin Tones: an Articulatory Phonology Account. Ph.D. dissertation, Yale University.
- Gao, M. (2009). Gestural Coordination among Vowel, Consonant and Tone Gestures in Mandarin Chinese. Chinese Journal of Phonetics, 2, 43.
- Katsika, A. (2012). Coordination of prosodic gestures at boundaries in Greek. Ph.D. Dissertation. Yale University.
- Katsika, A., Krivokapic, J., Mooshammer, C., Tiede, M., & Goldstein, L. M. (2014). The coordination of boundary tones and their interaction with prominence. Journal of Phonetics, 44, 62-82.
- Keating, P. (2006). Phonetic encoding of prosodic structure. In Speech production: Models, phonetic processes, and techniques, edited by J. Harrington & M. Tabain (Eds.). Macquarie Monographs in Cognitive Science, Psychology Press, New York and Hove. Pp. 167-186.

- Keating, P., Cho, T., Fougeron, C., & Hsu, C. (2004). “Domain-initial articulatory strengthening in four languages”, in Phonetic Interpretation (Papers in Laboratory Phonology VI), edited by J. Local, R. Ogden and R. Temple. (Cambridge, UK: Cambridge University Press). Pp. 143-161.
- Krivokapić, J. (2014). Gestural coordination at prosodic boundaries and its role for prosodic structure and speech planning processes. Communicative rhythms in brain and behaviour. Theme Issue of the Philosophical Transactions of the Royal Society B (Biology), 369, 20130397.
- Krivokapić, J., & Byrd, D. (2012). Prosodic boundary strength: An articulatory and perceptual study. *Journal of Phonetics,* 40, 430–442.
- Ladd, D. R. (1988). Declination “reset” and the hierarchical organization of utterances. *Journal of the Acoustical Society of America*, 84, 530-544.
- Ladd, D. R. (1996/2008). *Intonational Phonology*. Cambridge: Cambridge University Press.
- Mücke, D., Nam, H., Hermes, A., & Goldstein, L. (2012). Coupling of tone and constriction gestures in pitch accents. In P. Hoole, M. Pouplier, L. Bombien, Ch. Mooshammer & B. Kühnert (Eds.), Consonant Clusters and Structural Complexity. Mouton de Gruyter. 205-230.
- Ramanarayanan, V., Bresch, E., Byrd, D., Goldstein, L. & S. Narayanan (2009). Analysis of pausing behavior in spontaneous speech using real-time magnetic resonance imaging of articulation, *Journal of the Acoustical Society of America Express Letters*, 126, (EL160-EL165)

- Ramanarayanan, V., Byrd, D., Goldstein, L. & S. Narayanan (2010). Investigating articulatory setting - pauses, ready position and rest - using real-time MRI, In: *Proceedings of Interspeech 2010*, Makuhari, Japan, Sept 2010
- Ramanarayanan, V., Goldstein, L., Byrd, D., & S. Narayanan (2013). A real-time MRI investigation of articulatory setting across different speaking styles, *Journal of the Acoustical Society of America*, 134, 510-519.
- Rochet-Capellan, A., Laboissière R, Galván A, & JL Schwartz (2008). The speech focus position effect on jaw-finger coordination in a pointing task, Journal of Speech Language and Hearing Research, 51, 1507-1521.
- Saltzman, E., Nam, H., Krivokapic, J., & Goldstein, L. (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. In P. A. Barbosa, S. Madureira, & C. Reis, (Eds.).(2008). Proceedings of the Speech Prosody 2008 Conference, Campinas, Brazil.
- Swerts, M. (1997). “Prosodic features at discourse boundaries of different strength”, J. Acoust. Soc. Am., 101, 514-521.