Bengali (and a few other) geminates [in the mental lexicon]

Aditi Lahiri, Sandra Kotzor, Adam Roberts, & Allison Wetterlin,
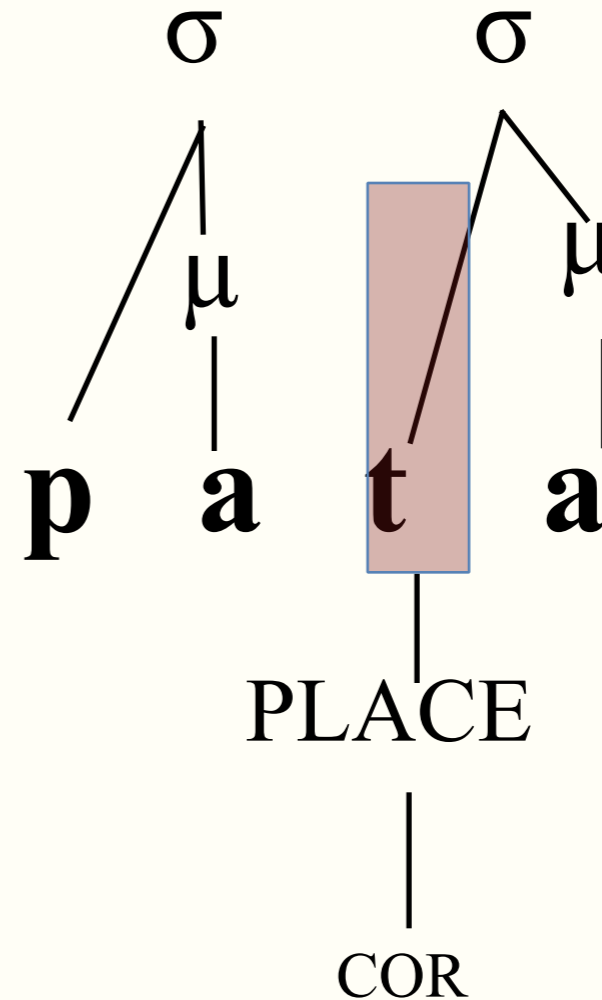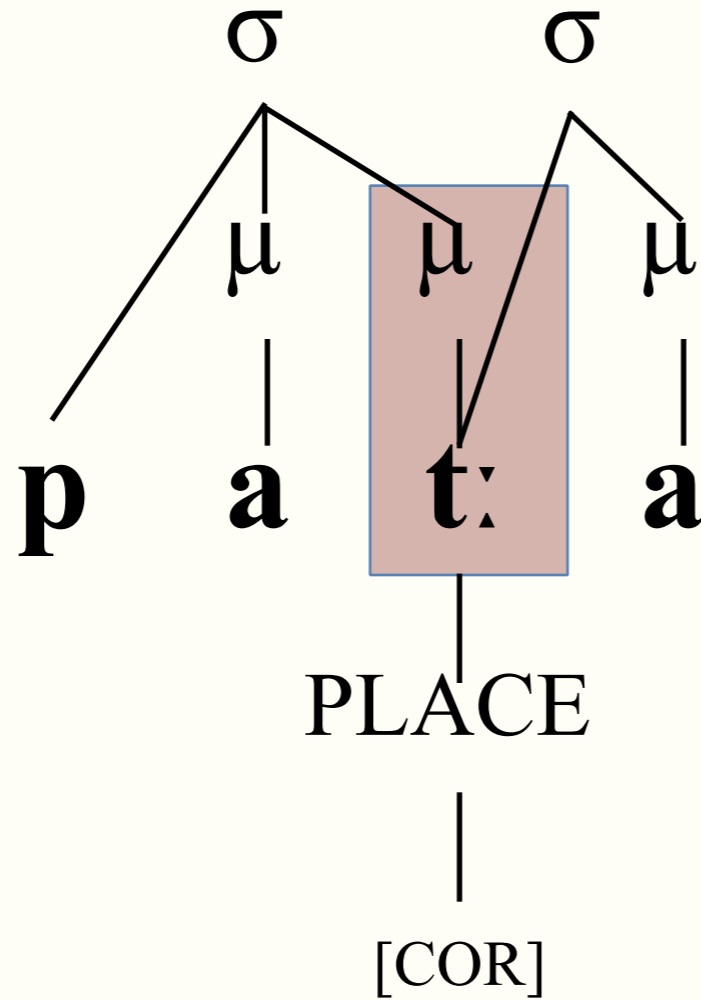*Language and Brain Laboratory, University of Oxford*

WORKSHOP
Geminate consonants across the world
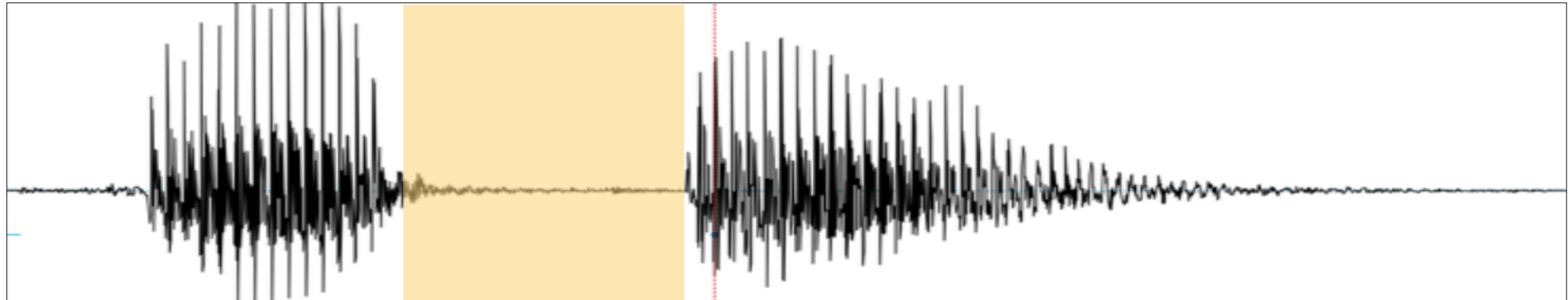ICPhS 18, 2015, GLASGOW

Bengali singleton & geminates: Phonological contrasts

METRICAL/ SKELETAL (TIMING & WEIGHT)

FEATURAL

σ   σ

μ   μ   μ

p   a   t: a
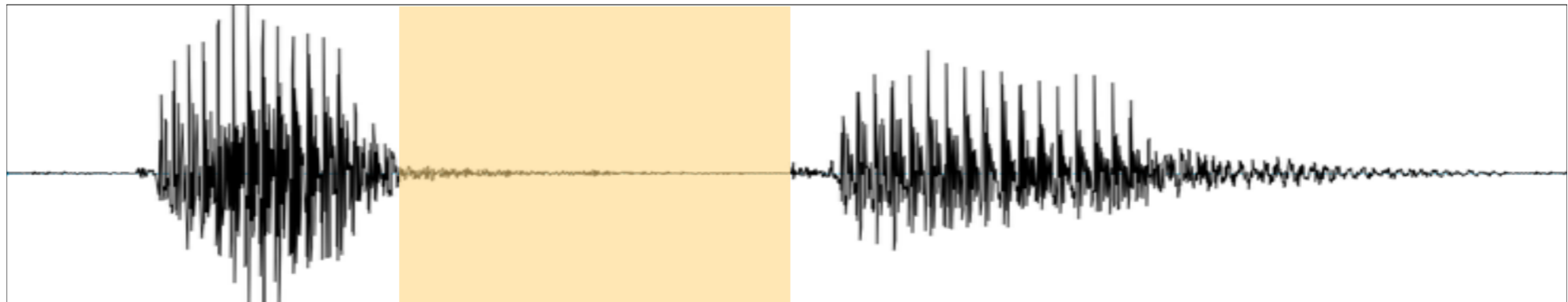
PLACE

[COR]

σ   σ

μ   μ

p   a   t   a

PLACE

COR

Acoustic differences: singleton ~ geminate voiceless stops

pata 'leaf'

patːa 'whereabouts'

Phonological contrasts

METRICAL

σ   σ   σ   σ

μ   μ   μ   μ   μ

p: a t: a   p: a t a
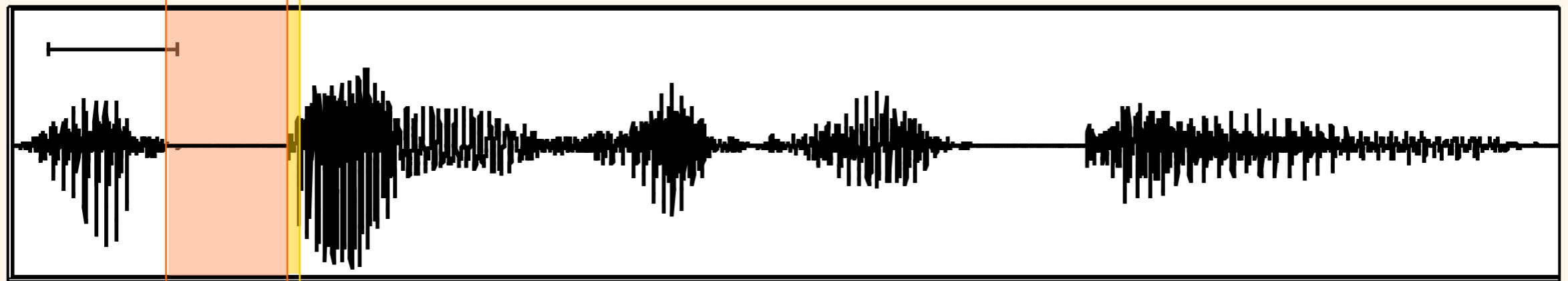
FEATURAL

PLACE   PLACE
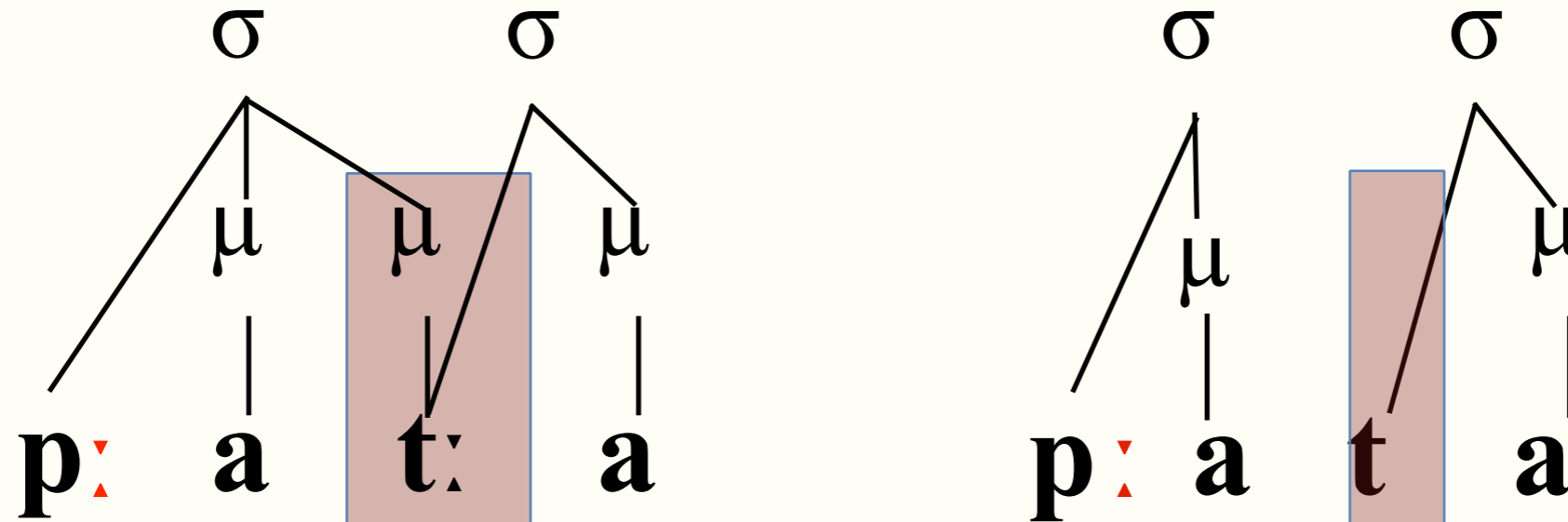
[COR]   COR

SWISS GERMAN - a caveat

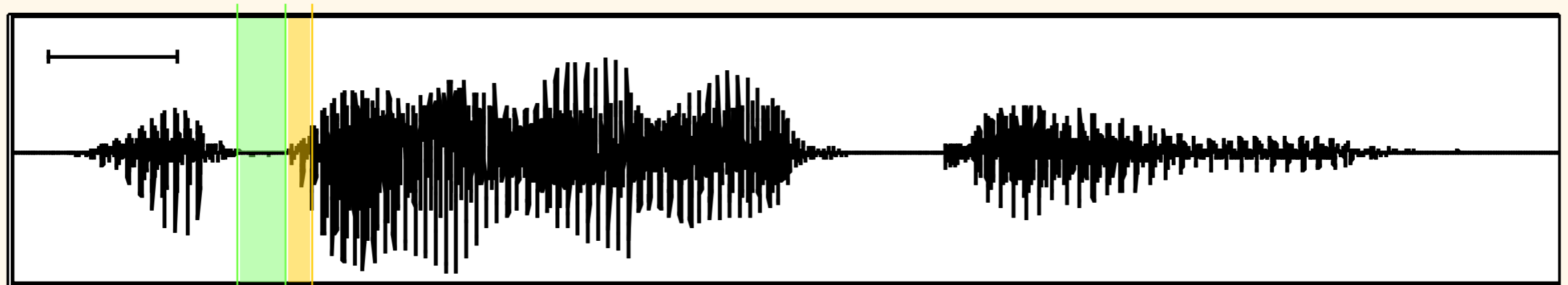# Initial stops in Swiss German in phrase meidal position

CD VOT

i ha pp o m f ʀ i tt n ∅ kk ɛ ʀ n

'I don't like French fries.'

CD VOT

i ha p oo n e n ∅ kk ɛ ʀ n

'I don't like beans.'

A mystery as to how & why such geminates should survive

Lahiri & Krahenmann (2004); Kraehenmann & Lahiri (2008)

Phonetic & phonological evidence for initial voiceless geminate stops 1000 years ago
Notker's *Anlautgesetz* Law of Initials and initial geminates in Swiss German

Notker Labeo, an Abbot of the monastery at St Gall, Switzerland "wrote as he spoke and heard"

He devised his own alphabet & the orthography reflects his pronunciation.

b   d   g   word-initially after **sonorants**, e.g. *vowels,* /n, m, l, r/

p   t   k/c   word-initially after **plosives & fricatives**, e.g. /p, t, k, h/ after a pause

ín díu óugen begínnet (Nc09720)
Íh pegínne (Nc03519)

díu súnna gât (Nc02311)
er férrost   kât (Nc10721)

Uńde dáz kelóuben so uúaz îh pefîndo fóne dîr
Uuás mag táz siñ?

Examples: *Martianus Capella* (Codex Sangallensis 872) early 11th century.

Words beginning with /p~b/
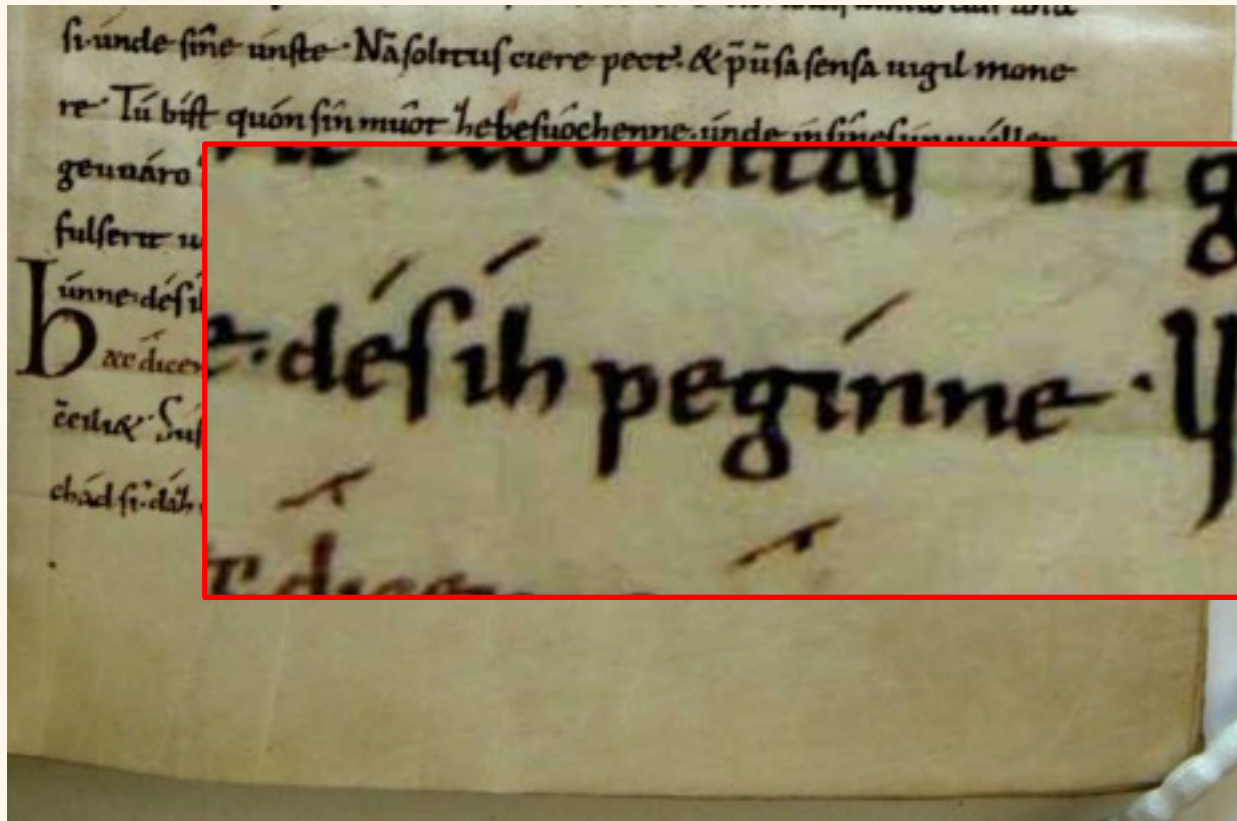
dés í**h** **p**egínne
that I   begin

in díu óuge**n** **b**egínnên.
In   the   eyes   begin (*3p.sg*)
'**it begins in the eyes**'

Words beginning with /k~g/
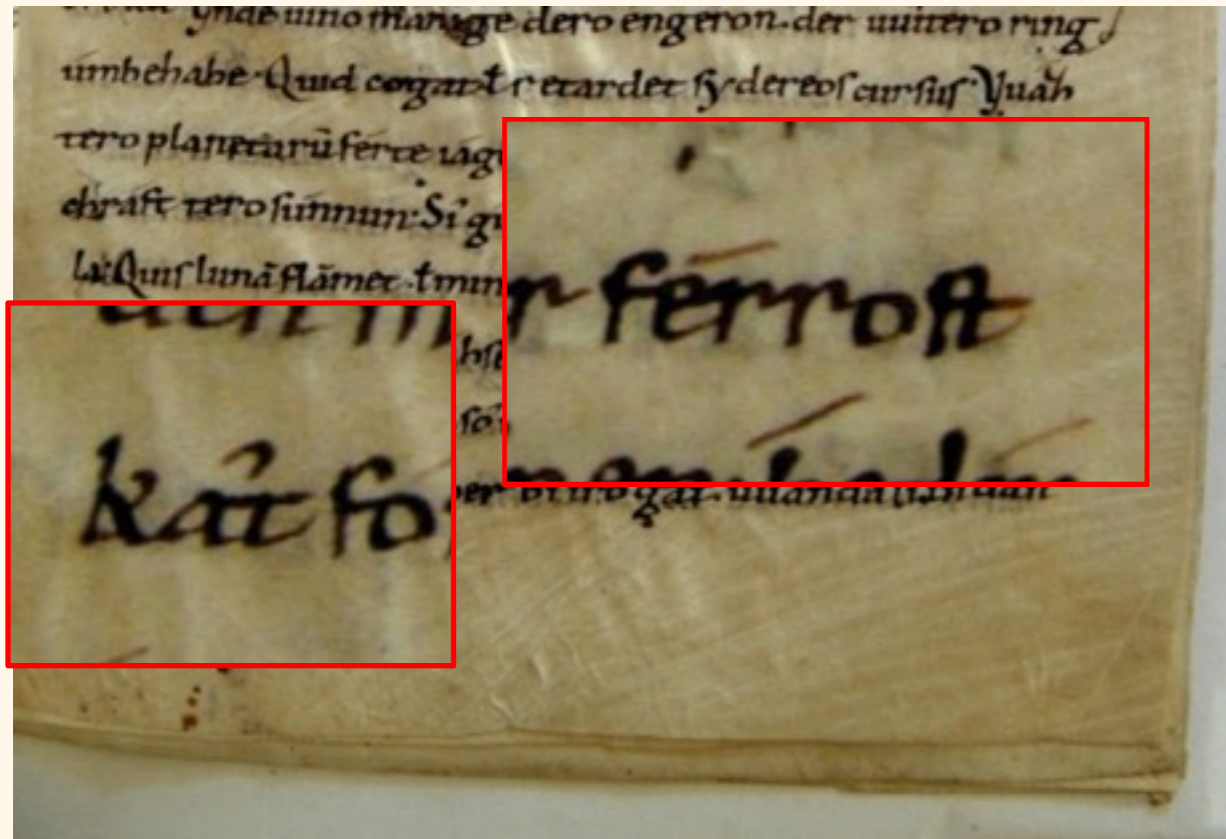
Sô er férrost kât
he furthest goes

diu súnna gât
the sun   goes

Words beginning with /t~d/

Uuás ma**g** **t**áz siñ?

Uńd**e** **d**áz kelóuben so uúaz
íh pefindo fóne dír
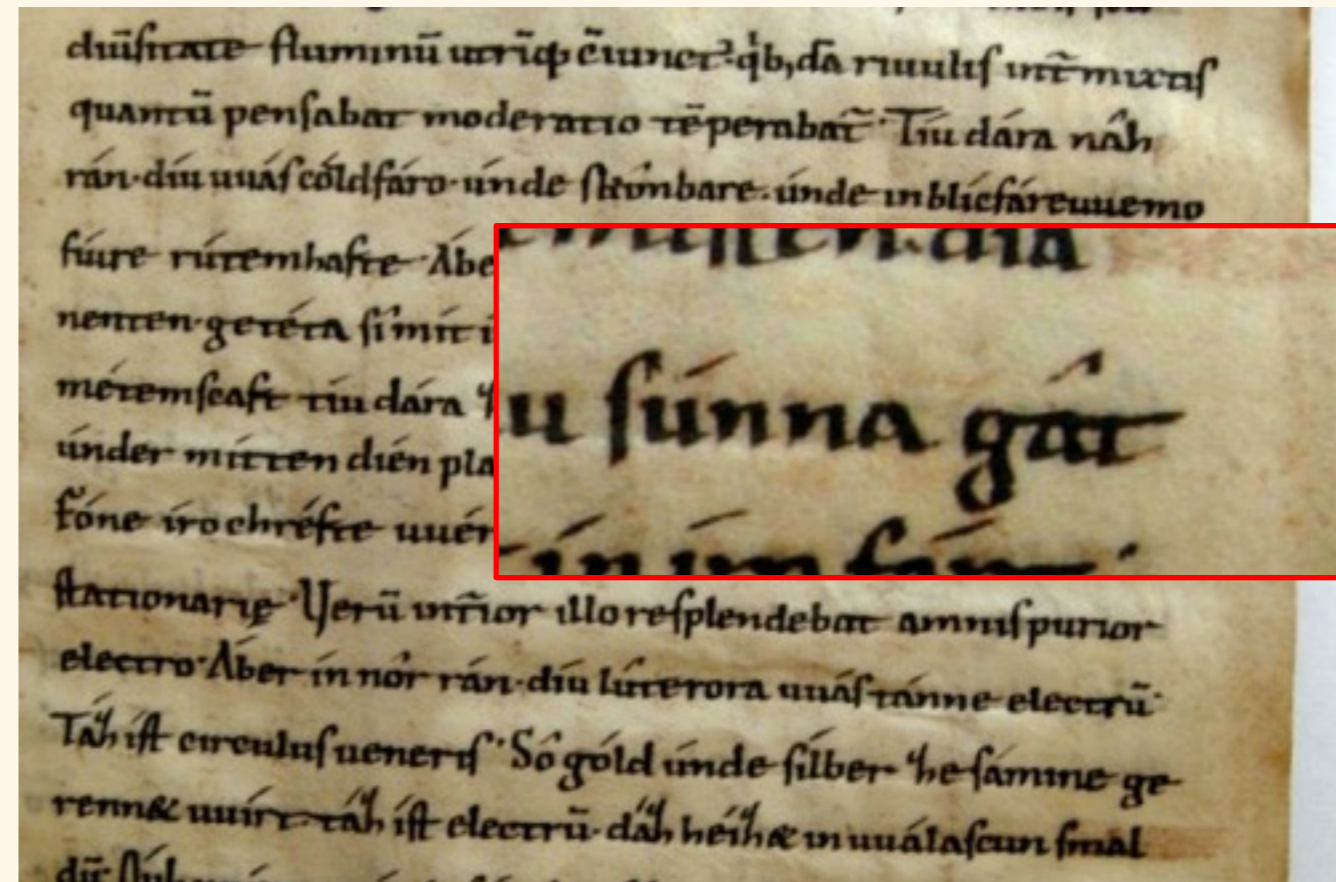
Asymmetry in word initial consonants: Notker's *Anlautgesetz* Law of Initials

Letters **b d g**   word-initially after sonorants, e.g. /*vowels*, n, m, l, r/
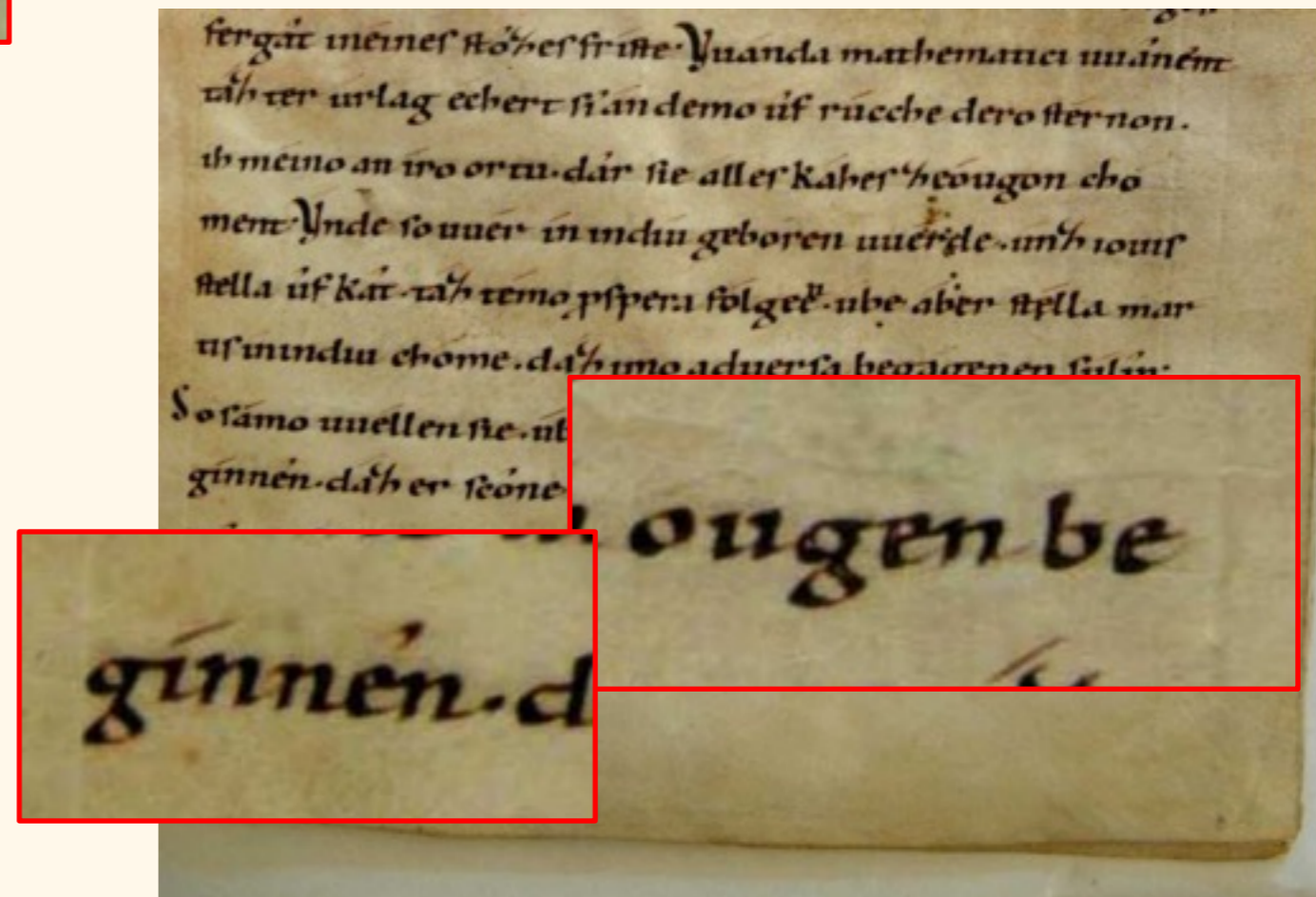Letters   **p t k/c**   word-initially after an plosives, e.g. /p, t, k, h/
        beginnen ~ peginnen, gat ~ kat, das ~ tas

This suggests that there was no meaningful contrast between word initial
<p t k> & <b d g>:   **the consonant alternation was entirely predictable.**

The phonemic inventory had no voiced /b d g/

In **CORONAL** consonants,
we find exceptions!

Asymmetry: Exceptions with some CORONAL <t>s

Sô mánig **t**ág ist in iâre

Expected: Letter <t> follows an obstruent

Sô uuárd táz ter **t**ág pegónda décchen die stérnen

Unexpected: Letter <t> follows a sonorant!

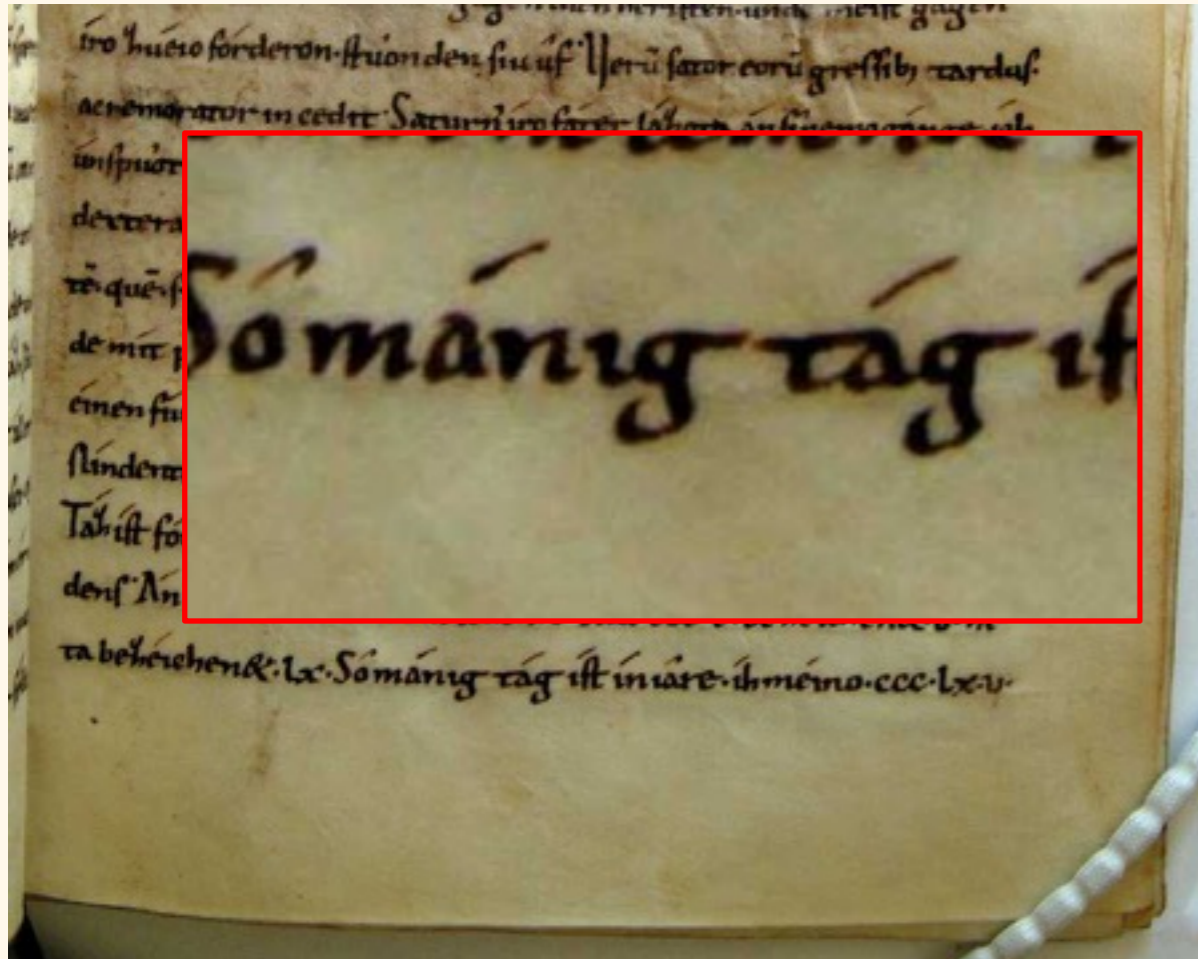# Asymmetry: always CORONAL \<t\> following obstruent & sonorant
no alternation

Sô máni**g** **t**ág ist in iâre
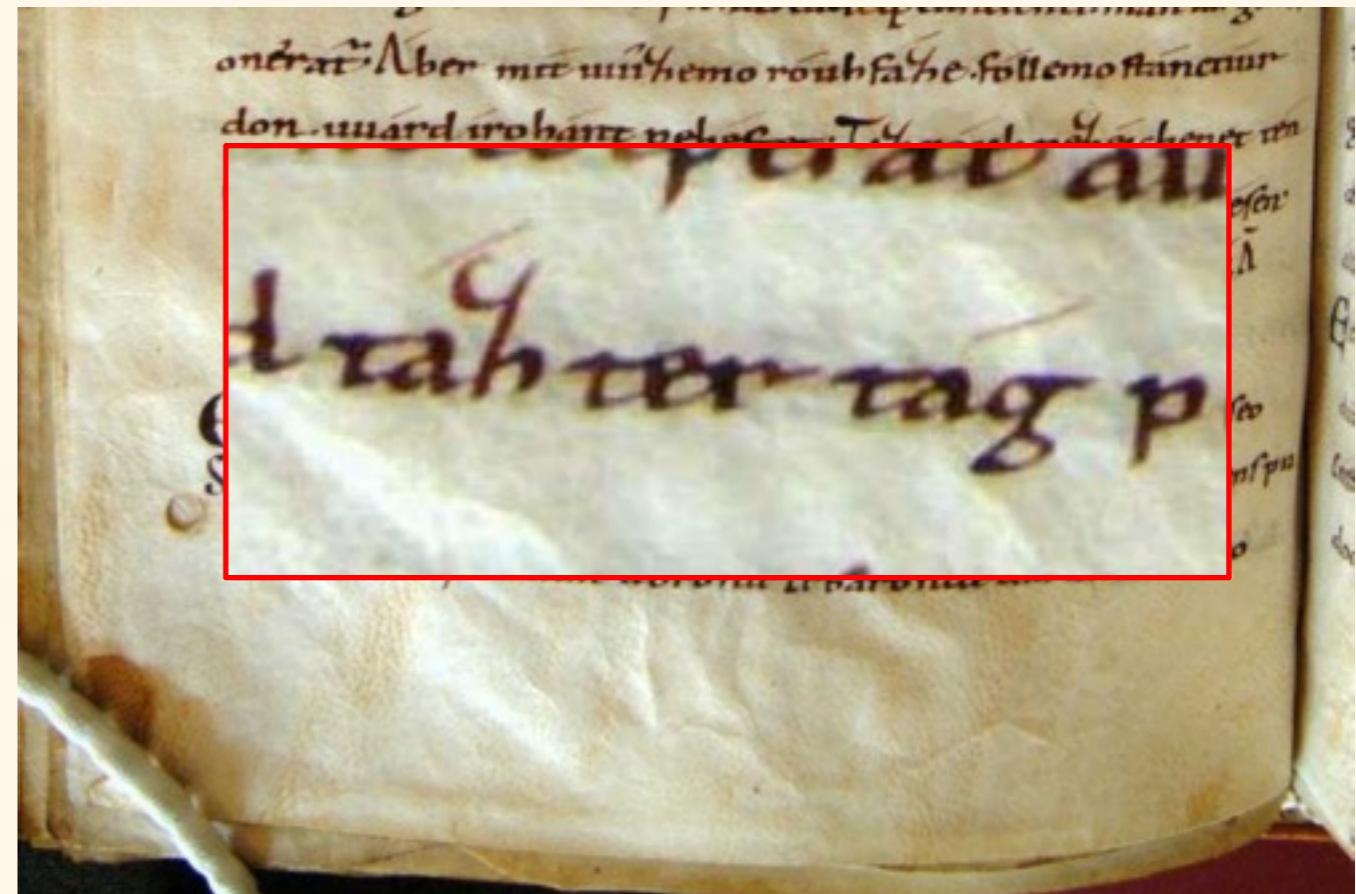
Sô uuárd táz te**r** **t**ág pegónda
décchen die stérnen

Asymmetry in word initial consonants

Letters **p t k/c** word-initially after an plosives, e.g. /p, t, k, h/
b d g word-initially after sonorants, e.g. /*vowels*, n, m, l, r/

Letters **b d g** word-finally (*There was no final devoicing*!)

Some words with initial **t** regardless of context - after sonorants and plosives

What was the phonological/phonetic contrast? Why did < **t** > differ? How were these words which did <u>not</u> alternate (such as *tag*) differ from those that did, such as *tas~das*.
Conclusion: The alternation and the contrast was in quantity not of voicing!

p **t** **k**

tt
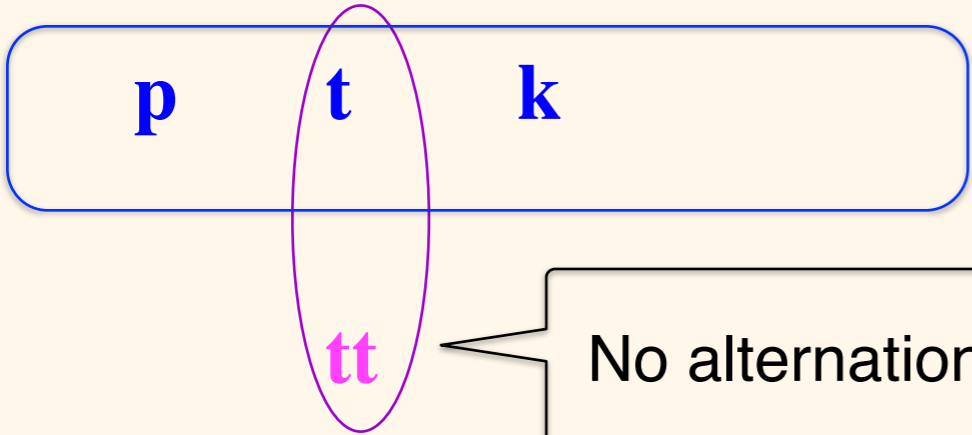
No alternation!

These consonants alternate word initially - quantity alternation.

**Word initial contrast CORONAL asymmetry —
restricted to t : tt**

Change from West Germanic to Notker's phonological system

West Germanic obstruent **phonemes** (ancestor of Old English & Old High German)

	LABIAL	CORONAL	DORSAL		
geminate	-pp	-tt	-kk	*voiceless*	PLOSIVES
singleton	p	t	k		

	LABIAL	CORONAL	DORSAL	
	b	d	g	*voi*
	-bb	-dd	-gg	
		þ		

FRICATIVE

Do we have further evidence? YES - compare the system a 1000 years later!

Pre-Old High German system

	LABIAL	CORONAL	DORSAL
	pf	ts	kx
	p	t	k
	-pp	-tt	-kk
		t	

CLASH

NOTKER's system resolves clash

	LABIAL	CORONAL	DORSAL
	pf	ts	kx
	p	**tt**	k
	-pp	-tt	-kk
		t	

Comparing Notker's words with modern Swiss German

Notker

Swiss German

predictable

$\left\{\begin{array}{l}\end{array}\right.$

<b ~ p>   p

<g ~ k>   k

<d ~ t>   t

non-alternating t   **tː**

p-   t-   k-

tː-

LABIAL CORONAL DORSAL

→

p-   t-   k-

pː-   tː-   kː-
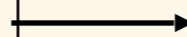
LABIAL CORONAL DORSAL

Modern Swiss German has increased the singleton-geminate contrast in all places of articulation! How?

Initial geminates in Swiss German Alemannic (dialect of Thurgau) : Word initial singleton-geminate contrast

p	t	k
pː	tː	kː

a) /pː/ – /p/: /t͡svai **pː**aaʁ/ 'two pairs' /k͡xai **p**aaʁ/ 'no bar'

b) /tː/ – /t/: /k͡xain **tː**ankk/ 'no tank' /k͡xain **t**ank/ 'no thanks'

c) /kː/ – /k/: /k͡xai **kː**ɔttlɛtt/ 'no cutlet' /k͡xai **k**ɔttə/ 'no godmother'

Why did the asymmetric contrast system change?

Incorporating loans & extending the geminate-singelton contrast

Proto Germanic	Middle Dutch	Latin/ Old French	Italian	English	Swiss German	
/**b**/*ôk-s*					/**p**uǝx/	
	/**b**/*auma*				/**p**omm/	*Inherited*
				/**b**/*rother*	/**p**ruder/	
		/**p**/*air,* /**p**/*ar*			/**p**ːaar/	
			/**p**/*izza*		/**p**ːit͡sa/	*Loans*
				/**p**/*ullover*	/**p**ːuli/	

- **The consonantal quantity contrast existed only in word initial coronals in Notker's dialect - the system was asymmetric**

- **Later generations extended this contrast to other places of articulation to incorporate a voicing contrast in loans which did not exist in the native dialect.**

- **This had the effect of enhancing a '*nonsesuch*' contrast in the language**

Geminates come and (very reluctantly) go

- Sanskrit, Apabhramsa, Pali and indeed all older Indo Aryan languages all had medial geminates

- Old English, Old High German, Old Norse also had medial geminates

- Modern West Germanic languages have largely lost them: Swiss German being a major exception

- *but* perhaps English has reluctantly kept a few

- Indo-Aryan geminates blossomed!

/j/-assimilation (like Germanic); /kʂ/ > /kʰː/; /dm/ > /dː/

OE, OHG	be**d**-**j**-es	>	be**dː**es	'bed-NOM.PLURAL'
Skt	sa**t**-**j**-a	>	ʃo**tː**o	'truth'
Skt	ra**k**-**ʂ**-as	>	ra**kʰː**oʃ	'monster'
Skt	pa**d**-**m**-a	>	pɔ**dː**o	'lotus'

Geminates come and (very reluctantly) go

/r/-Assimilation: CORONAL **consonants assimilate to a preceding rhotic**

across morphemes

ko**r**-**tʃʰ**-i >   ma**tʃʰ**ː-i   'do-PROGRESSIVE.PRESENT-1P'

po**r**-**t**-am >   po**t**ːam   'read-HABITUAL PAST-1P'

across words

bɔro dada   >   bɔ**r**ø **d**ad̶a̶   >   bɔ**d**ːa   'older brother'

ghɔ**r** **dʒ**amai >   ghɔ**dʒ**ːamai   'stay-at-home son-in-law'

Concatenation

kʰu**l**-**l**-o   >   kʰu**l**ːo 'open-SIMPLE PAST-3P'

bʰa**b**-**b**-o >   bʰa**b**ːo 'think-FUTURE-1P'

pa**t**-**t**-am >   pa**t**ːam 'lay down-HABITUAL PAST-1P'

pũ**tʃ**-**tʃʰ**-i >   pũ**tʃʰ**ː-i 'wipe-PROGRESSIVE.PRESENT-1P'
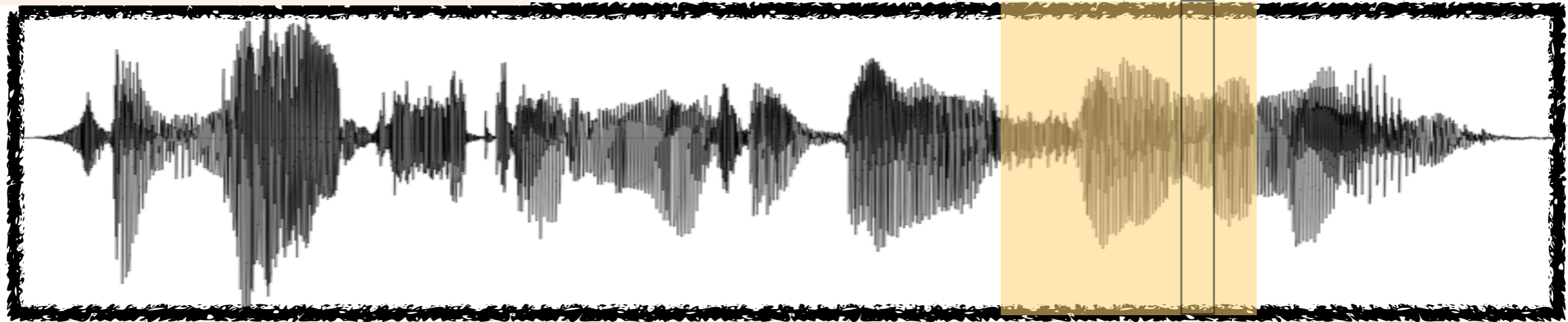

Loss of geminates occurs ONLY when there are metrical constraints (complex foot structure, constraint against trimoraic syllables, syllable edges). Nevertheless, they go reluctantly!
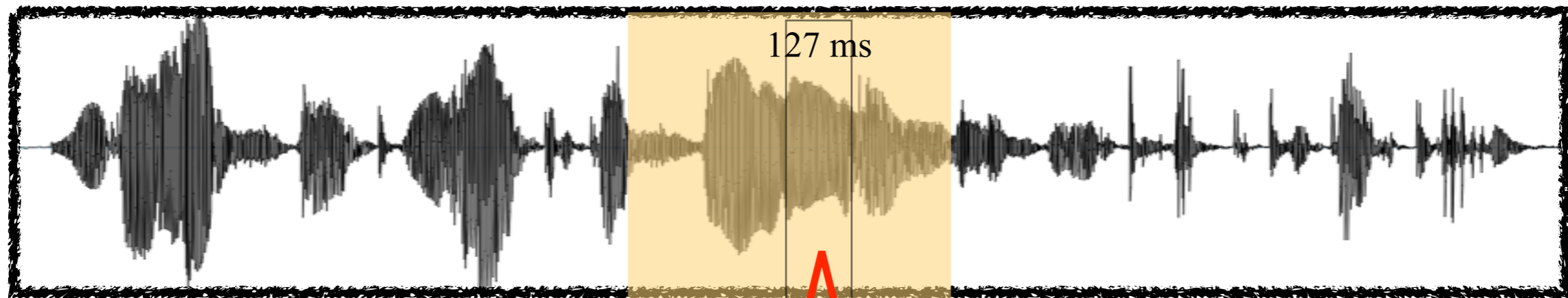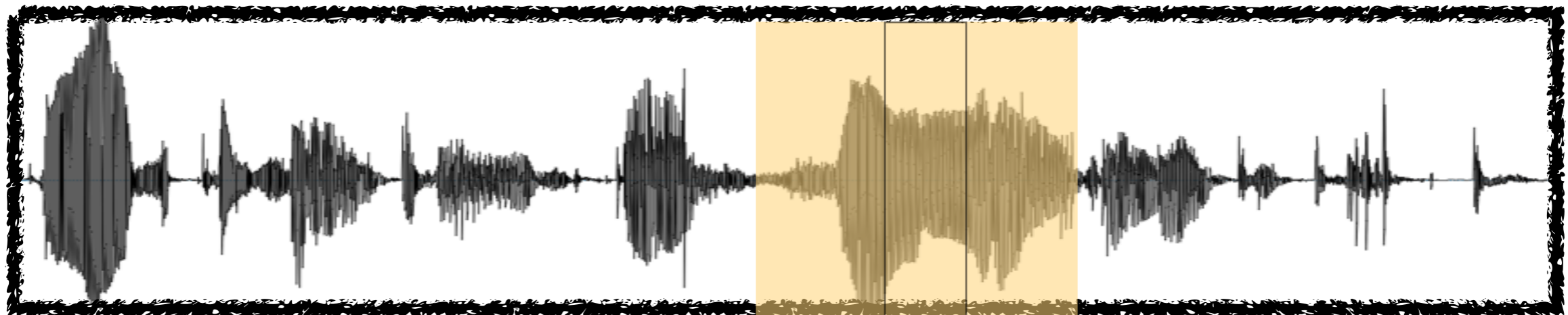
ENGLISH: *hole, whole*

St Wulfstan was a very **ho**ly man.

50 ms

These facilities are wholly inadequate. {*who**le - ly***}

120 ms

127 ms

Winifred read the who**le lea**se from top to bottom.   [hoʊl = liː]

Bengali consonantal inventory

OBSTRUENTS

LABIAL	CORONAL LOW	CORONAL HIGH	CORONAL STRIDENT	DORSAL
p, b, pʰ, bʰ	t, tʰ, d, dʰ	ʈ, ʈʰ, ɖ, ɖʰ	tʃ, tʃʰ, dʒ, dʒʰ	k, kʰ, g, gʰ
pː, bː, pʰː, bʰː	tː, tʰː, dː, dʰː	ʈː tʰː, ɖː, ɖʰː	tʃː, tʃʰː, dʒː, dʒʰː	kː, kʰː, gː, gʰː
(s)		ʃ, ʃː		

SONORANTS

LABIAL	CORONAL	CORONAL	DORSAL
m, mː	n, nː	l, lː	ŋ, ŋː

Bengali Geminates

Singletons & Geminates

- Underlying geminates are represented by a single set of features and a single release
- Medial geminates are part of two syllables
- Never treated as two separate entities which undergo separate phonological processes

```
   σ  σ            σ      σ
 / | \ / |        /|     / |
  p a t: a       pa .  t  a
```

Acoustic Cues:

- The predominant acoustic cue for gemination is consonant (closure) duration (e.g. Hankamer et al. 1989, Ridouane 2010)
- Neither differences in the preceding vowel nor release properties reliably distinguish geminates from singletons

Expressions of short/long contrasts in languages

- Languages do not have monomorphemic words like "little long", "a bit short", etc.

- Once geminates disappear, there is only one set of consonants left and they are treated as metrically short - no language has only geminates

Duration is obviously variable

- How long does the closure duration have to be, to be perceived as a long consonant? How short does it have to be to be short?

- To what extent do speakers tolerate durational changes in words in a language with a geminate/singleton contrast?

- If the segmental information is accurate, is mispronunciation in duration tolerated?

Possible hypotheses:

(A) No mispronunciations with durational changes are accepted

(B) All mispronunciations are accepted if only durational information is changed

(C) Durational contrasts are asymmetric - i.e. 'long' vs. 'short' are not the same; the asymmetry is a consequence of the representation

Expressions of short/long contrasts in languages

- Languages do not have monomorphemic words like "little long", "a bit short", etc.
- Once geminates disappear, there is only one set of consonants left and they are treated as metrically short - no language has only geminates
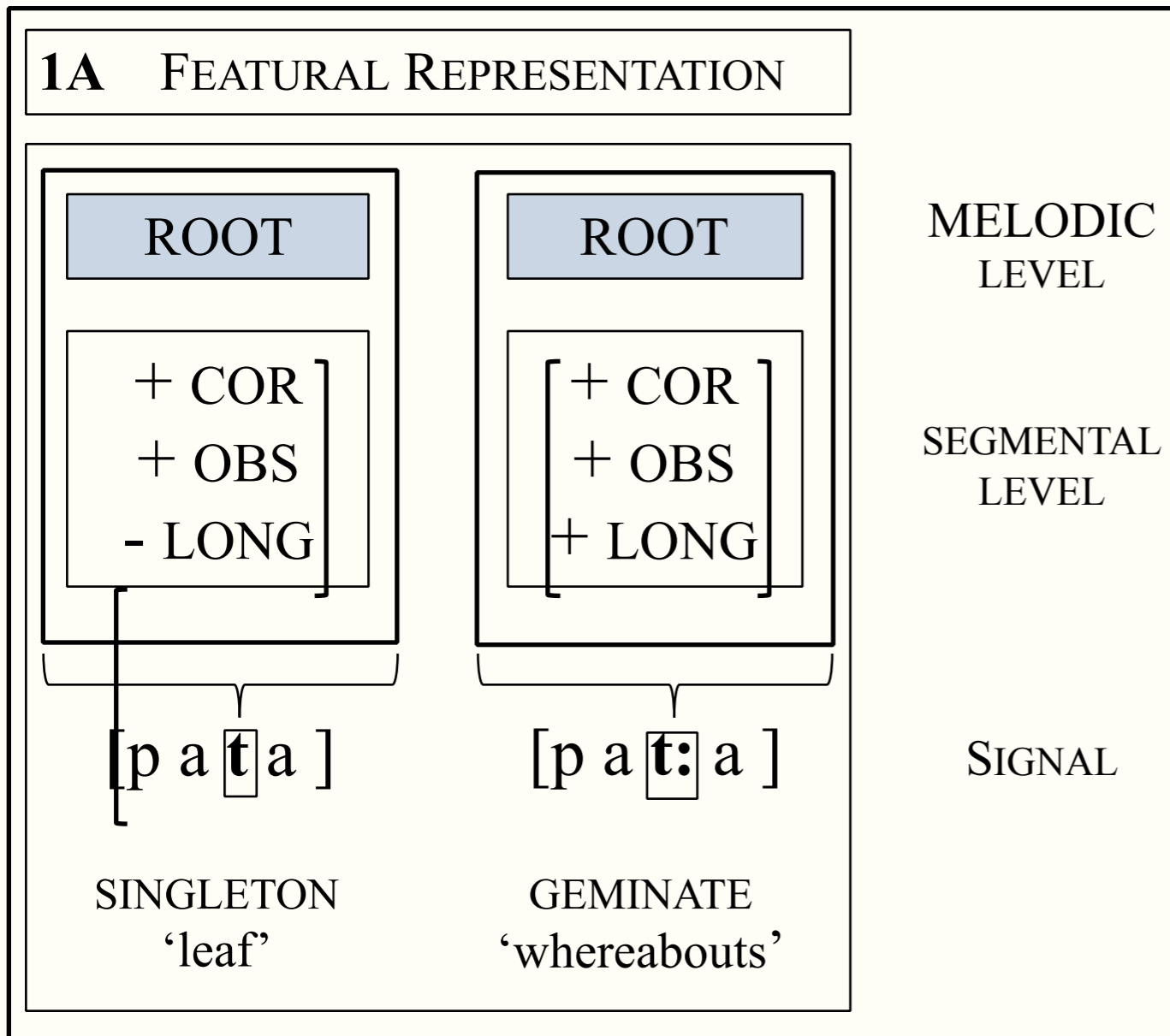
Duration is obviously variable

- How long does the closure duration have to be, to be perceived as a long consonant? How short does it have to be to be short?
- To what extent do speakers tolerate durational changes in words in a language with a geminate/singleton contrast?
- If the segmental information is accurate, is mispronunciation in duration tolerated?

Possible hypotheses:

(A) No mispronunciations with durational changes are accepted

(B) All mispronunciations are accepted if only durational information is changed

(C) **Durational contrasts are asymmetric - i.e. 'long' vs. 'short' are not the same; the asymmetry is a consequence of the representation**

Lexical representation

1A   FEATURAL REPRESENTATION

ROOT	ROOT	MELODIC LEVEL
$\begin{bmatrix} + \text{COR} \\ + \text{OBS} \\ - \text{LONG} \end{bmatrix}$	$\begin{bmatrix} + \text{COR} \\ + \text{OBS} \\ + \text{LONG} \end{bmatrix}$	SEGMENTAL LEVEL
[p a t a]	[p a t: a]	SIGNAL
SINGLETON 'leaf'	GEMINATE 'whereabouts'	

Duration differences can be considered to be "cues" to featural properties such as [voice]. If so, perhaps geminates and singletons could also be characterised as [± LONG]?

If the contrast is purely symmetric, and both 'features' are represented, then one might assume a symmetry in access and recognition.

If the representation is asymmetric - i.e. geminates are specified in their representation but singletons are not, then we may predict an asymmetry.

Prediction

Lengthening would be tolerated

Shortening would not be tolerated.

Semantic priming: Latency and Electrophysiological measures

TASK: Lexical Decistion - Is the Target a word or a nonword

Prime
(auditory)

Related
bullet

Unrelated CONTROL
milk

RT 1

RT 2

Reaction Time (RT) to Target GUN is measured.

If RT 1 is **less** than RT2, then GUN has been facilitated by *bullet*.

Target
(semantically related)

GUN

Target
Unrelated nonword

*LUN

DESIGN

Task: Cross-modal Lexical decision with semantic priming

24 singletons and mispronunciations dʰama 'large basket' *dʰamːa

Semantically matched targets বুড়ি dʒʰuri   'forgiveness'

24 geminates and mispronunciations gramːo 'country person' *gramo
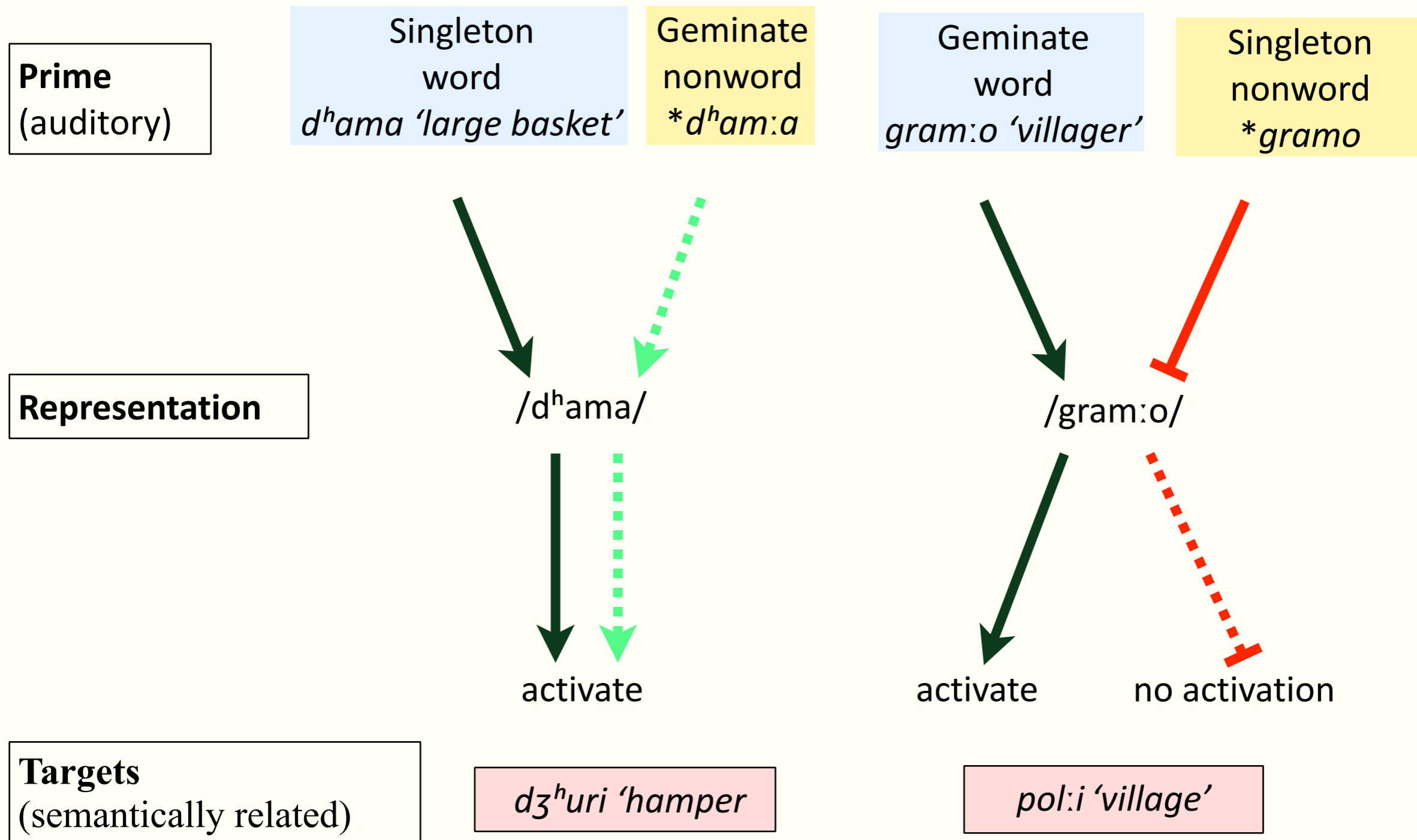
Semantically related target পল্লী polːi   'village'

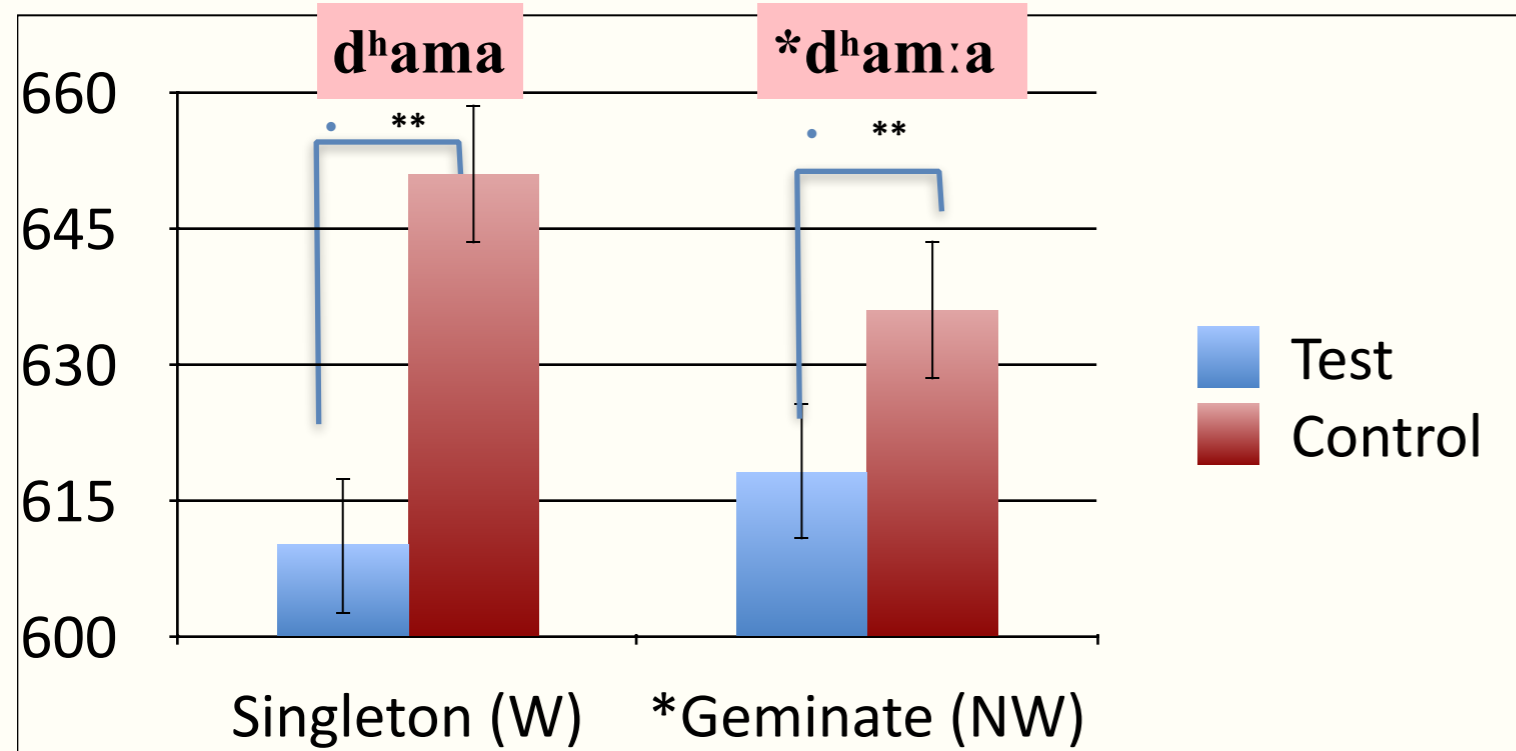Equal numbers of fillers, words & nonwords

Prime-Target semantic relationships, familiarity, frequency of usage ratings - all checked by independent questionnaires.

Average length for singleton (**89ms**) & geminate (**207ms**)

Geminate/Singleton mispronunciation task

Prime (auditory)

| Singleton word *dʰama 'large basket'* | Geminate nonword **dʰamːa* | Geminate word *gramːo 'villager'* | Singleton nonword **gramo* |

Representation

/dʰama/

/gramːo/

activate

activate   no activation

Targets (semantically related)

dʒʰuri 'hamper

polːi 'village'

Semantic Priming: Geminates vs. Singletons

dʰama   ***dʰamːa***

660
645
630
615
600

Singleton (W)   *Geminate (NW)

** **

Test
Control

SHORT > *LONG

dʰama > *dʰamːa —> dʒʰuri

Same amount of facilitation for both singleton (W) and geminate (NW) primes

➔ Geminate (NW) prime leads to lexical access

---

gramːo   ***gramo***

630
618
605
593
580

Geminate (W)   *Singleton (NW)

** **

Test
Control
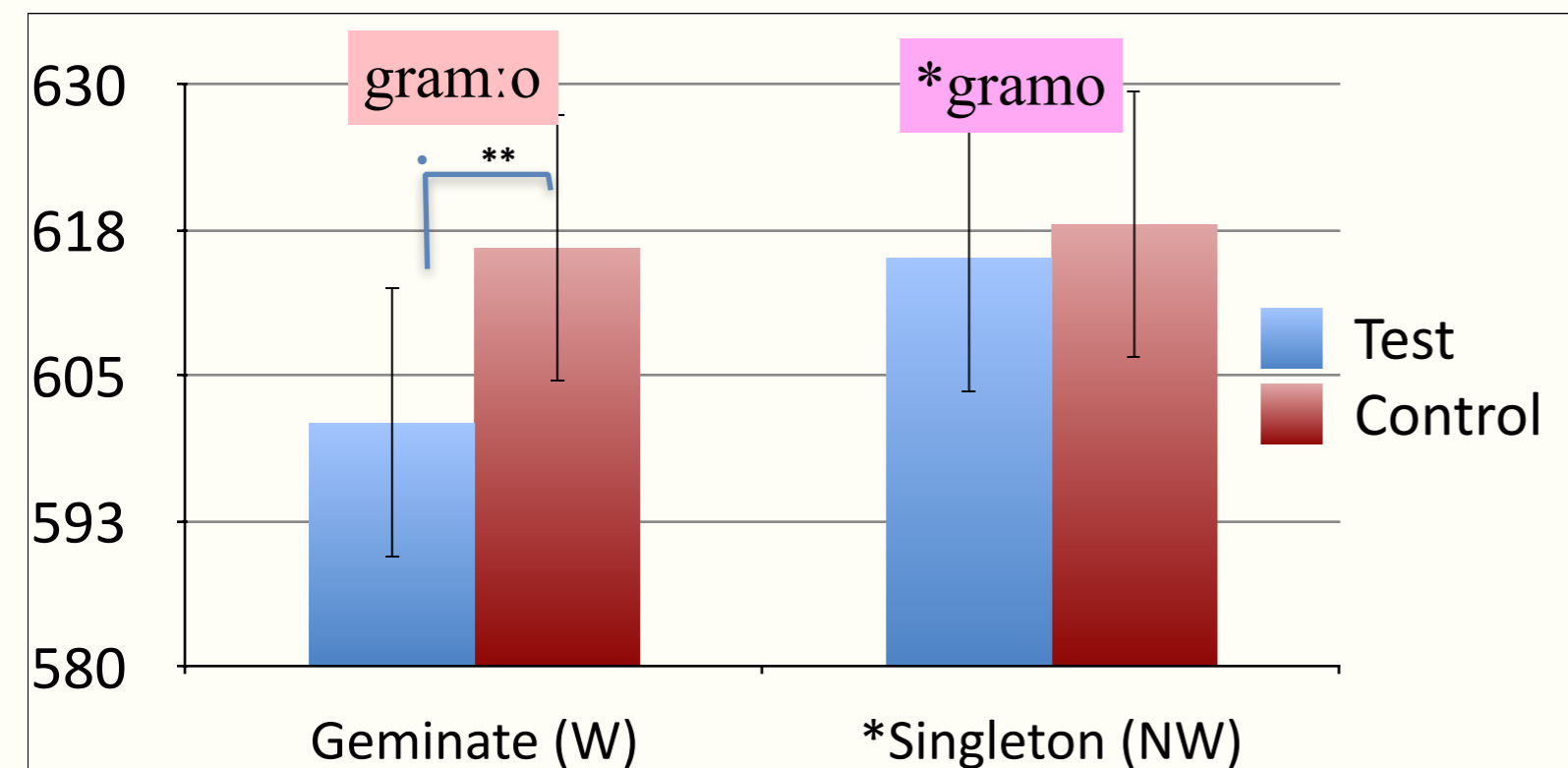
LONG > *SHORT

gramːo > *gramo —> polːi

Facilitation effect only for geminate (W) primes

➔ Singleton (NW) prime **does not** lead to lexical access

Kotzor, Wetterlin, Roberts & Lahiri, *Language & Speech 2015*

Semantic Priming: Geminates vs. Singletons
Event Related Potentials

Event related potentials are a direct measure of brain activity.

The most obvious component to look for is the N400 which is involved in semantic integration.
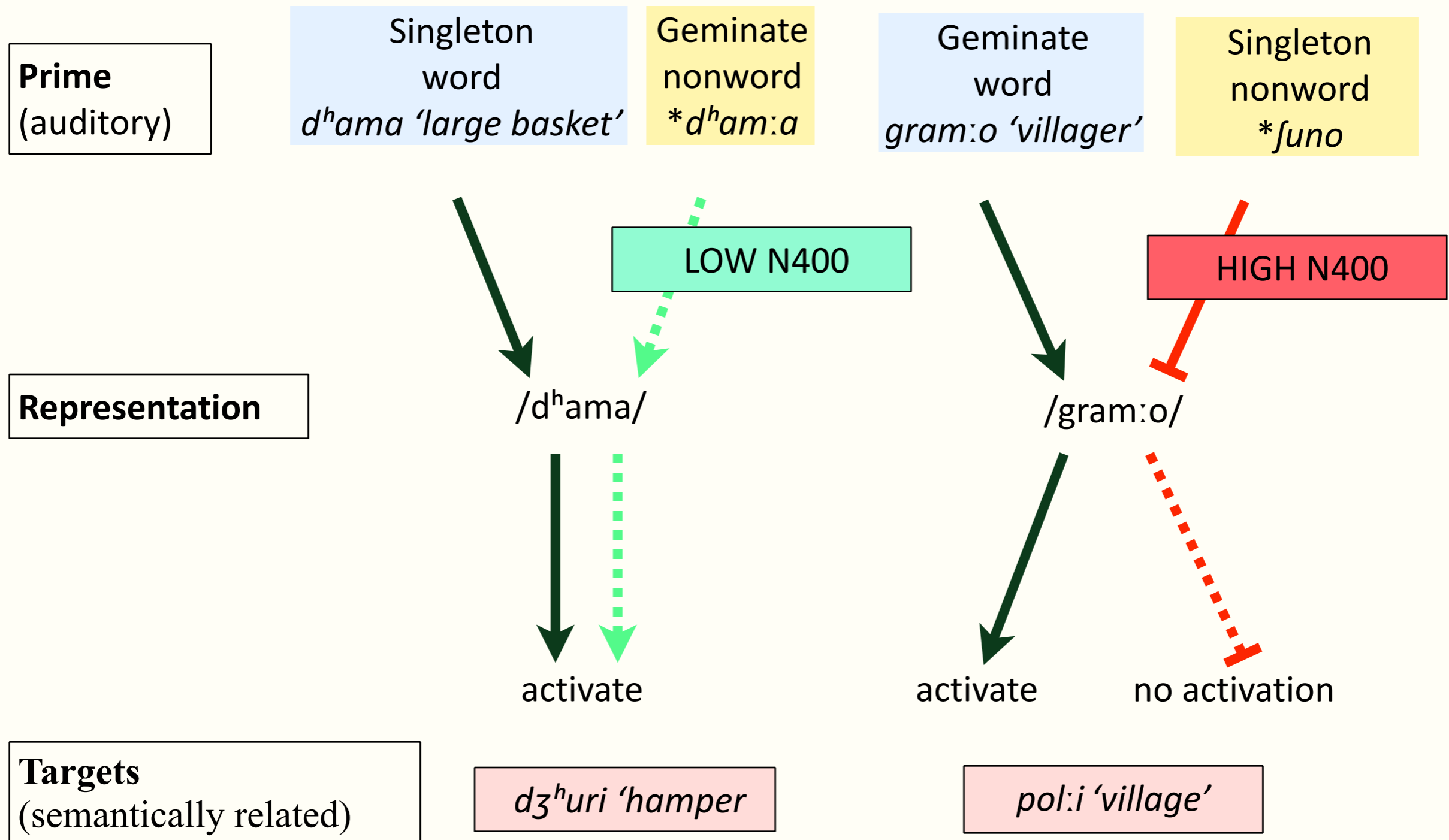
If a mispronunciation is successful in lexical integration, we expect a low N400.

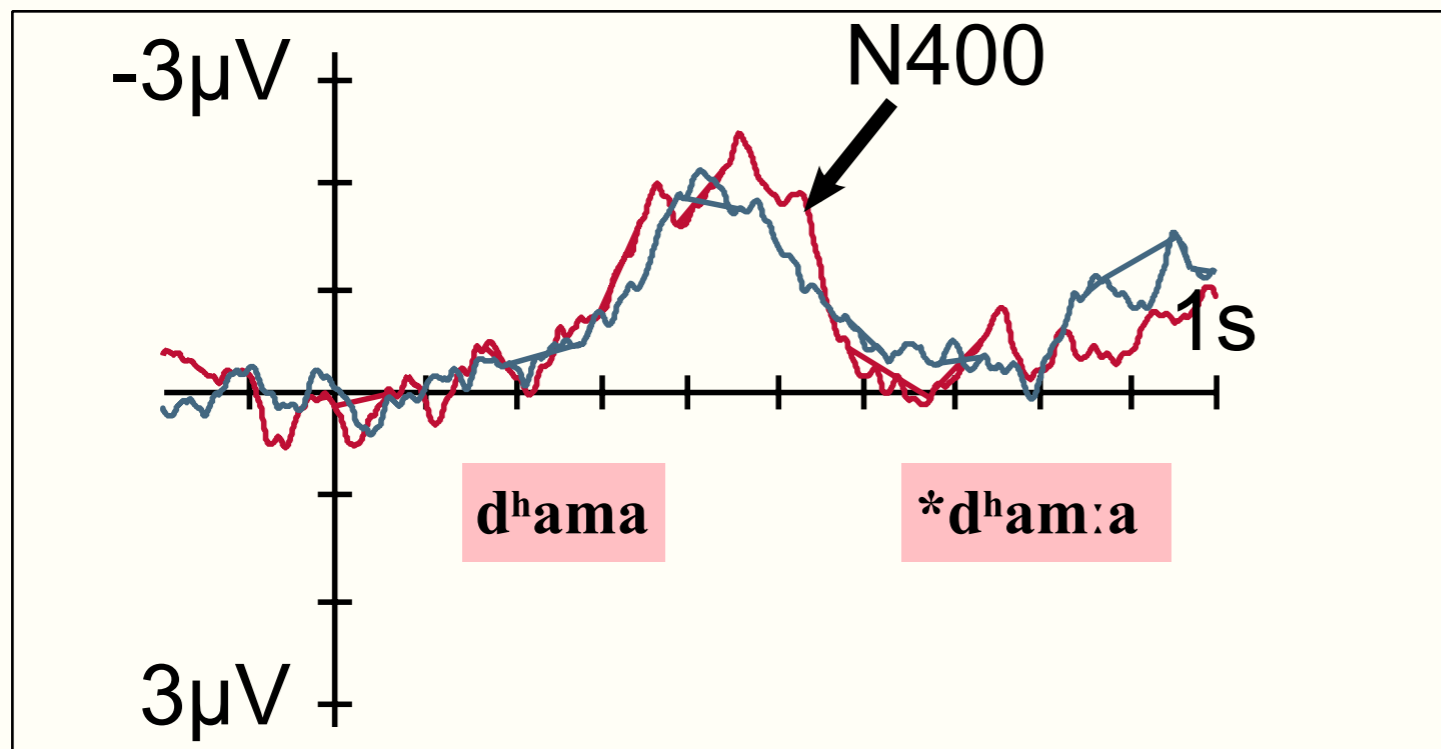If a mispronunciation is NOT accepted, we expect a large N400.

The EEG experiments were run with our portable system in Calcutta .

# Semantic Priming: Geminates vs. Singletons
Event Related Potentials

Prime (auditory)

| Singleton word *dʰama 'large basket'* | Geminate nonword *dʰamːa* | Geminate word *gramːo 'villager'* | Singleton nonword *ʃuno* |

LOW N400

HIGH N400

Representation

/dʰama/

/gramːo/

activate

activate   no activation

Targets (semantically related)

dʒʰuri 'hamper   *polːi 'village'*

Semantic Priming: Geminates vs. Singletons
Event Related Potentials

SHORT > *LONG

No difference in N400 response between singleton (W) and geminate (NW) primes

→ Geminate (NW) prime leads to lexical access
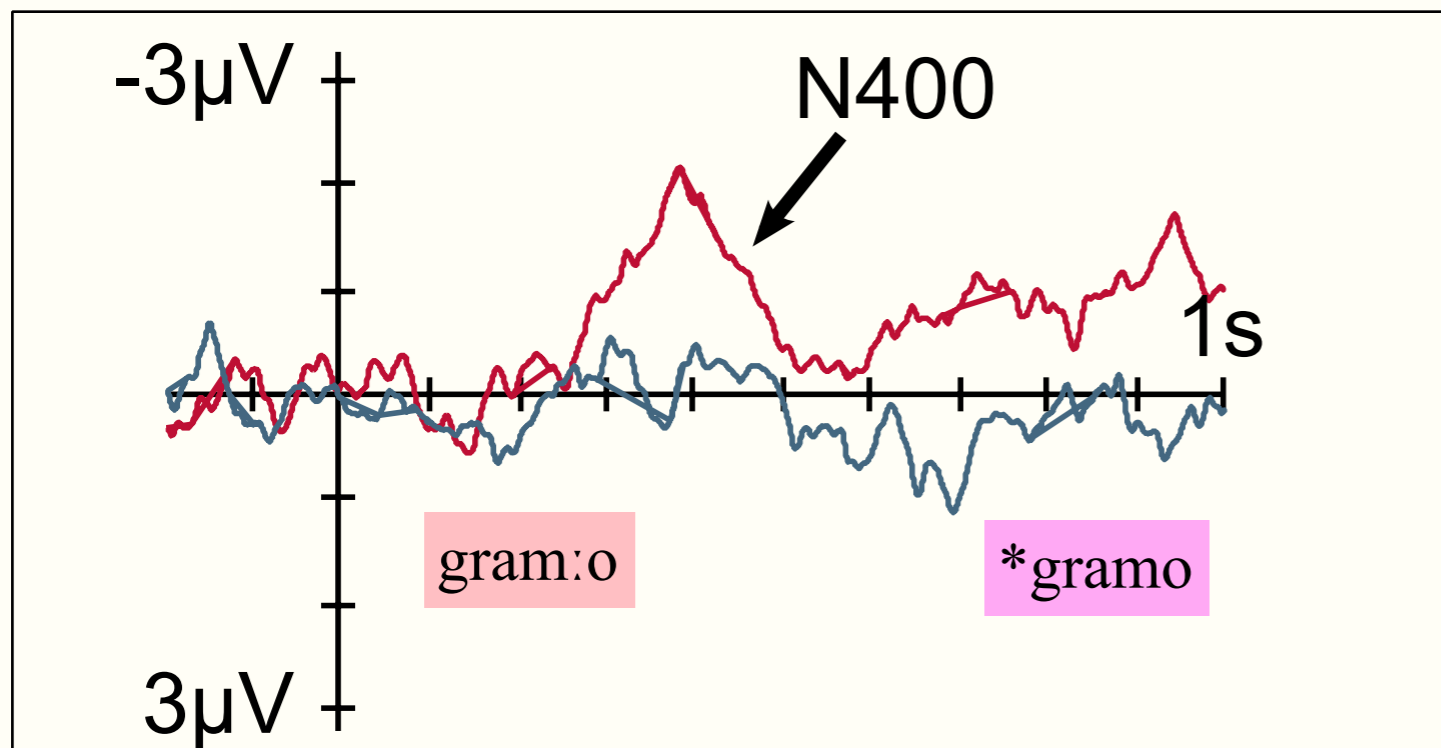
LONG >*SHORT

No N400 for geminate (W) but N400 for singleton (NW) prime

→ Singleton (NW) prime **does not** lead to lexical access

Roberts, Kotzor, Wetterlin & Lahiri (2014) *Neuropsychologia*

In conclusion…

- Gemination is an active process in many langugaes

- Degemination is usually constrained by metrical constraints

- In terms of lexical contrast, our assumption is that only 'long' consonants are specified, giving us an asymmetry in in lexical representations : μ vs. [   ]

- This asymmetry is reminiscent of the featural asymmetries we find which depends on specified and underspecified features.

- When singletons and geminates are manipulated to give the opposite duration, (long-to-short and short-to-long) we find that **lengthening a singleton does not hinder lexical access.**

- However, **shortening a specified geminate, blocks lexical access.**

- The evidence comes from reaction time latencies as well as from electrophysiological measure.

If geminates are considered as nonesuches, they are quite nice ones!
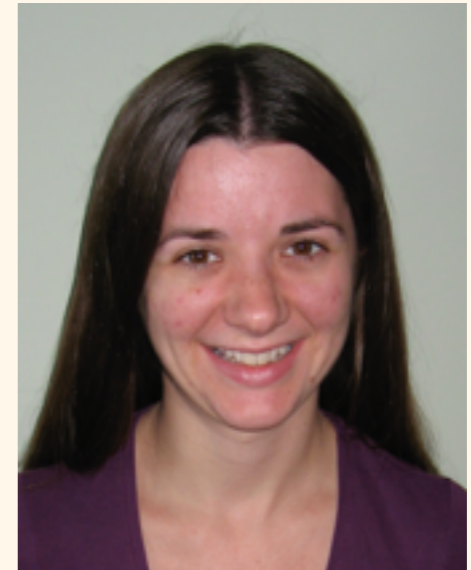
And we would also like to thank:

Astrid Kraehenmann

Allison Wetterlin

Sandra Kotzor

Holly Kennard

Research Collaborators

Adam Roberts

Henning Reetz