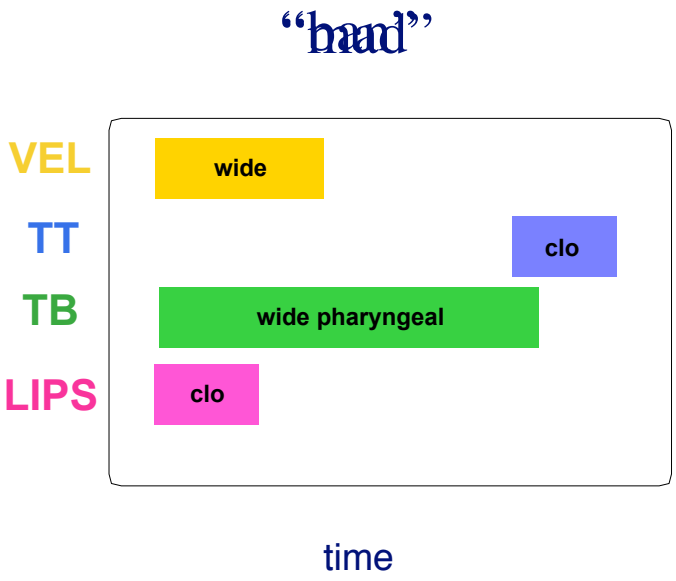
Gestural Scores and Phonetic Transcription

Coordination of gestures in time

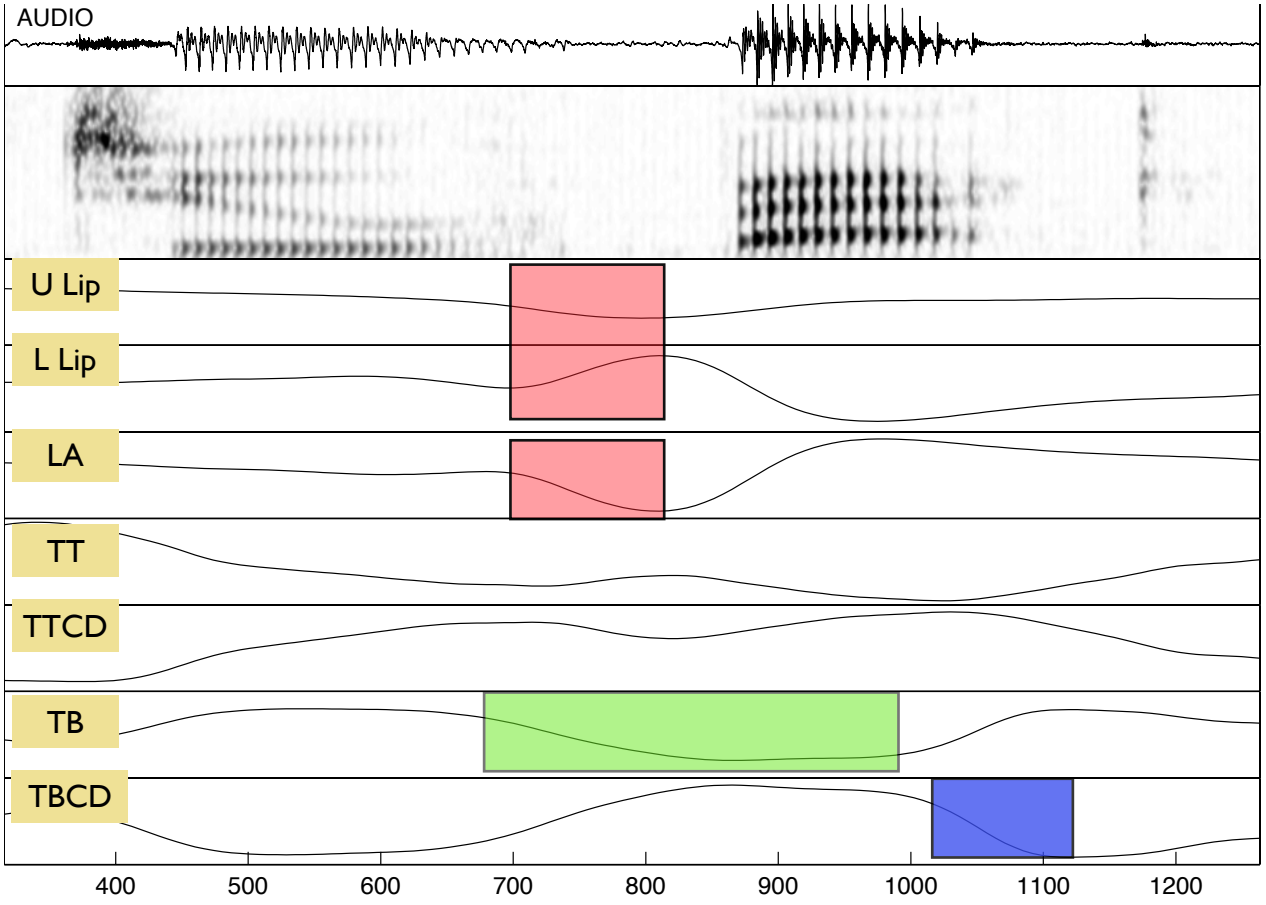


- The consonant and vowel gestures that form a word are each active for a fixed interval in time.
- The multiple gestures associated with a given consonant or vowel may not be synchronous with each other.
- Relative timing of gestures carries information.
- What is the appropriate timing?
- How do we find out?

Finding Gestures in Time

- To find when gesture is active in time, examine the movements of the constricting device that forms the constriction for that gesture.
    - When it begins to move towards the gesture's constriction target, this is the moment of gestural activation.
    - When it begins to move away the gesture's constriction target, this is the moment of gestural deactivation.

Gesture Activation Times



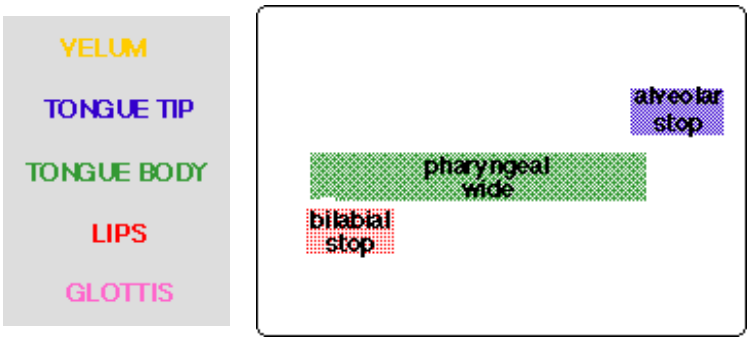
"two back"

Principles underlying Gesture Timing

- Gestures for word-initial C and the V begin at the same time!
  - Initial C and V are co-produced
- The V gesture continues about two and a half times as long as the C gesture
- The gesture for the final C begins roughly when the V gesture ends.

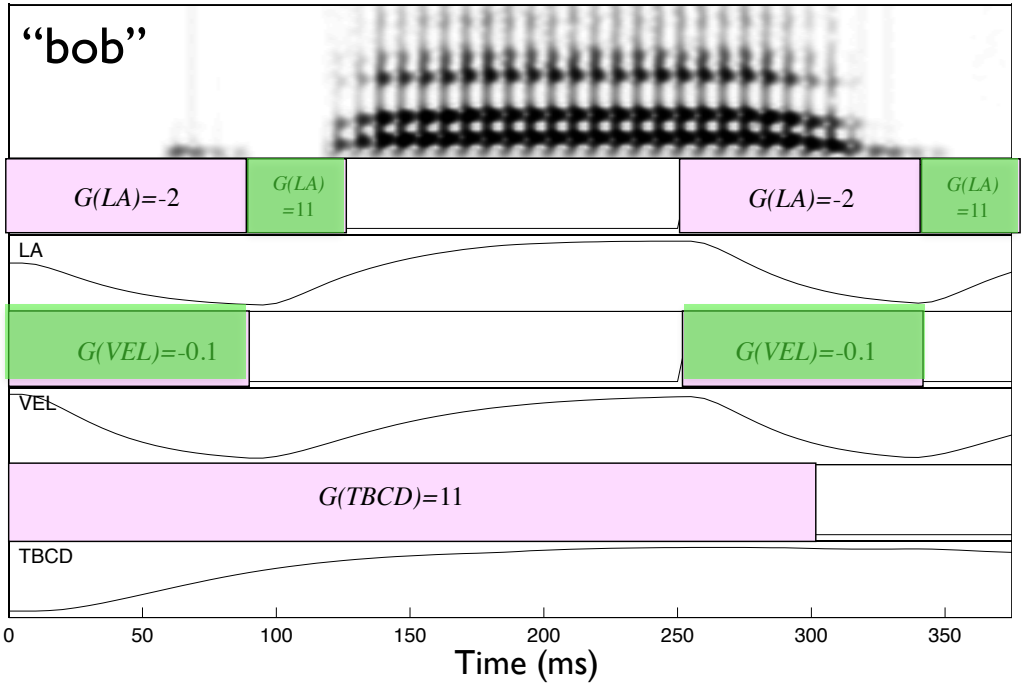
Gestural Scores

- Representation of the temporal organization of gestures
  - Time along horizontal dimension
  - Boxes represent intervals of time during which gestures are active in the vocal tract.
  - Gestures of oral constrictors, velum, glottis are displayed on different rows, e.g., "bad":



  - Labels on the boxes indicate the constriction degree (and location) of the gesture.
  - Default Simplifications (gestures left out of displays):
    Glottal narrowing for voicing
    Velic closure for oral stops

Gestural Score for “bob”

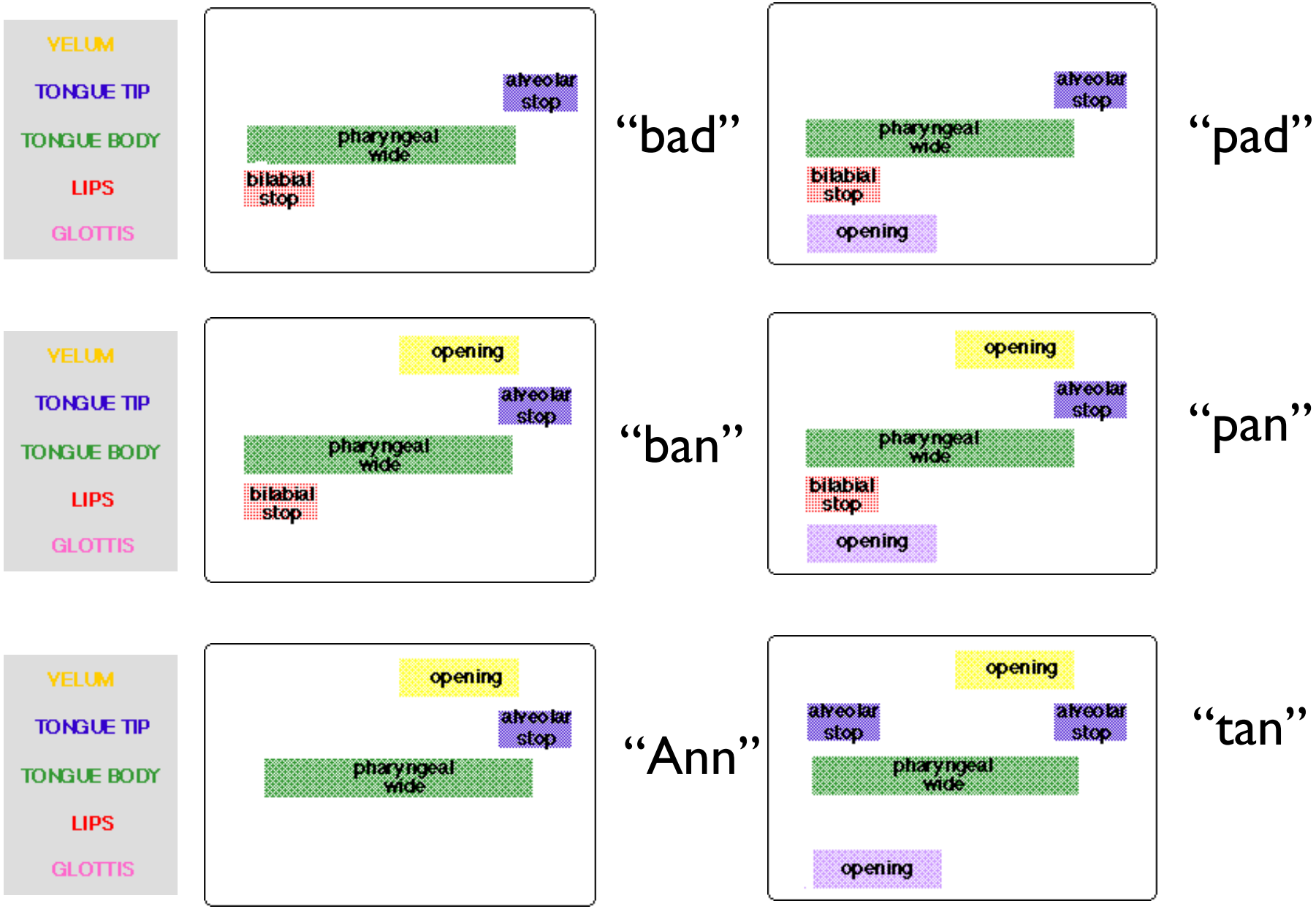



```
%/b/
'LA' 0 0 9 0 -2 8 1 JA=8,UH=5,LH=1 100 0.01
'LA' 0 9 13 0 11 8 1 JA=8,UH=5,LH=1 1 1
'VEL' 0 0 9 0 -0.1 8 1 NA=1 0 0

%/aa/
'TBCD' 0 0 30 0 11 4 1 JA=1,CL=1,CA=1 1 1
'TBCL' 0 0 30 0 170 4 1 JA=1,CL=1,CA=1 1 1

%/b/
'LA' 0 25 34 0 -2 8 1 JA=8,UH=5,LH=1 100 0.01
'LA' 0 34 37 0 11 8 1 JA=8,UH=5,LH=1 1 1
'VEL' 0 25 34 0 -0.1 8 1 NA=1 0 0
```

Contrast among gestural scores

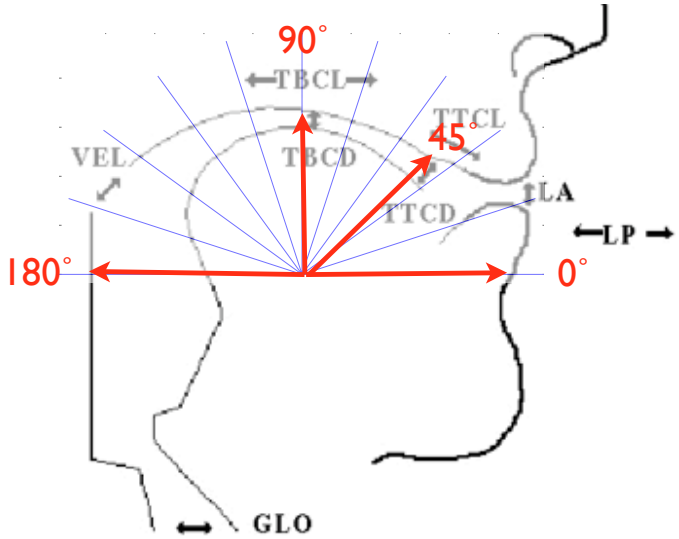
- Differences in gestural scores that can count as different words:
  - presence or absence of particular gestures
  - gestures' values of CD and CL
    - Can be abbreviated with “names” but are quantitative values (of target and stiffness of appropriate task variables)
  - organization of gestures in time

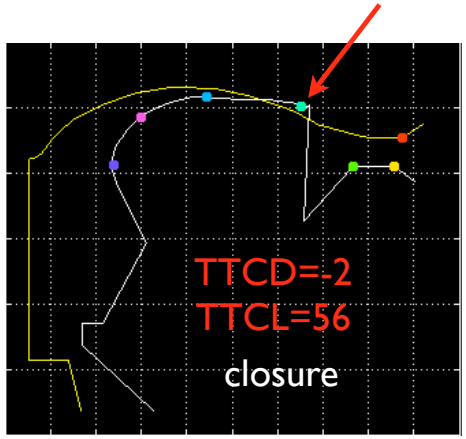
Presence or absence of gestures: compositionality

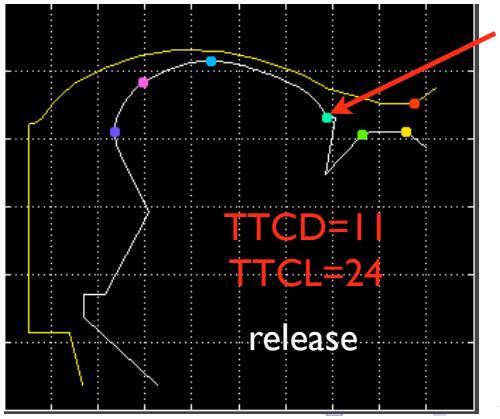


Tongue Tip Tasks

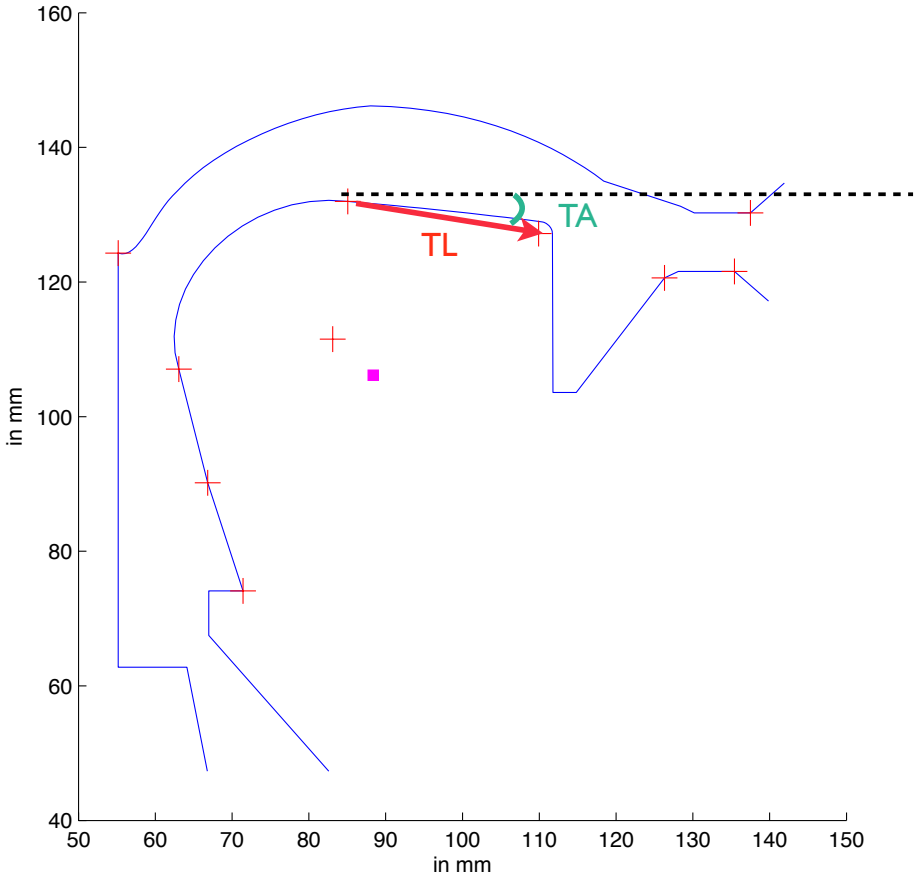
for /t,d,n/





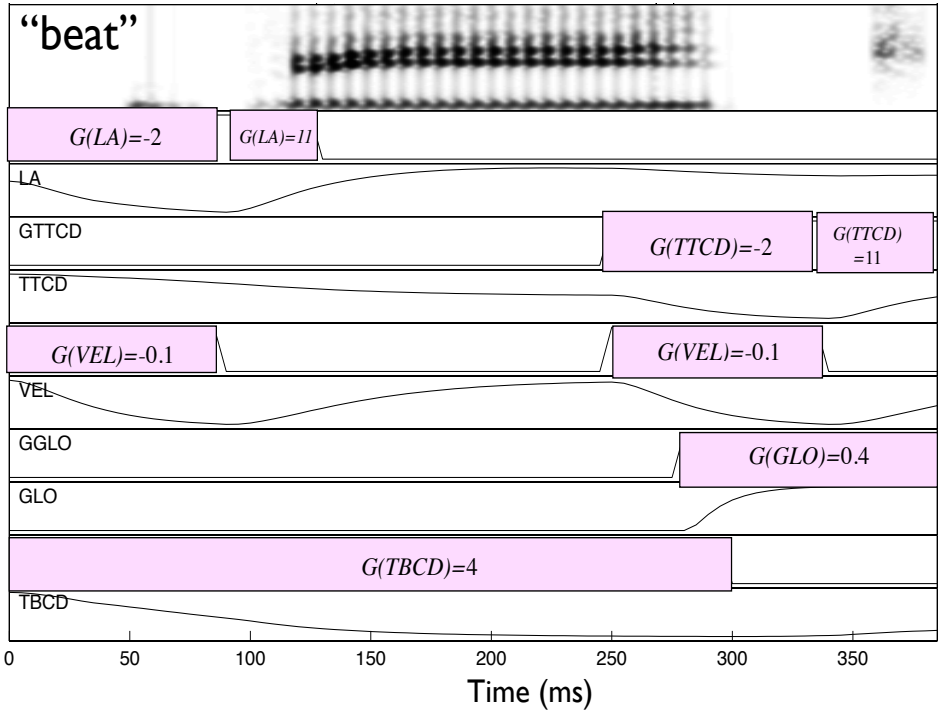


Articulators for TT Tasks



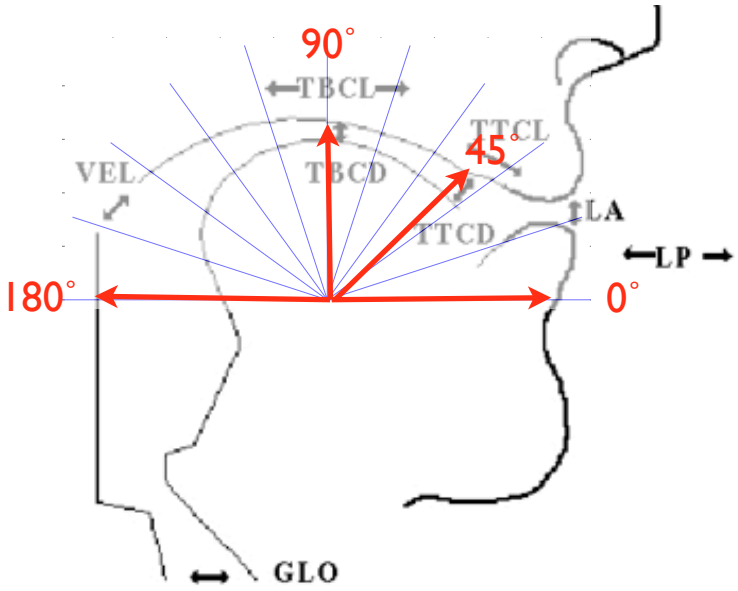
- And also Jaw Angle (JA), Tongue Center (CL, CA)

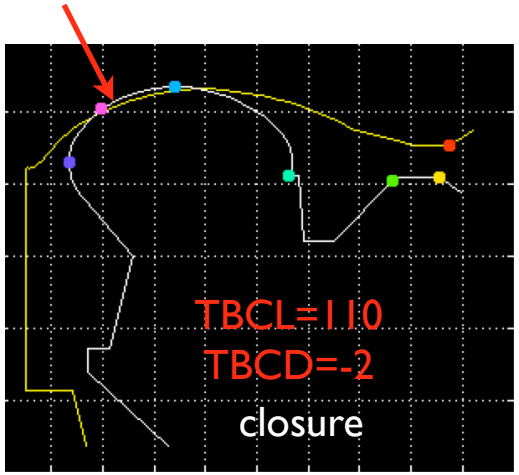
“beat”

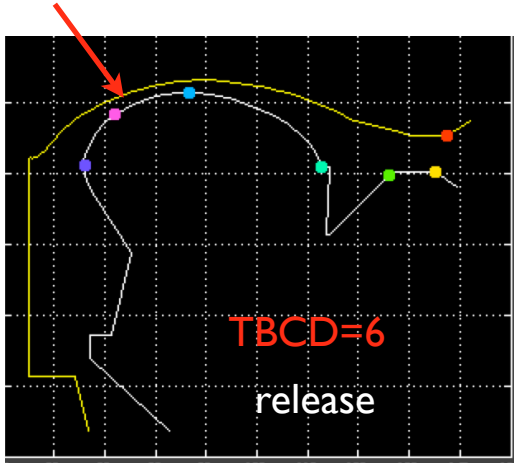



```
%/b/
'LA' 0 0 9 0 -2 8 1 JA=8,UH=5,LH=1 100 0.01
'LA' 0 9 13 0 11 8 1 JA=8,UH=5,LH=1 1 1
'VEL' 0 0 9 0 -0.1 8 1 NA=1 0 0

%/iy/
'TBCD' 0 0 30 0 4 4 1 JA=1,CL=1,CA=1 1 1
'TBCL' 0 0 30 0 95 4 1 JA=1,CL=1,CA=1 1 1


%/t/
'TTCD' 0 25 34 0 -2 8 1 JA=32,CL=32,CA=32,TL=1,TA=1 100 0.
'TTCL' 0 25 34 0 56 8 1 JA=32,CL=32,CA=32,TL=1,TA=1 1 1
'TTCL' 0 34 39 0 24 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'TTCD' 0 34 39 0 11 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'VEL' 0 25 34 0 -0.1 8 1 NA=1 0 0
'GLO' 0 28 39 0 0.4 16 1 GW=1 0 0
```

Tasks for velar stops?







Tasks for stop consonants

Lip (Labial)		T Tip (Alveolar)		T Body (Velar)	
CLO	REL	CLO	REL	CLO	REL
LA=-2	LA=11	TTCL=56 TTCD=-2	TTCL=24 TTCD=11	TBCL=110 TBCD=-2	TBCD=6
b VEL=-.1		d VEL=-.1		g VEL=-.1	
P VEL=-.1 GLO=.4		t VEL=-.1 GLO=.4		k VEL=-.1 GLO=.4	
m VEL=.2		n VEL=.2		nx VEL=.2	

Stops: Oral Constriction Gestures

```
% /b/

% clo
'LA' ... -2 8 1 JA=8,UH=5,LH=1 100 0.01

% rel
% 'LA' ... 9 13 0 11 8 1 JA=8,UH=5,LH=1 1 1

% /d/

% clo
'TTCL' ... 56 8 1 JA=32,CL=32,CA=32,TL=1,TA=1 1 1
'TTCD' ... -2 8 1 JA=32,CL=32,CA=32,TL=1,TA=1 100 0.01

% rel
'TTCL' ... 24 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'TTCD' ... 11 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1

% /g/

% clo
'TBCL' ... 110 8 1 JA=10,CL=1,CA=1 10 0.1
'TBCD' ... -2 8 1 JA=100,CL=1,CA=1 100 0.01

%rel
'TBCD' ... 6 8 1 JA=10,CL=1,CA=1 1 1
```

Glottal and Velic gestures for stops and fricatives

```
% Velic closure for stops and fricatives

'VEL' ... -0.1 8 1 NA=1 0

% Velic opening for nasals

'VEL' ... 0.2 8 1 NA=1 1 1

% Glottal opening for voiceless stops and fricatives

'GLO' ... 0.3 16 1 GW=1 0
```

Gesture Combinations

- English words can begin with combinations of Oral, Laryngeal and Velic constriction gestures.
- The resulting combinations are usually analyzed as consonants or consonant segments.
- From the gestures we illustrated, we can form 9 combinations (consonants) in English.

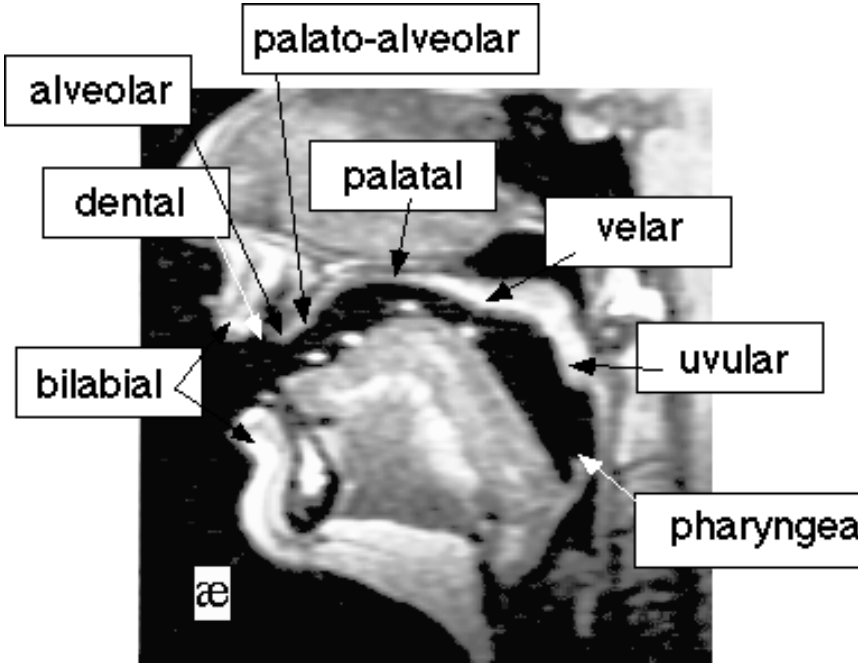
VELIC	closed	closed	open
LARYNX	narrow	open	narrow
LIPS	"bought"	"pot"	"Mott"
TT	"dot"	"tot"	"not"
TB	"got"	"cot"	"pong"

- But there are more than 9 consonants in English. Where do the rest come from?

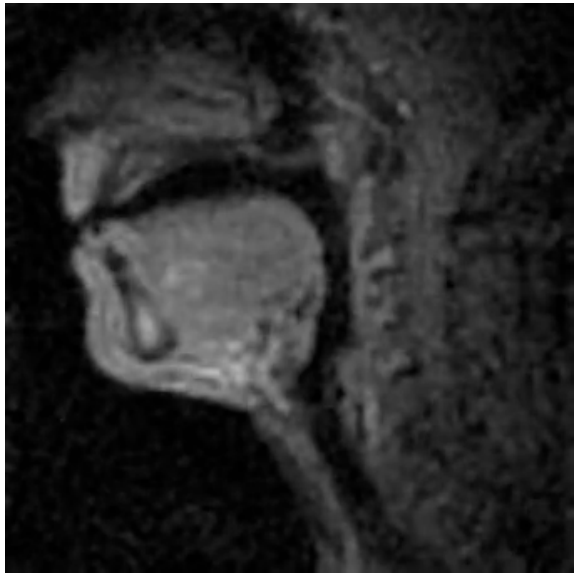
Differentiating oral constriction gestures

A given constrictor can produce several different distinctive gestures by varying:

- Constriction Degrees
  (how narrow is the constriction?)
  - stop ("dip, tip")
    complete obstruction of tube
    generates "pop" sound source
  - fricative ("zip, sip")
    narrowing to create jet noise source
  - approximant ("rip")
    narrowing with no source change
- Constriction Locations
  (exactly where is is it?)



Constriction Locations for TT fricatives



dental

"thick"

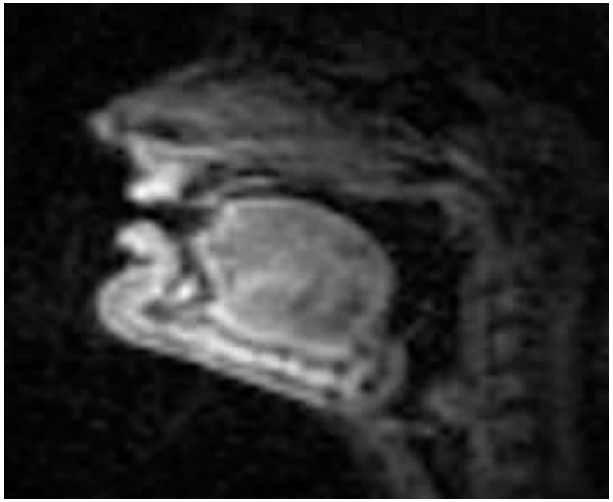
alveolar

"sick"

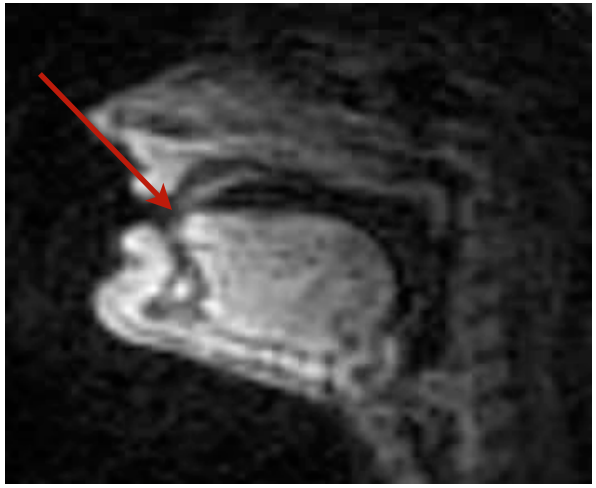
palatoalveolar

"Shick"

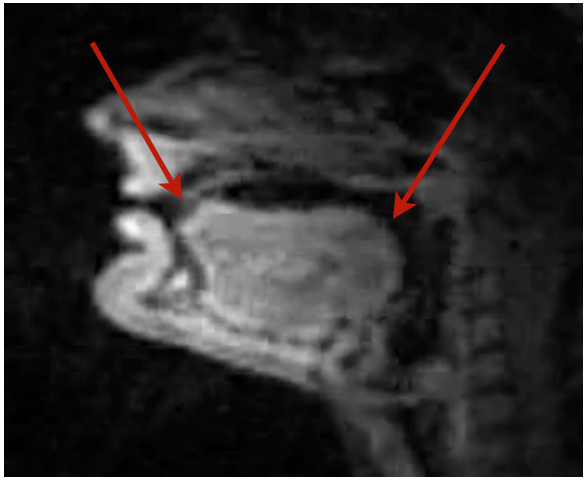
Fricatives: Oral Constriction Tasks



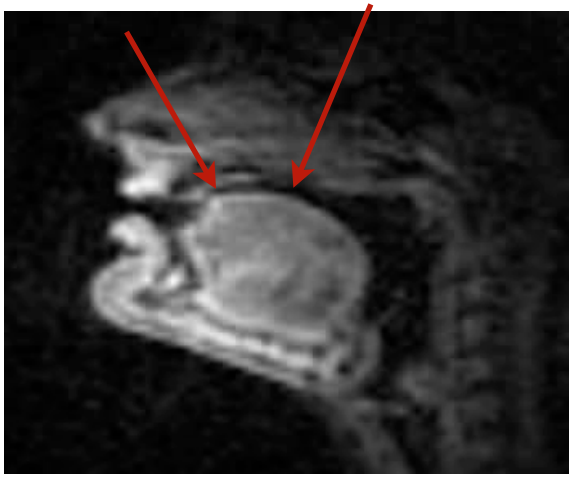
θ



z



ʃ



Fricatives: Oral Constriction Tasks

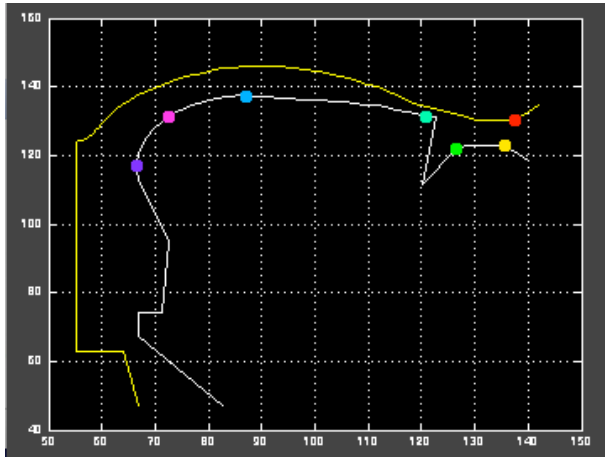
	θ		s		∫	
	clo	rel	clo	rel	clo	rel
TT	TTCL=40 TTCD=1.2	TTCL=24 TTCD=11	TTCL=56 TTCD=1.2	TTCL=24 TTCD=11	TTCL=60 TTCD=1.2	TTCL=40 TTCD=11
TB			TTCL=110 TBCD=10		TBCL=95 TBCD=8	

Fricatives: Oral Constriction Gestures

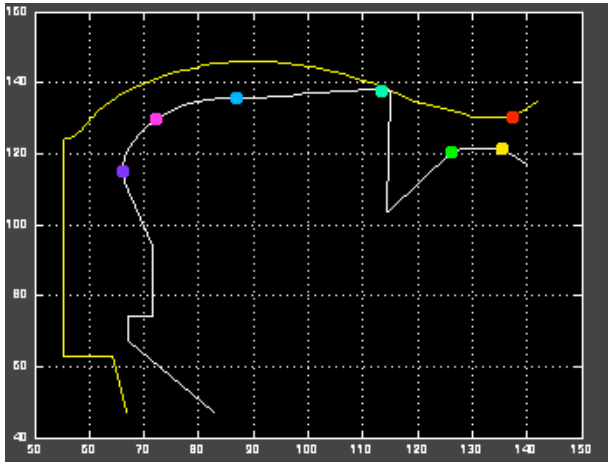
```
% /th/

% clo
'TTCL' ... 40 10 1 JA=32,CL=32,CA=32,TL=1,TA=1 1 1
'TTCD' ... 1.2 10 1 JA=32,CL=32,CA=32,TL=1,TA=1 10 0.1

% rel
'TTCL' ... 24 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'TTCD' ... 11 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1

% /s/

% clo
'TTCL' ... 56 10 1 JA=640,CL=32,CA=32,TL=1,TA=1 1 1
'TTCD' ... 1.2 10 1 JA=640,CL=32,CA=32,TL=1,TA=1 10 0.1

'TBCL' ... 110 8 1 JA=10,CL=1,CA=1 10 0.1
'TBCD' ... 10 8 1 JA=10,CL=1,CA=1 10 0.1

% rel
'TTCL' ... 24 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'TTCD' ... 11 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1

% /sh/

%clo
'TTCL' ... 60 10 1 JA=640,CL=32,CA=32,TL=1,TA=1 1 1
'TTCD' ... 1.2 10 1 JA=640,CL=32,CA=32,TL=1,TA=1 10 0.1

'TBCL' ... 95 8 1 JA=10,CL=1,CA=1 100 0.01
'TBCD' ... 8 8 1 JA=10,CL=1,CA=1 100 0.01

%rel
'TTCL' ... 40 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
'TTCD' ... 11 8 1 JA=512,CL=512,CA=512,TL=1,TA=1 1 1
```

Fricatives: examples

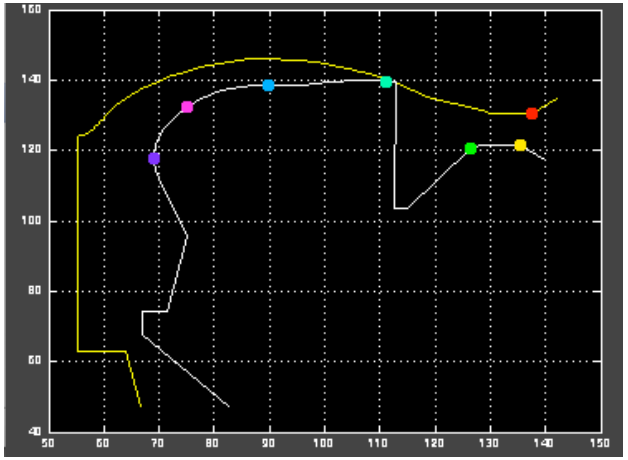
θ



s



ʃ



Glides: Oral Constriction Gestures

```
% /w/

%clo
'LA' ... 1 8 1 JA=8,UH=5,LH=1 1 1

'TBCL' ... 125 8 1 JA=10,CL=1,CA=1 10 0.1
'TBCD' ...   2 8 1 JA=10,CL=1,CA=1 100 0.01

%rel
'LA' ... 11 8 1 JA=8,UH=5,LH=1 1 1

% /j/

'TBCL' ... 95 8 1 JA=10,CL=1,CA=1 100 0.01
'TBCD' ...  2 8 1 JA=10,CL=1,CA=1 100 0.01
```

Multiple oral constrictions

"lie"

- Tongue Tip
  CD: stop
- Tongue Body
  CD: approximant, CL: uvular

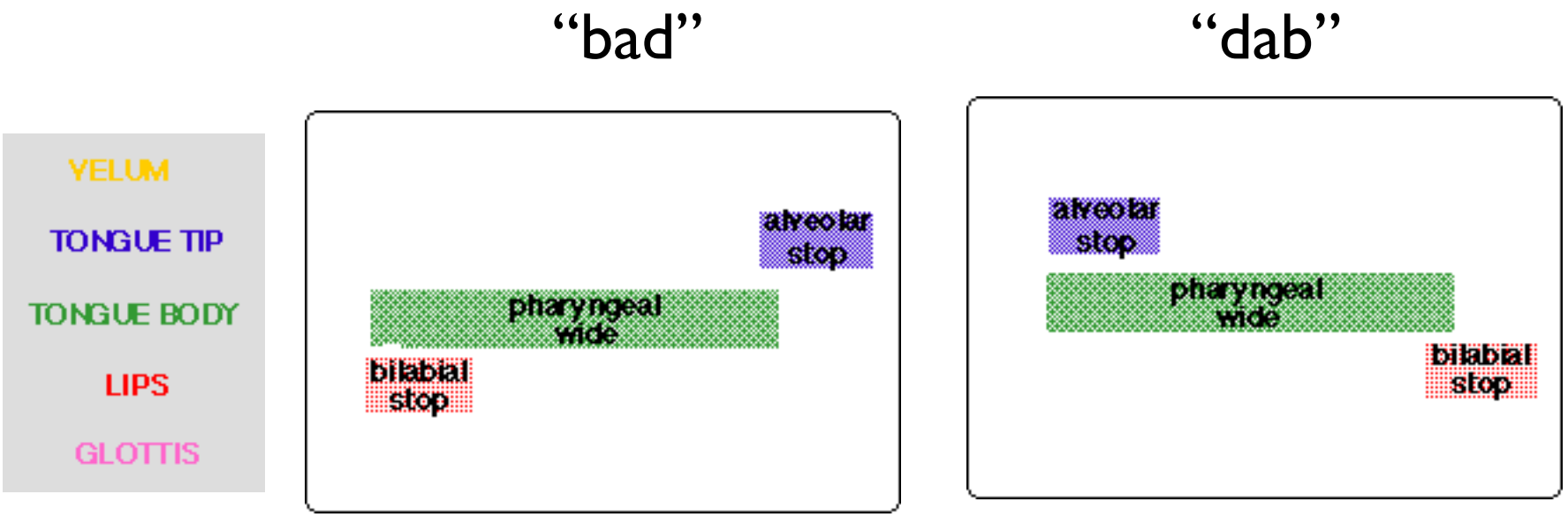
"rye"

- Lips
  CD: approximant
- Tongue Tip/Body
  CD: approximant, CL:palatal
- Tongue Root
  CD: approximant

Consonant Phonemes & Gestures

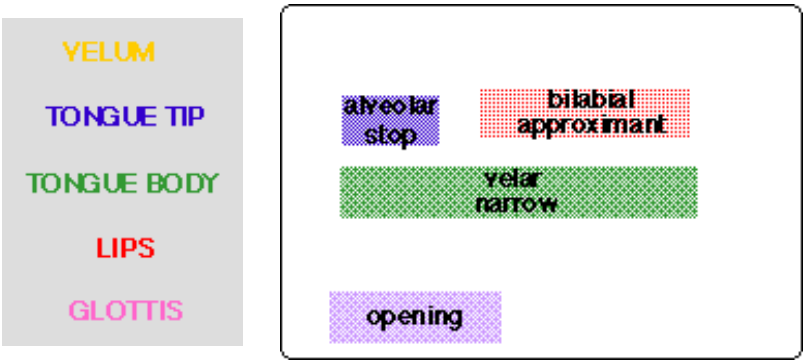
	lips	tip	body	velum	glottis
b	bilabial stop				
p	bilabial stop				opening
m	bilabial stop			opening	
t		alveolar stop			
d		alveolar stop		opening	
n		alveolar stop			opening
k			velar stop		
g			velar stop	opening	
ŋ			velar stop		opening
s		alveolar fric	(velar approx)		opening
z		alveolar fric	(velar approx)		
l		alveolar stop	(uvular approx)		
h					opening

Contrast: organization in time

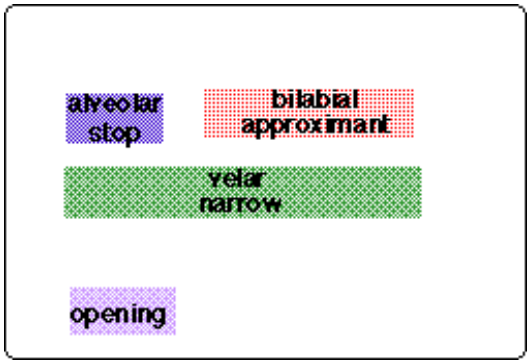


Cross-language timing differences: English vs. French

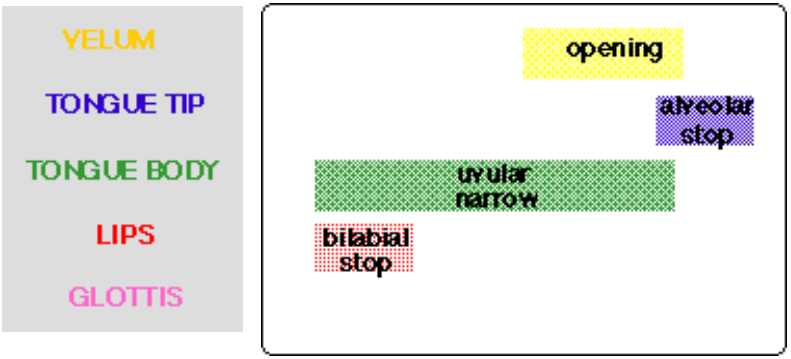
"two"



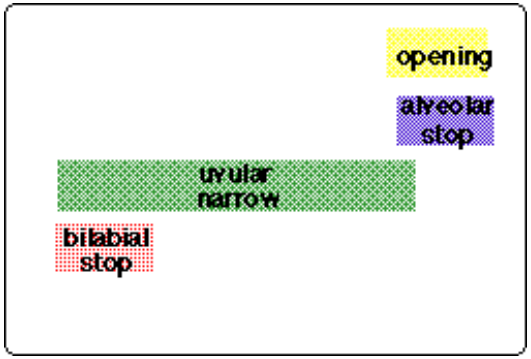
"tous"



"bun"



"bonne"



Gestural Scores & Phonetic Transcription

- Gestural scores
  - how how words differ from one another (contrast) in a given language
  - show how (similar) words in different languages are produced differently
- Phonetic transcription
  - intended to fulfill the same functions
- What is the relation?

Phonetic Transcription as Gestural Annotation

- Phonetic transcription can be viewed as a system (developed by phoneticians) for annotating the gestural score, using an string of alphabetic symbols.

- Humans can attend to (and become aware of) patterns of gesture in their own vocal tracts: which words are comprised of which gestures, in what rough organization.

    - This awareness must have been the basis for the invention of alphabetic writing systems, which represent gestural structures using strings of symbols.

- Analogy with chemistry

Transcription Types

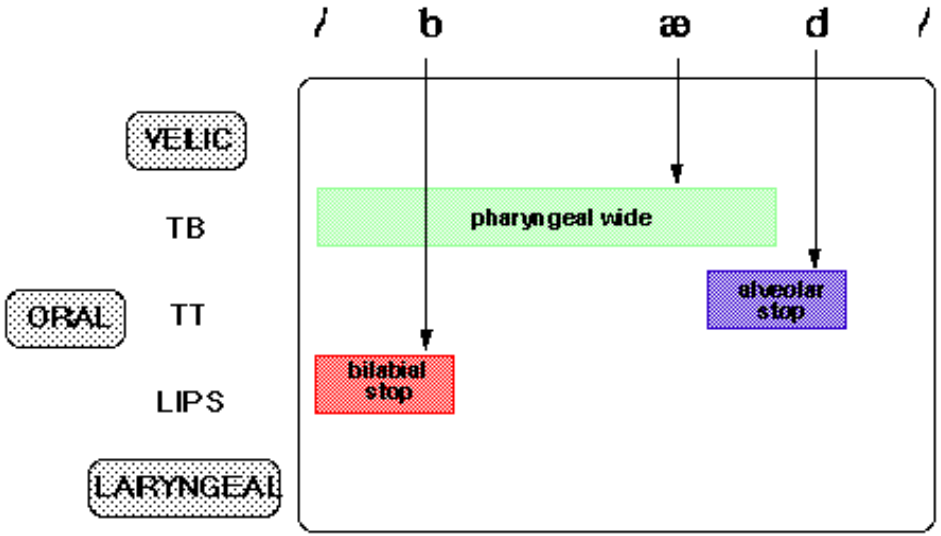
- Broad Phonemic
  - Each phoneme is a symbol for a contrastive gesture or a set (combination) of gestures
  - The order of phonemes symbolizes contrastive aspects of gestural organization.
  - Two transcriptions that differ in at least one symbol or one ordering are potentially contrastive.
- Narrow Phonetic
  - Annotates non-contrastive details of gestural score
  - Two transcriptions that differ in at least one symbol represent utterances that may or may not contrast

Phonemic Transcription: Annotation Principles

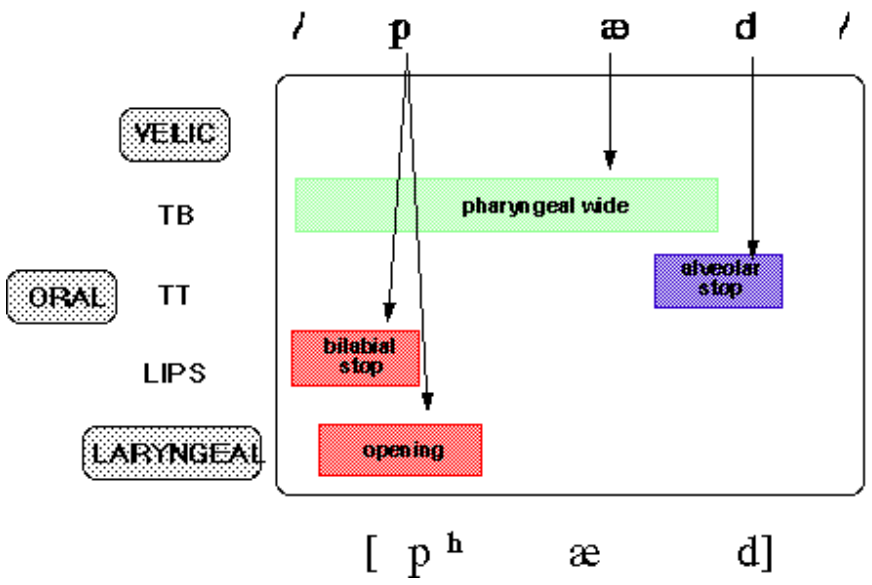
1. Each distinctive oral constriction gesture is annotated by a distinct symbol.

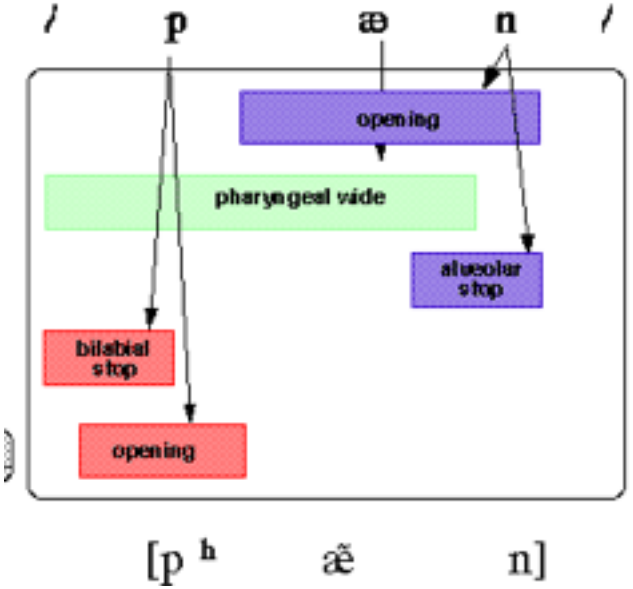
bæd vs dæd

2. The ordering of symbols represents the order in which the the corresponding gestures (or sets of gestures) reach their targets.



3. When a distinctive Laryngeal or Velic gesture overlaps an oral constriction gesture, a single symbol is used for the gestural combination.





Consonant Phonemes & Gestures

	lips	tip	body	velum	glottis
b	bilabial stop				
p	bilabial stop				opening
m	bilabial stop			opening	
t		alveolar stop			
d		alveolar stop		opening	
n		alveolar stop			opening
k			velar stop		
g			velar stop	opening	
ŋ			velar stop		opening
s		alveolar fric			opening
z		alveolar fric			
l		alveolar stop	uvular approx		
h					opening

Vowels

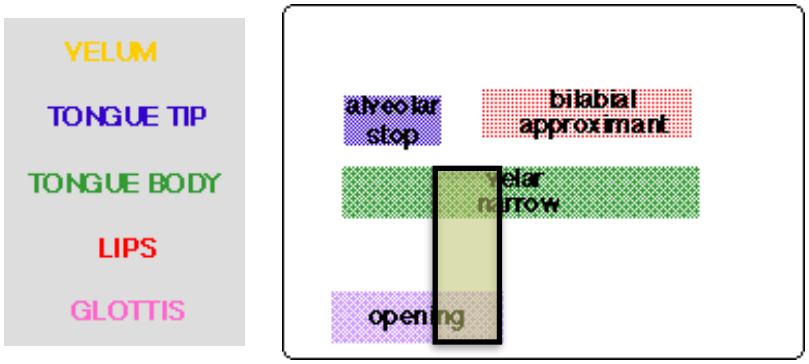
- A single symbol is used for distinctive combinations of tongue and lip gestures for vowels.
  - e.g., /bʊt/
- Diphthongs have two symbols: one for each tongue gesture
  - e.g. /baɪt/

Narrow Transcription

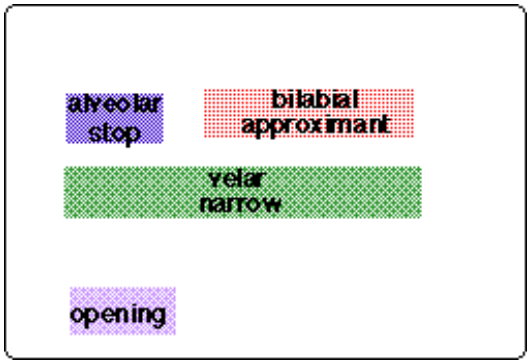
- Annotation of details of the gestural score
- Alphabetic symbols between square brackets e.g., [bæ̃n]
- Two transcriptions that differ in at least one symbol represent utterances that may or may not contrast
- English:
  - [tɛ̃n] in "ten times"
  - [tɛ̃n̪] in "ten things"

Language-specific details of gestural score

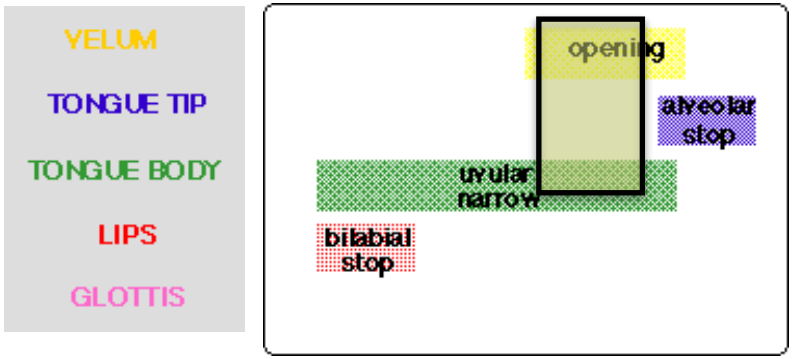
"two"



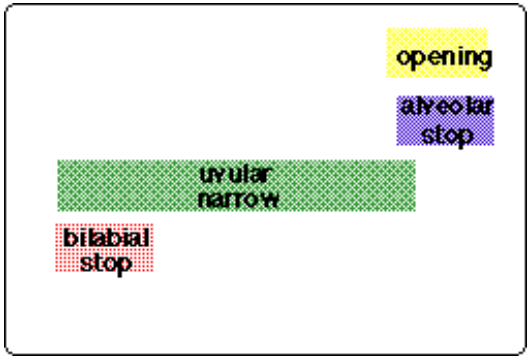
"tous"



"bun"



"bonne"



Types of Details

1.temporal overlap between gestures

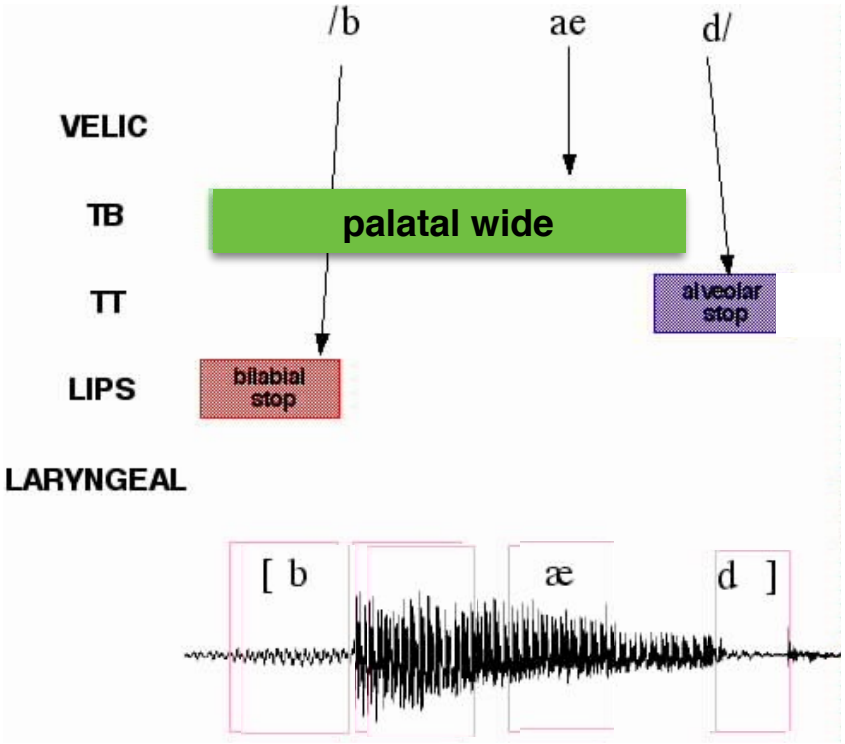
- aspiration of initial stops in English
- nasalization of vowels in English before nasals.

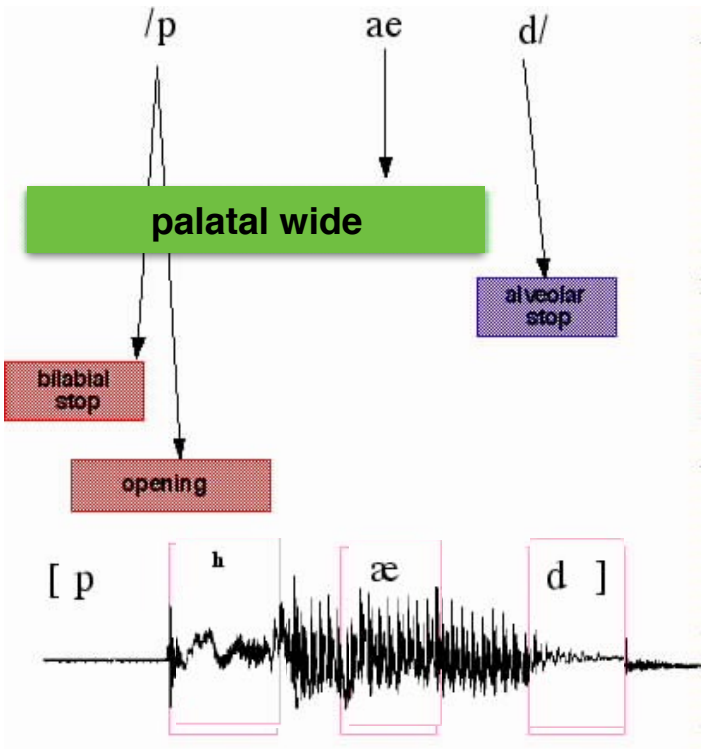
2. constriction properties

- degree:
  flapping in American English: Coronal stops and the laryngeal opening gestures "shrink" between stressed and unstressed vowels, and become approximants or "flaps".
  "latest" [leɪɾɪst]

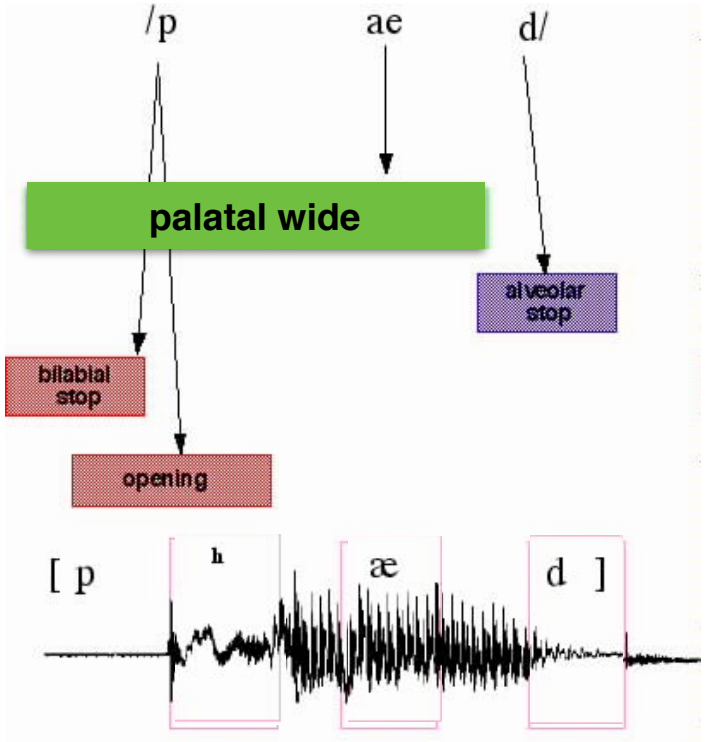
- location:
  place assimilation in American English: Final nasals may be assimilated to the place of a following stop.
  "miss you" [mɪʃju]

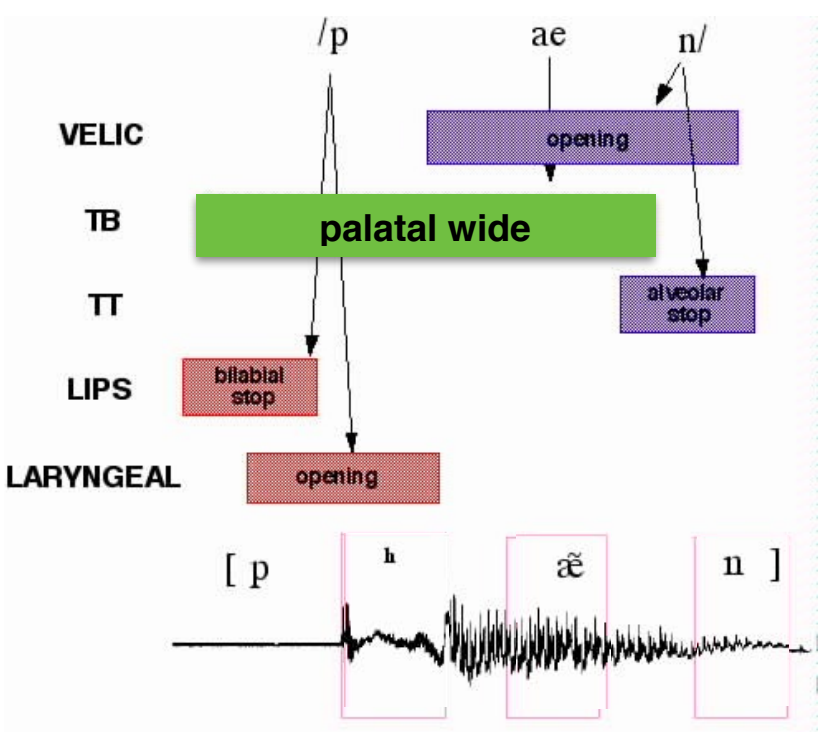
Aspiration of initial voiceless stops





Nasalization of vowels before nasals

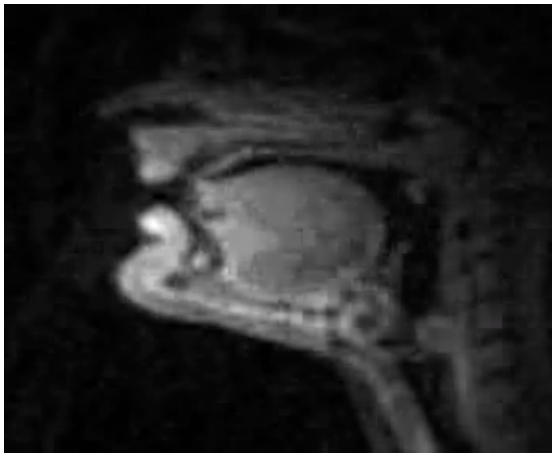




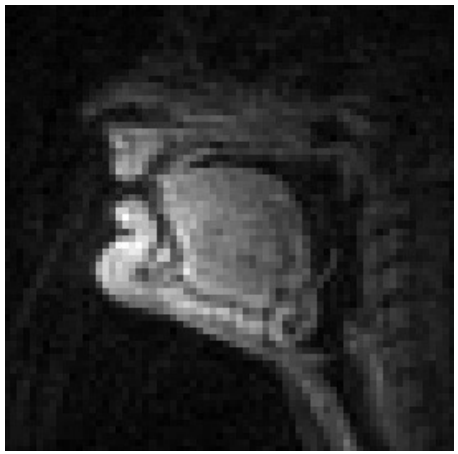
Flapping

"Tim takes"

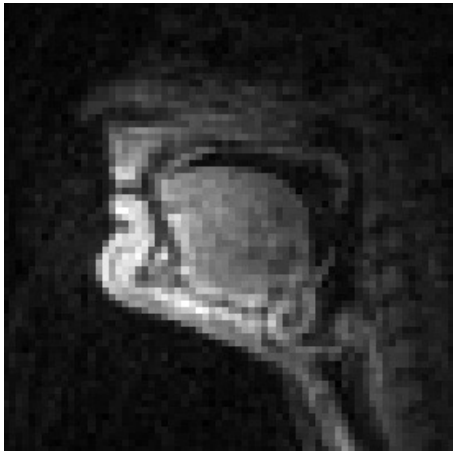
"latest"



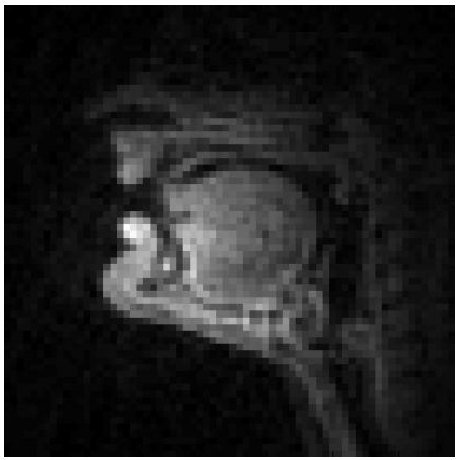
[t̪ʰɪm]



[tʰeɪks]



[leɪɾɪst]



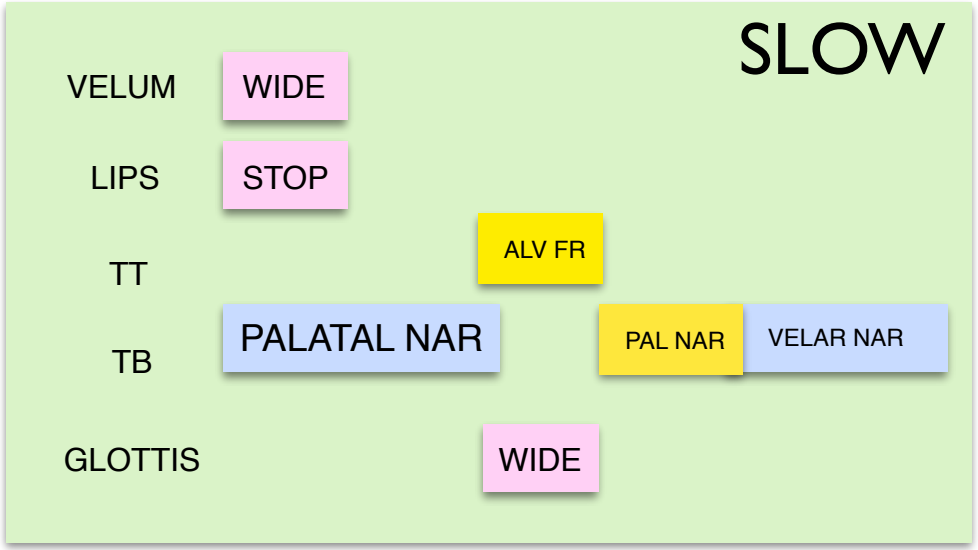
Place Assimilation

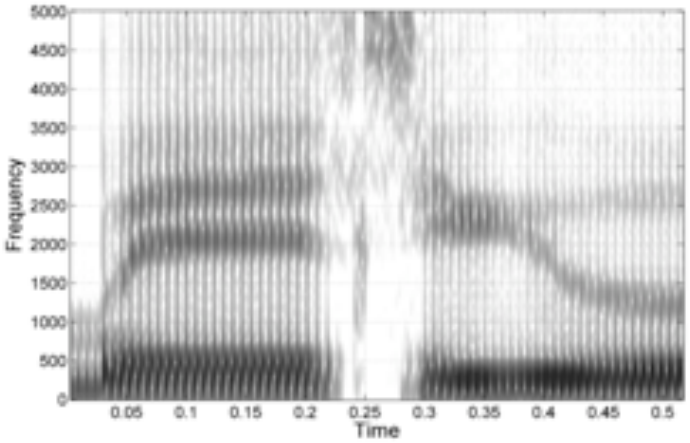
- Contextual variation in narrow transcription of the same word:
  - "miss it" [mɪs]
  - "miss you" [mɪʃ]
- Example
  - "I'm sure I'm gonna miss you"

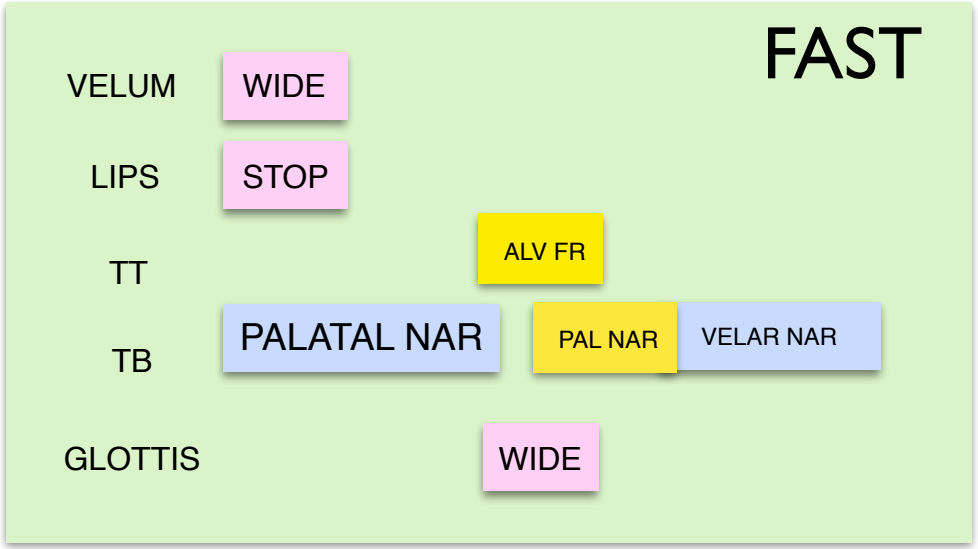
slow fast

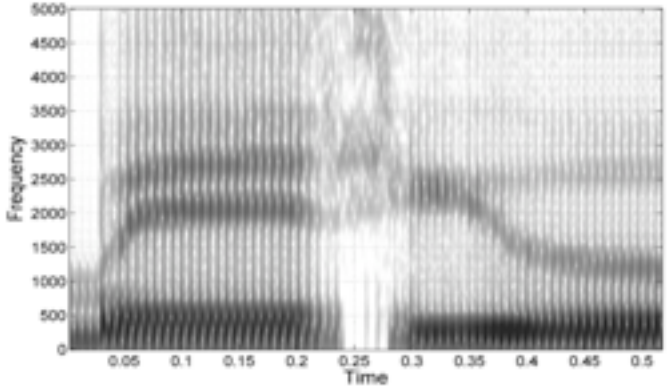
- What is going on here?
  - We change alveolar fricative to palatoalveolar before [j]?
  - Gestures overlap in time?

Change in Gestural overlap: Synthesis



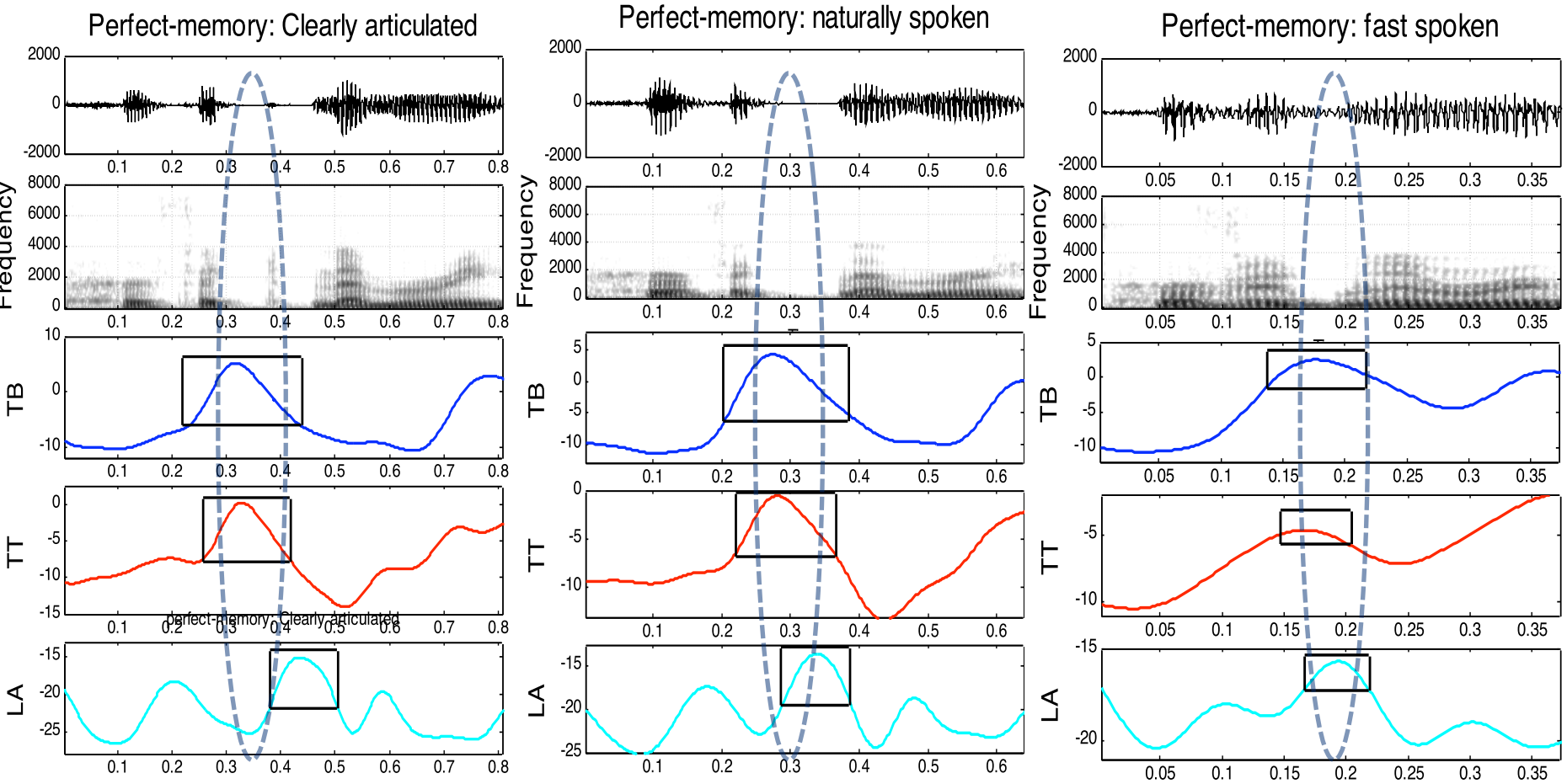






Example: /t/ "deletion"

from Tiede *et al.* 2004.



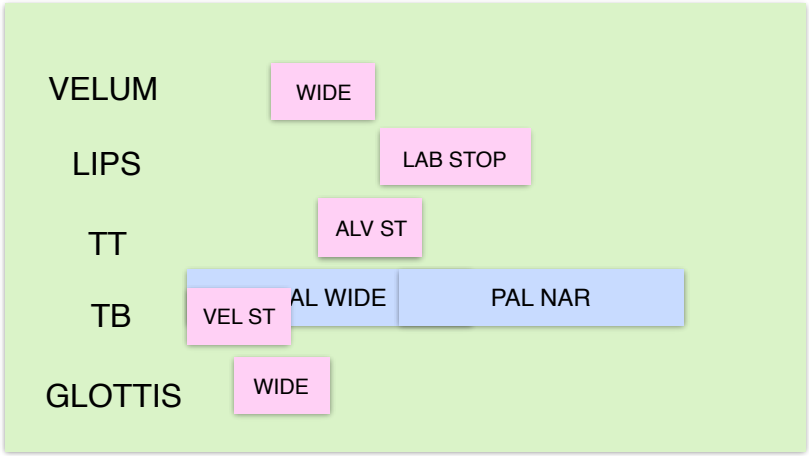
Due of change of speaking style, the degree of overlap between the gestures is altered

Place Assimilation: nasal

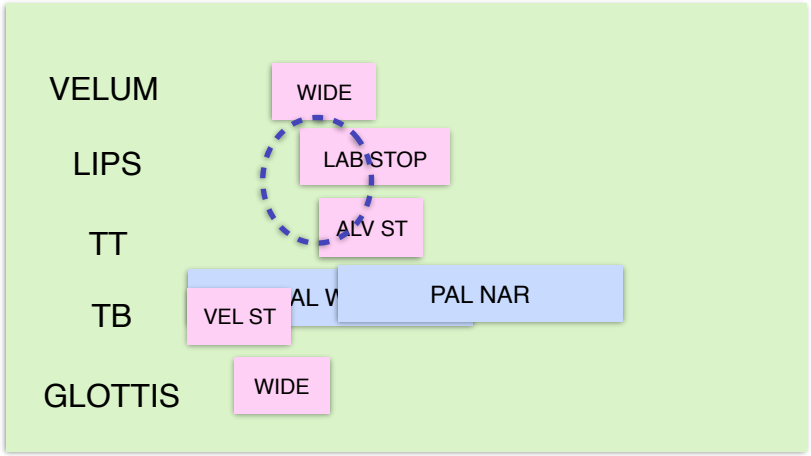
- Final /n/ is sometimes assimilated to the place of a following labial or dorsal stop:
- “can be”
  [kæ̃nbi] slow vs. [kæ̃mbi] fast

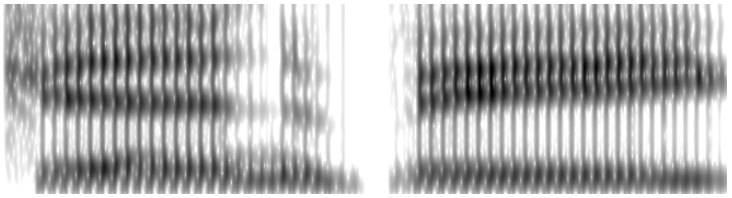
Nasal Assimilation: Synthesis

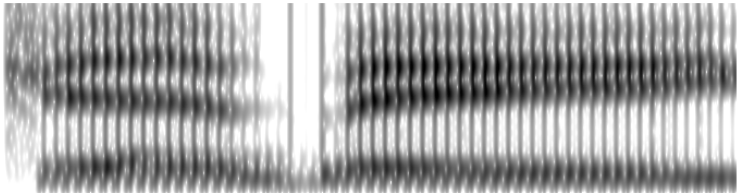
"can be" SLOW



"can be" FAST

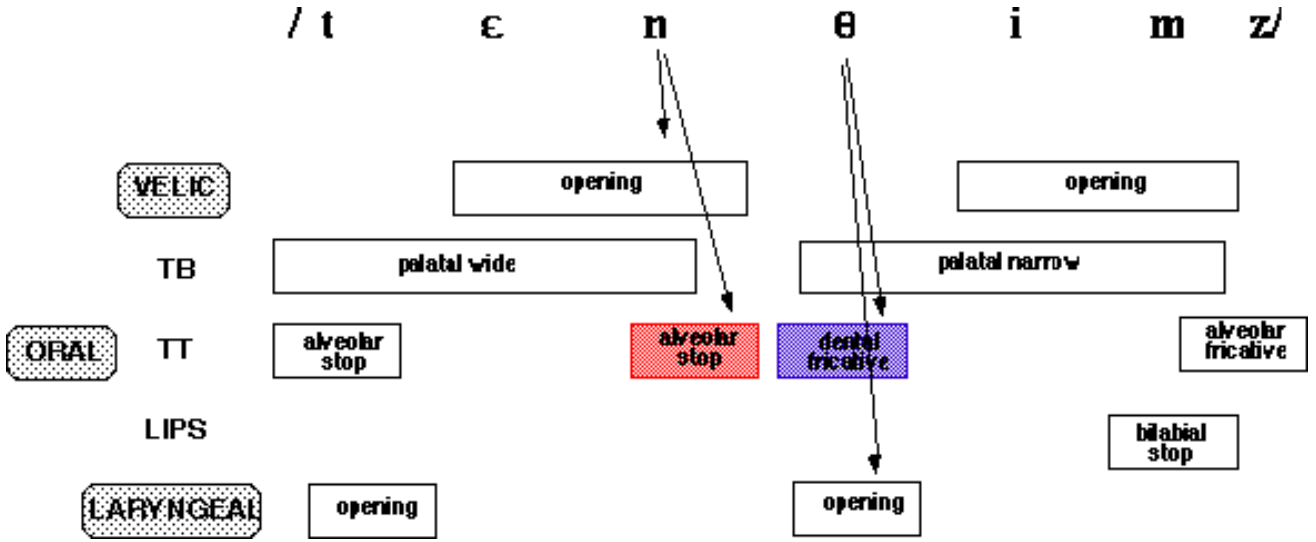


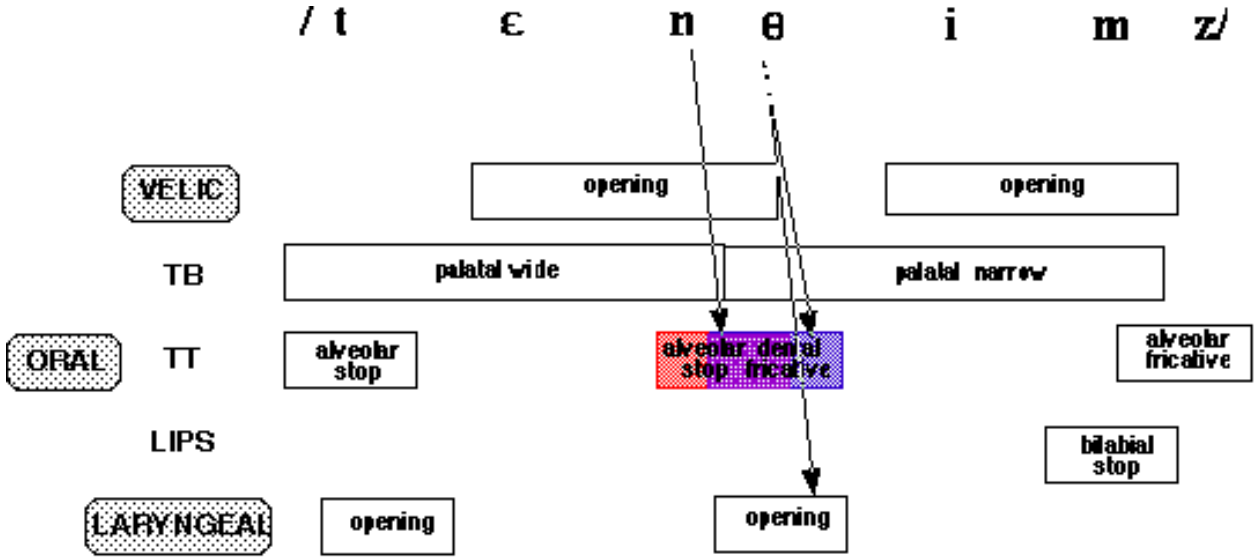




Nasal Assimilation to a following following coronal

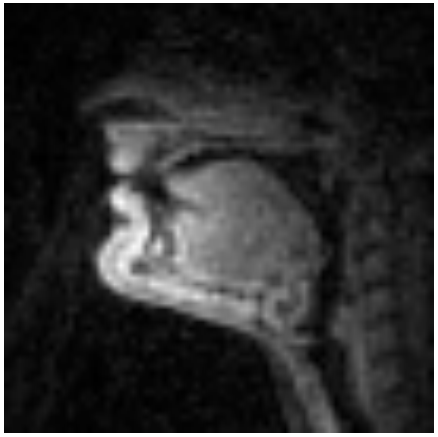
- “ten times” [tɛ̃n] vs “ten things” [tɛ̃n̪]
- overlap of alveolar nasal and dental fricative results in blending of the two TT gestures


/ t ɛ n θ i m z/
VELIC
opening
opening
TB
palatal wide
palatal narrow
ORAL
TT
alveolar stop
alveolar stop
dental fricative
alveolar fricative
LIPS
bilabial stop
LARYNGEAL
opening
opening

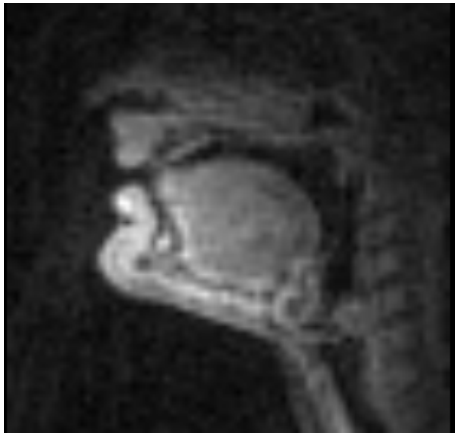

/ t ɛ n θ i m z/
VELIC
opening
opening
TB
palatal wide
palatal narrow
ORAL
TT
alveolar stop
alveolar stop
dental fricative
alveolar fricative
LIPS
bilabial stop
LARYNGEAL
opening
opening

[t ʰ ɛ̃ n̪ θ i m z]

MRI evidence for blending

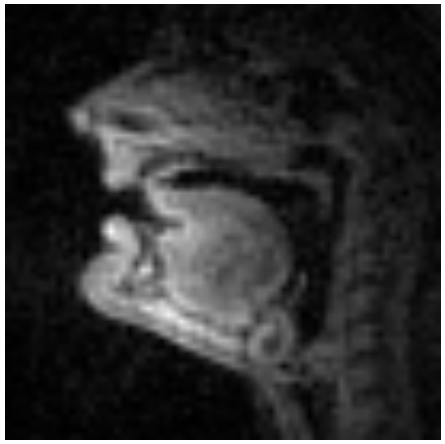


“shorten this”

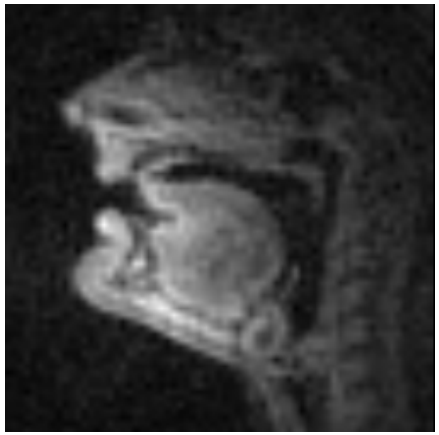


[n̪]

[ð]



“open every”



[n]