Modeling Emergence in Phonological Space

Rebecca L. Morley

Penn Linguistics Conference
Interplay between Linguistics and Biology Panel
March 22, 2019

Biology & Linguistics: Transmission & Change

- Two kinds of biological thinking in linguistics
  - Evolutionary model
  - Cognitive Model : Language is part of general cognition, subserved by (auditory) perception
- Two kinds of sound change
  - Genesis of a new phoneme (loss of an allophone)
  - Change in the realization of an existing phoneme

Evolutionary Phonology

Synchronic phonological alternations mirror diachronic changes

[Blevins(2004); Ohala(1971, 1974,1981,1990,1993,etc.)]

"Recurrent synchronic patterns have their origins in recurrent phonetically motivated sound change." (Blevins; p.8)

"…[they] can be shown to emerge naturally from the imprecise transmission of language across generations." (Blevins; p.18)

"This process of transmission involves a speaker providing input to a listener, with the listener attempting to internalize the speaker's grammar in order to understand speech. The process of transmission takes place in a sea of noise and starts from a point where the human infant listener has no knowledge of any sound-meaning associations in the speaker's language." (Blevins; p.31)

Evolutionary Phonology

Synchronic phonological alternations mirror diachronic changes

[Blevins(2004); Ohala(1971, 1974,1981,1990,1993,etc.)]

	Synchronic	Diachronic
Nasal assimilation:	$/\mathrm{n}/\rightarrow [\mathrm{m}]/\_[\mathrm{labial}]$	$\mathrm{n} > \mathrm{m}/\_[\mathrm{labial}]$
Final devoicing:	$/\mathrm{d}/\rightarrow [\mathrm{t}]/\_\#$	$\mathrm{d} > \mathrm{t}/\_\#$
Vowel nasalization:	$/\mathrm{V}/\rightarrow [\tilde{V}]/\mathrm{N}\_\#$	$\mathrm{VN} > [\tilde{V}]$

Evolutionary Phonology

Synchronic phonological alternations mirror diachronic changes

[Blevins(2004); Ohala(1971, 1974,1981,1990,1993,etc.)]

	Synchronic	Diachronic
Nasal assimilation:	/n/→ [m]/_[labial]	n > m/_[labial]
Final devoicing:	/d/→ [t]/_#	d > t/_#
Vowel nasalization:	/V/→ [$\tilde{V}$]/N_#	VN > [$\tilde{V}$]

And phonological patterns mirror gradient, variable phenomena that occur in fast speech, casual register, high frequency word, etc.

Exemplar Theory:

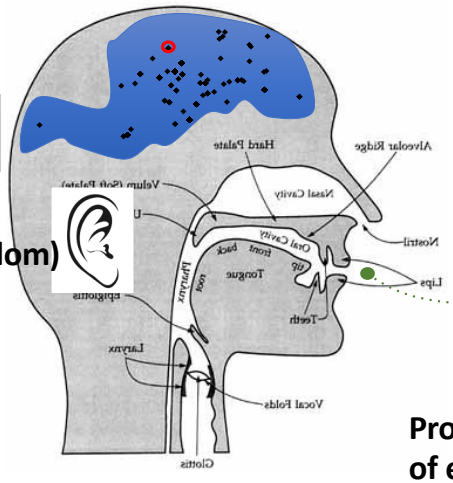
Goldinger(1996);Bybee(2001);Pierrehumbert (2001)

Wedel (2006): "…the data from which a learner abstracts a language system can be understood as a population of variants."

Population of stored experiences

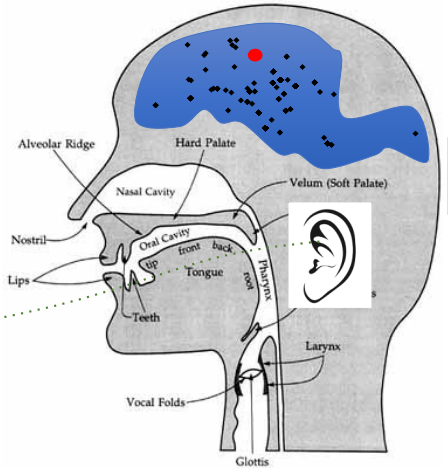
VARIATION

Select stored exemplar (at random) for production



REPRODUCTION

Produce exemplar with some degree of error and/or articulatory bias



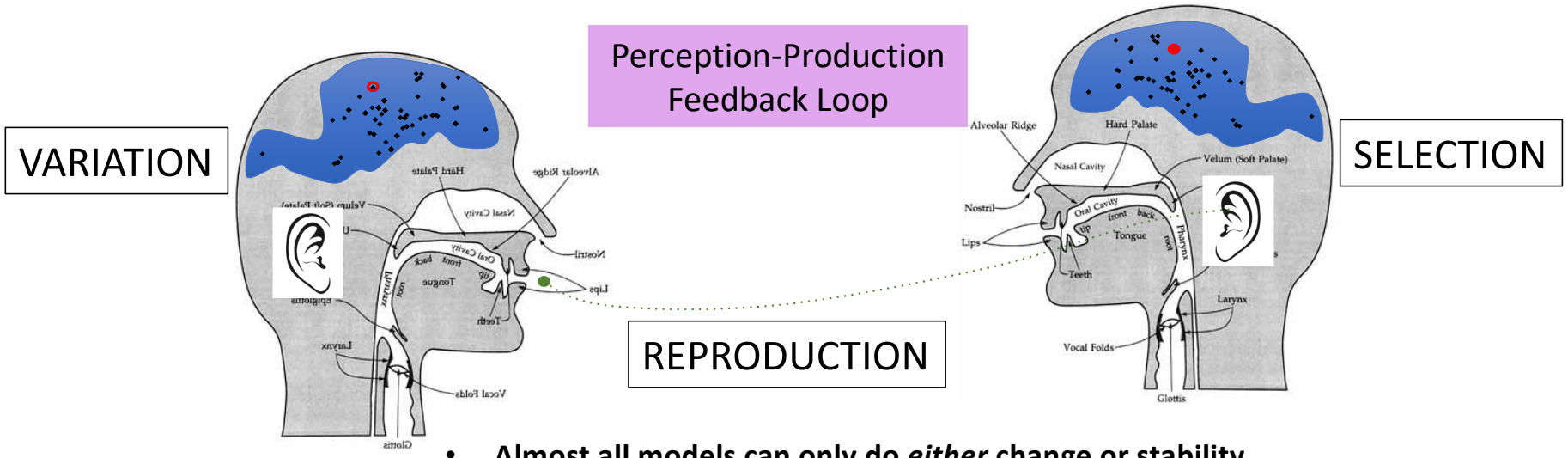
SELECTION

Store perceived exemplar if it is a good enough match

Exemplar Theory:

Goldinger(1996);Bybee(2001);Pierrehumbert (2001)

Wedel (2006): "…the data from which a learner abstracts a language system can be understood as a population of variants."



- **Almost all models can only do *either* change or stability**
- **Perception to production is typically treated as an identity mapping**
- **Cumulativity results in collapse**

(Pierrehumbert 2001;Wedel 2006;2007; Blevins & Wedel 2009; Garrett & Johnson 2013; Kirby 2014; Tupper 2014)

Modeling Emergence

- Rather than hard-wired constraints on learnability, or expressability
  - Incremental, on-line changes
  - Arising from synchronic variation
- Universality is produced by same set of forces acting on all languages
  - Physical articulators
  - Perceptual system
  - General cognition: memory, learning
  - Social/cultural systems

- Add source of change in the translation from perception to production (Evolutionary Phonology)
- Take phonological structure to originate in the speech signal
- And speech processing to consist of
  - Simultaneous segmentation and mapping onto abstract categories
  - Using all available cues (predictive material)
  - Selecting the candidate with the highest probability
  - Evaluated relative to the other available mappings (as opposed to a category prototype)
- And “change” to be a property of of the distribution as a whole, rather than an individual token

Sound Change I:
Co-articulation based phoneme genesis

Chinantec

haa	'so,such'
hãã	'foam,froth'

Phonemic vowel nasalization:
Unpredictable sound difference that carries a meaning difference

English

æn	indefinite article
æd	"a paid announcement, as of goods for sale"

Allophonic vowel nasalization:
Predictable sound difference that does *not* carry a meaning difference

Sound Change I:
Co-articulation based phoneme genesis

Traditional Representations

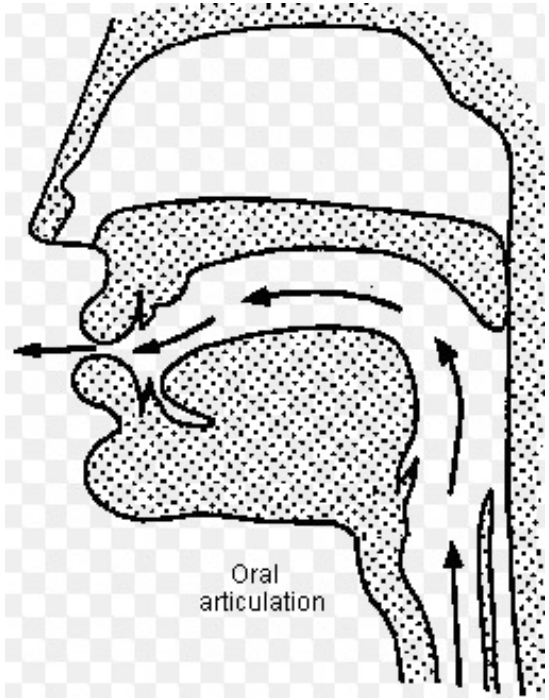
Synchronic		Diachronic
$V + N \rightarrow \tilde{V}N$	N > 0	$\tilde{V}N > \tilde{V}$
V+N		$\tilde{V}$
$\tilde{V}$		V+N
Predictable Allophonic Variants Derived by Rule		Loss of Predictability

Model I

Desiderata:

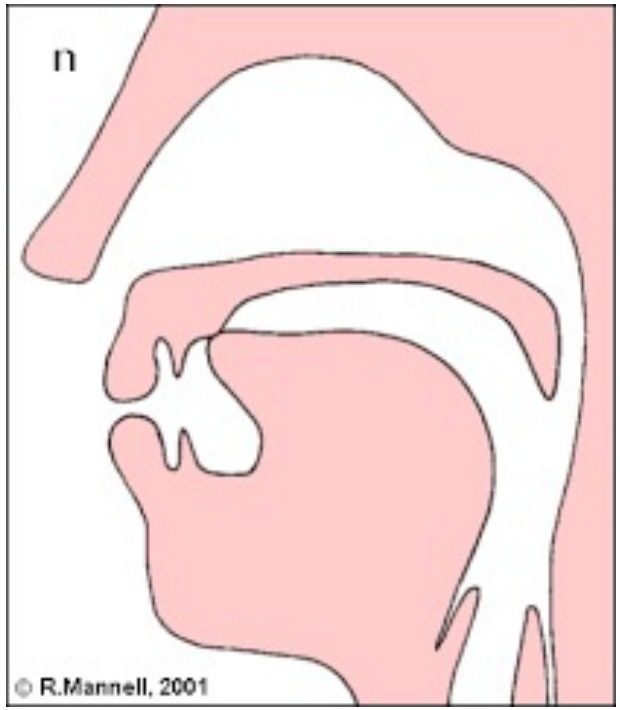
1. Link phonetic source of nasalization with sound change
2. Link loss of nasal consonant with “phonologization” of nasal vowel
3. Explain why nasalization “shifts” to vowel
4. Generate both “change” and “no change” outcomes

1. Coarticulatory Nasalization



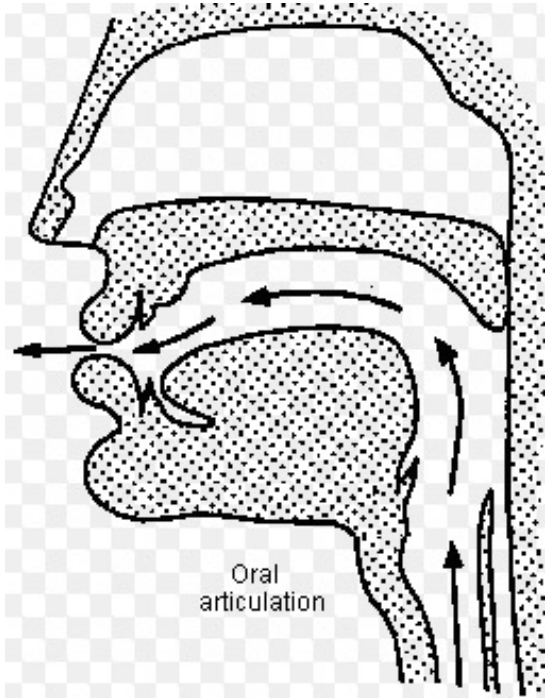
æ

+

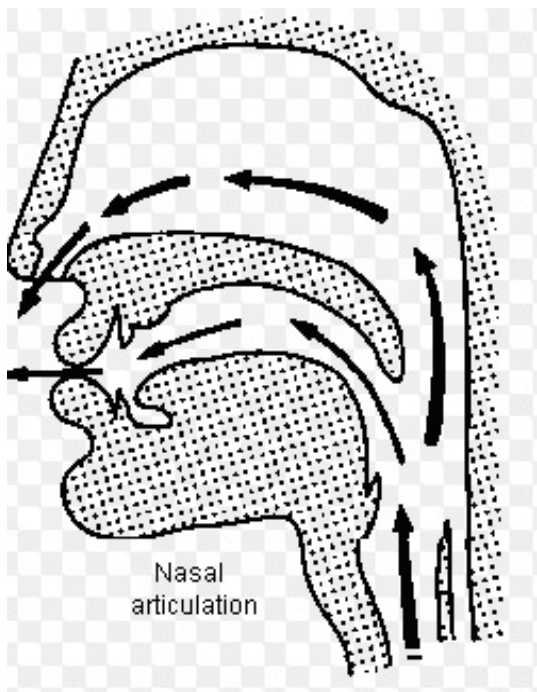


n

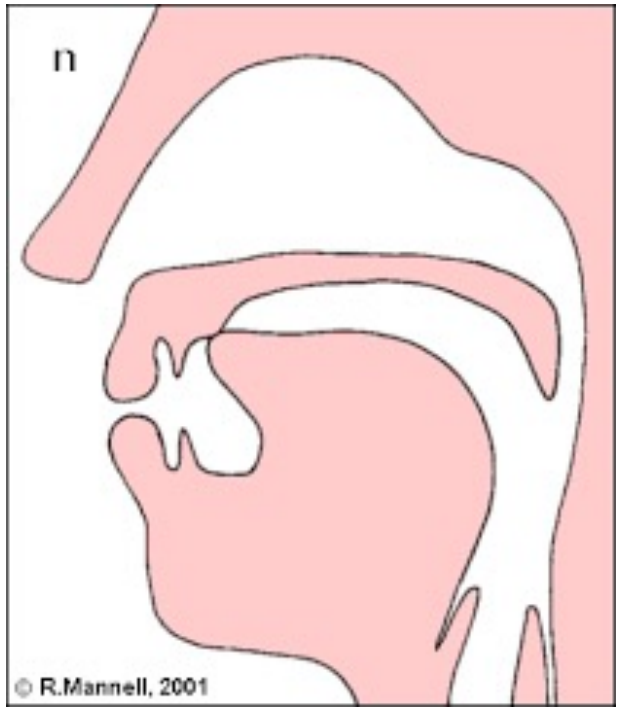
1. Coarticulatory Nasalization



æ

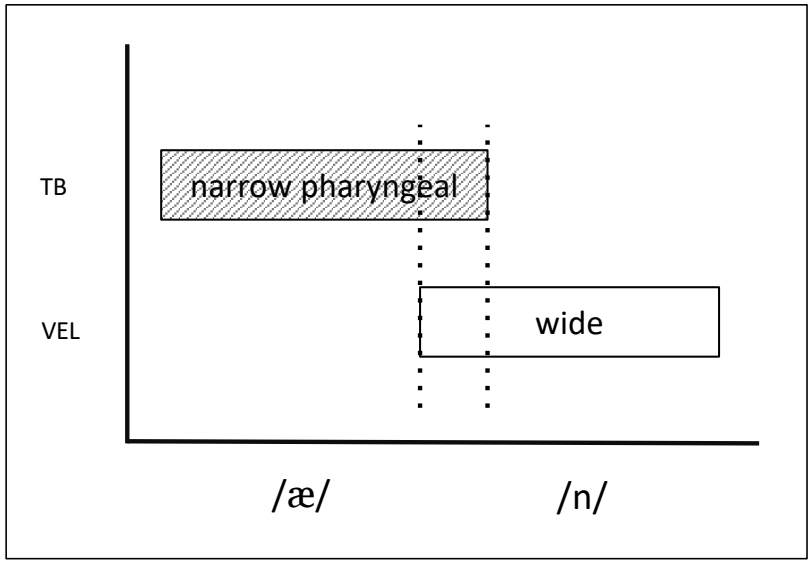


æ̃



n

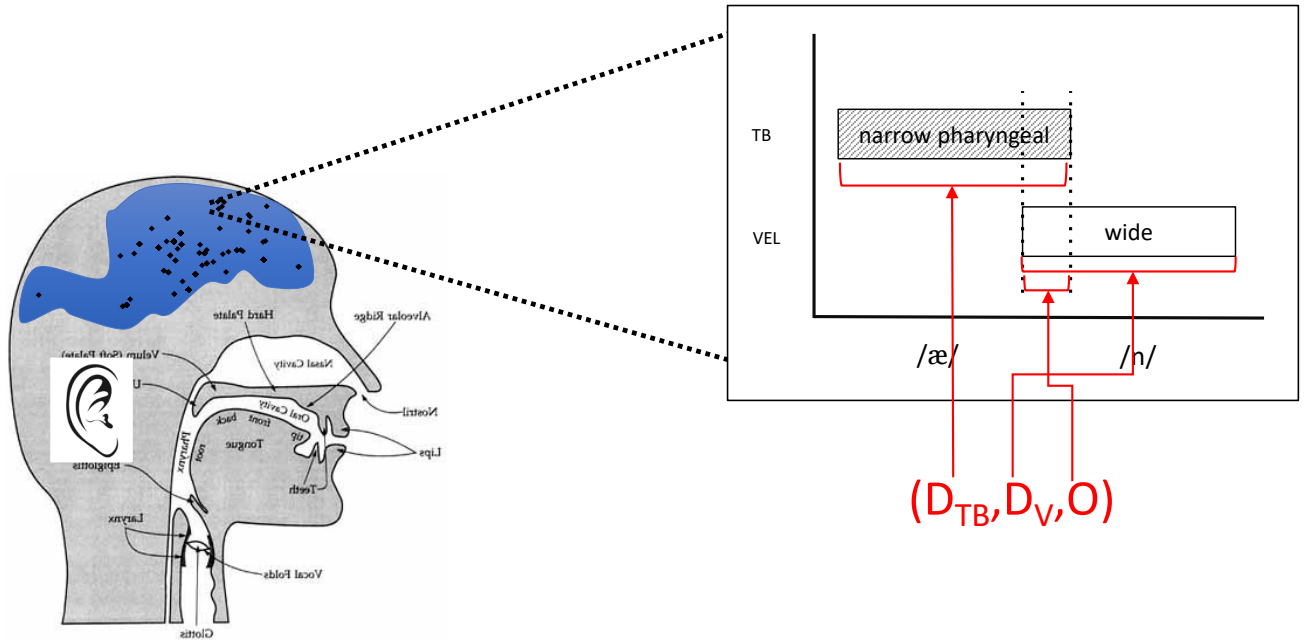
2. Articulatory Phonology (Browman & Goldstein 1989)

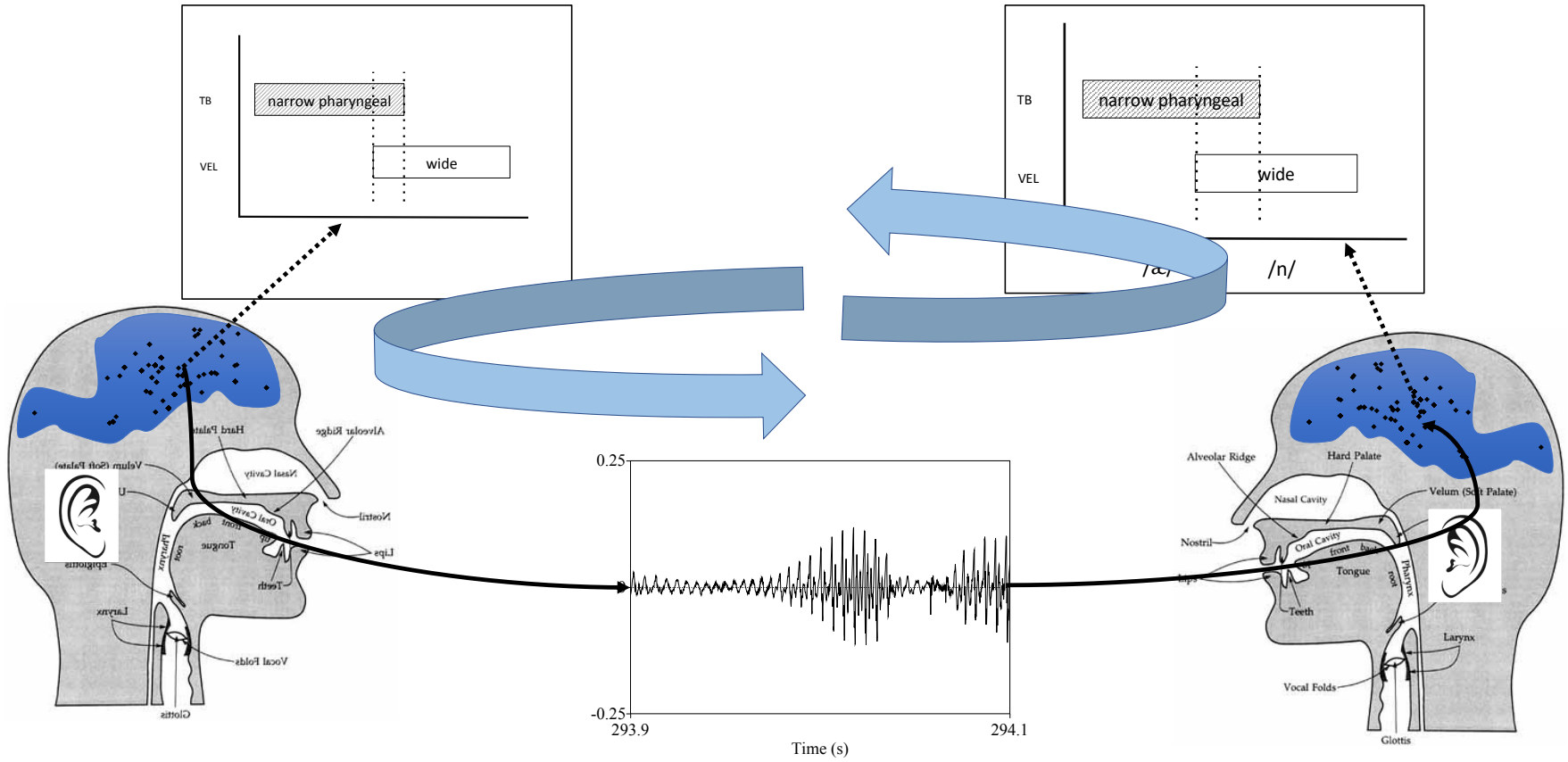


- Velum opening gesture overlapping with tongue body gesture produces nasalization on vowel
- Greater overlap results both in more vowel nasalization and "less" nasal consonant (Beddor 2009)

3a. Representations

Instead of treating nasalization as the product of a rule that applies at production, encode the amount of overlap as part of the representation of each word token

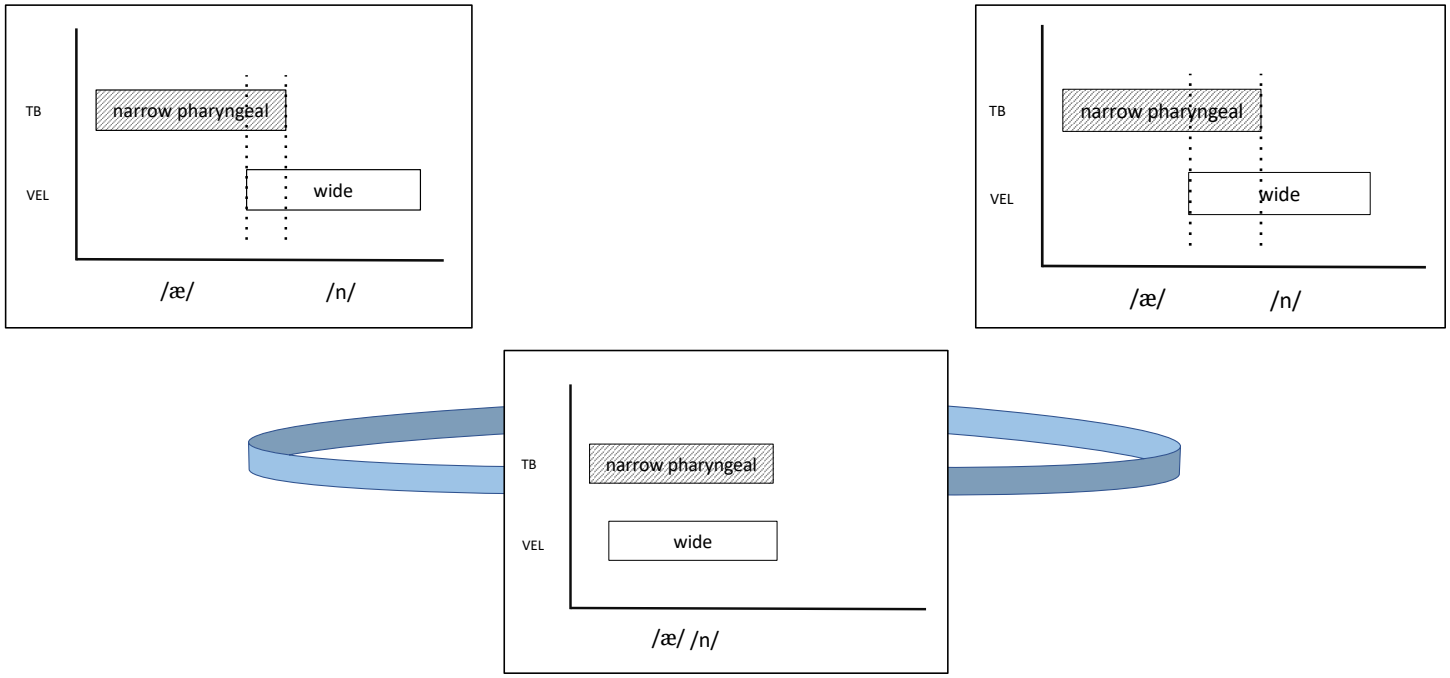




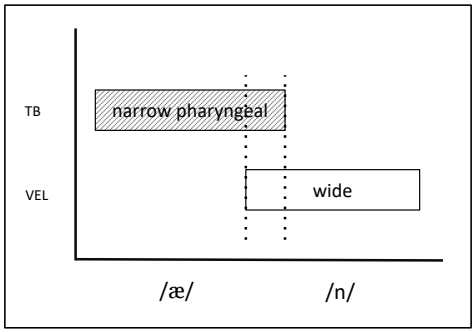
3b. Frequency (Fluency, tighter coordination, efficiency)

(e.g., Soskuthy 2011, 2015)

Model I


TB
narrow pharyngeal
VEL
wide
/æ/
/n/
TB
narrow pharyngeal
VEL
wide
/æ/
/n/
TB
narrow pharyngeal
VEL
wide
/æ/ /n/

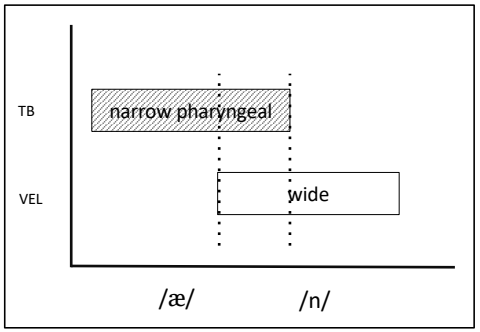
Model I



Gestural overlap increases a certain amount on each production:

$$O = O + \beta(D_V - O)$$

β parameterizes relative frequency



4. Speaking rate

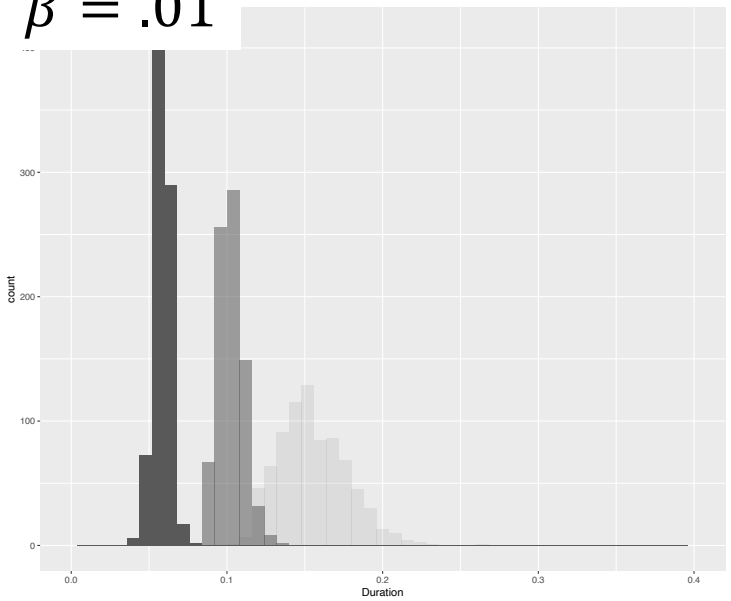
Bi-directional force that disrupts frequency feedback loop

Slowing down:
individual gestures get longer, and less overlapped
Speeding up:
individual gestures get shorter, and more overlapped

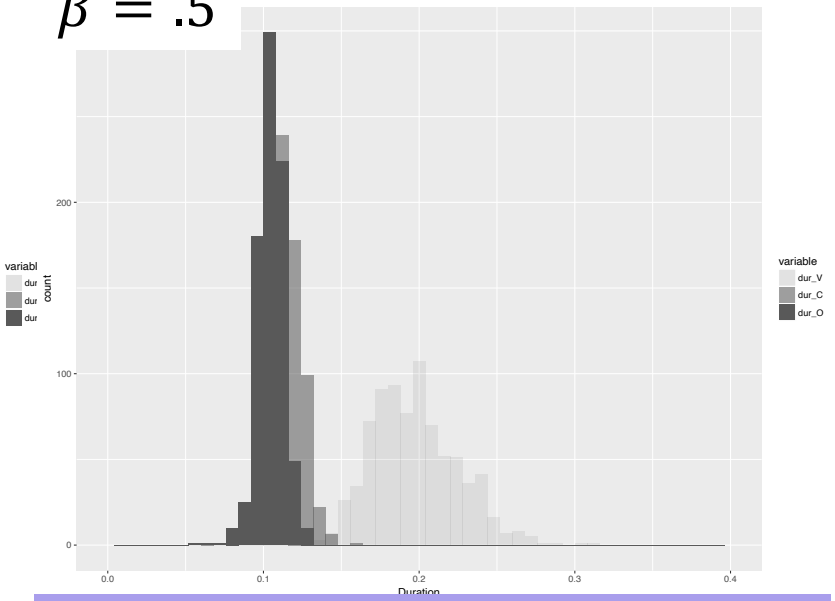
Allows different values of β to lead to different stable states

Emergence of nasal vowels

$\beta = .01$



$\beta = .5$



On average, the duration of the nasal gesture is roughly the same as the duration of overlap

Sound Change 2:
contrastive feature change (transphonologization)

English

$p^hɪt$	'hole or cavity'
pɪt	'mouthpiece of a bridle"

/p/ vs. /p^h/

French

[bɛ̃]	bath
[pɛ̃]	bread

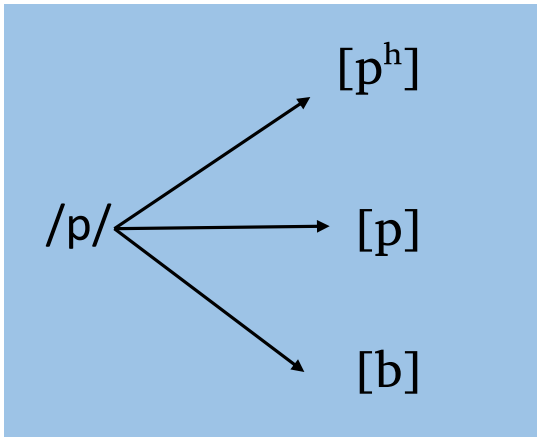
/p/ vs. /b/

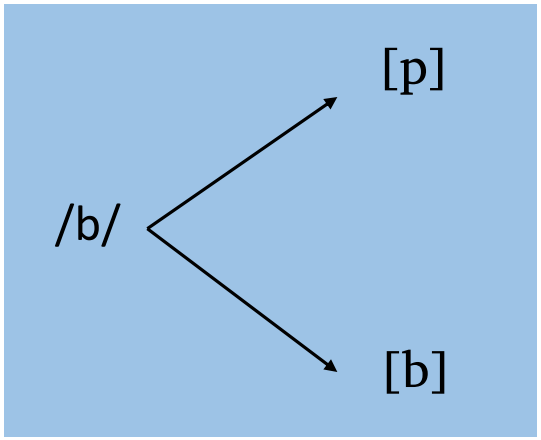
Hindi

bal	'hair'
p^hal	'knife blade'
pal	'take care of'
b^hal	'forehead'

Sound Change 2:
contrastive feature change (transphonologization)

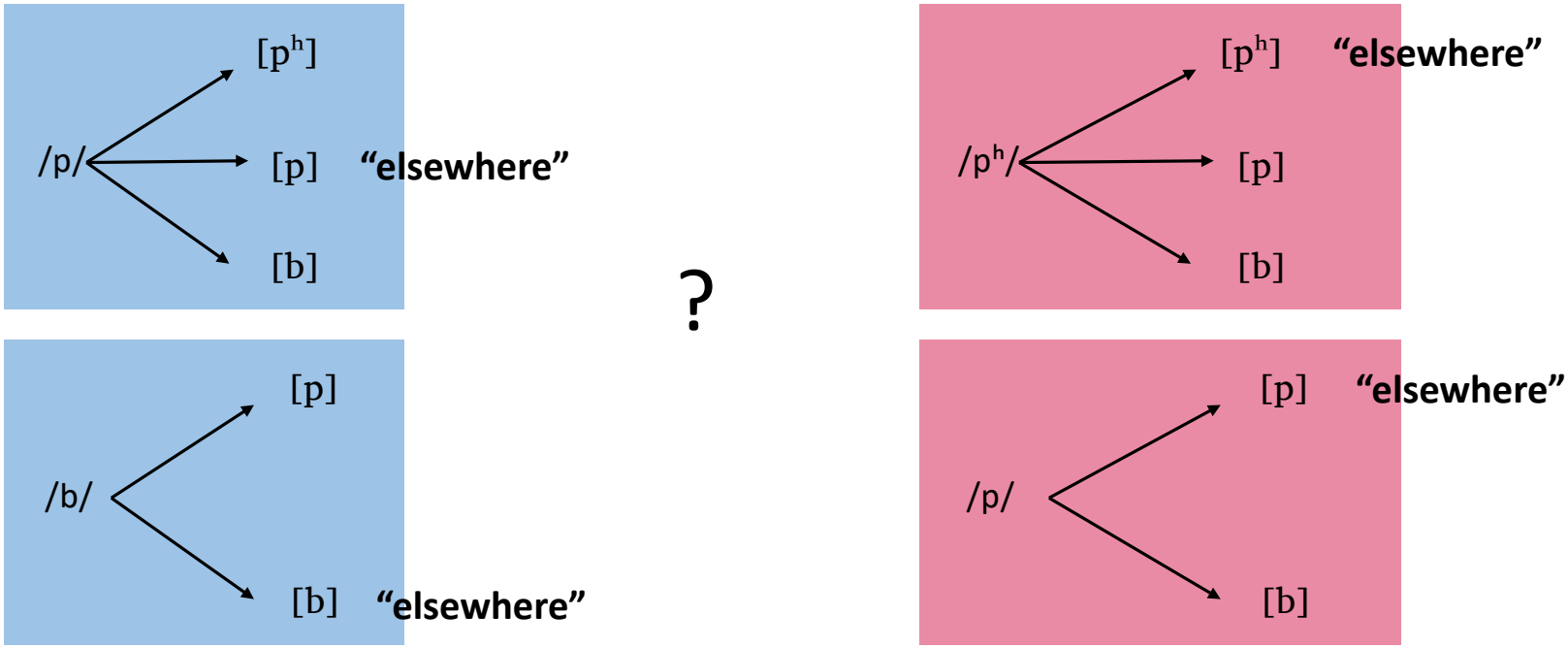
Traditional Representations





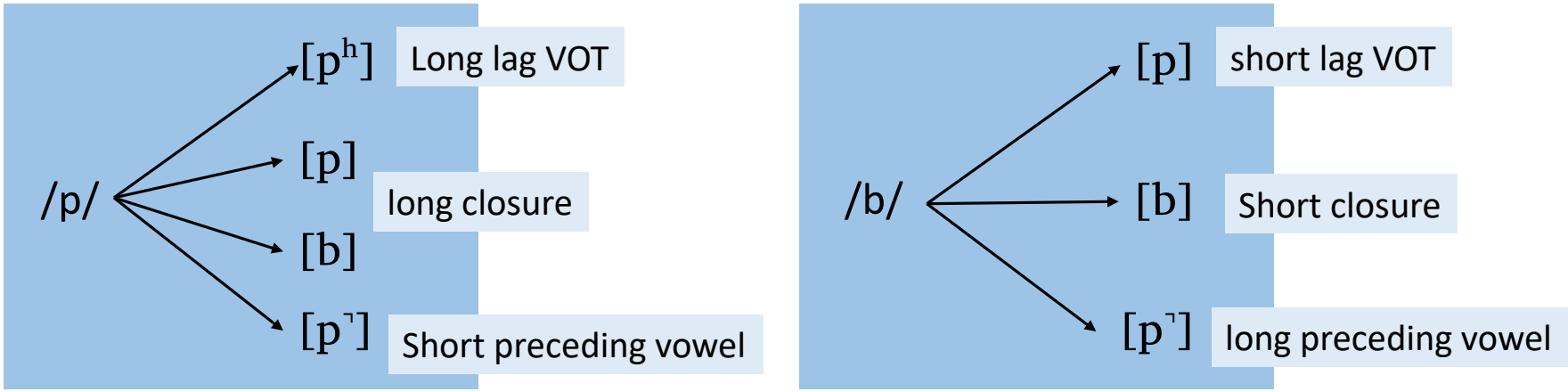
Predictable Allophonic Variants Derived by Rule

Sound Change 2:
contrastive feature change (transphonologization)



Predictable Allophonic Variants Derived by Rule

Sound Change 2:
contrastive feature change (transphonologization)



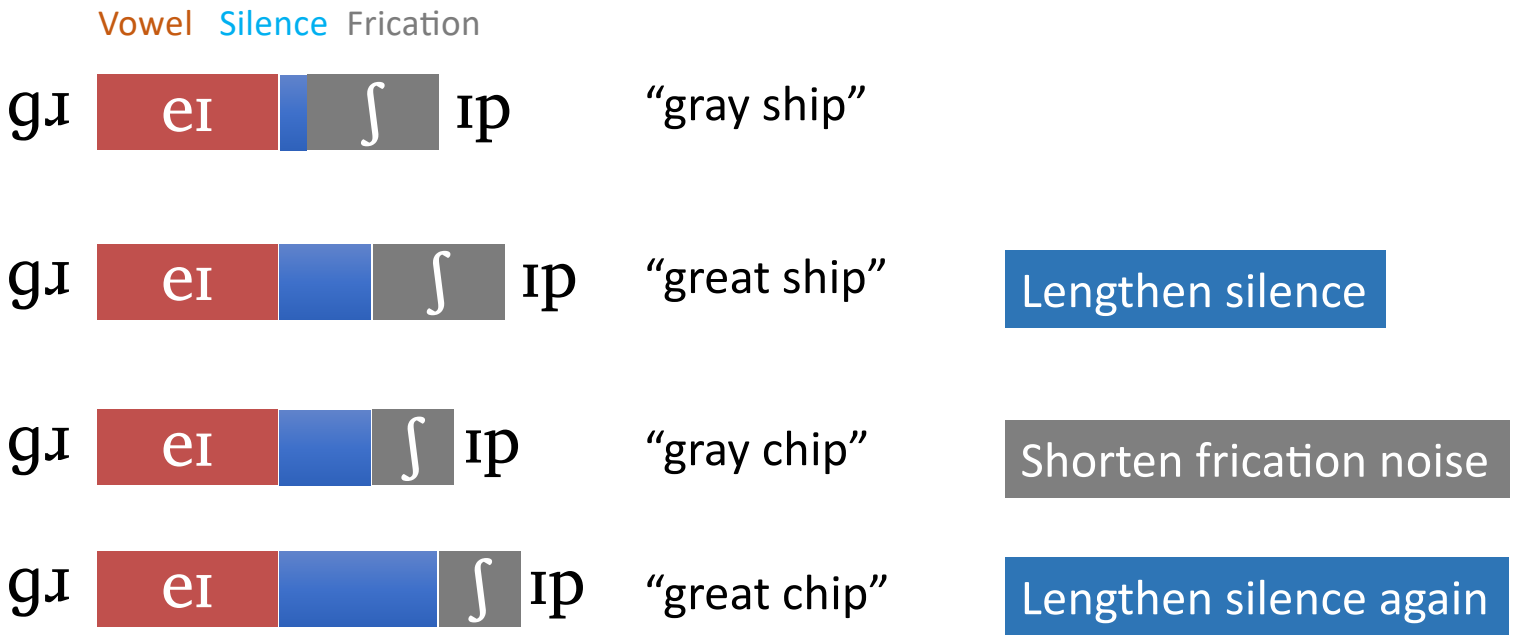
Predictability

Model II

Desiderata:

1. Link phonological allophones to phonetic allophones to featural “trading relations”
2. Explain how acoustic cues become associated with segments in the first place
3. Explain why “contrastiveness” shifts to different cue
4. Generate both “change” and “no change” outcomes

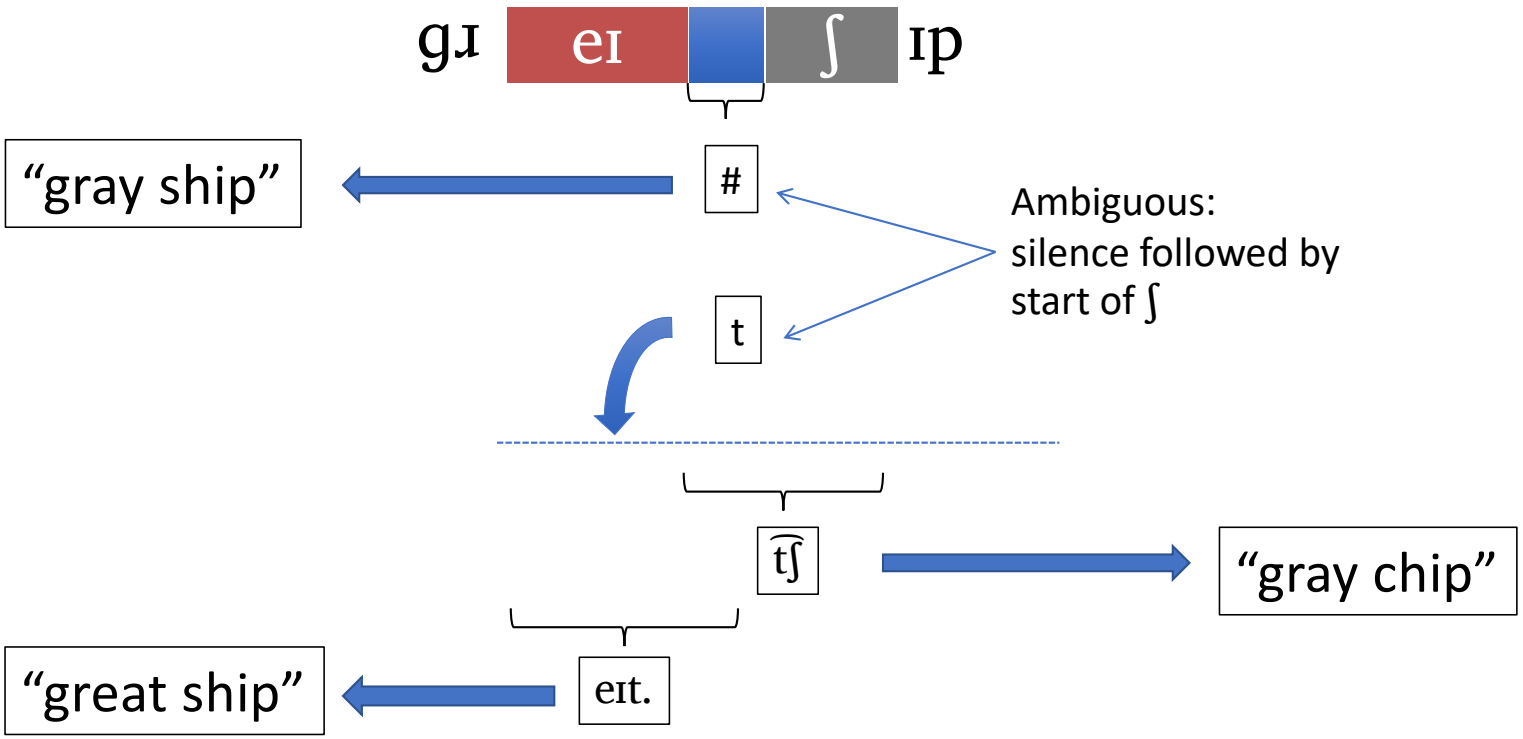
Speech Processing



Repp et al. (1978)

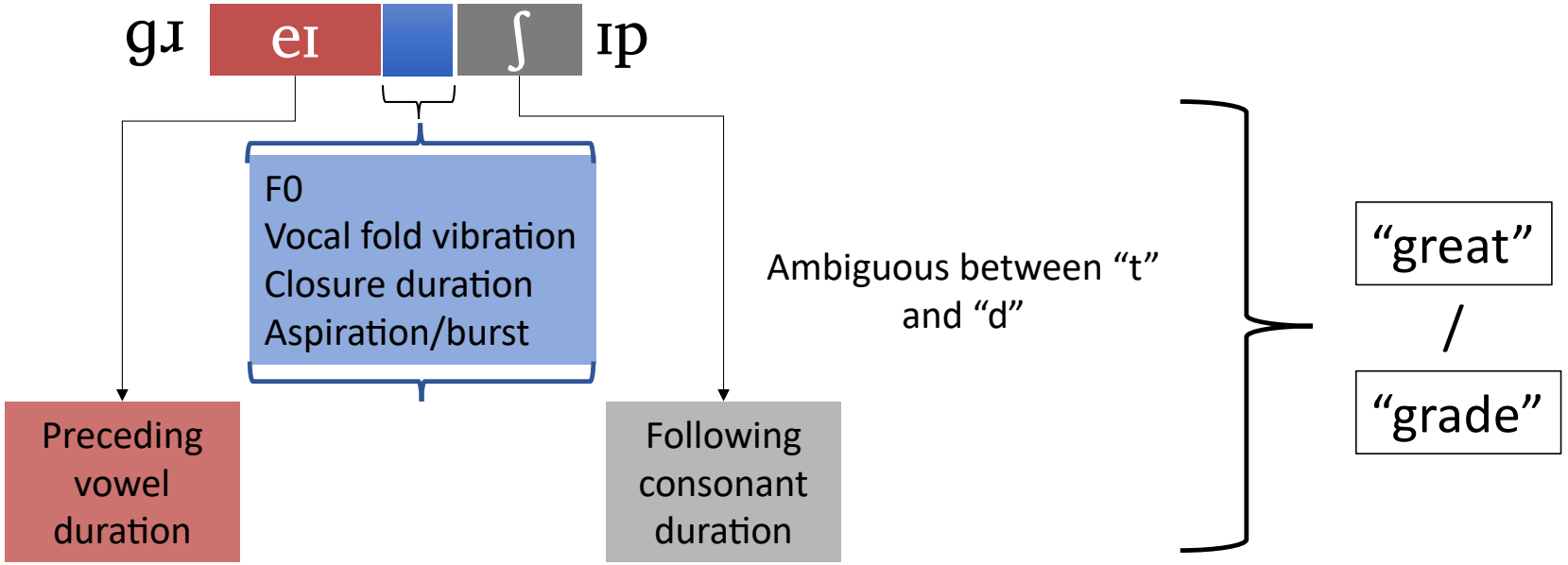
Speech Processing

1. Highest probability parse



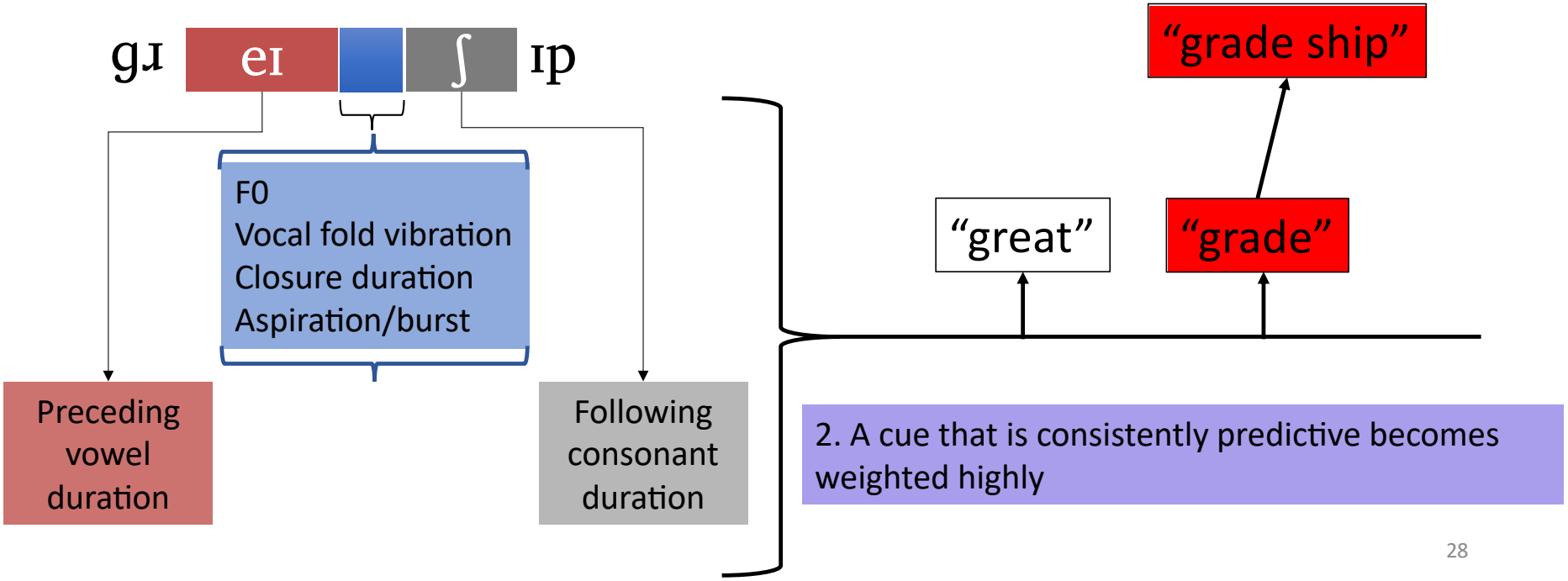
Speech Processing

2. Statistical Learning: pattern discovery



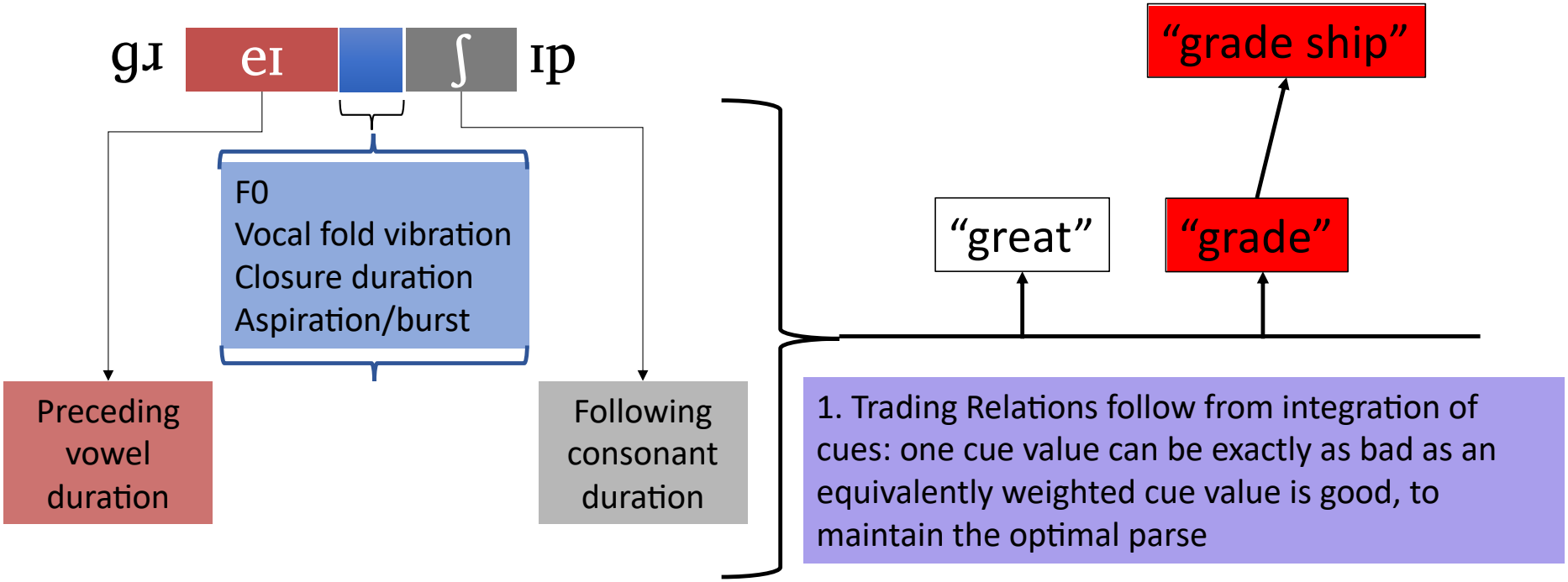
Lexical Access

2. Once word is successfully retrieved, the contribution of each cue to that successful retrieval can be used to weight that cue



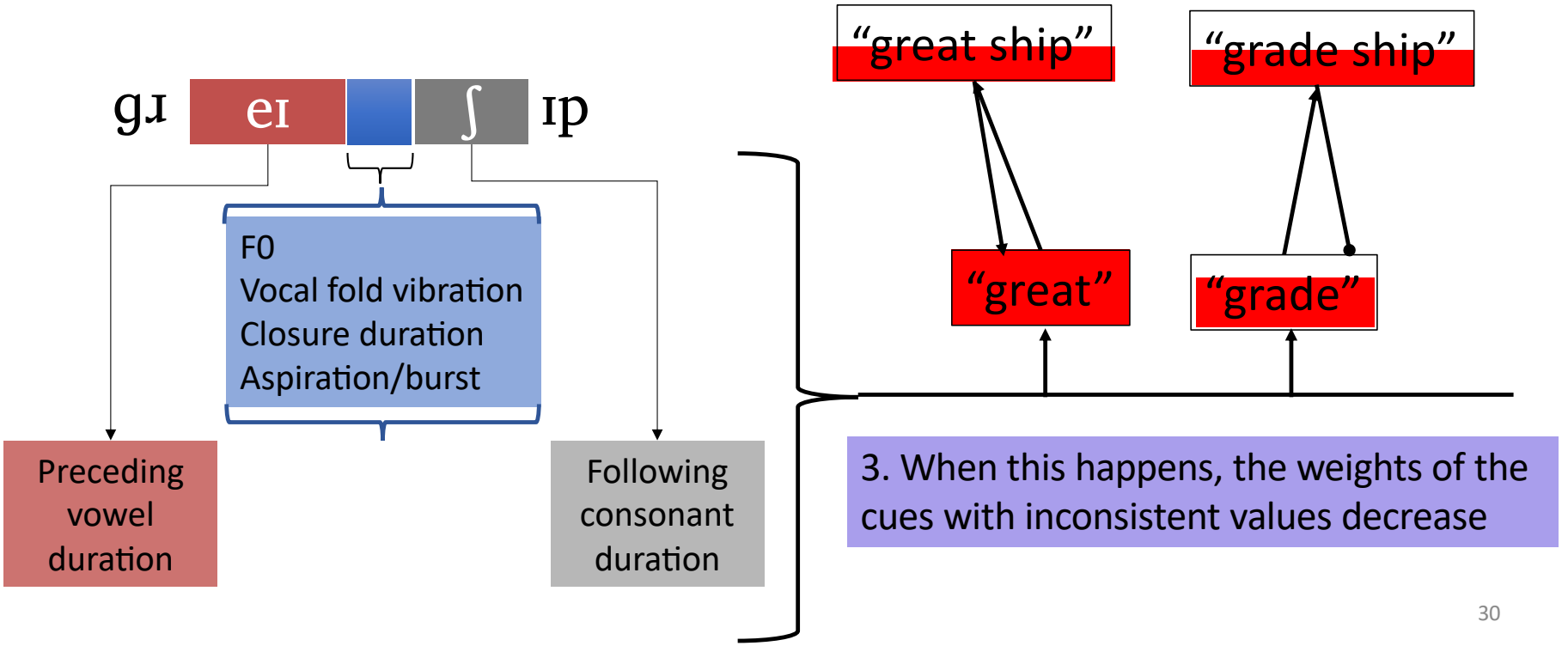
Lexical Access

2. Once word is successfully retrieved, the contribution of each cue to that successful retrieval can be used to weight that cue



Lexical Access

Even if acoustic values remain ambiguous, word can be recovered from external information.



3. When this happens, the weights of the cues with inconsistent values decrease

Perception-Production Feedback Loop

/p/ vs. /b/ → [p] vs. [b]

Some tokens of b are acoustically devoiced (or voicing is masked)

However, the identity of the segment is recoverable from other cues

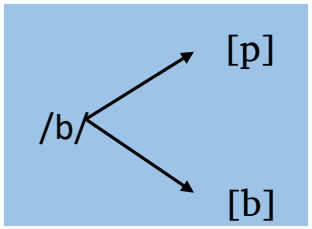
4. This happens consistently with respect to perceived voicing

Even if acoustic values remain ambiguous, word can be recovered from external information.

3. Lowered perceptual cue weight leads to less articulatory control of/attention to that cue in production

Perception-Production Link

In environments in which voicing is difficult to maintain less effort is made to preserve voicing



Allophony

4. Limit of coarticulation

4. Limit to subset of environmentss

3. Lowered perceptual cue weight leads to less articulatory control of/attention to that cue in production

3. Less articulatory control leads to a greater effect of phonetic context (e.g., coarticulation)

3. Context-specific changes in cue value make cue less predictive globally

Biology & Linguistics: Transmission & Change

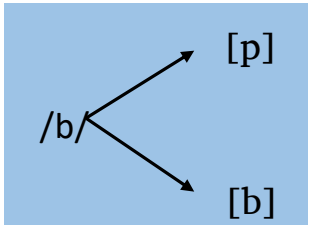
- Evolutionary Model:
  - incremental, “blind” processes can lead to complex structure
  - Macro-structure from micro-changes

  ❖Re-think long-held assumptions about the underlying representations and processes
- Cognitive Model:
  - Language module is likely to recruit existing cognitive structure
  - Phonological Theory has to, ultimately, link up with the acoustic signal

  ❖Leverage what is known about auditory, visual perception

Thank You!

Perception-Production Link



- Loss of predictivity in voicing cue will lead to increased reliance on other cues
- Subject to fluctuations in predictability
- Self-reinforcing: cue inferred to be predictive will become so via articulatory contgrol

3. Lowered perceptual cue weight leads to less articulatory control of/attention to that cue in production

3. Less articulatory control leads to a greater effect of phonetic context (e.g., coarticulation)

3. Context-specific changes in cue value make cue less predictive globally