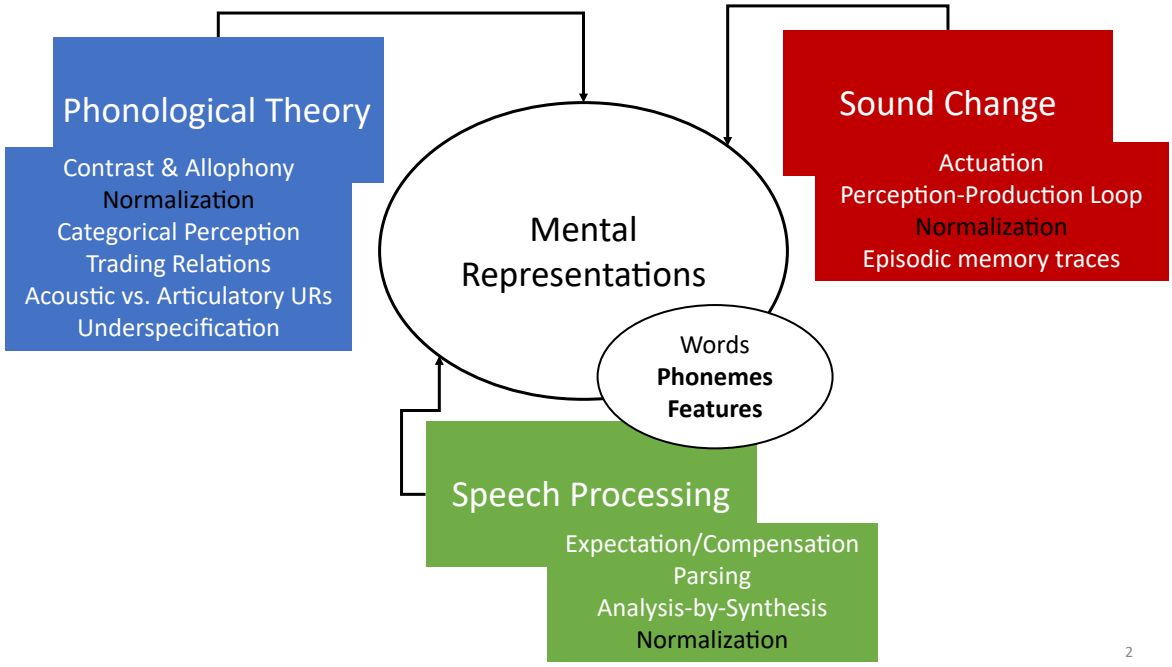
Towards A Domain-General Theory of Phonological Contrast

Rebecca L. Morley
The Ohio State University

November 19, 2020
Cornell
Ithaca, NY (Virtually)

1

1



2

Phonological Theory

3

3

Contrast

The Classical Model:

Words	Phonemes	Features
"bit" vs. "bead"	Non-minimal contrast	
"bit" vs. "bid"	/t/ vs /d/	[+coronal, +obstruent, -sonorant, -voice, ., ., .] [+coronal, +obstruent, -sonorant, +voice, ., ., .]
"spit" vs. "sbit"	Lack of contrast	

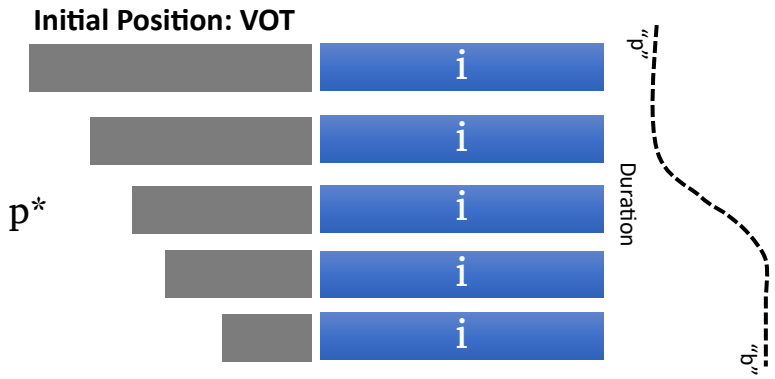
4

4

Diagnostics of Contrast

Word Recognition	Phoneme Recognition: Categorical Perception Experiment

"pea" vs. "bee"



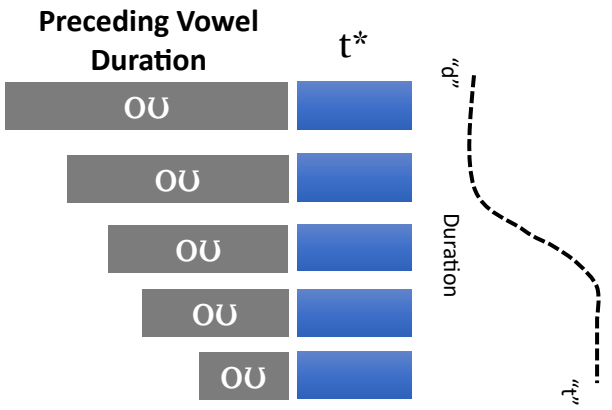
5

5

Diagnostics of Contrast

Word Recognition	Phoneme Recognition: Categorical Perception Experiment

"coat" vs. "code"



6

6

Allophony

What is the underlying contrastive feature?

- /p/→[pʰ]/ˈ__V
- /b/→[p]/#__

pat/bat:
/p/ has longer VOT than /b/, word-initially

- /p/→[b]/ˈV__V
- /b/→[p]/__#
- /p/→[p̚]/__#

rapid/rabid:
/p/ has longer closure duration than /b/, post-stress, word medially

tap/tab:
Vowel preceding /b/ is longer than vowel preceding /p/, in coda position

/b/ ≠ [b]

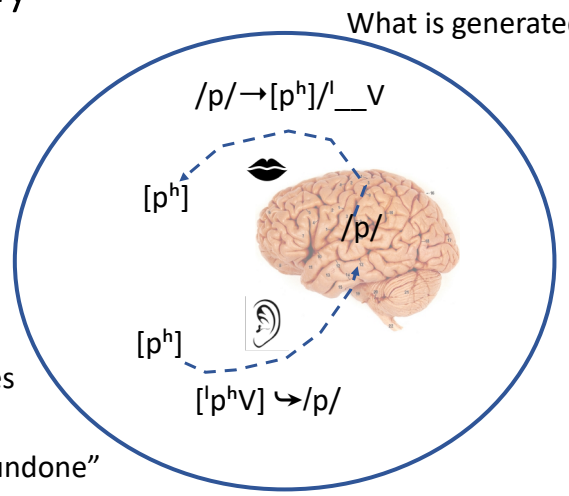
/χ/ ≠ [b]

Paradox 1:
phonetic symbol system is identical to phonological symbol system

7

7

Allophony



Allophony implies normalization

That has to be "undone"

8

8

Trading Relations

The (lack of) invariance problem

- Longer VOT changes category membership

- shorter following vowel shifts category boundary to shorter stop

Summerfield (1981); Miller & Volaitis (1982)

- longer preceding vowel shifts category boundary to longer stop

Summerfield (1981); Miller & Volaitis (1982)

9

9

Cues that could be exploited by listeners

Word-Final stops:

- duration of voicing
- intensity of voicing
- Aspiration
- F0 contour
- length of vowel formant transitions with respect to steady state duration
- F1 offset frequency
- speed of jaw lowering
- jaw offset position

Continuous "trade offs"

…/acoustic/articulatory/phonetic/ phonological/…

But why does allophony occur in the first place?

Paradox 2:
Contrast should prevent allophony!

Fitch 1981; Crowther & Mann 1992; Van Summers 1987
cf. Bailey & Summerfield 1980

10

10

Specification and Underspecification

Variation is allowed on dimensions that are not specified

Control of articulation is evidence for specification

[±voice]

[voice]

or not

Ø

- aspirated/un-aspirated
- long VOT/short VOT
- Released/un-released
- voiced/voiceless/ partially voiced

voicing present

voicing present when context is appropriate

p#d

VpV

11

11

What are the Underlying Features?

Articulatory rather than acoustic?

Fitch 1981; Crowther & Mann 1992; Van Summers 1987; Bailey & Summerfield 1980

English:
[spread glottis] vs. Ø
("p" vs. "b")
Kim 1970; Iverson & Salmons 1995

Russian:
[voice] vs. Ø
("b" vs. "p")
Keating 1984; Beckman et al 2013

Specifications are not necessarily parsimonious in real human languages

Regressive assimilation to
s
al 1992;Jansen 2004

g of passive voicing
12

n:
[spread glottis] vs. [voice]
Beckman et al 2001

12

12

Important Implications: [wake up here]

- Categorical perception does not entail a *necessarily* contrastive feature, only a possible one (when all other information is absent). See Trading Relations.
- There is a very large number of possible allophonic rules (at least on the acoustic side).
- Allophonic rules imply normalization
- The relationship between underspecification and variation implies an inverse relationship between contrast and allophony
- Specification is likely to be continuous, and not necessarily parsimonious [this is an empirical question]

13

13

Sound Change

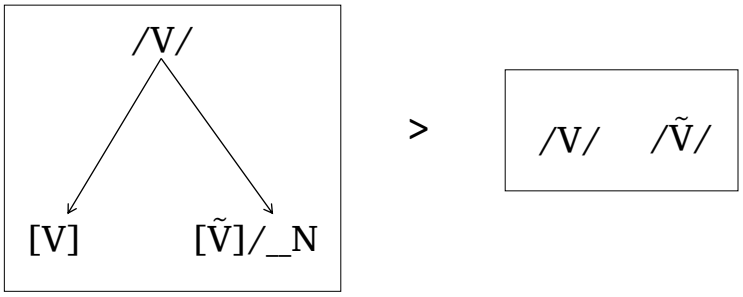
14

14

The Actuation Paradox

Paradox 3

A direct result of the normalization assumption



Normalization: No Change

Lack of Normalization: Change

In collusion with the discrete sequencing assumption k+æ+t

15

15

The Actuation Paradox

Allophonic Vowel Nasalization Rule [synchronic]:

$V \rightarrow \tilde{V} / \_N$

Loss of coda nasals [diachronic]:

$N > \varnothing / \_.$

Loss of allophonic rule/predictability [synchronic]

$V \rightarrow \tilde{V} / \_??$

$V \rightarrow V / \_??$

Paradox 3:
loss of allophonic context
should lead to
loss of allophone

Minimal Pair Test

/V/ vs. /Ṽ/

16

16

Models of Sound Change

Phonetically ambiguous/outlier tokens are discarded

- As a result of competing contrastive categories
  (Blevins & Wedel 2009;Wedel 2006, 2007; Tupper 2014)
- In speech mode, but not non-speech mode
  (Garrett & Johnson 2013)
- For some (neurotypical) individuals, but not others
  (Yu 2013)
- Except when misparsing occurs
  (Kirby 2014)

Sometimes you normalize
Sometimes you don't

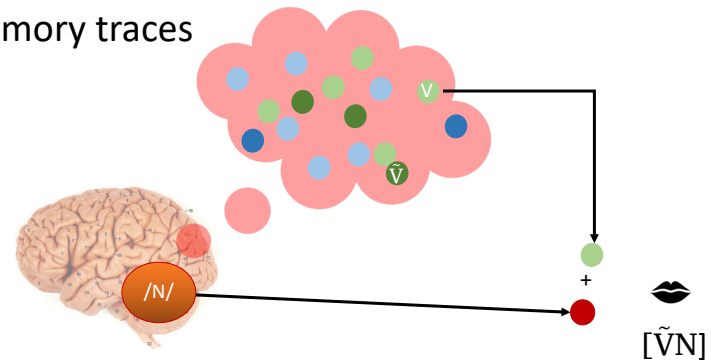
17

17

No Normalization = Change?

Exemplars: episodic memory traces

Paradox 4: unnormalized tokens are not consistent with allophonic rules



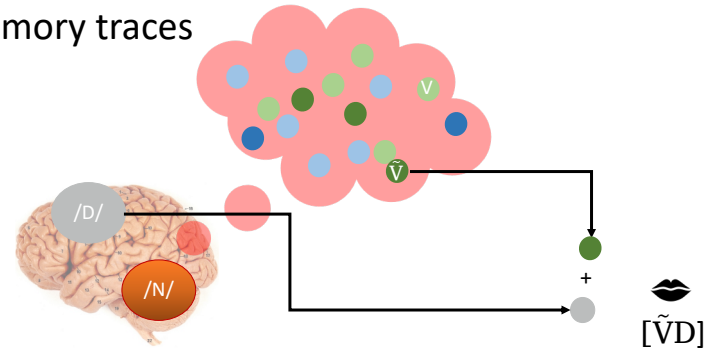
18

18

No Normalization = Change?

Exemplars: episodic memory traces

Paradox 4: unnormalized tokens are not consistent with allophonic rules



19

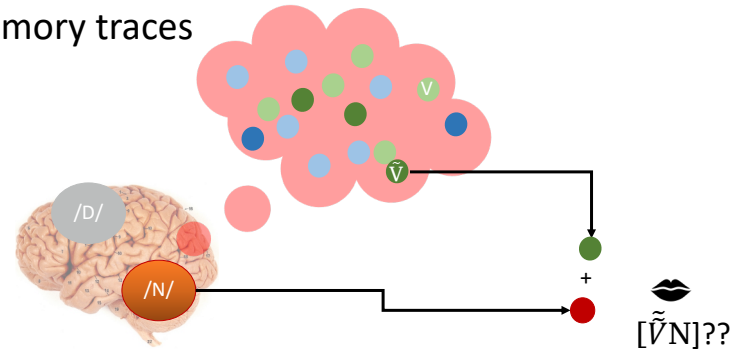
19

No Normalization = Change?

Exemplars: episodic memory traces

Paradox 4: unnormalized tokens are not consistent with allophonic rules

Paradox 5: the same symbol systems are used for acoustic and articulatory representations!

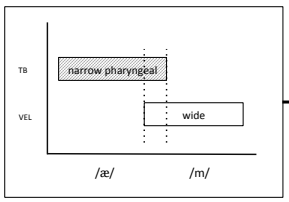


20

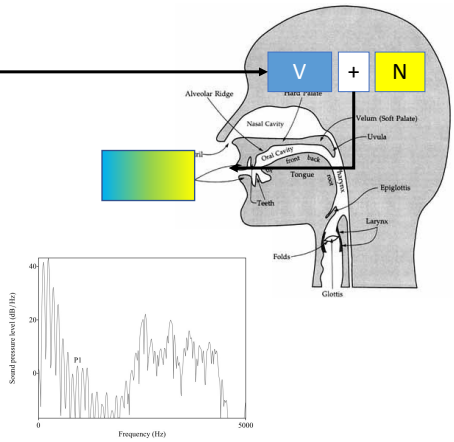
20

Why do allophones exist?

Coarticulation!



Browman & Goldstein 1988

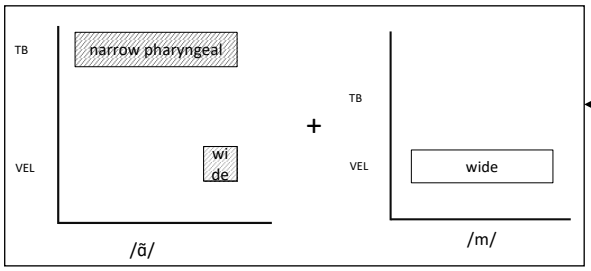


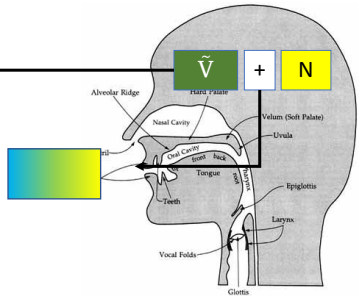
21

21

Why do allophones exist?

Coarticulation!

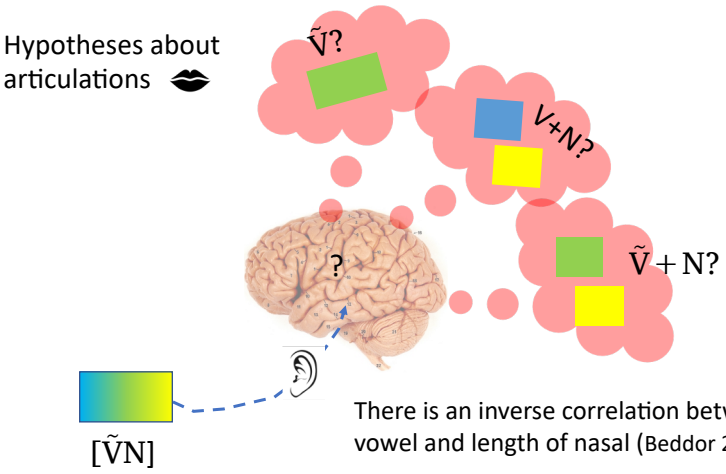




22

22

Perception ⇔ Production



Model (Morley 2019):

- Ambiguity in normal speech processing
- Competing Hypotheses (Parses) for each acoustic input
- Parse depends on acoustic variables
- Not nasal loss, but nasal merger

There is an inverse correlation between degree of nasalization on vowel and length of nasal (Beddor 2013).

23

23

Important Implications: [wake up here]

- Speech perception must involve:
  - mapping acoustic input to articulatory categories (targets)
  - Segmentation of ambiguous input
- Lack of normalization implies lack of allophony
- Unless allophony is an emergent property of articulation
- Incremental change is possible with distributions of stored, unnormalized acoustic tokens (exemplars)

24

24

Speech Processing

25

25

Is Normalization Required?

1. Longer vowel creates expectation for following voiced (rather than voiceless) stop:

oʊ

Sounds like "coat"

oʊ

Sounds like "code"

2. Nasalized vowel spliced into non-nasal syllable sounds more like lower vowel (higher F1): Krakow et al 1988

sɛ̃nd → /ɛ/

sɛ̃nd / sɛd → s ɛ̃ d → /æ/

3. Expected degree (and direction) of place-of-articulation assimilation facilitates word recognition: Gow and McMurray 2007

gɹin bin , gɹin dɔg >> gɹi n dɔg , gɹi n bin

26

26

Is Normalization Required?

Acoustic Normalization:

- attributing distributed acoustic cues to individual (source) segments
- adjusting acoustic cue values based on context

[sɛ̃nd] ↬/ɛ/

oʊ d ↬ /oʊ/

gɹinʷ bin ↬ /n/

Implies a more abstract/invariant underlying representation that must be closely matched

e.g., Gow (2002); Beddor et al. (2002); Kirby (2014); McMurray et al (2011)

27

27

Is Normalization Required?

- Trading Relations/Categorical Perception/(In)Variance Problem
- Expectation/Compensation/Prediction
  - No acoustic cues are absolute
  - Classification depends strongly on other cues present, context
  - Previous input, current context, current knowledge generate predictions about upcoming material
  - listeners are highly sensitive to correlations among acoustic cues

28

28

The Phonological Parser

Vowel Silence Frication

gɹ eɪ ∫ ɪp "gray ship"

gɹ eɪ ∫ ɪp "great ship" Lengthen silence

gɹ eɪ ∫ ɪp "gray chip" Shorten frication noise

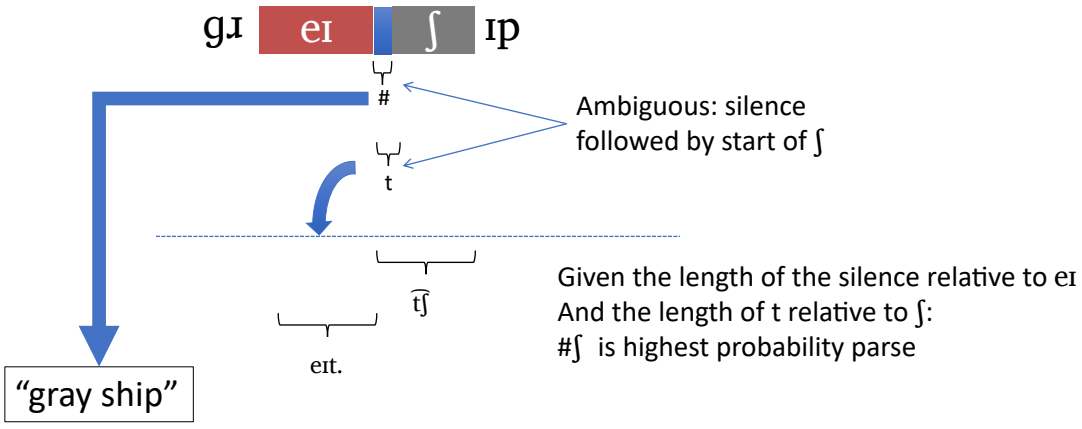
gɹ eɪ ∫ ɪp "great chip" Lengthen silence again

Repp et al. (1978)

29

29

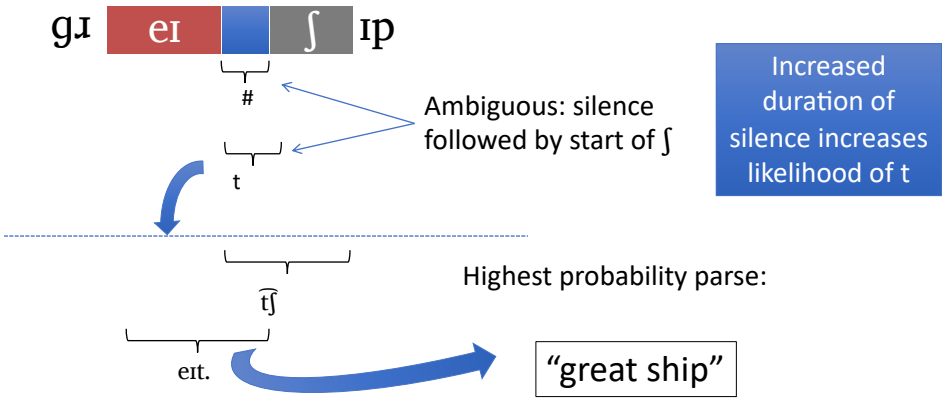
The Phonological Parser



30

30

The Phonological Parser



31

31

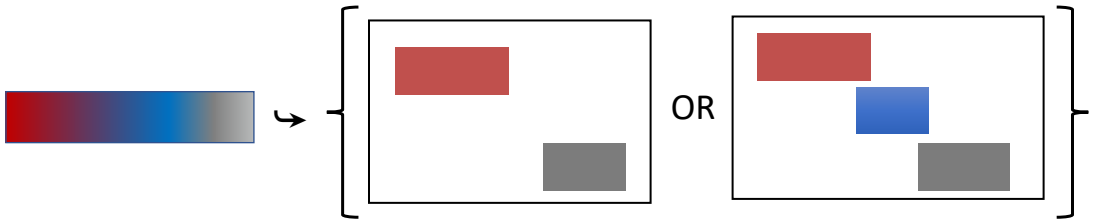
Is Normalization Required?

Rather than matching abstract acoustic symbols,

speech perception can be described as choosing the optimal phonological parse, that is,

the likeliest string of **articulatory symbols** that would result in the observed acoustic input

Remember Paradox 5!



32

32

Is Normalization Required?

Rather than matching abstract acoustic symbols,

speech perception can be described as choosing the optimal phonological parse, based on:

- All available contextual information
- Self-consistency

Local Optimum

- Relative feature values

- In most cases, from a small candidate set (~ 2 words)
- In most cases, involving local comparisons within a 2-3 segment window

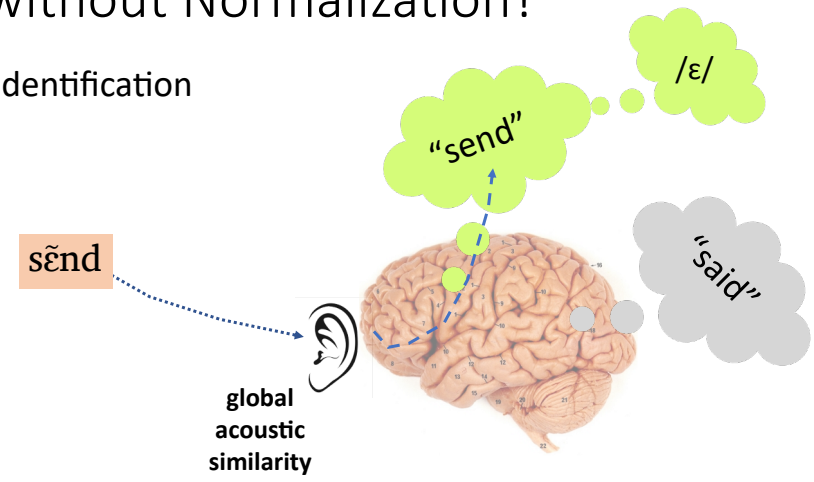
33

33

A World without Normalization?

Classification/Identification

Word-level



Krakow et al. 1988

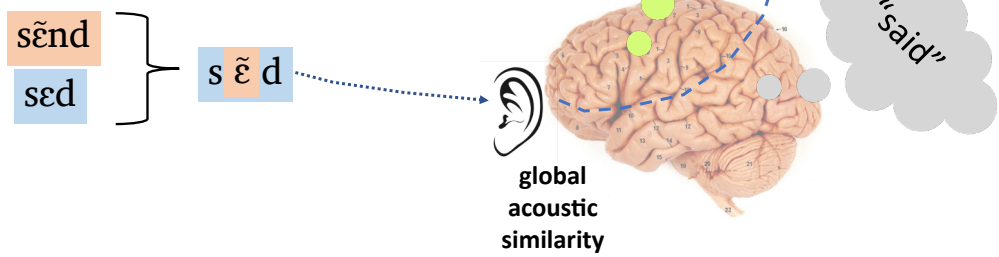
34

34

A World without Normalization?

Classification/Identification

Word-level



Krakow et al. 1988

35

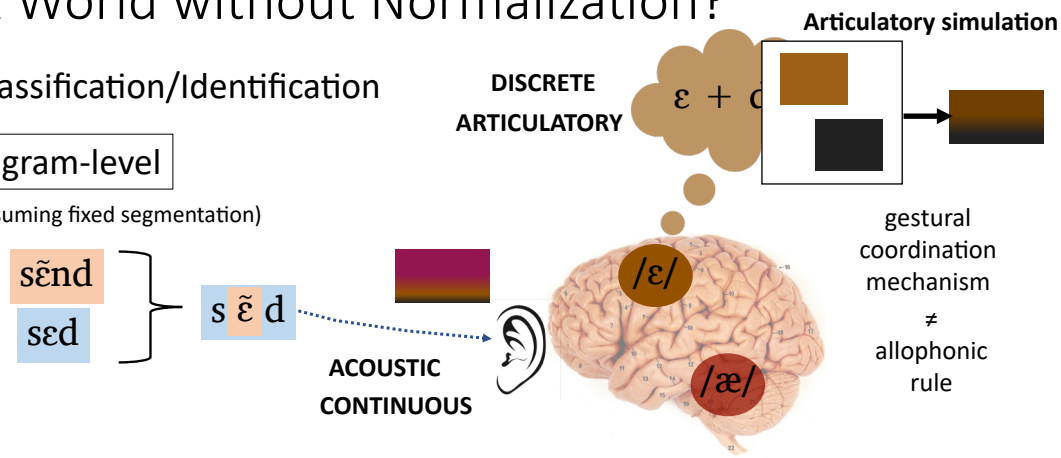
35

A World without Normalization?

Classification/Identification

Bigram-level

(assuming fixed segmentation)



Analysis-By-Synthesis: Halle & Stevens (1967); Poeppel & Monahan (2011)

36

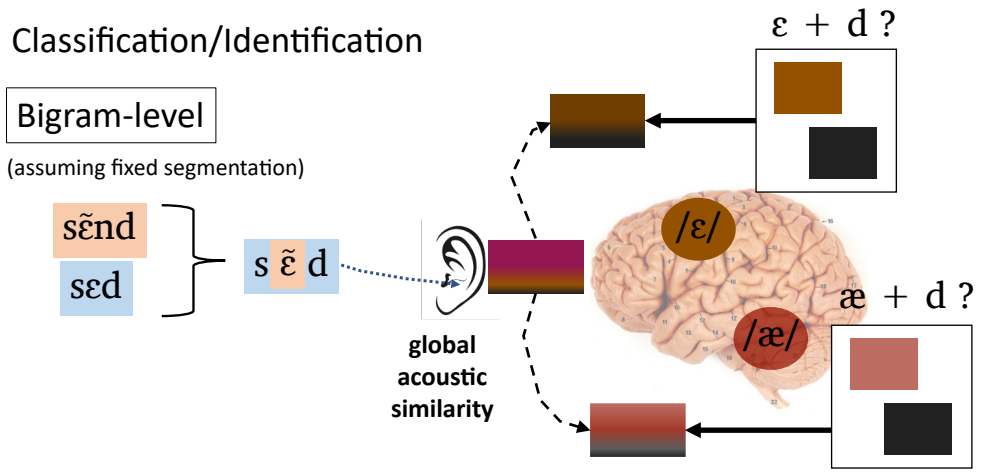
36

A World without Normalization?

Classification/Identification

Bigram-level

(assuming fixed segmentation)



Analysis-By-Synthesis: Halle & Stevens (1967); Poeppel & Monahan (2011)

37

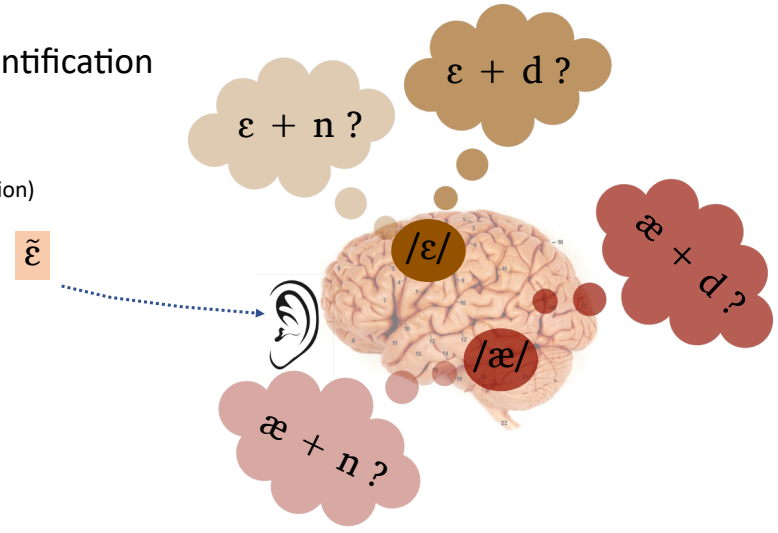
37

A World without Normalization?

Classification/Identification

Phoneme-level

(assuming fixed segmentation)



Krakow et al. 1988

38

38

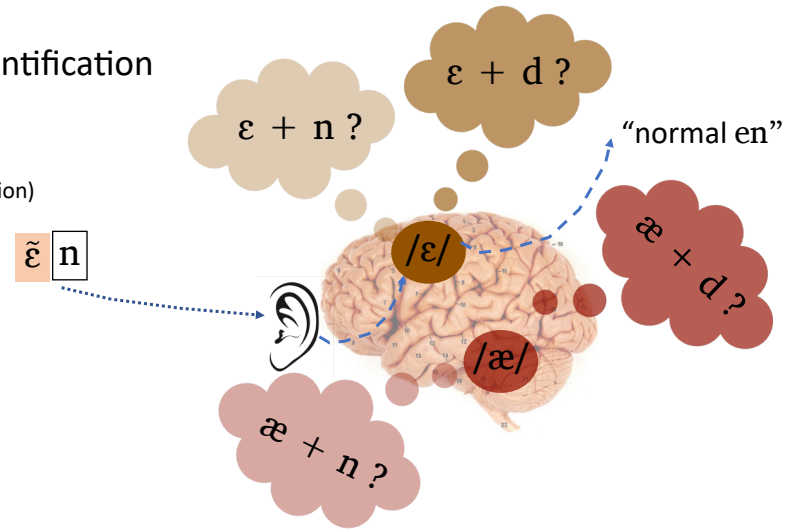
A World without Normalization?

Classification/Identification

Phoneme-level

(assuming fixed segmentation)

ɛ̃ n



Krakow et al. 1988

39

39

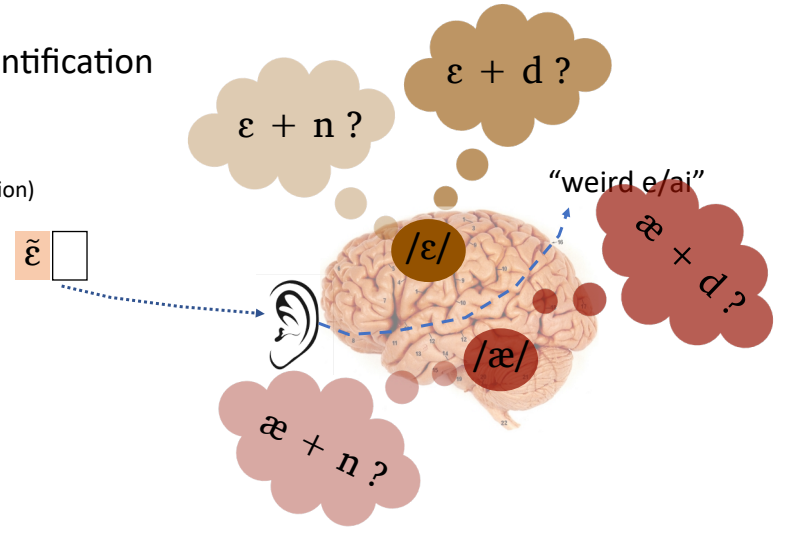
A World without Normalization?

Classification/Identification

Phoneme-level

(assuming fixed segmentation)

ɛ̃



Krakow et al. 1988

40

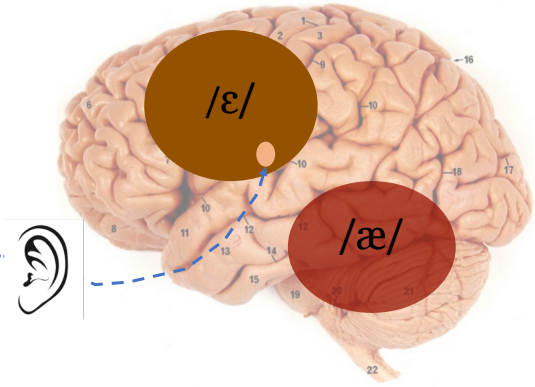
40

A World without Normalization?

Storage

Phoneme-level

ɛ̃ n



41

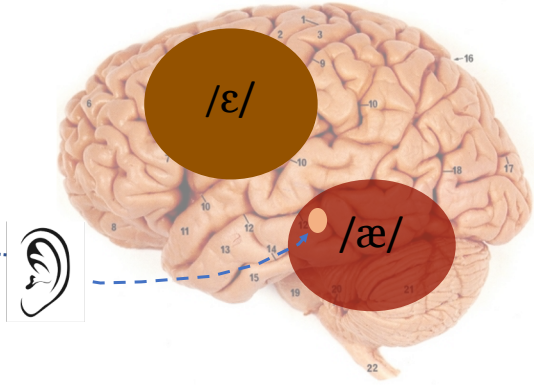
41

A World without Normalization?

Storage

Phoneme-level

ɛ̃



"I said I would do the dishes"

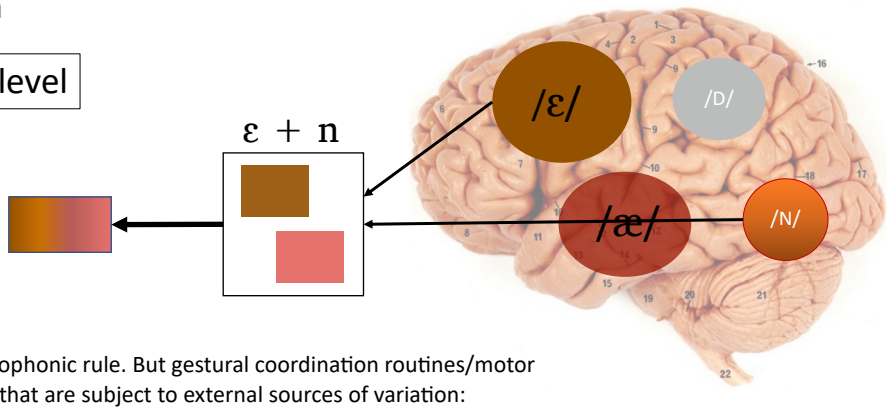
42

42

A World without Normalization?

Production

Phoneme-level



No allophonic rule. But gestural coordination routines/motor plans that are subject to external sources of variation:
Speaking rate, word predictability/frequency, ease of articulation…

43

43

The Model (implementation is crucial)

- Synchronic variation is the result of normal speech processing
- Sound change is present in synchronic variation
  - allophony of continuously fluctuating degree
- Allophones are emergent from the interaction of
  - articulatory specification
  - gestural coordination
  - acoustic specification

44

44

The Model

- Speaker/Listeners store:

  acoustic targets

  check/constrain output

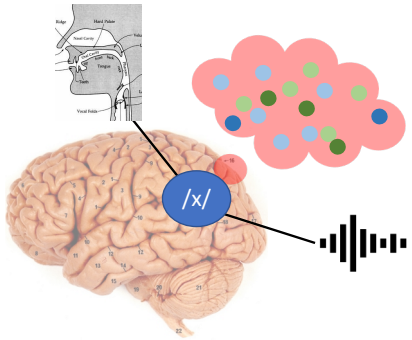
  articulatory targets

  for production to be possible

  unnormalized acoustic experiences

  short-term adaptation/accomodation

  incremental sound change



45

45

The Model

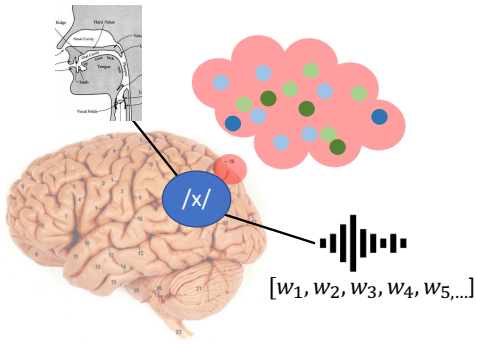
- Speaker/Listeners store acoustic targets, articulatory targets, and unnormalized acoustic experiences

  Short-term adaptation/accommodation

  Incremental sound change

- Degree of specification for each acoustic feature depends on its informativity wrt the contrast

  Phonological Specification



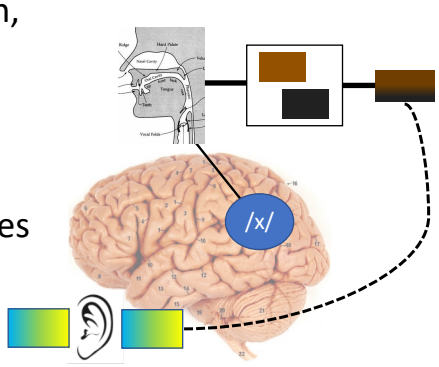
46

46

The Model

Perception

- Acoustic input is evaluated by simulation, using underlying articulatory categories
  - Local context sensitivity — Parsing window
  - Avoids normalization
- Best overall acoustic similarity determines category membership
  - Avoids normalization



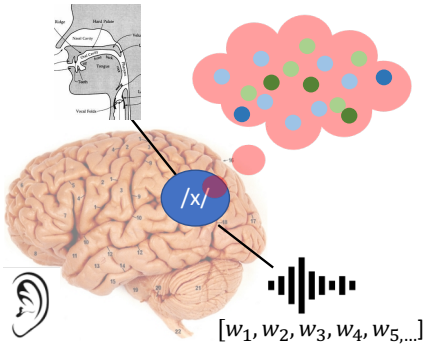
47

47

The Model

Perception

- Acoustic input is evaluated by simulation, using underlying articulatory categories
- Best overall acoustic similarity determines category membership
- Targets are directly affected by changes to the distribution of stored acoustic exemplars
  - Incremental sound change
  - Avoids actuation paradox



48

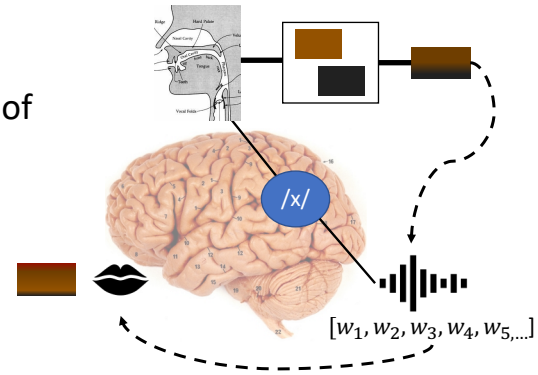
48

The Model

Production

- Articulatory targets are perturbed by: production error and gestural overlap of adjacent units
- Ultimate acoustic realization is constrained by degree of specification of each individual feature

Emergent Allophony



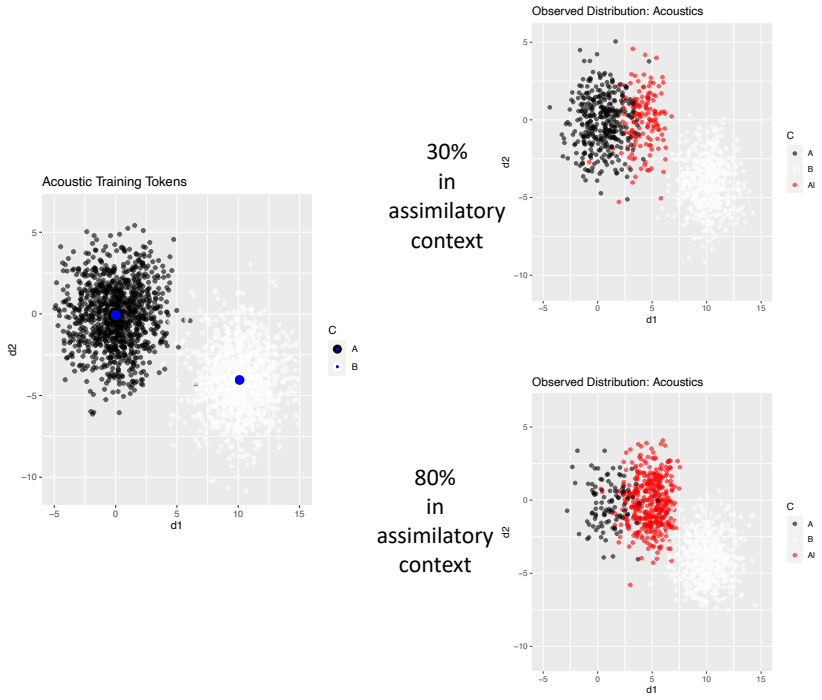
49

49

Thank You!

50

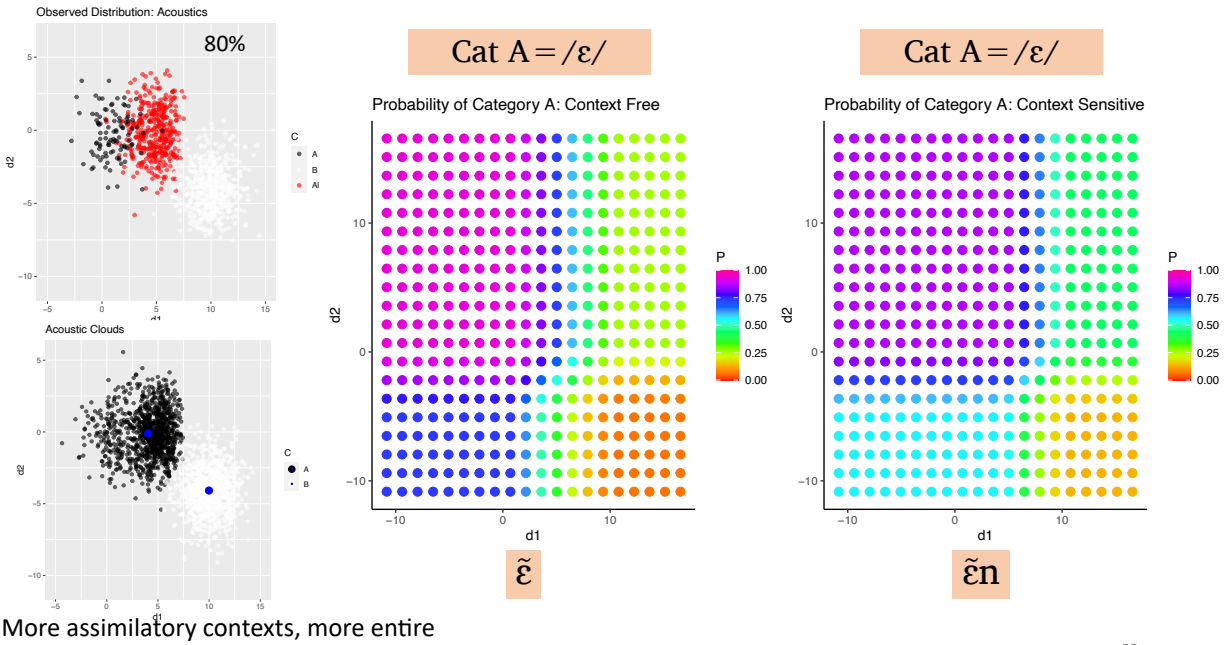
50



- No acoustic normalization
- Acoustic targets updated from acoustic clouds
- Weights updated from acoustic clouds
- Fixed articulatory targets
- Range of allophony constrained by weight values

51

51



More assimilatory contexts, more entire category is shifted: bug or feature?

52

52