Ioana Chitoran
Dartmouth College, Hanover, NH, USA

Dynamic representations
The syllable

- Dynamic representations in phonology
  - Are symbolic dynamic representations plausible?
- Focus on the syllable
- Case studies – including Romance languages

Lecture 1

Introduction to gestures and Articulatory Phonology

Speech has been described in two ways:

- Phonological description
  - Sequence of discrete symbols from a finite set
  - Combinatorial properties
- Physical description
  - As gradient, continuous, context-dependent variation in several parameters – acoustic, articulatory, aerodynamic

Are the two descriptions compatible?

- The relationship between phonetics and phonology
  Two views:
  - Separate representations – implementation / mapping of discrete units onto continuous parameters in time and space *(modular)*
  - Shared representations in phonetics and phonology *(unidimensional)*
    - Reductionist – phonetic detail directly encoded in phonology
    - Non-reductionist – continuous physical speech measurements can be decomposed into discrete events

Cognitive and physical properties of speech

- Articulatory Phonology – AP *(Browman & Goldstein 1988, 1992, 1995, 2000; Goldstein & Fowler 2003; Goldstein, Byrd, Saltzman 2006)*
- Phonological and physical structure of speech constrain each other:
  - constraints on the system, determined by physical properties
    - Gaps in inventories explained by reference to physical properties
  - constraints on physical properties, determined by the structure of the system
    - Contextual variation may be constrained by the system: Context-dependent variation in vowels varies as a function of the number of vowels in the inventory (Manuel 1990)

Representations needed to characterize phonological systems

- Encode contrast
- Combine

Proposal: articulatory gestures

- Units of *contrast* and *combination*
- Encode symbolic properties of speech
- Linked to continuous physical structure by virtue of being *dynamic*

Unified representation

- Dual nature
  - symbolic — *"units of information"*
  - physical — *"units of action"*

- Earlier proposals – *Feature Theory*

  Same features shared by two domains – *Jakobson, Fant, Halle; Halle; Stevens*

Gestural hypothesis

Cf. Fowler et al. (1980)

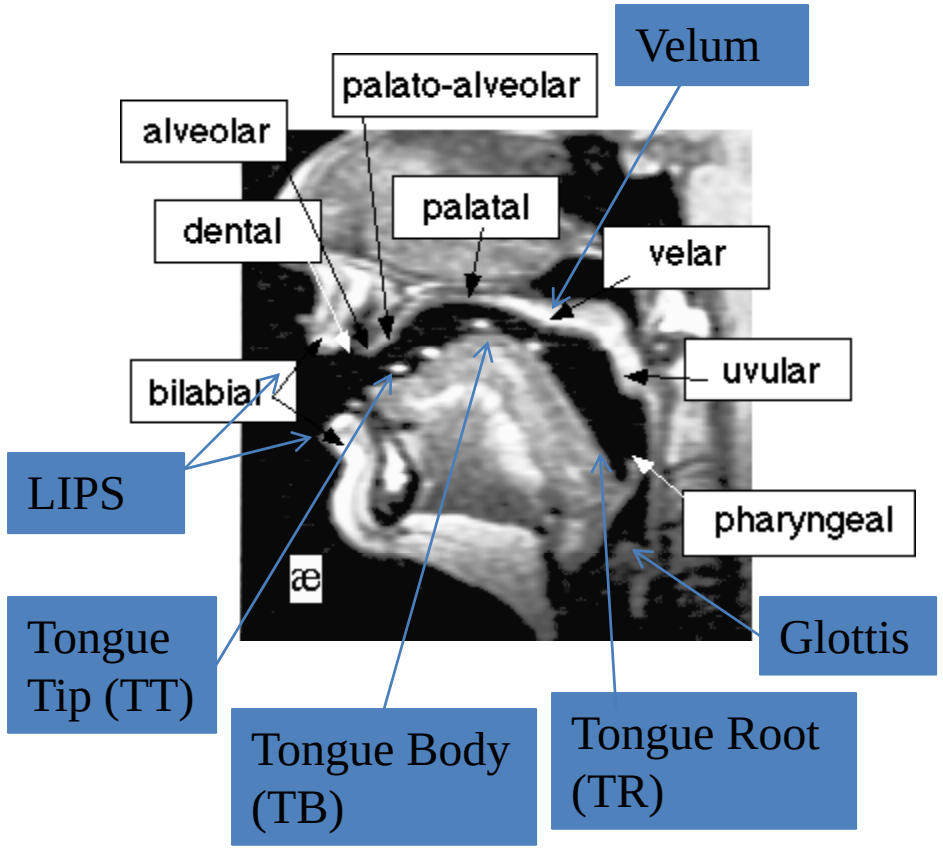
- *Products* of speech are continuous and context-dependent
- The act of speech can be decomposed into discrete, context-independent *actions of the vocal tract*

Actions of the vocal tract

Organs that can act independently
Mechanically coupled

Gesture – a constriction action of an organ

Unit of action



Hypothesis:

- The time-varying and context-dependent properties of speech result from the unfolding of the units' dynamics:
  - producing the same unit in different contexts yields different movements and sound

Decomposition of speech production into dynamical units – Main empirical evidence

- Experiments with articulator perturbation – Organ-specific reaction *(Kelso et al., 1984)*

  Lip closure gesture for [p] involves coordinated movements of three articulators:

  – Lower lip moves up, upper lip moves down, jaw raises – *coordinative structure, functional synergy*

  – Prevent lip closure by pulling lower lip down → upper lip reacts instantly by lowering further

  Prevent TT closure → no reaction from lips

Evidence from facial imitation by infants

- Neonates show sensitivity to the partitioning of the oro-facial system into distinct organs *(Meltzoff & Moore 1977)*
  - Infant cannot see its own face
  - Infant has no proprioception from the model's face
- Imitation is specific to the organ involved

Evidence from speech production errors

- The most frequent sublexical units involved in errors are single segments *(Shattuck-Hufnagel 1983)*

coffee pot > poffee cot

- Errors unlikely to be transcribed if partial
- Evidence for gradient errors from acoustic analysis *(Frisch & Wright 2002, Goldrick & Blumstein 2006)*

- Error Elicitation: Repetition of phrases with "alternating" consonants

cop top, kip tip, bad bang

- EMMA, 10-15 seconds, synchronized to metronome
- Variations in rate, stress, order
- Non-alternating controls

(Goldstein, Pouplier, Chen, Saltzman, Byrd 2006)

- During repetition task many errors involve the production of an “extra”, intruding gesture, along with the intended one:

  e.g., in **c**op **t**op, tongue dorsum ([k]-like) raising gesture during [t]

- Errors vary continuously in magnitude
- Small movements are perceived as “normal” (report: /t/)
- Large movements perceived as errors (report: /k/)

Where do errors occur?
Planning or execution phase? Evidence for both…

- Assuming that units of both planning and execution are dynamical units of action *(gestures)* could account for such errors.

Proposal:

- Errors arise from competition between:
  - *gestures as planning units* (lexical constraints on coordination of gestures)
  - *gestures as execution units* (intrinsic dynamical constraints on coordination)

A word as gestural structure:

bæn

Organs

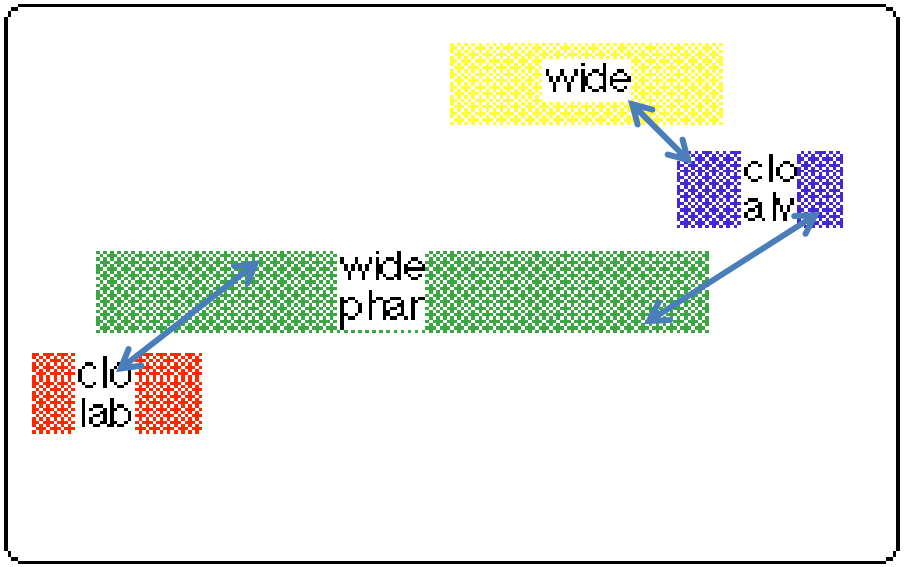
VELUM

TONGUE TIP

TONGUE BODY

LIPS

GLOTTIS



Coupling graph

Arrows → coordination relations between gestures

Lecture 2

Questions about phonotactics (combinatorial properties)

Answers from gestures (units of action)

Georgian stop-stop sequences

C_1C_2	Word-initial sequences	Word-medial sequences
Front-to-back	bgera 'sound' p^ht^hila 'hair lock' dg-eb-a 'stands up'	abga 'saddle bag' ap^ht^har-i 'hyena' a-dg-eb-a 'will stand up'
Back-to-front	g-ber-av-s 'is inflating you' t^hb-eb-a 'it is warming up' gd-eb-a 'to be thrown'	da-gbera 'say the sounds' ga-t^hb-a 'it has become warm' a-gd-eb-a 'throw in the air'

Georgian consonants

- Stop inventory: b d dz ʤ g
  p^h t^h ts^h $tʃ^h$ k^h
  p’ t’ ts’ tʃ’ k’ q
- “Harmonic clusters” : ***bg, dg, p’k’, t’k’, p^hk^h, t^hk^h***
  - share laryngeal specifications
  - labial-dorsal and coronal-dorsal only
  - claimed to pattern as single segments
    - Morphology: max 3 stops per root, except if one is a harmonic cluster
    - Cannot be broken by a syllable boundary

- Acoustic evidence:

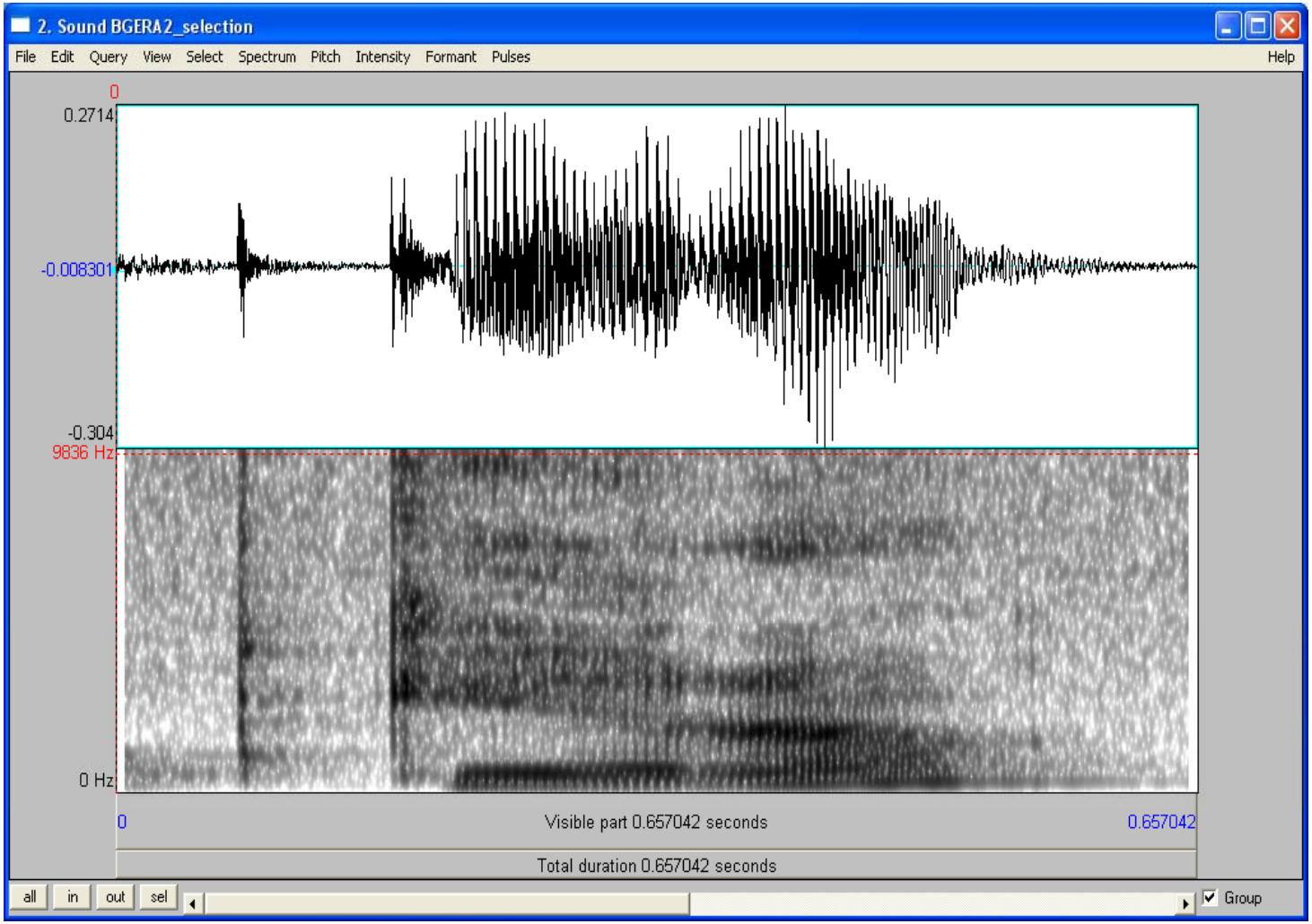
  No structural difference from stop-stop sequences.
  Each stop is clearly released.

  (Chitoran 1998, McCoy 1999)

- Mixed intuitions on syllabification:

  a . dge . ba ~ ad . ge . ba **adg . e . ba*

… *[bgera]*…


2. Sound BGERA2_selection
File Edit Query View Select Spectrum Pitch Intensity Formant Pulses Help
0
0.2714
-0.008301
-0.304
9836 Hz
0 Hz
0
Visible part 0.657042 seconds
0.657042
Total duration 0.657042 seconds
all in out sel
Group

Questions about Georgian

- Does sonority play a role in Georgian?
- What is special about “harmonic clusters”?
- Issues of gestural overlap and perceptual recoverability
  - Variation in speech production may be constrained by the limits of the listener, who must be able to recover the linguistic units and the message from the signal *(Kingston 1985, 1990; Silverman & Jun 1994; Silverman 1995; Byrd 1996; Wright 1996)*

Perceptual recoverability in stop sequences

- Sequences of stops (C_1C_2) are especially sensitive to variation in articulatory timing:
  - formant transitions only occur on one side of each stop
  - high degree of gestural overlap will obscure C1 release burst
  - high degree of gestural overlap will decrease the specificity of formant transitions

Hypotheses

- Stop sequences will be less overlapped
  - Word-initially than word-medially
  - In back-to-front than in front-to-back order of constriction location

Based on :

Effect of Word Position

- Sequences of consonant gestures exhibit less temporal overlap in a word onset than in a coda or across syllables. *(Byrd 1996, Hardcastle 1985, Wright 1996; preliminary acoustics Chitoran 1999)*
- Threat to perceptual recoverability is particularly problematic in utterance-initial position.
  - Potentially no VC transitions
  - Important in lexical access

Order of Place of Articulation

(front-to-back vs. back-to-front)

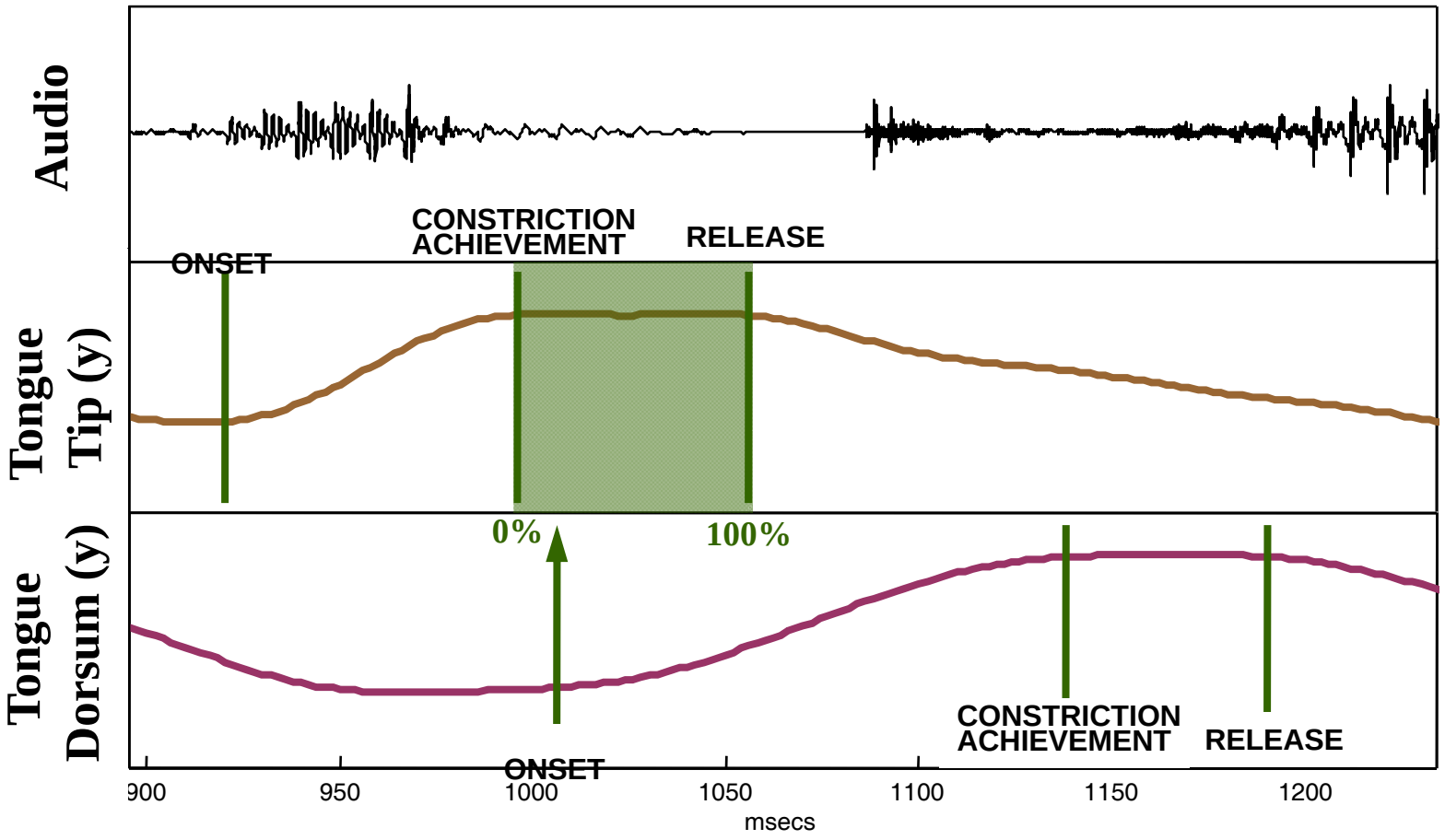
- No acoustic manifestation of C_1 release if:
  - the constriction for C_2 is already formed
  - and C_2 is **anterior to** C_1.
- A more limited degree of overlap is predicted for such a back-to-front sequence.
- Previous results demonstrating this effect:
  - perceptual studies (labials & coronals): *Byrd, 1992; Surprenant & Goldstein, 1998*
  - articulatory studies (coronals & dorsals): *Hardcastle & Roach, 1979; Byrd, 1996; Zsiga, 1994*
  - acoustic studies: *Wright, 1996; Chitoran, 1999*

Experimental Hypotheses:
Oral constriction gestures

H1: Word-initial stop-stop sequences will be less overlapped than like word-internal sequences.

H2: Stop-stop sequences with a back-to-front order of constriction location will have less gestural overlap than stop-stop sequences with a front-to-back order.

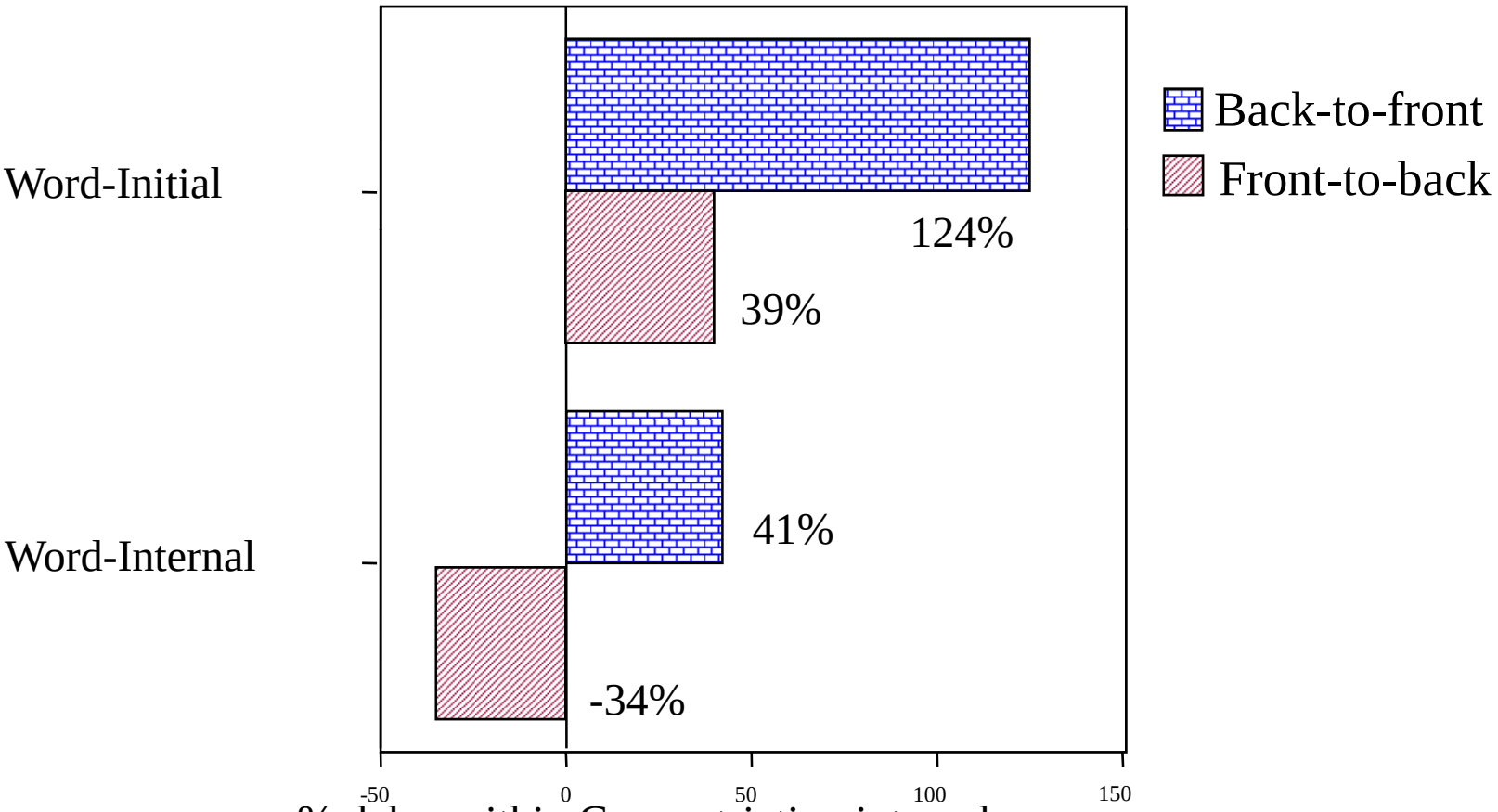
Example:
Analyzed Articulatory Events



Utterance: *'...dgeba...'*

Results (speaker 1)

How early does C_2 onset occur within the constriction 'plateau' interval of C_1?



Discussion

- Articulatory coordination is sensitive to both:
  - an efficiently coproduced motor behavior that encodes and transmits information at a high rate;
    - *and*
  - Optimizing the chance of successful perceptual recoverability.

    Speakers control the timing of coproduced articulations in order to make their acoustic effects more clear

- In certain contexts, the recoverability requirement appears to be paramount —
  - e.g., when the units are in a prosodically important position (i.e. word initially) and/or when acoustic information would be obscured with substantial overlap.

Implications for sonority

- The sonority scale: O < N < L < G < V
- Phonetic definition: increased perceptibility of segments/gestures *(Mattingly 1981, Ohala 1990)*
- Sonority sequencing is a way of achieving "parallel transmission".
  - C sequences allowing substantial overlap *while* maintaining recoverability are more common
  - C sequences requiring more limited overlap to preserve information are less common

Implications for the status of "harmonic" clusters

- Substantial overlap in front-to-back sequences can account for their laryngeal homogeneity.
  - Less overlap = possible mixed voicing
    - homogeneous, back-to-front: ***gd****eba*
    - hon-homogeneous, back-to-front: ***t'b****a*
  - More overlap = laryngeal homogeneity
    - homogeneous, front-to-back ("harmonic"): ***dg****oma*
    - non-homogeneous, front-to-back: unattested ***(*bk^h)***
- Assumption: one laryngeal gesture per onset cluster, coordinated with C1.

Questioning perceptual recoverability

- Stop-liquid / liquid-stop clusters
- Perception of stop clusters *(preliminary results)*

HYPOTHESIS: STOPS AND LIQUIDS

If the order of place effect is due exclusively to perceptual recoverability, combinations of stops and liquids should not show this effect.

- Overlap in [pl] front-to-back, and [kl] back-to-front, should be comparable

  The acoustic release of the stop is never completely hidden, regardless of the amount of overlap

- Overlap in [rk] front-to-back, and [rb] back-to-front, should be comparable

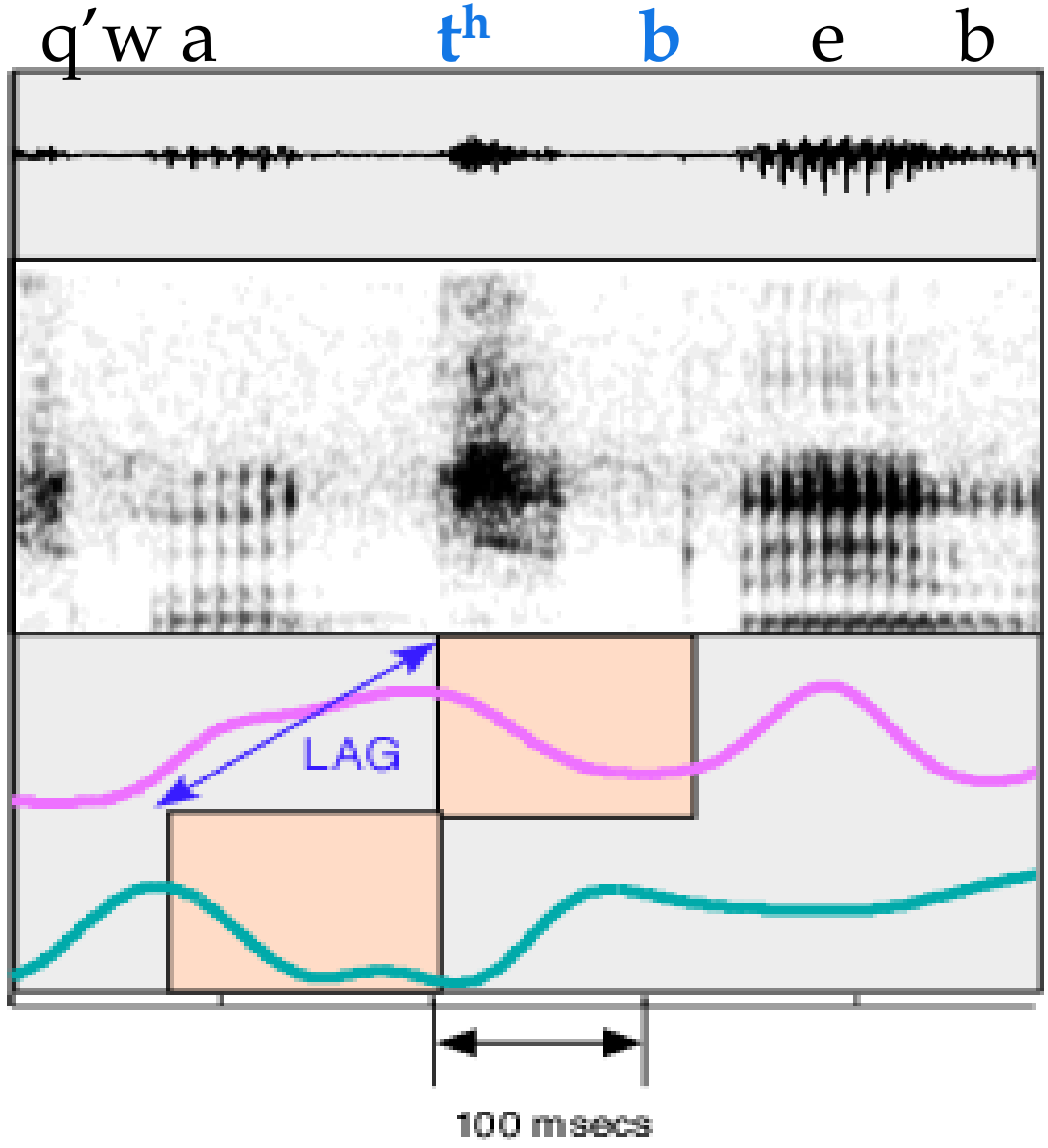
  The acoustic release of the liquid is not necessary to its perception

Overlap measure: ONSET LAG

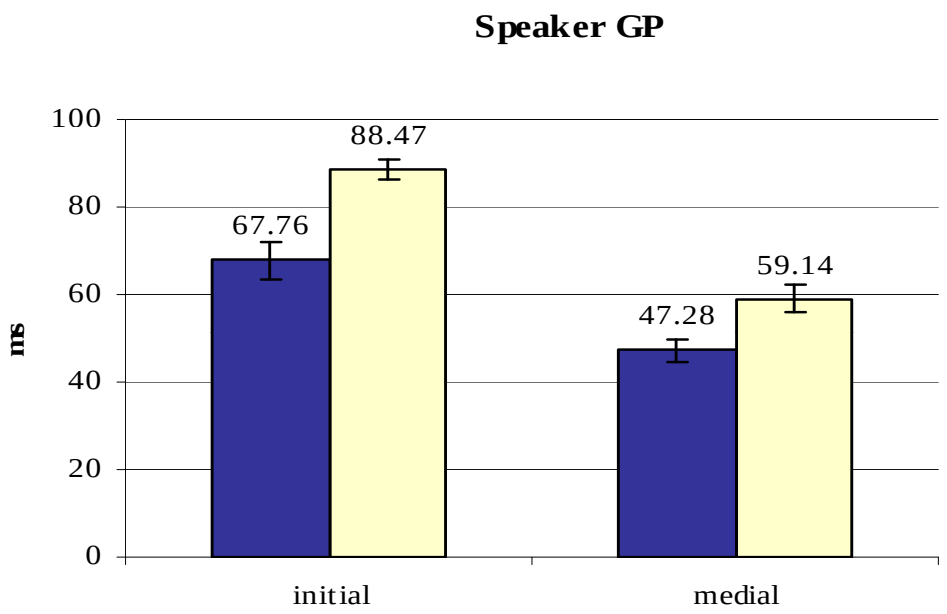
Time between onset of C1 gesture and onset of C2 gesture

[…q'wa#t^hbeb…] in

[sit'q'wa#t^hbeba]



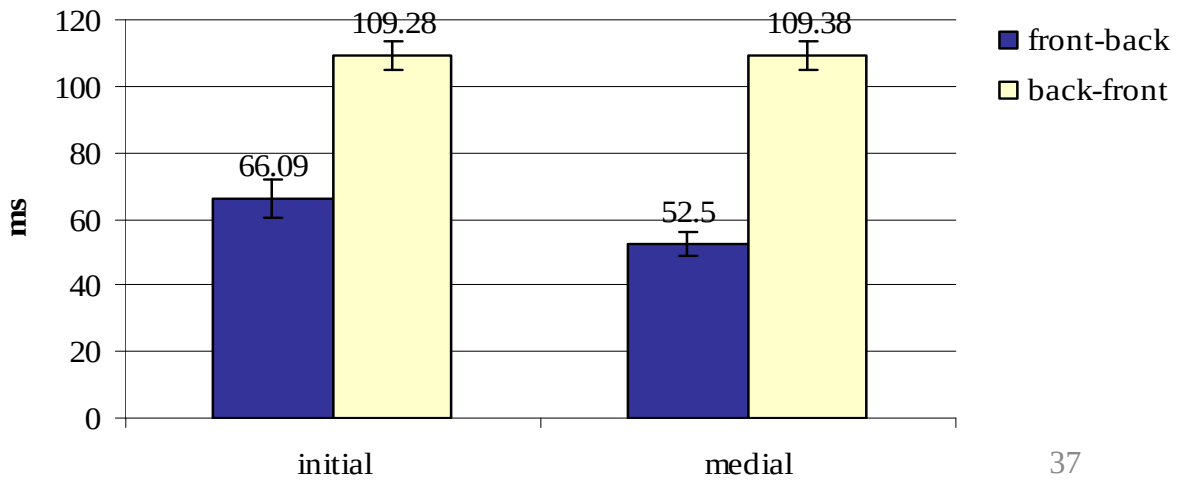
Replication of stop-stop results



→ Less overlap
word-initially and
in back to front sequences
$p < .0001$

Less overlap - back
to front sequences

$p < .0001$ →

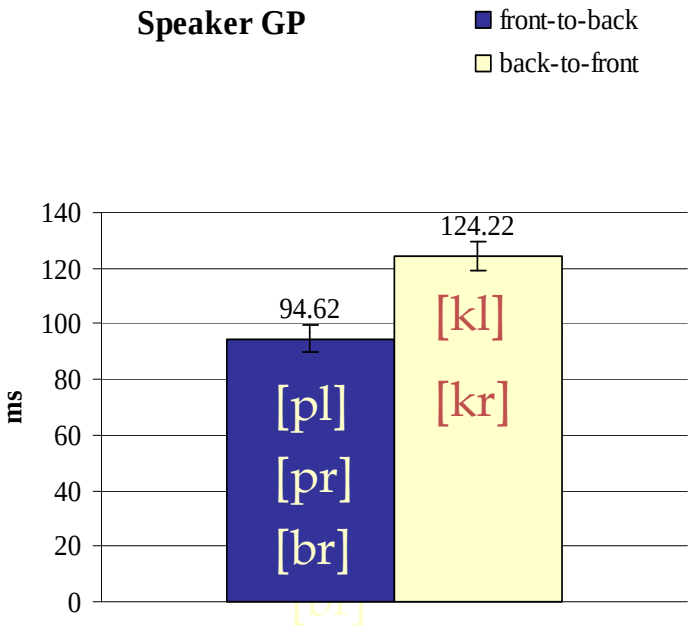


Experiment 2: Stop-liquid sequences

C_1C_2	Front to back (labial-coronal)	Back to front (dorsal-coronal)
	p^hlet^h-a ‘to tear up’	k’leb-a ‘reducing’
	p’res-a ‘press’	k’reb-a ‘to meet’
	p’rasa ‘leek’	k’rep^h-a ‘picking’
	braz-i ‘anger’	k’rav-i ‘lamb’

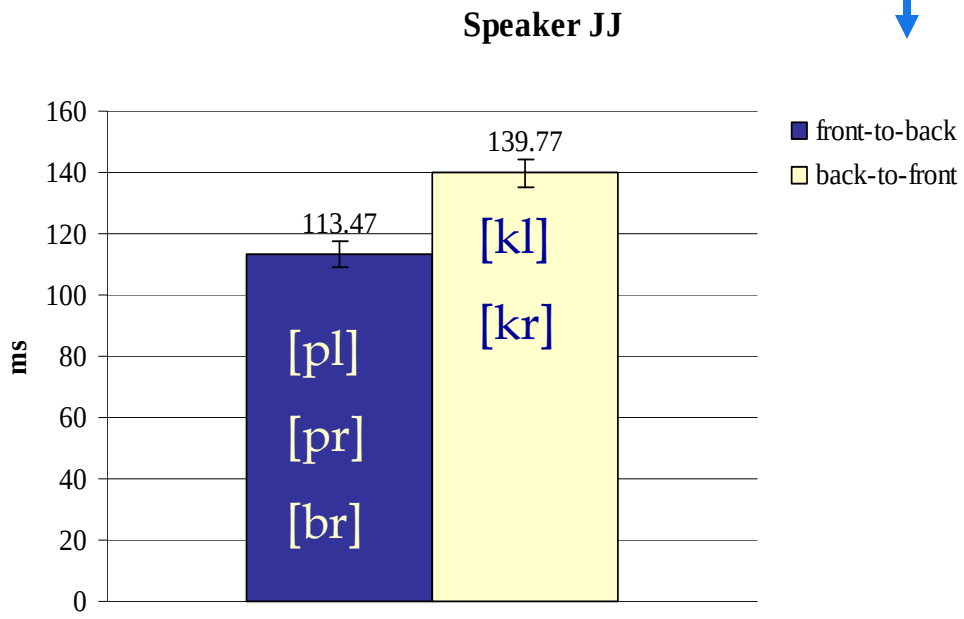
- 7 repetitions, randomized, in carrier phrase

Results



Longer lag (less overlap) in back to front [kl, kr] than front to back [pl, pr, br]

$p < .0001$

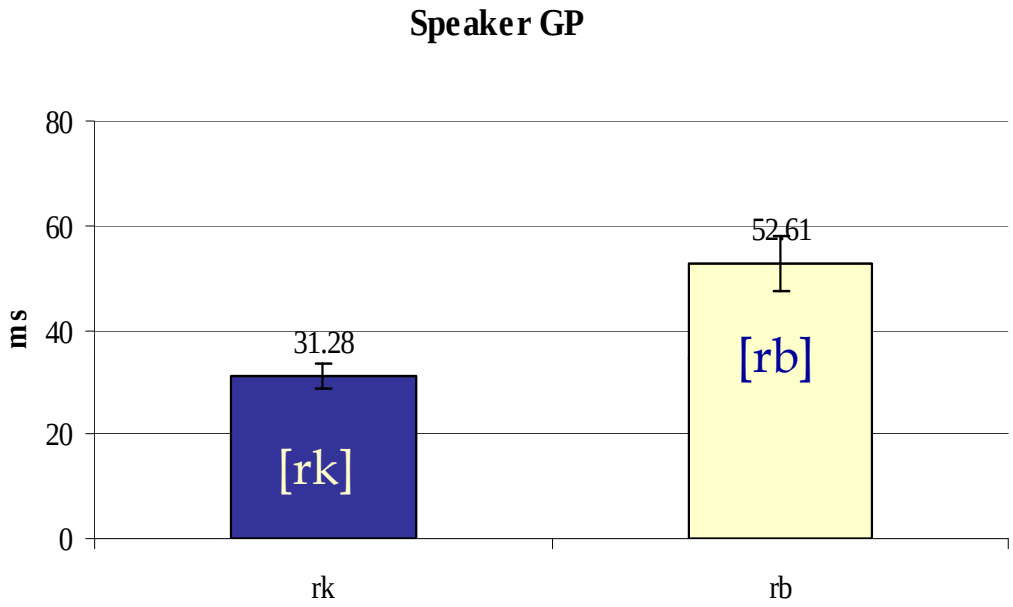


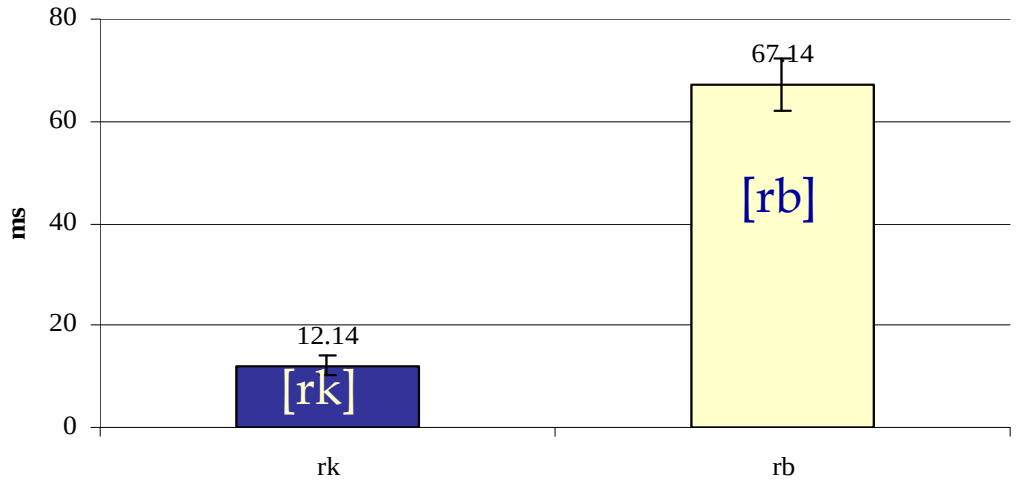
Experiment 3: Liquid-stop sequences

C_1C_2	Front to back (coronal-dorsal)	Back to front (coronal-labial)
	rk'al-i 'arc' rk^h-eb-i 'horn' pl.	rbev-a 'to raid' rben-a 'to run'

- 7 repetitions, randomized, in carrier phrase

Results – Onset lag





Discussion

- An order of place effect is present in:
  - stop-liquid clusters
  - liquid-stop clusters
- Therefore the order of place effect previously found in stop-stop clusters does not support the direct involvement of perceptual recoverability in the grammar.

Rather,

Coordination pattern specific to Georgian?

Cross-linguistically more support is found for the

WORD POSITION EFFECT

Hardcastle 1985 (English); Wright 1996 (Tsou); Kochetov 2006 (Russian); Yanagawa 2003 (Modern Hebrew); Gafos et al. 2006 (Moroccan Arabic)

Less so far for the ***ORDER OF PLACE EFFECT***

Byrd 1992, 1996; Zsiga 1994, Surprenant and Goldstein 1998 (English)

Perception

- Speaker 2 has so little overlap that very often he produces vowels between clusters – e.g., *[k'Vbili]* for *k'bili*.

HYPOTHESIS:

- If the pattern of reduced overlap is lexicalized in Georgian, then native speakers should perceive CCV sequences as CCV even when the speaker produces an epenthetic vowel *($C^{V}CV$)*. The presence or absence of those vowels should not make a difference to them, perceptually.

Preliminary results from 9 listeners

- An epenthetic V was reported 13.6% of the time
- C1 or C2 was not perceived 6.7% of the time

- Sequences were V was reported are predominantly ***back-to-front***
- Sequences where C1/C2 were not perceived are predominantly ***back-to-front***

Not predicted by perceptual recoverability:

- Even though the stimuli have long lag, still listeners miss C1 or C2. For some listeners there are many more missed Cs than reported epenthetic vowels.
  - But are Cs not perceived in any stimuli with epenthetic Vs?
- This happens more in B-to-F clusters, where the lag is even longer than in F-to-B.

- Variation in speech may be partly constrained by perceptual variation
- But results also suggest presence of language-specific patterns of coordination that are learned

Prediction:

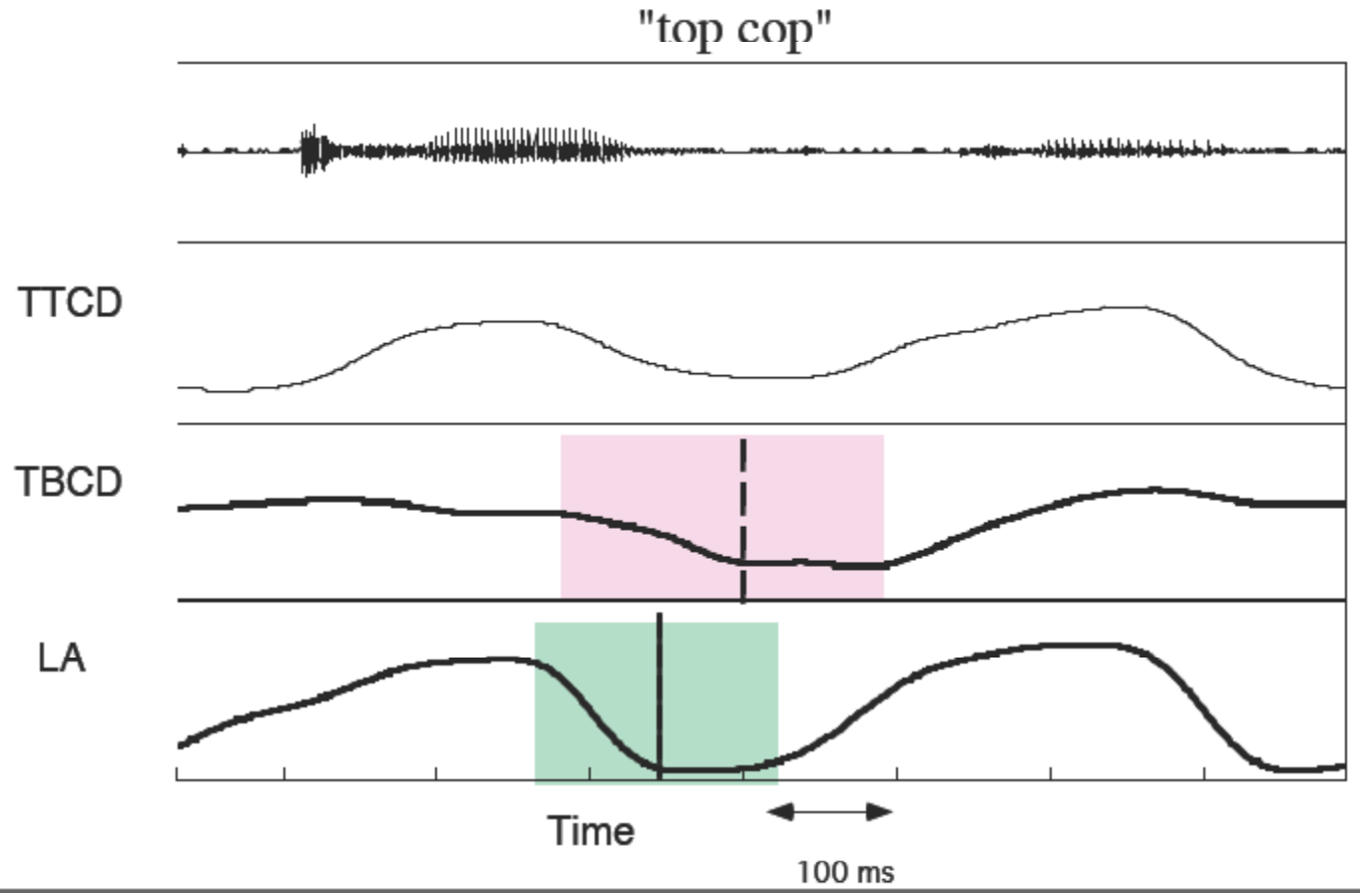
Across languages

- Languages that are hypothesized to differ in patterns of overlap are predicted to exhibit multiple corresponding differences in articulatory and acoustic consequences.

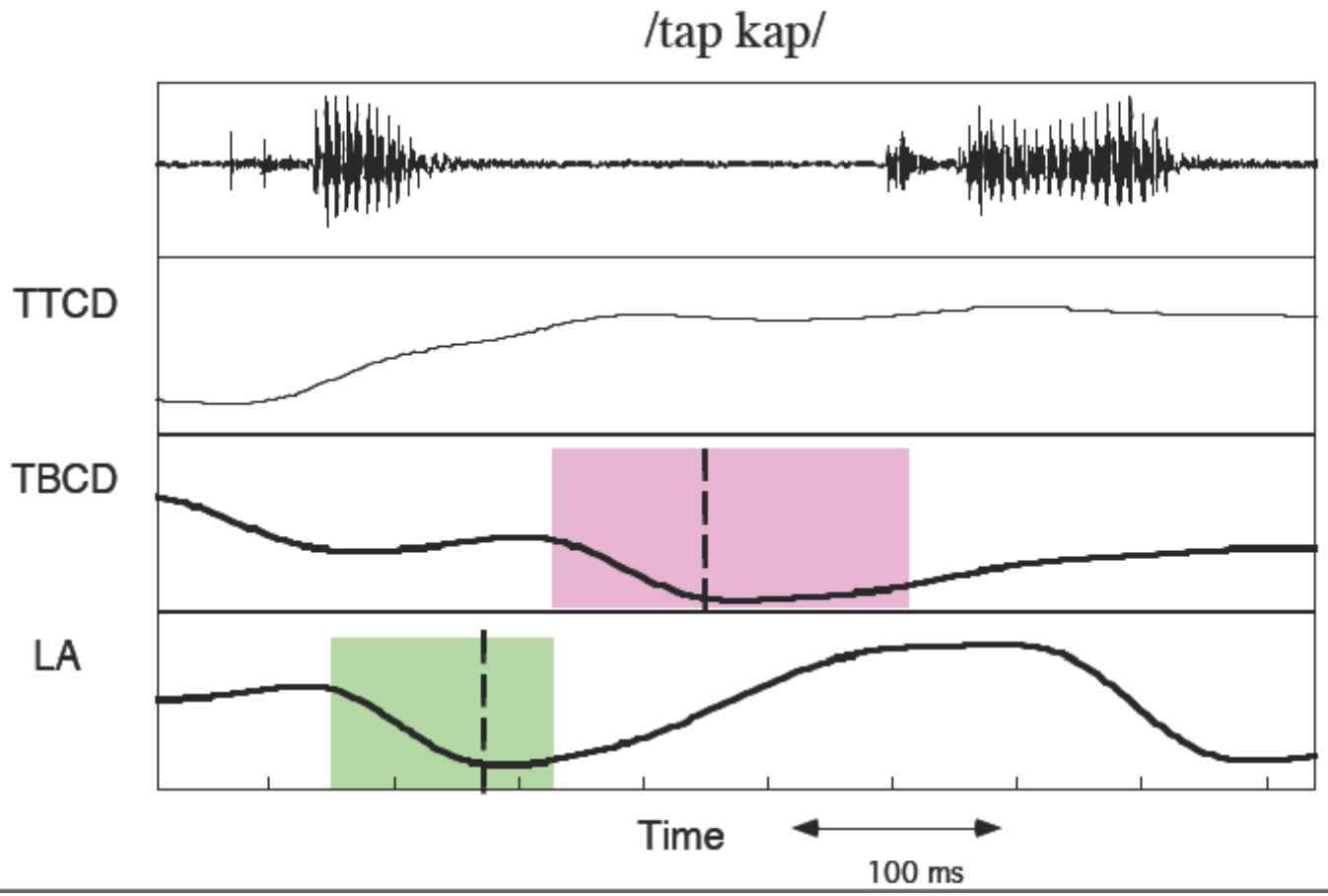
Evidence for different patterns of overlap

- /s + j/ sequences in English *(miss#you)* and in Russian *(Zsiga 2000)*
- Stop sequences in English and Russian

English – *overlapped gestures*



Russian – *no overlap*



Lecture 3

Implications for the syllable

1. The syllable in a gestural coupling model
2. Language-specific differences in syllable structure – *Georgian vs. Tashlhyit*

Phonetic correlates of syllable structure

- Gestural coordination is governed by linguistic structure – includes prosodic structure, syllable affiliation *(Fougeron & Keating 1997, Turk & Shattuck-Hufnagel 2000)*
- Differences in temporal coordination between onset and coda position
  - Liquids – *Giles & Moll 1975, Sproat & Fujimura 1993, Gick 2003, Gick et al. 2006*
  - Nasals – *Krakow 1989, 1993*
  - Kühnert et al. 2006 (French), Hermes et al. 2008 (Italian), Shaw et al. 2009 (Moroccan Arabic), Goldstein et al. 2007 (Georgian, Tashlhyit), Goldstein et al. 2008 (English, Georgian)

Empirical observations

- Both C-V timing and C-C timing in clusters differ in onset and coda position.

- Onset consonants overlap less with each other and overlap more with V, compared to coda consonants.

Proposal *(Browman & Goldstein 1988, 2000)*

Onsets

- Onset consonants are timed ***globally*** with the following V, as an ensemble of gestures
- The midpoint of the cluster ("c-center") maintains a stable relationship with V, regardless of onset composition (one C or more)

Codas

- Coda Cs are timed ***locally*** with preceding V. The left edge of the cluster is in a stable relationship with V regardless of the number of Cs.

- Distinct timing patterns hypothesized to arise from specific *coupling modes* which can be observed in other domains of motor control – finger tapping, limb coordination

- Two intrinsic coupling modes require no learning and can be stably maintained: *in-phase* and *anti-phase* *(Turvey 1990)*

Syllable structure and coupling modes

Hypothesis

- If a C constriction gesture and a V constriction gesture are to be coordinated in a spontaneously available mode, the possibilities are:

  In-phase

  – Hypothesized for C-V (onset) – ***most stable***

  Anti-phase

  – Hypothesized for V-C (coda)

In-phase (onset)

- Onset C and V gestures begin ***synchronously***
- If they have the same frequency they remain synchronous throughout

Anti-phase (coda)

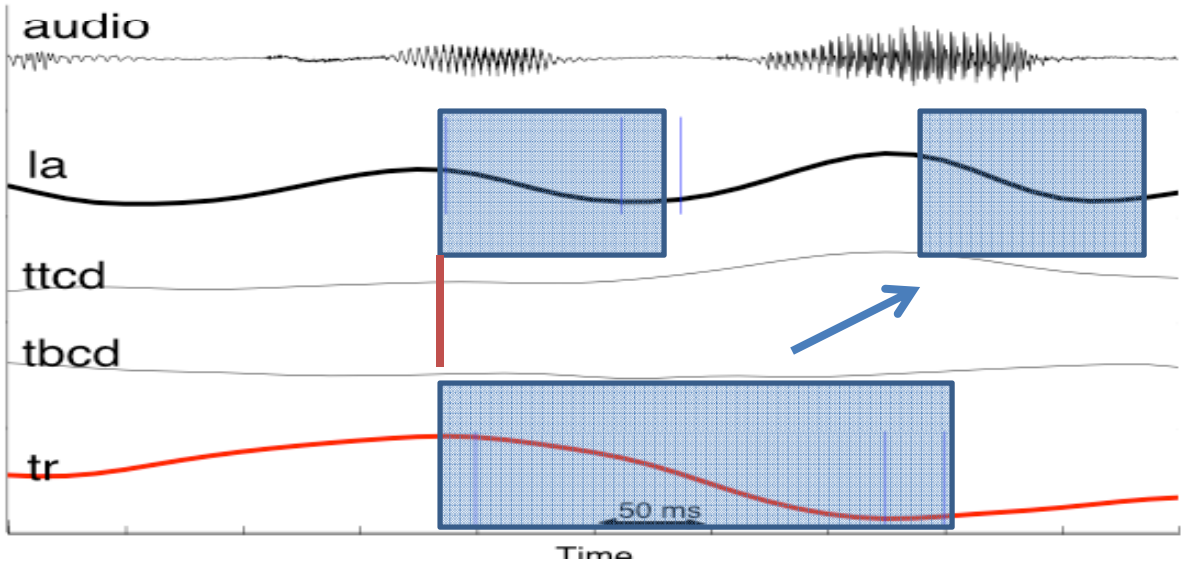
- Coda C begins later, as V reaches its target – ***sequential***

C-V and V-C modes

/p i p a p/

LIPS

Tongue Root



CC clusters in onset

- If onset is defined by an in-phase relation between C gesture and V, then all onset C gestures should be synchronous with V (and therefore with each other).

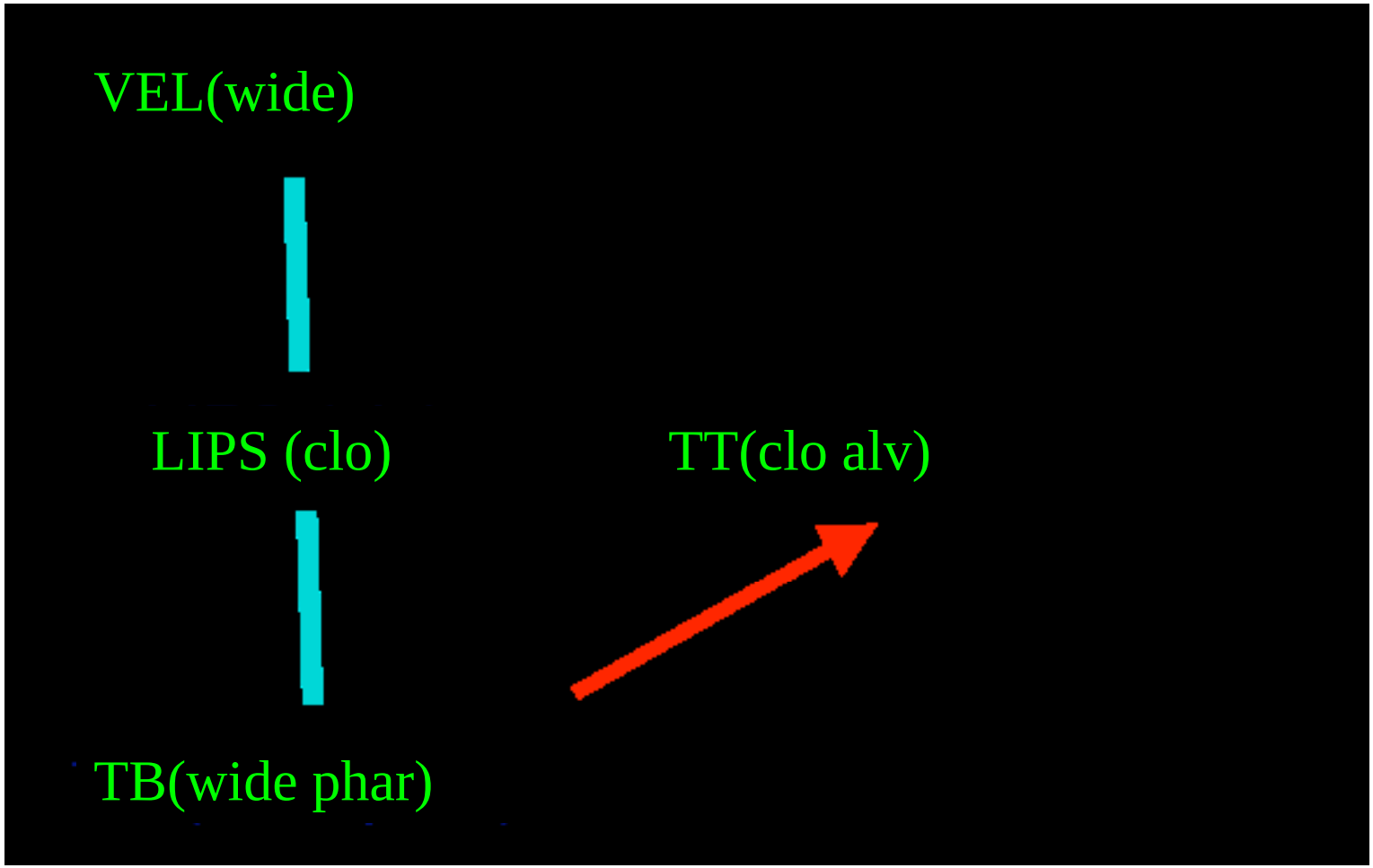
- Multiple constriction gestures in onset cluster (/spat/):
  - Gestures must be at least partially sequential to afford perceptual recoverability and to allow order contrasts (e.g., /spa/ vs /psa/)

Onset – competitive coupling

- Onset C gestures are synchronous with the V and sequential with each other
- As more Cs are added to the onset, rightmost C shifts rightward, toward V, leftmost C shifts leftward, away from it.

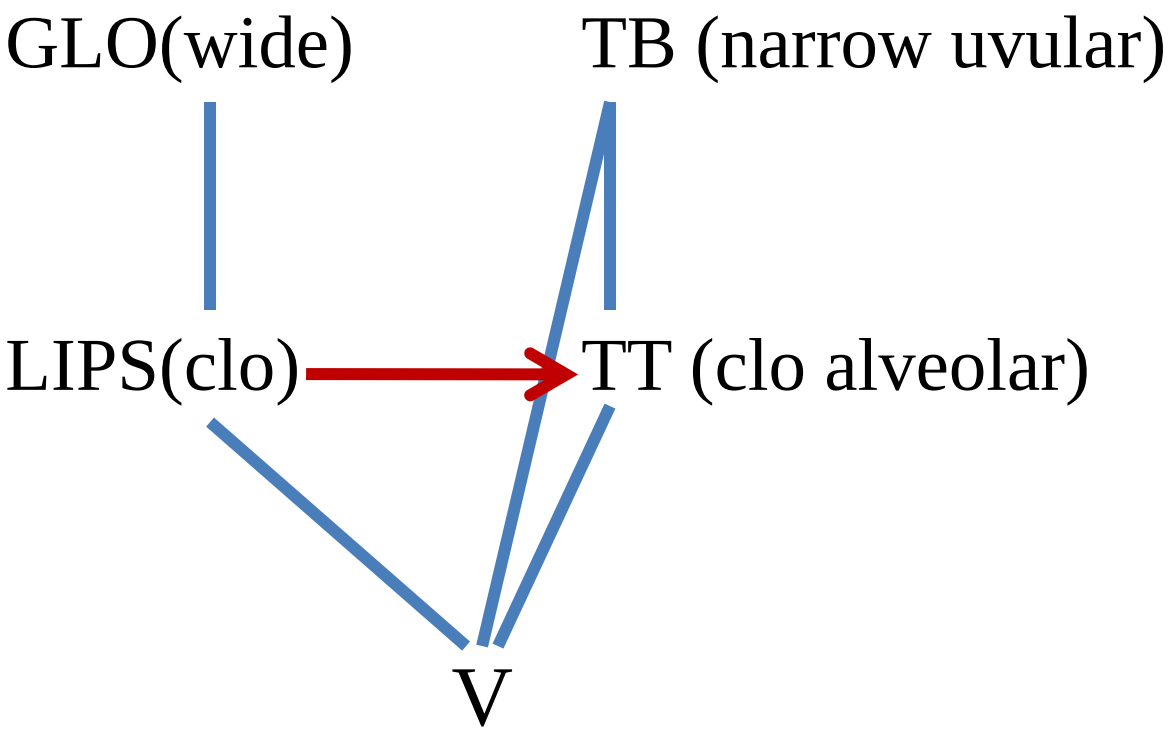
- What in the coupling graph identifies them as both in the onset?

Coupling graph – ***mad***



Onset /plV/ based on experimental data

(Goldstein et al., 2008)



Based on TADA simulation, as well

- /l/ is composed of two gestures: a tongue tip constriction and a body constriction at the uvula.
- Are both gestures coupled with V, or just TT gesture?

Predictions:

- The /pl/ graph with the extra link should result in a *tighter coupling* of the /l/ with respect to the vowel (multiple links)
- This tighter coupling should also cause a *reduction in rightward shift*.
- This is confirmed in the quantitative modeling

Georgian onsets

based on experimental data and simulation

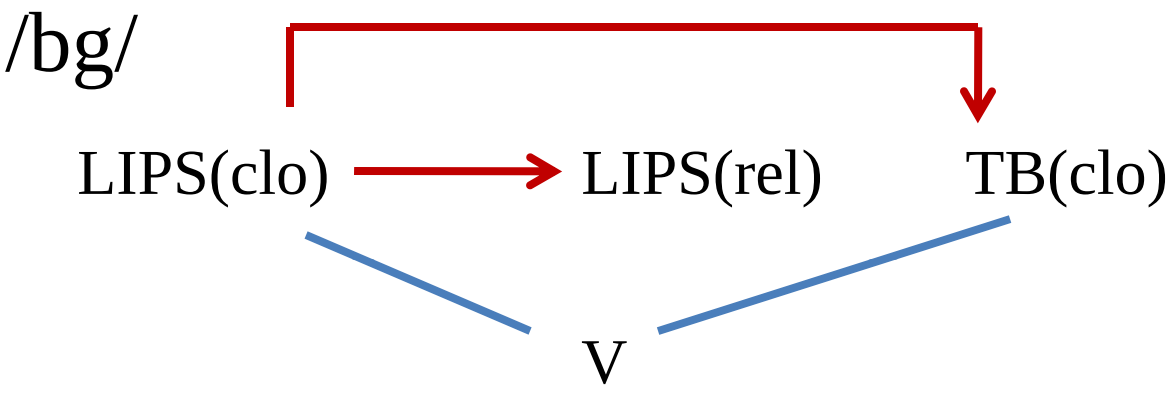
(Goldstein et al., 2008)

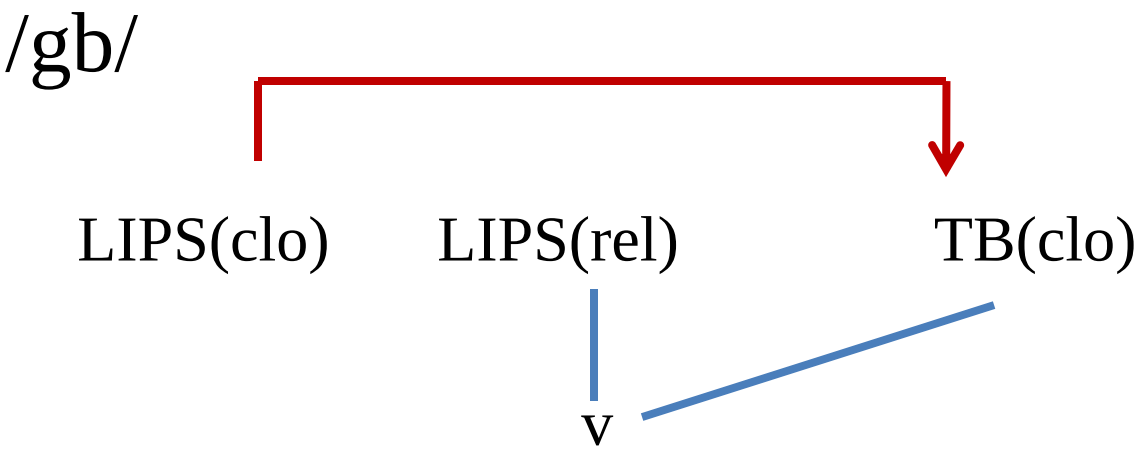
- Front-to-back sequences are produced with a shorter lag between the onset gestures than back-to-front sequences – *systematic*, *phonologically relevant*

Proposal: modeling releases as separate gestures

- LIPS and TB closures both coupled in-phase with the V, and anti-phase with each other.
- Release gestures are coupled only to their corresponding closures, so their presence does don't affect the relative timing of the other gestures in the graph.

Georgian *bg* vs. *gb* onsets





Georgian and Tashlhyit *(Goldstein, Chitoran, Selkirk 2007)*

If a consonant sequence is syllabified as part of an onset, then it should exhibit rightward shift.

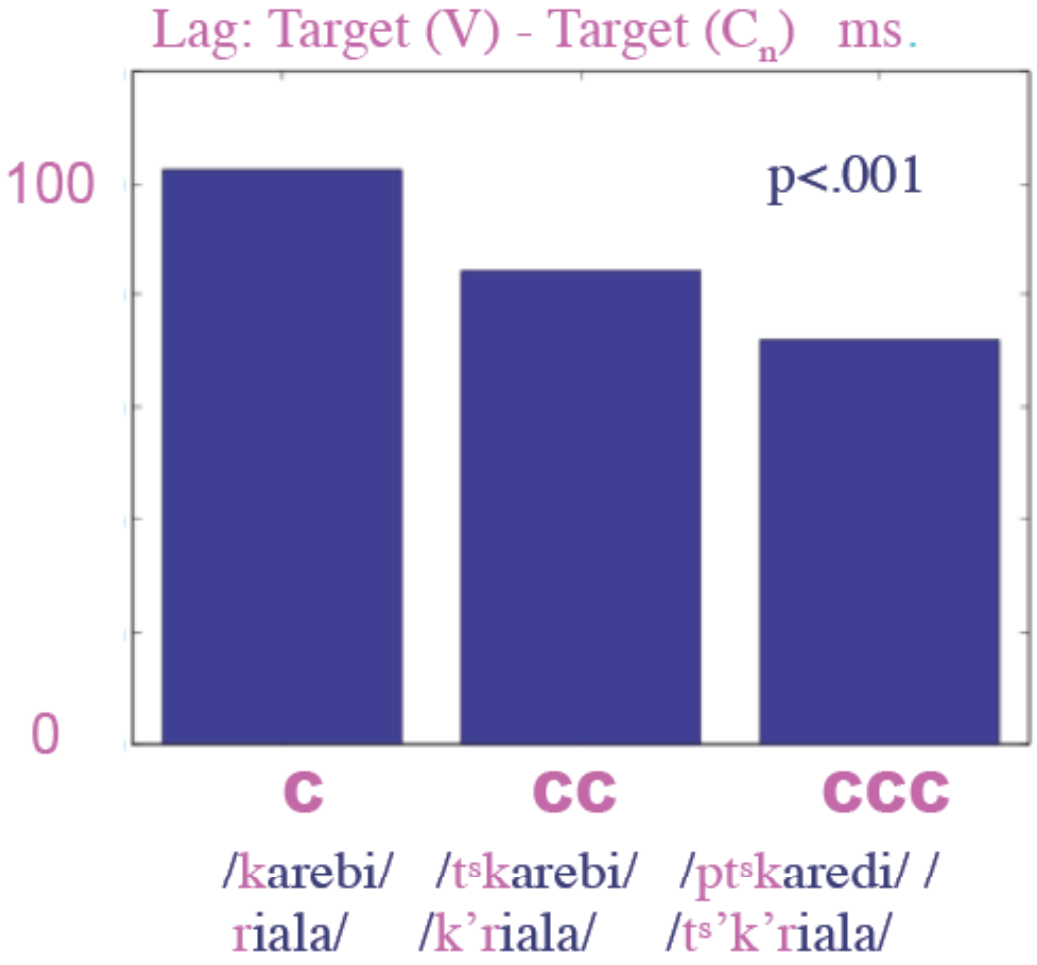
Georgian and Tashlhiyt Berber are languages in which words can begin with sequences of 3 obstruents.

Languages differ as in syllabification of such words:

- Georgian Cs are complex onsets
- Berber only allows a single C in onset, other Cs constitute nuclei of additional syllables.

Do Georgian and Berber differ in rightward shift?

Georgian – rightward shift



/p/ Lip Aperture

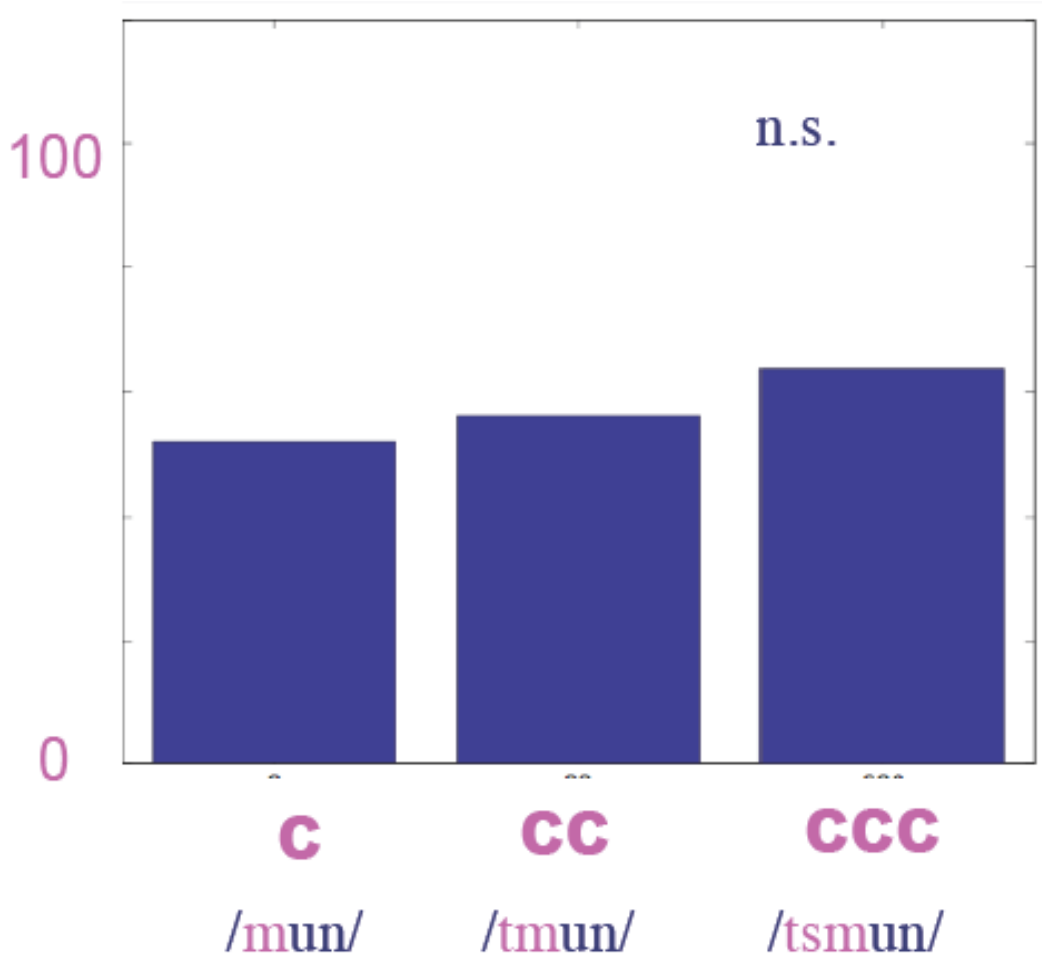
/t/ Tongue Tip Constriction Degree

/k/ Tongue Dorsum Constriction Degree

/V/ Tongue Body Constriction Degree

2 speakers

Tashlhyit – no rightward shift



/m/ Lip Aperture

/t/ Tongue Tip Constriction Degree

/s/ Tongue Tip Constriction Degree

/V/ Tongue Body Constriction Degree

2 speakers

- Browman, C. & L. Goldstein. 1988. Some notes on syllable structure in articulatory phonology. *Phonetica* 45: 140-155.
- Browman, C. & L. Goldstein. 1992 Articulatory Phonology: an overview. *Phonetica 49: 155-180.*
- Browman, C. & Goldstein, L. 1995. Dynamics and articulatory phonology. In T. van Gelder & R.F. Port (Eds.), *Mind as Motion. Cambridge, MA:* MIT Press. 175-193
- Browman, C. & L. Goldstein. 2000. Competing constraints on intergestural coordination and self organization of phonological structures. *Bulletin de la Communication Parlée 5: 25-34.*
- Fowler, C.A., Rubin, P.E., Remez, R., & Turvey, M. (1980). Implications for speech production of a general theory of action. In B. Butterworth (Ed.), *Language production (pp. 373-420). New York: Academic Press*
- Goldstein, Pouplier, Chen, Saltzman, Byrd. 2007. Dynamic action units slip in speech production errors. *Cognition* 103. 386-412
- Kelso, J.A.S., Tuller, B., Vatikiotis-Bateson, E., & Fowler, C.A. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance, 10, 812-832.*
- Kühnert, B; Hoole, P.; Mooshammer, C., 2006. Gestural overlap and C-center in selected French consonant clusters. *Proc. 7th Int. Seminar on Speech Production, Ubatuba, 327-334.*
- Manuel, S.Y. 1990. The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *JASA 88, 1286-1298.*
- Meltzoff, A.N., & Moore, M.K. 1977. Imitation of facial and manual gestures by human neonates. *Science, 198, 75-78*
- Shattuck-Hufnagel, S. 1983. Sublexical units and suprasegmental structure in speech production planning. In P.F. MacNeilage (Ed.), *The production of speech (pp.*109-136). New York: Springer-Verlag
- Studdert-Kennedy, M. 1998. The particulate origins of language generativity: from syllable to gesture, in Hurford, et al., (eds.) *Approaches to the Evolution of Language*. Cambridge UP. 202-221.