

# Segments in Articulatory Phonology

Carol A. Fowler  
University of Connecticut  
and Haskins Laboratories

# Why should a phonology be articulatory?

Part of the answer: There is something wrong with these ideas:

“The concept of E [external]-language, *however construed*, appears to have no significance” (Chomsky, 1986 p. 31; italics added)

“Externalization [of language] by the SM [sensorimotor] system appears to be a secondary property of language” (Chomsky, 2011, p. 275)

# What (among other things) is wrong with these ideas:

- Language evolved to be spoken
- Language forms are the very means by which linguistic messages can be publicly shared
- So, they should be adapted to public language use.

# Phonological systems show evidence of adaptedness to public use

- Popularity of vowels that are easy to distinguish perceptually
- Popularity of consonants that are easy to say (e.g., Lindblom and Maddieson, 1988)
- Existence of phonological processes that reflect articulatory constraints: e.g., popularity of vowel harmony; unpopularity of consonant harmony

## Accordingly, some questionable ideas

- Phonological entities have their primary home in the mind/brain of a language user
- Properties: they are discrete, static, context-free
- Their implementation in talking involves distortion and even destruction of these properties, thanks largely to coarticulation

# Articulatory phonology

(Browman & Goldstein, 1986, 1992, 1995...; Goldstein, Byrd & Saltzman, 2006; Goldstein & Fowler, 2003)

- Language forms have their primary home in the vocal tract
- They are adapted to public use
- Properties: they are discrete, but not temporally discrete, dynamic, context-free
- Coarticulation does not distort or destroy these properties

# Phonological primitives: Phonetic gestures

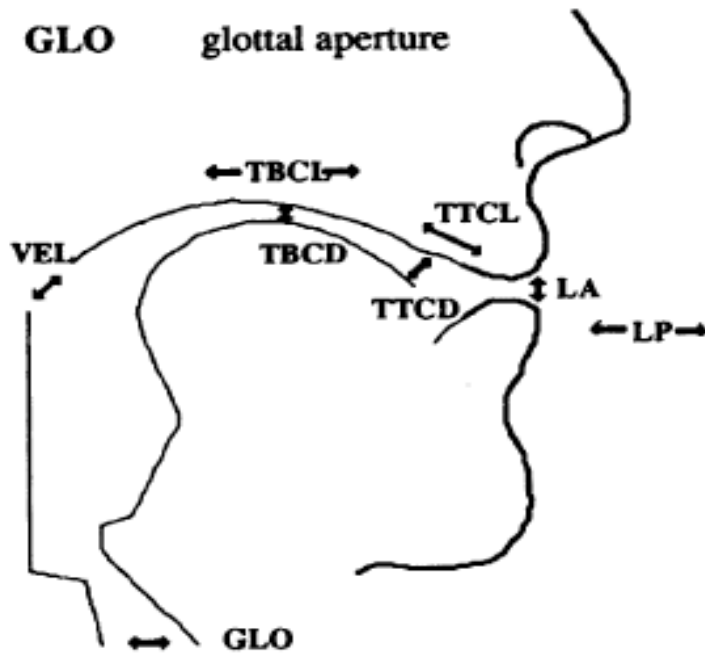
These are coarse-grained actions of the vocal tract  
(not movements of individual articulators)  
implemented by synergies (dynamical systems)

They create and release constrictions in the vocal  
tract

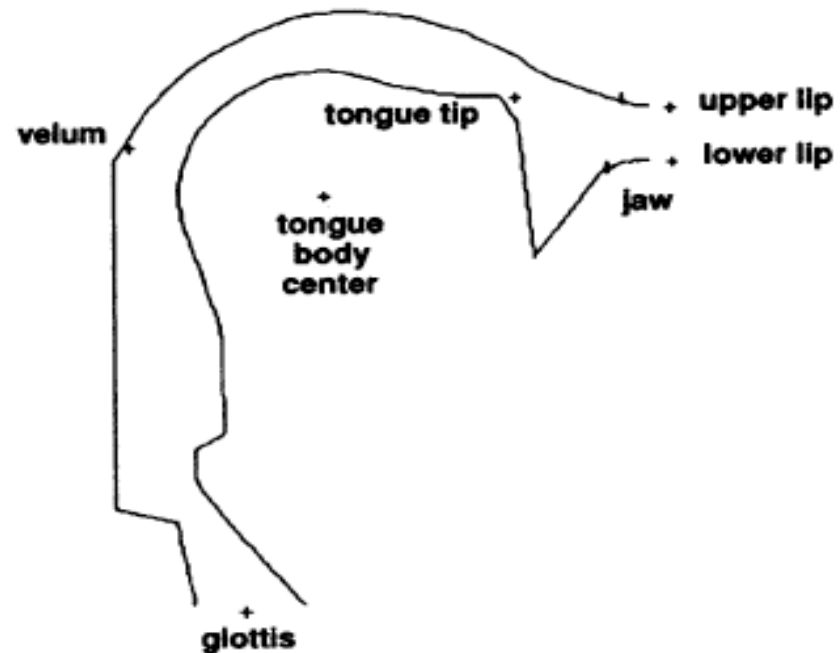
Distinguishing properties: constriction location,  
degree (The dynamical parameters that realize these  
properties are context-free)

# Gestures in articulatory phonology

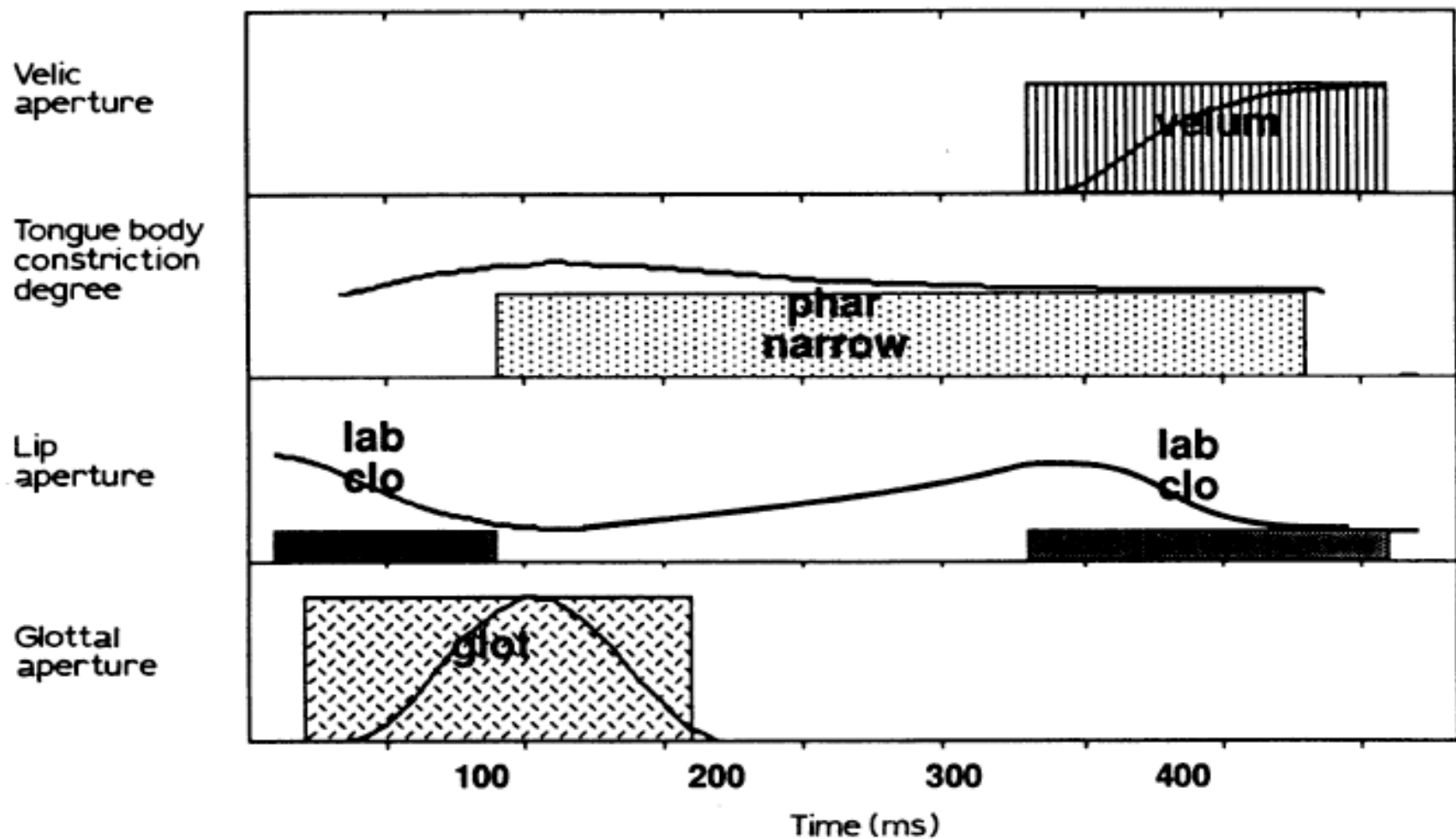
- TTCL** tongue tip constrict location
- TTCD** tongue tip constrict degree
- TBCL** tongue body constrict location
- TBCD** tongue body constrict degree
- VEL** velic aperture
- GLO** glottal aperture



- tongue tip, tongue body, jaw
- tongue tip, tongue body, jaw
- tongue body, jaw
- tongue body, jaw
- velum
- glottis

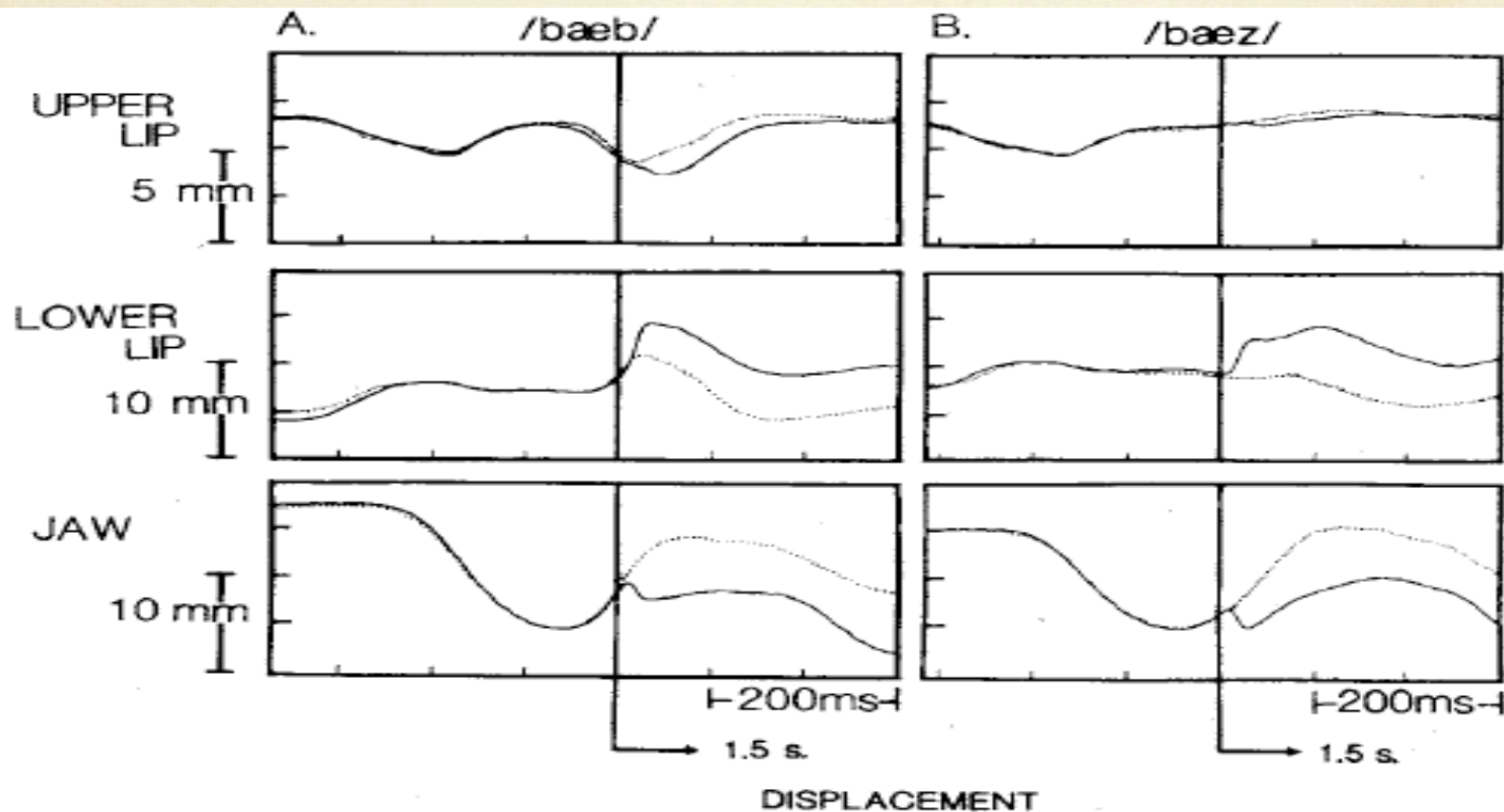


# A gestural score: "palm"

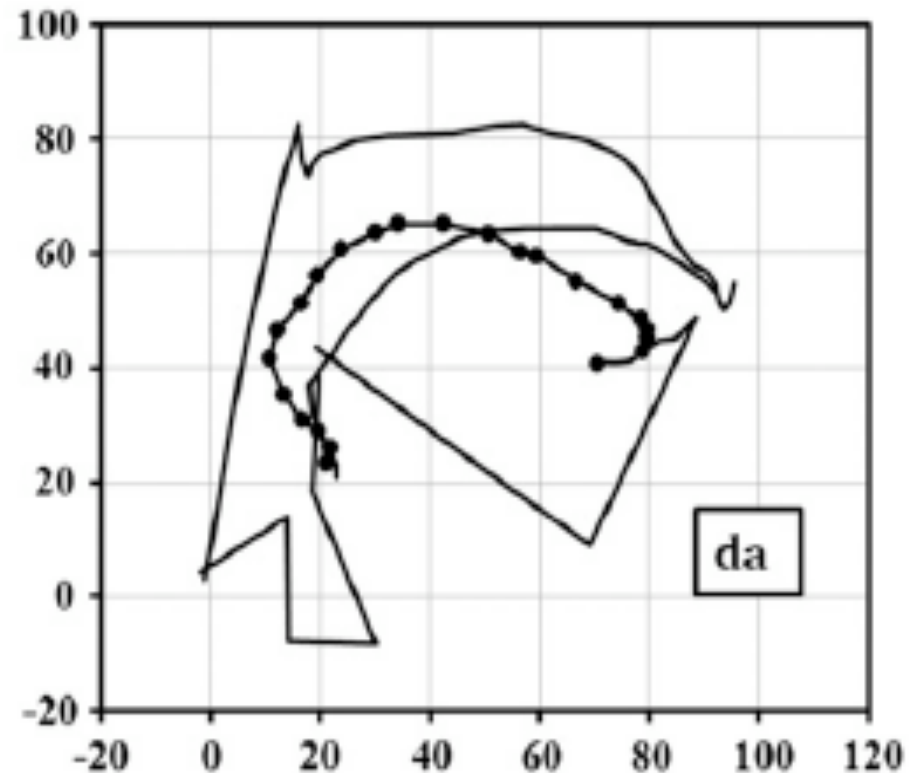
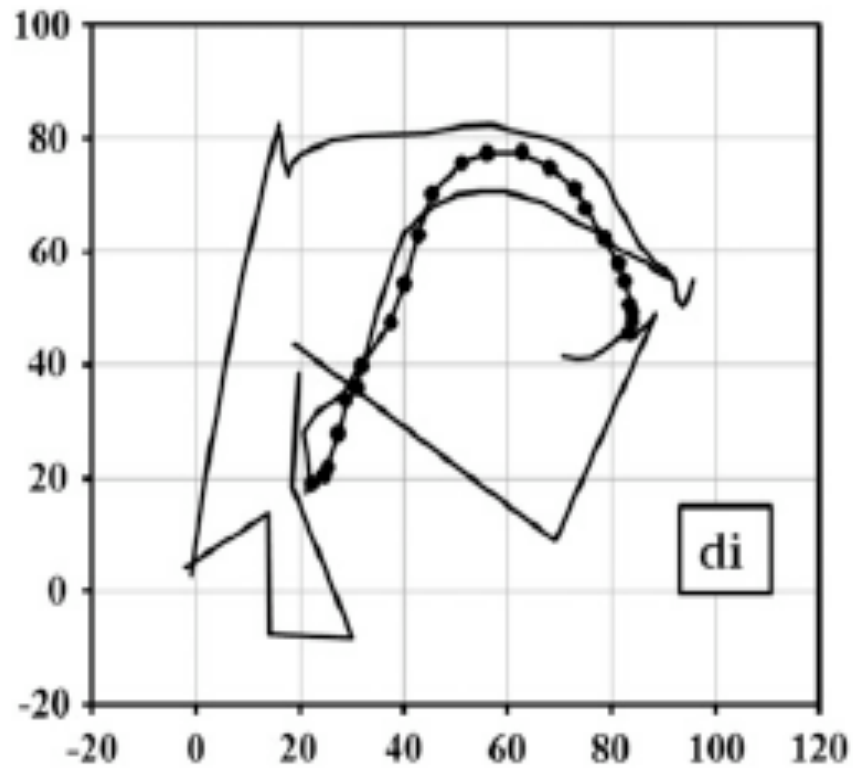


# Gestures are dynamic, overlapping, context-free

Equifinality over perturbation:



# Relevance of equifinality to coarticulation: /di/, /da/

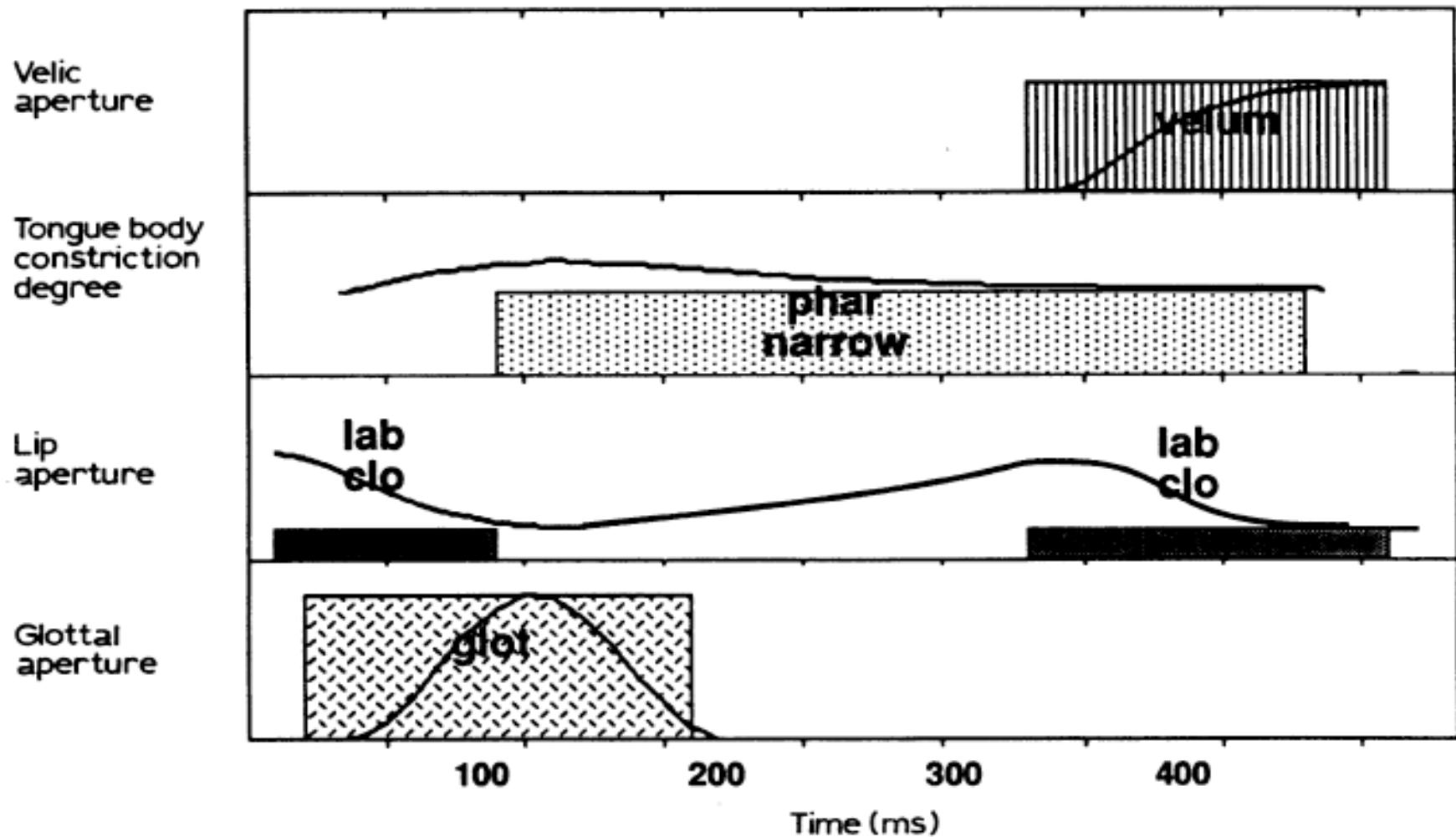


X ray tracings; Lindblom & Sussman, 2012

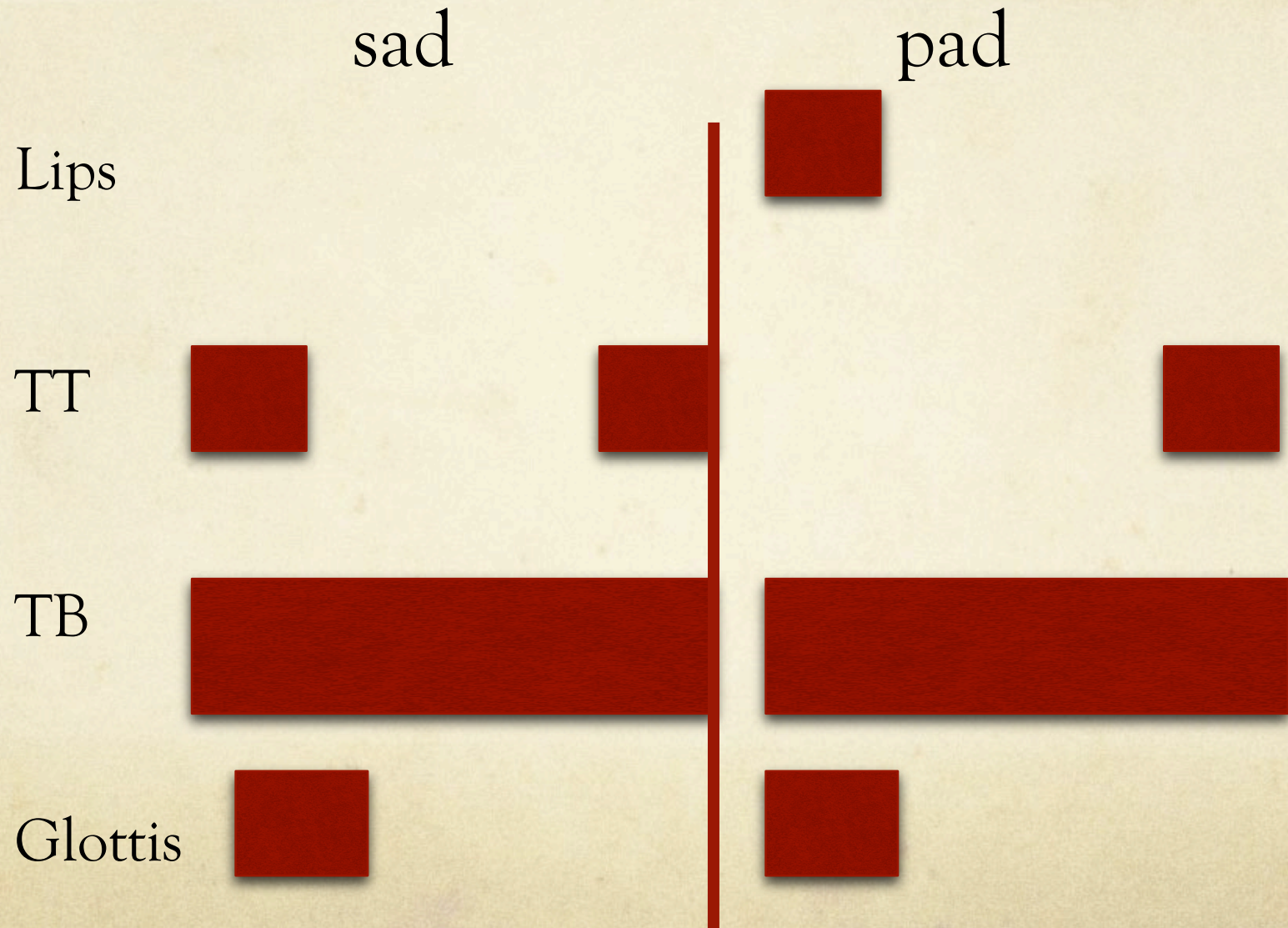
Note that the gestural score  
does not necessarily partition  
into segments

- /p/ is associated with two gestures: lip closing, devoicing (glottal)
- /m/ is associated with two gestures: lip closing, velum lowering
- There is nothing in the gestural score that affiliates these gestures more with each other than, say, with the vocalic gesture

# A gestural score: “palm”



See also: sad, pad, spad



spad: just one glottal gesture: Does it belong to lip gesture for /p/ or to TT gesture for /s/?

Lips



TT



TB



Glottis



Does it matter that segments  
do not naturally emerge in AP?

A common criticism of articulatory phonology:  
It is not a phonology

It is rather, an articulatory phonetics: just an  
illuminating way of looking at speech articulation

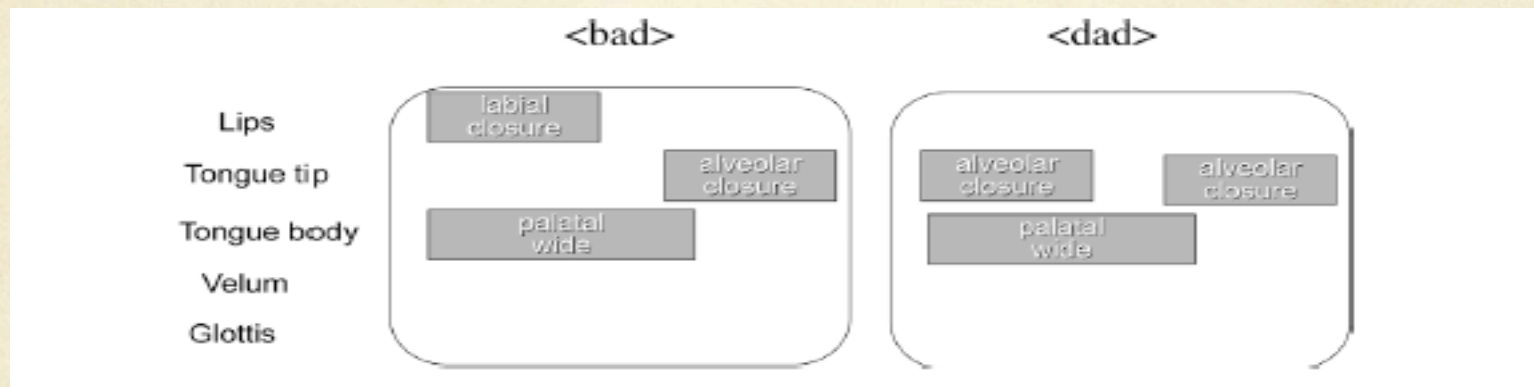
So it does not matter if segments are apparent in  
this description

Au contraire

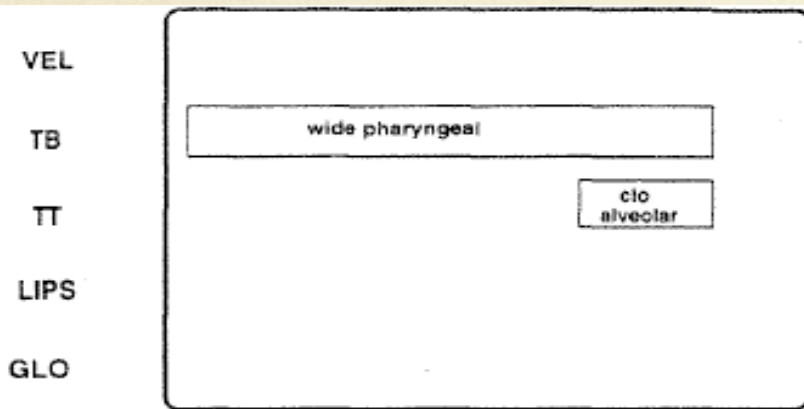
Articulatory phonology is the right  
name!

# What makes this a phonology, not merely a phonetics

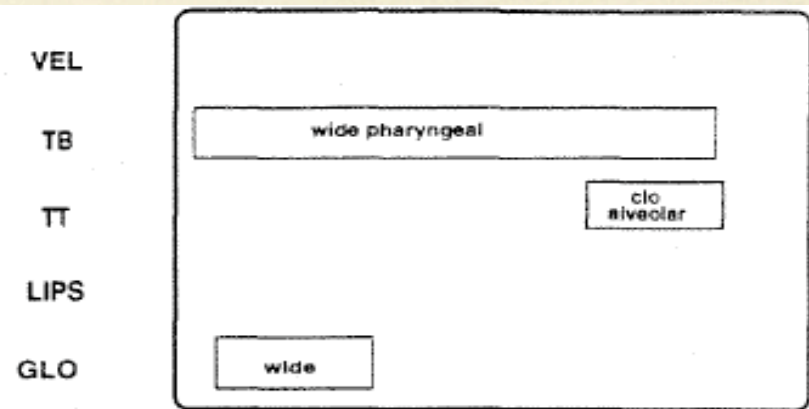
## 1. Contrast: substitution of gestures



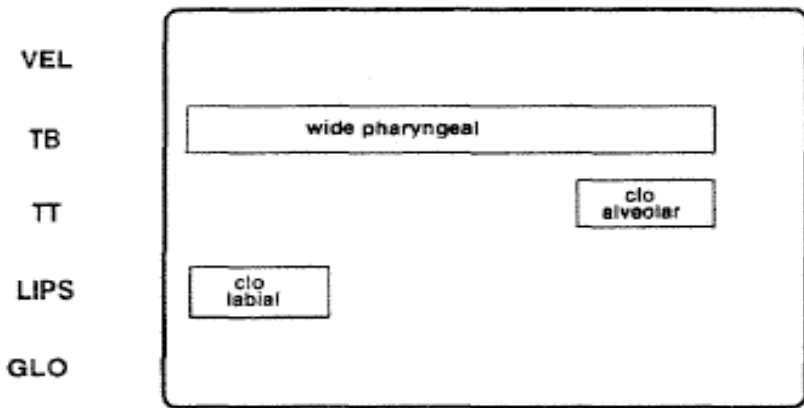
# Contrast: addition, deletion of a gesture



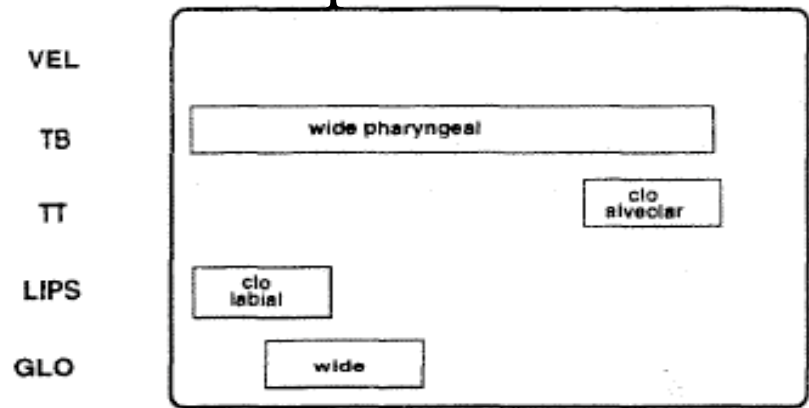
add (a)  
bad



had (b)  
pad



(c)



(d)

# What makes this a phonology, not merely a phonetics

2. Phonological processes: Articulatory phonology provides a uniquely illuminating perspective on some phonological processes:

- a) Statement of some phonological processes must make reference to the temporal properties of segments/gestures (Moroccan Arabic, Gafos , 2002)

Phonological units only *have* temporal properties in articulatory phonology

b) There is no true action at a distance in phonological processes (the “locality” principle, e.g., vowel harmony vs consonant harmony: Gafos, 1999 )

Vowel harmony emerges from vowel-vowel coarticulation and, like vowel-vowel coarticulation, requires articulatory adjacency

C-C coarticulation “across” a vowel is rarely feasible; compatibly, C harmony is rare. It occurs just where C-C coarticulation is feasible

c) Some phonological processes apply in a *gradient* fashion and even can be differentially gradient in different speaking styles (e.g., final devoicing in German: Gafos & Benus, 2006)

This is naturally captured in an account in which multiple OT constraints (here, markedness: coda Cs must be unvoiced, faithfulness: of input to output segments) compete in the dynamical system framework of articulatory phonology

# Segments in articulatory phonology?

- As noted, segments do not emerge naturally in gestural scores for bi- or multigestural segments (“palm,” “sp” clusters)
- For Browman & Goldstein, the existence of coupling relations among gestures such that gestures affiliate into segmental structures is, at best, an empirical issue:

## Browman & Goldstein, 1990:

“[W]e would argue that the basis for [segments] seems to be their utility as a practical tool rather than their correspondence to important informational units of the phonological system

(Fowler: if they are fictional, why do they have practical utility?)

# More Fowler, not B&G:

Some considerations regarding the reality or not of segments:

- Duality of patterning (the “particulate principle of self-diversifying systems”)
- Speech errors
- The remarkable success of alphabetic writing systems
- Natural languages are not tidy

# Duality of patterning

To account for language generativity at the level of the lexicon, we have to invoke SOME meaningless entities that can combine in indefinitely many ways to constitute word forms

Gestures might do the job in place of segments  
BUT NB we only combine gestures in the restricted ways that tempt us to think that there are segments (e.g., no devoiced nasals in coined English words.)

# Speech errors

In error corpora collected *outside* the laboratory, the majority of sublexical errors have been identified as segment errors (e.g, Shattuck-Hufnagel: 60%):

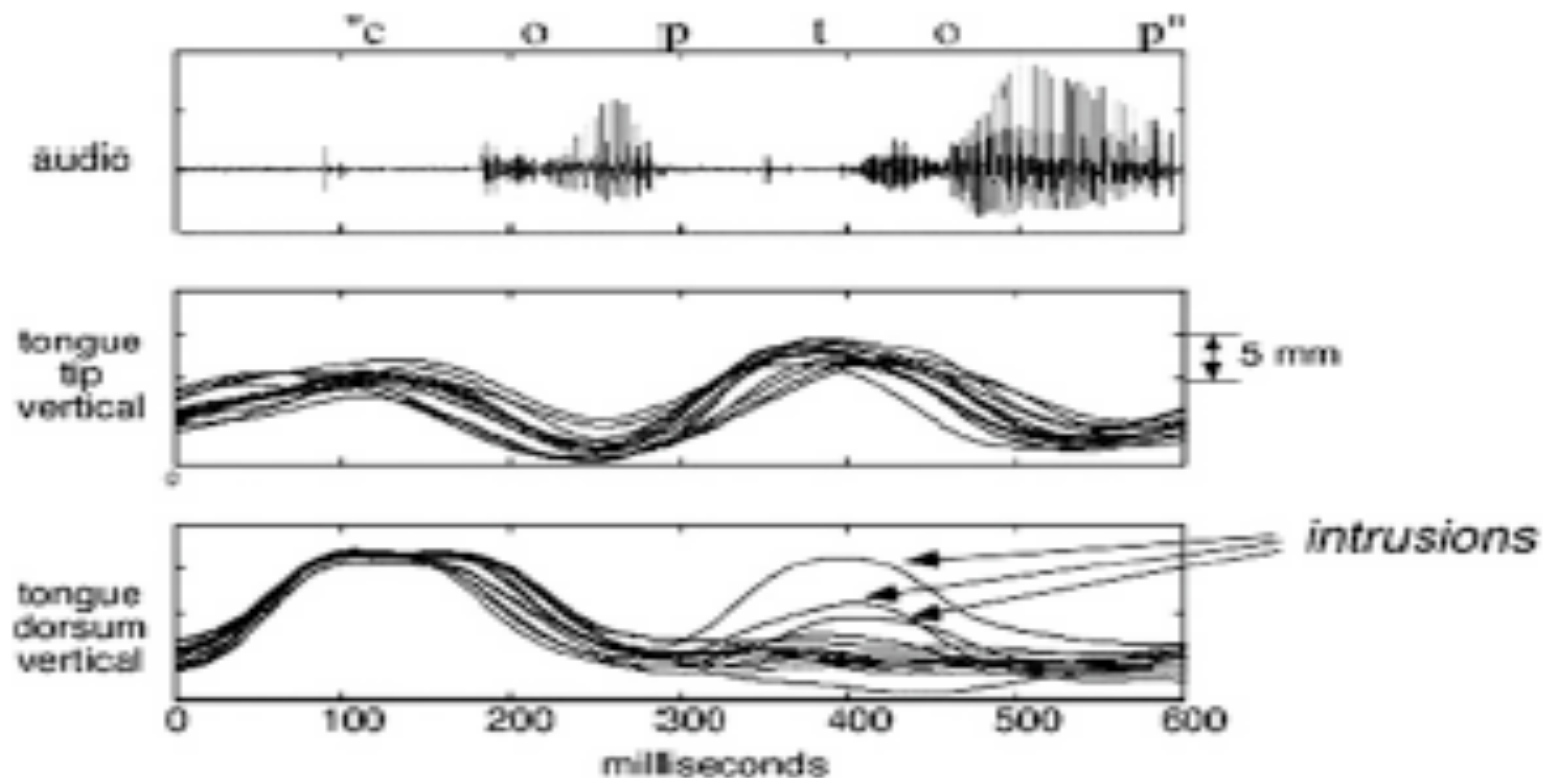
mardon me padam (the Rev Wm Spooner)

beef needle soup (Dell, 1986)

Doesn't this validate the segment as a unit with psychological reality?

Errors collected INSIDE the laboratory (repetition of word pairs, e.g., *top cop* or SLIP procedure) complicate the picture somewhat:

*L. Goldstein et al. / Cognition 103 (2007) 386–412*



Errors can be gradient; contra expectations if there is mis-selection of an abstract segmental symbol

Gestures for more than one consonant can be selected at the same time.

Crucial, but unknown as yet from lab research:  
when one gesture of a multi-gesture segment moves in an error, does it tend to take its co-gestures with it?

E.g., if the lip closing gesture of /m/ moves, does the velar gesture move with it with greater-than-chance likelihood?

Relevant: Shattuck-Hufnagel & Klatt, 1979: in transcription errors, there is evidence for segmental cohesion:

For example, for an intended sequence such as:

“sell door”

in which the initial consonants differ by two feature/gestures (manner, place)

The exchange: “dell soor” (in which both features swap together) is MUCH more common than “tell zoor” or “zell toor” (in which individual features move): 67/70 errors

# The alphabetic principle

Alphabetic writing systems map letters to segments.

They support skilled, very fluent reading.

I find it implausible that letters represent fictional units of the language, because, if so:

- a) on what basis were alphabets invented?
- b) why do they work so well?

# languages are not tidy

Linguistic structures arise in public, between-person language use

They are not wholly descriptively tidy.

Perhaps there are MOSTLY structures that conform to a segmental description

E.g., there may be no right answer to whether /sp/ is one segment or two: likewise: “church”: how many segments? No right answer?

Back to Goldstein and colleagues:

Couplings among gestures:

Articulatory data suggest to Goldstein, Byrd and Saltzman, 2006 that:

Gestures in syllable onsets ALL are coupled in-phase to the vowel gesture; they are coupled anti-phase to each other

The first consonantal gesture in the coda is coupled anti-phase to the vowel. Subsequent coda gestures are coupled anti-phase to one another. No natural emergence of segments

# Couplings among gestures:

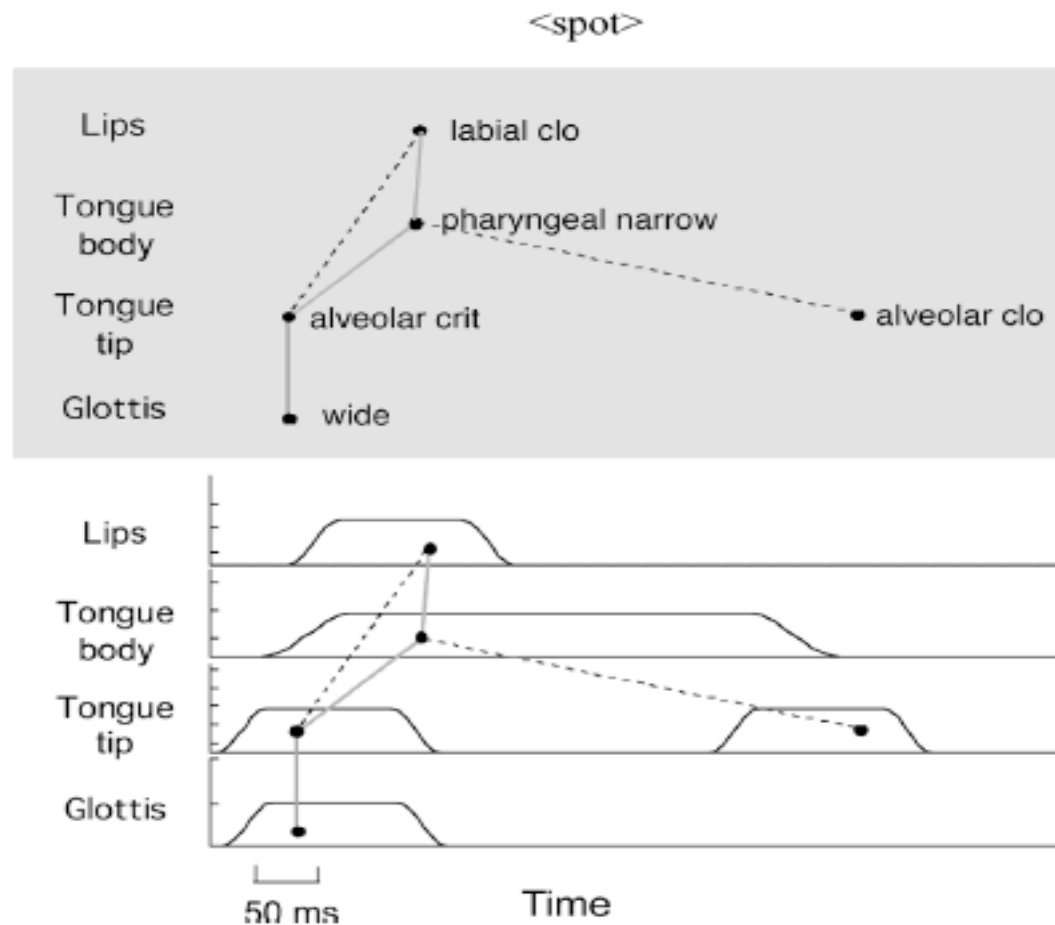


Figure 7.8 The coupling graph for “spot” (top) in which the tongue tip (fricative) gesture and the lip closure gesture are coupled (in-phase) to the tongue body (vowel) gesture, while they are also coupled to one another in the anti-phase mode. The pattern of gestural activations that results from the planning model is also shown (bottom). Lines indicate coupling relationships between pairs of gestures – solid and dashed are different in-phase and anti-phase coupling modes, respectively.

# Gestural affiliation in articulatory phonology

Segments do not emerge in any obvious way in these couplings:

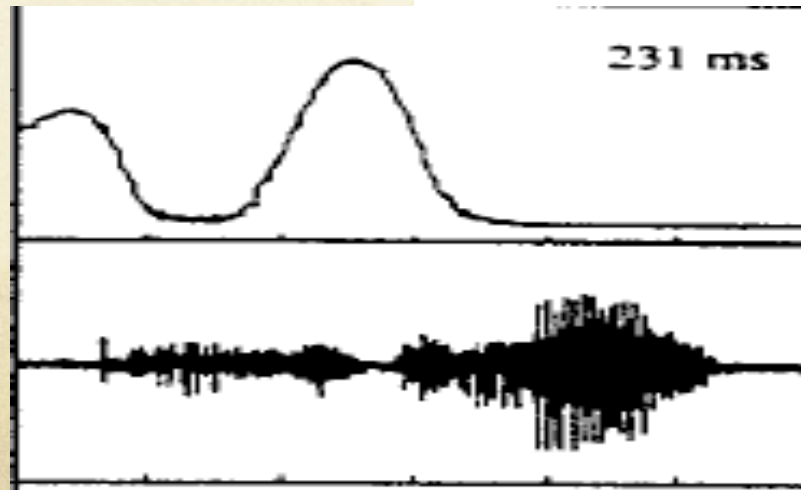
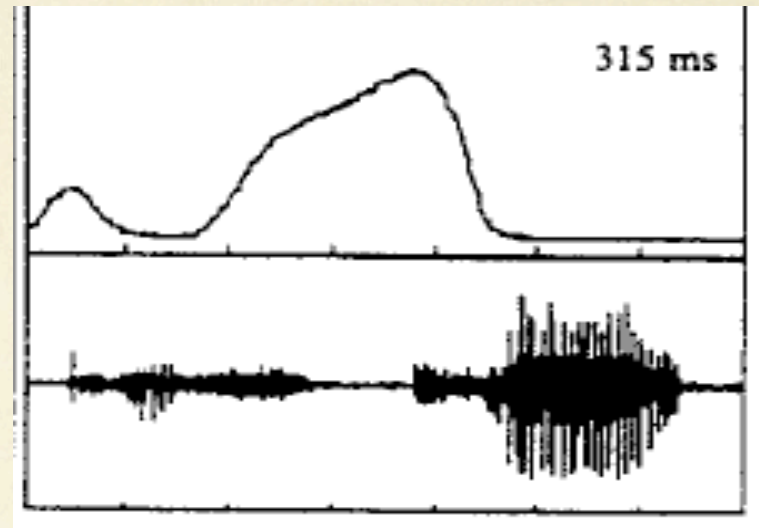
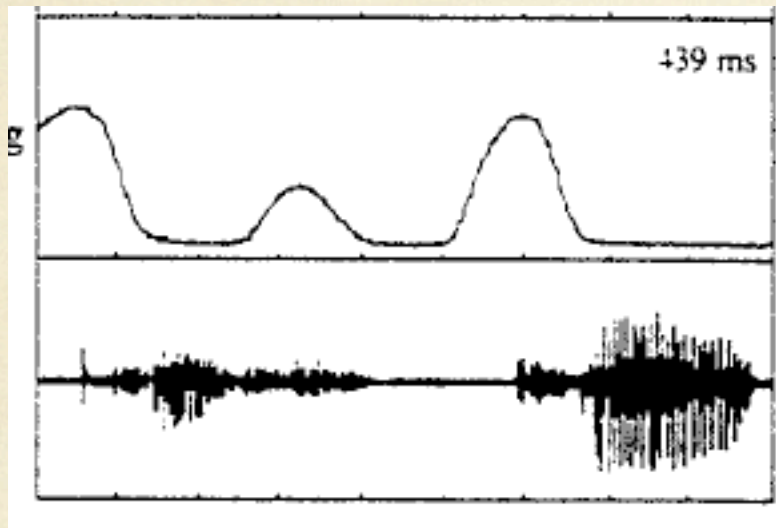
But: Munhall, et al: 1993: when the lip gesture for /p/ in /ipi/ was perturbed, closing the lips AND vocal fold gesture were both delayed; the vowel gesture was NOT delayed

Possibly: there is *stronger* coupling between gestures of a segment than between gestures of different segments

What about the single glottal gesture in spad or spot?

Possibly there are *two* glottal gestures (one for /s/, one for /p/) *but they completely overlap temporally:*

# Kiss Ted; 3 speaking rates (Munhall)



## Even so.....

Much evidence suggests a different phasing of the two gestures of nasals (Krakow), /l/ (Sproat and Fujimura), and /w/ (Gick) in the onset and coda:

They are temporally nearly synchronous in the onset; the more open of the two gestures leads in the coda:

So why should we count them as the same two-gesture segment occurring in both onset and coda? Dunno.

# Upshot

Much more research is needed in articulatory phonology that investigates intergestural coupling

My guess: alphabetic writing systems and speech errors reflect the fact that the coupling relations yield segment-like units for the most part. *But languages are no more tidy than interlocutors need them to be.*

Stay tuned....