Localising and Understanding the Neural Systems for Processing Spoken Words

Matt Davis

MRC Cognition & Brain Sciences Unit

Cambridge, UK

MRC Cognition and Brain Sciences Unit

UNIVERSITY OF CAMBRIDGE

@MattDavis@fediverse.science

1

Maya Angelou (1928-2014)

"Words are things, I'm convinced… Someday we'll be able to measure the power of words. I think they are things. I think they get on the walls, they get in your wallpaper, they get in your rugs, in your upholstery, in your clothes. And, finally, into you."

2

Processing Spoken Words

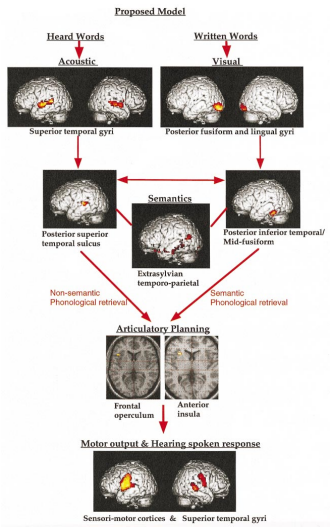
Localising vs explaining spoken word recognition

Bayesian inference in speech perception

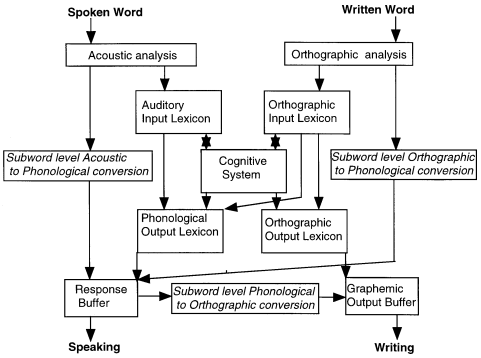
Predictive computations for word recognition

3

Box & Arrow Models of Word Recognition



Price (2000, *J Anatomy*)

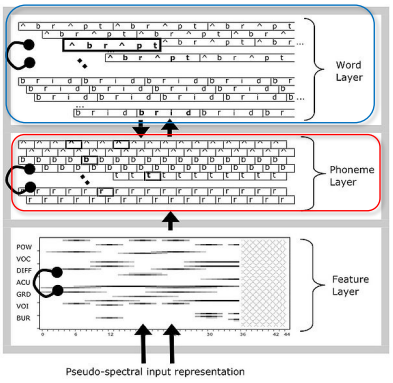


Morton (1969) Logogen Model
Patterson & Shewell (1987)

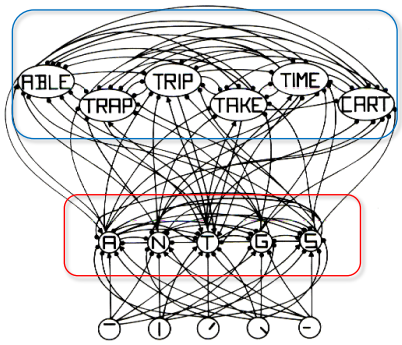
4

Computational accounts of Lexical Processing

TRACE model of speech perception
(McClelland & Elman, 1986)



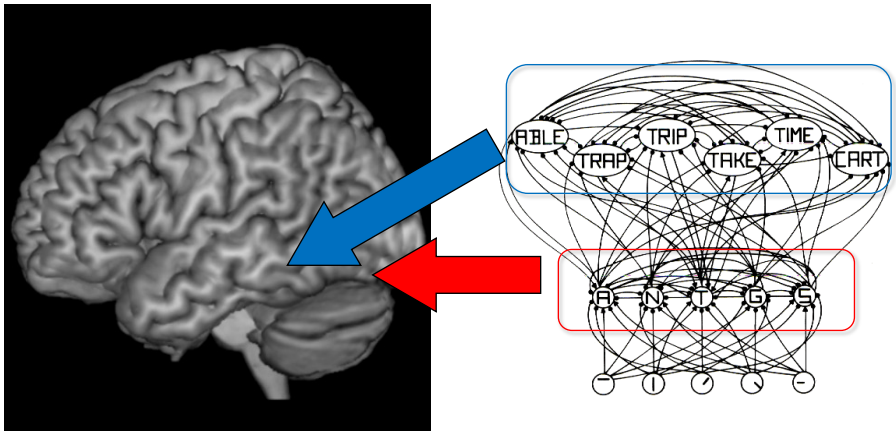
Interaction Activation Model of letter perception
(McClelland & Rumelhart, 1981)



5

Mapping Computational Accounts onto the Brain

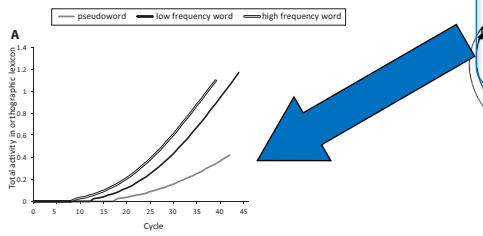
Interactive Activation Model
of Letter Perception
(McClelland & Rumelhart, 1981)

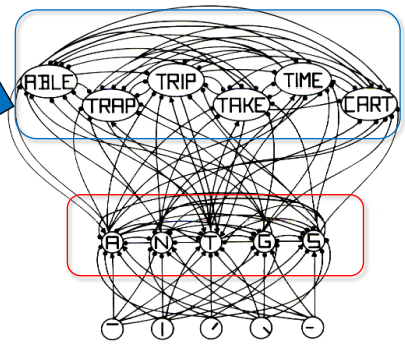


6

Mapping Computational Accounts onto the Brain

Interactive Activation Model
of Letter Perception
(McClelland & Rumelhart, 1981)



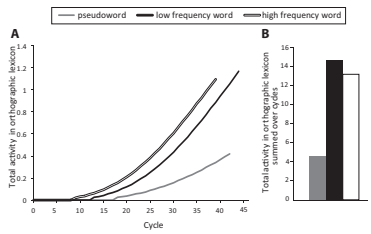


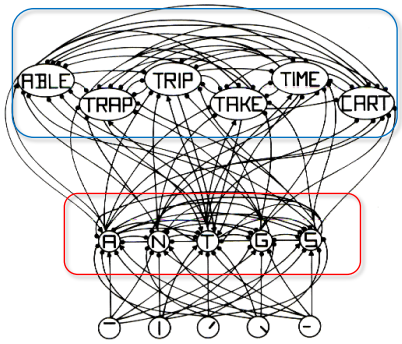
Taylor, Rastle, & Davis (2013)
Psychological Bulletin

7

Mapping Computational Accounts onto the Brain

Interactive Activation Model
of Letter Perception
(McClelland & Rumelhart, 1981)





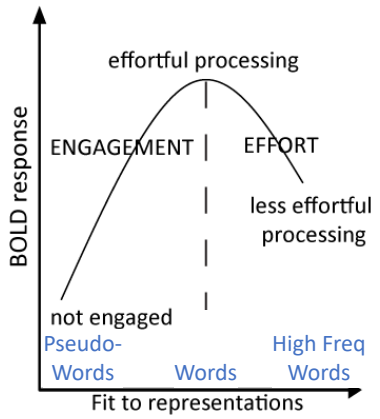
Taylor, Rastle, & Davis (2013)
Psychological Bulletin

8

Linking Computational Accounts to the Brain

1. **Engagement:**
   Stimuli that are represented by a region lead to greater neural activity
   (e.g. words > pseudowords)

2. **Effort:**
   Stimuli that are a good fit to representations lead to less effort during neural processing
   (e.g. low > high frequency words)

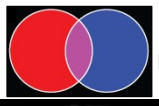


Taylor, Rastle, & Davis (2013)
Psychological Bulletin

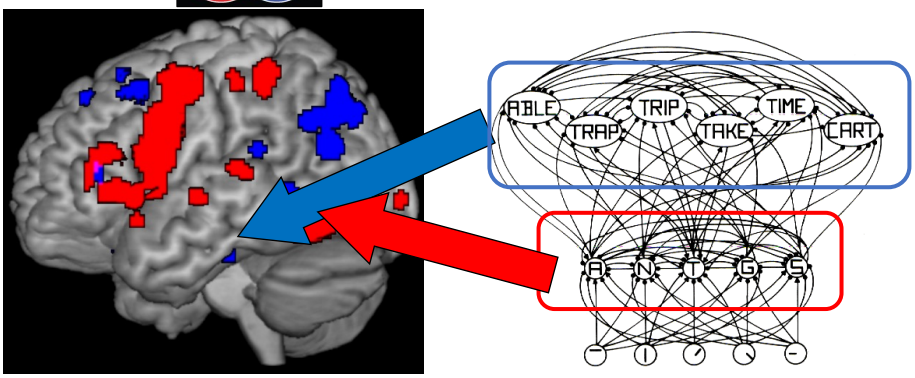
9

fMRI Meta-analysis: Written Words vs Pseudowords

pseudowords > words



words > pseudowords

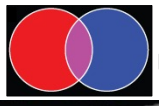


Taylor, Rastle, & Davis (2013)
Psychological Bulletin

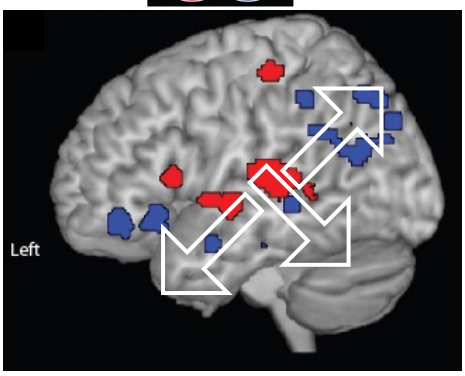
10

fMRI Meta-analysis: Spoken Words vs Pseudowords

pseudowords > words

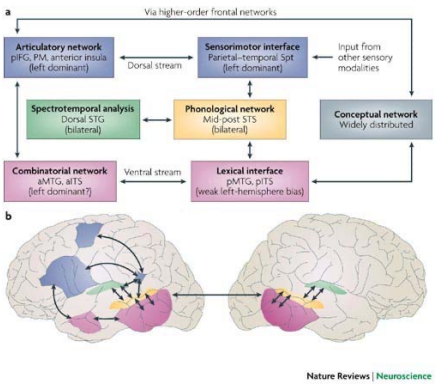


words > pseudowords



Davis & Gaskell (2013) Phil Trans Roy Soc B

Dorsal vs Ventral Pathways

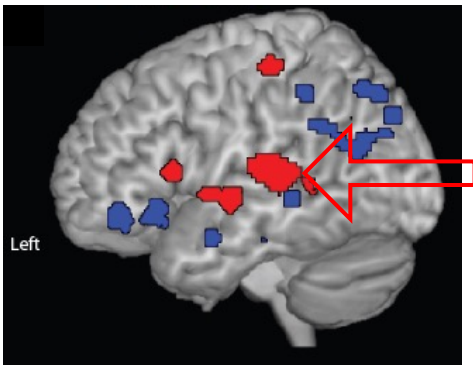


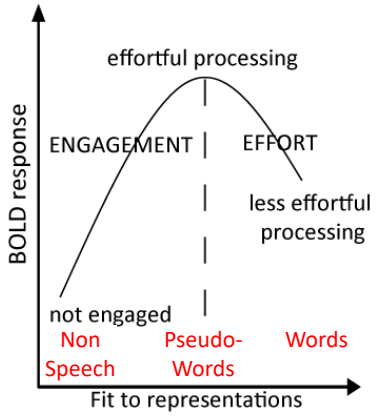
Hickok & Poeppel
(2007, *Nature Reviews Neuroscience*)

11

Processing Spoken Words

Localising vs explaining spoken word recognition



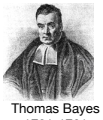


13

Processing Spoken Words

Localising vs explaining spoken word recognition

Bayesian inference for spoken words


Thomas Bayes
1701-1761


Peter Kay
1973-

Predictive computations for word recognition and learning

14

Bayesian Inference in Speech Perception

Posterior
How probable is each word given the sound heard

Likelihood
How probable is hearing that sound when that word is said?

Prior
How probable was each word *before* hearing any sound?

$$P(\text{Word}|\text{Sound}) = \frac{P(\text{Sound}|\text{Word}) \times P(\text{Word})}{P(\text{Sound})}$$

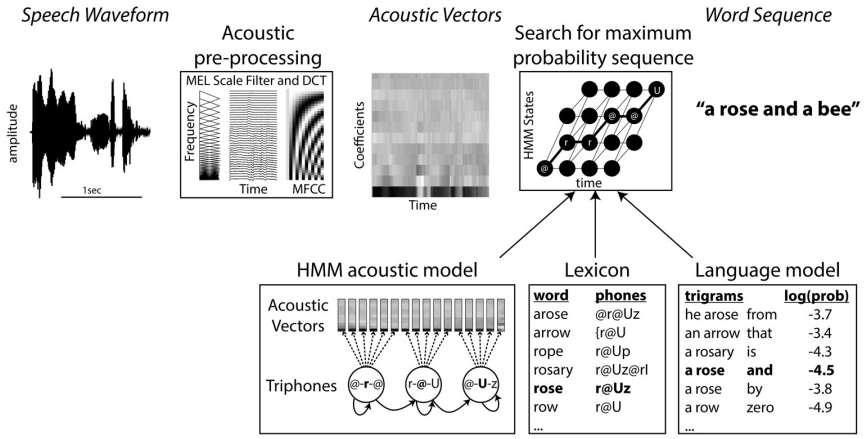
Marginal
How probable is hearing that sound

Shortlist B: Norris & McQueen (2008, *Psychological Review*)
Davis & Scharenborg (2016, *"Speech perception by humans & machines"*)


Thomas Bayes
1701-1761

16

Speech perception by machines

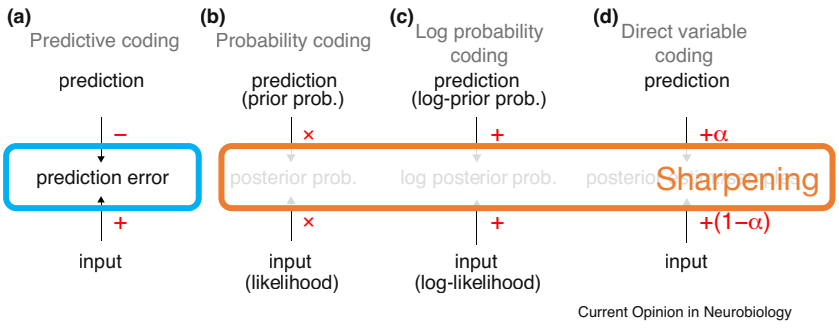


Traditional ASR System

from: Davis & Scharenborg (2016, in Gaskell & Mirkovic: *Speech Perception & Spoken Word Recognition*)

17

Neural Implementations of Bayesian Inference

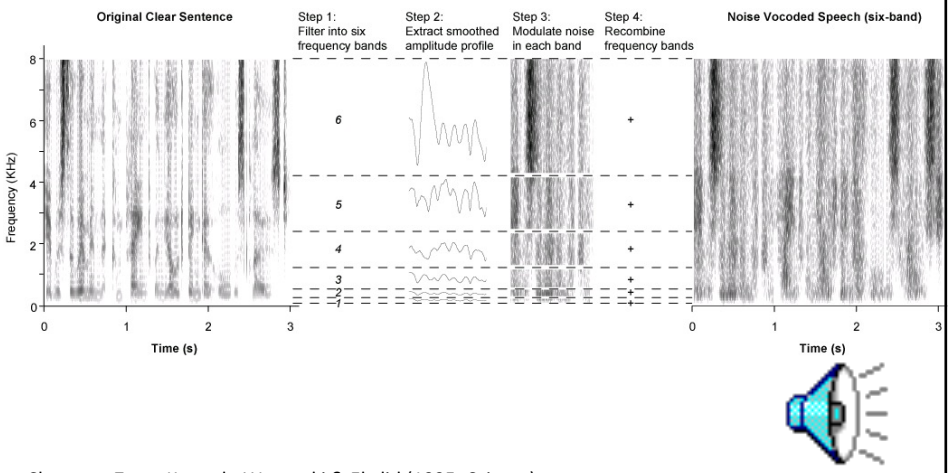


Current Opinion in Neurobiology

Aitchison & Lengyel (2017, *Current Opinion in Neurobiology*)

19

Vocoded Speech
(Simulation of a Cochlear Implant)

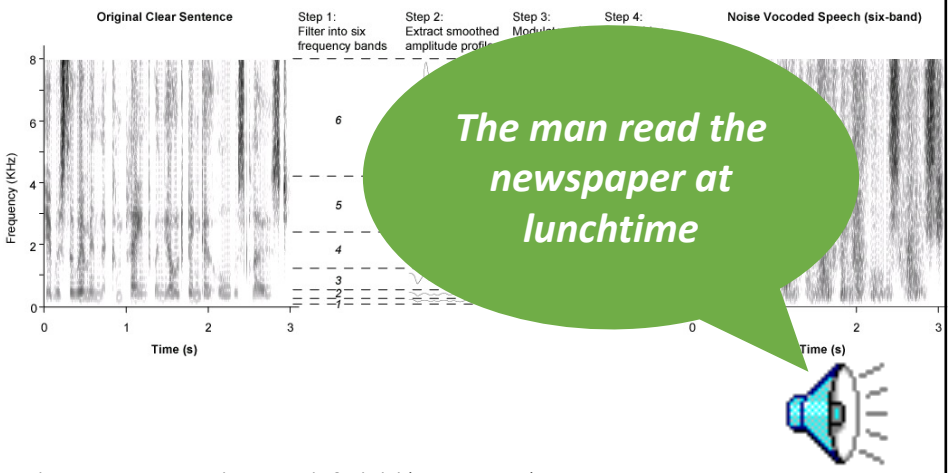


Shannon, Zeng, Kamath, Wygonski & Ekelid (1995, *Science*)

Figure from: Davis et al (2005, *Journal of Experimental Psychology: General*)

21

Vocoded Speech
(Simulation of a Cochlear Implant)



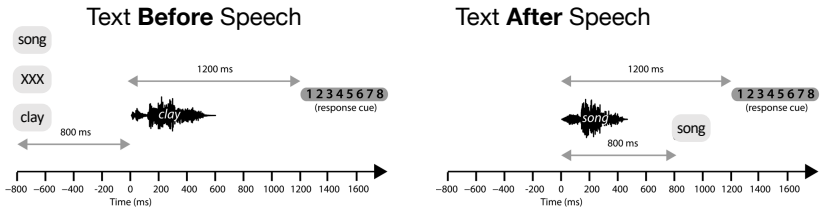
Shannon, Zeng, Kamath, Wygonski & Ekelid (1995, *Science*)

Figure from: Davis et al (2005, *Journal of Experimental Psychology: General*)

22

Prior knowledge enhances speech clarity

Ed Sohoglu



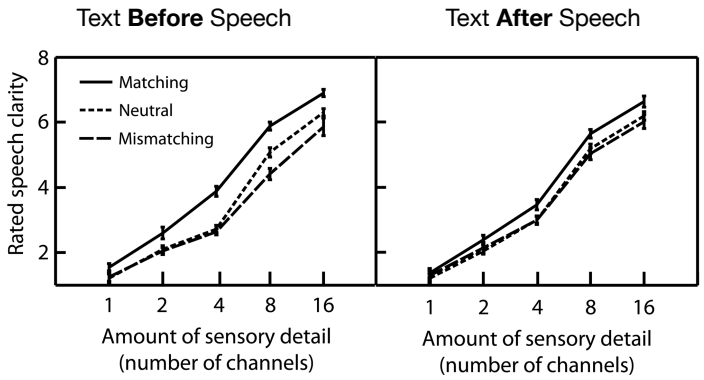
Rate clarity of 1/2/4/8/16-channel vocoded words
Paired with matching/neutral/mismatching text

Sohoglu, Peelle, Carlyon & Davis (2014, *JEP:HPP*)

23

Prior knowledge enhances speech clarity

Ed Sohoglu

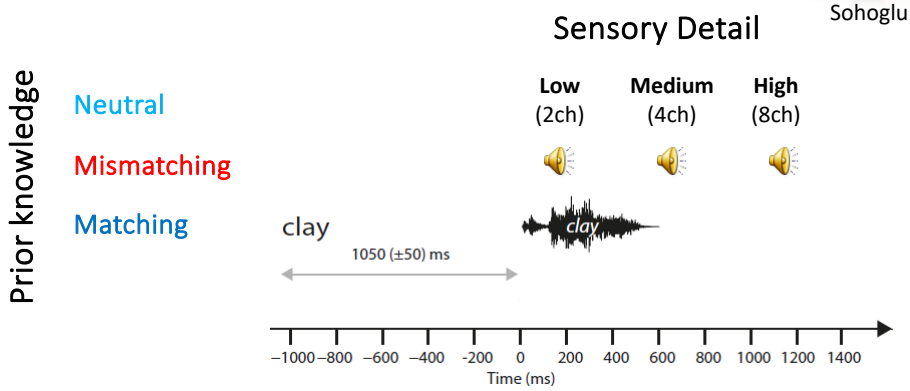


Sohoglu, Peelle, Carlyon & Davis (2014, *JEP:HPP*)

24

Prior knowledge and perception of speech

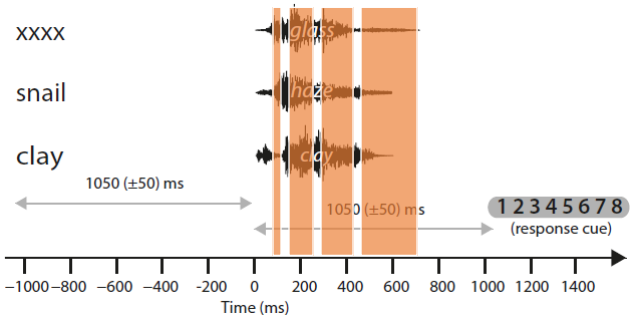
Ed Sohoglu



Sohoglu, Peelle, Carlyon & Davis (2012, *J. Neuroscience;* 2014, *JEP:HPP*)

26

Prior knowledge and perception of speech



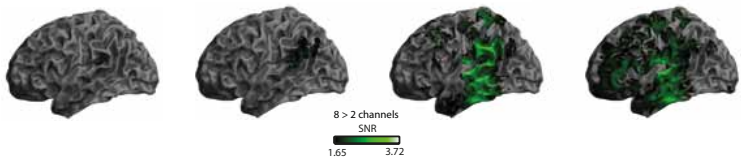
- Sohoglu, Peelle, Carlyon & Davis (2012, *J. Neuroscience*)

27

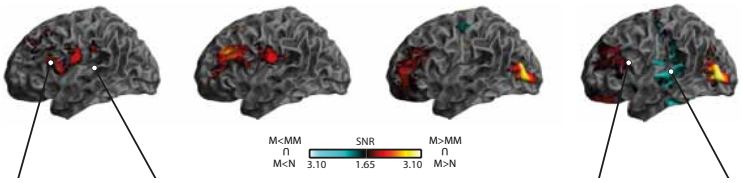
Prior knowledge and perception of speech

90-130 ms 180-240 ms 270-420 ms 450-700 ms

8 channel > 2 channel



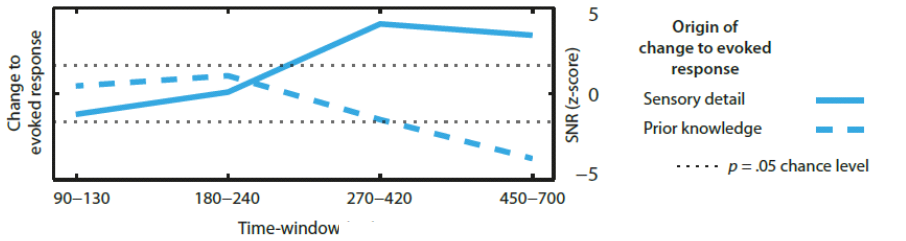
Match > Mismatch | Neutral Match < Mismatch | Neutral

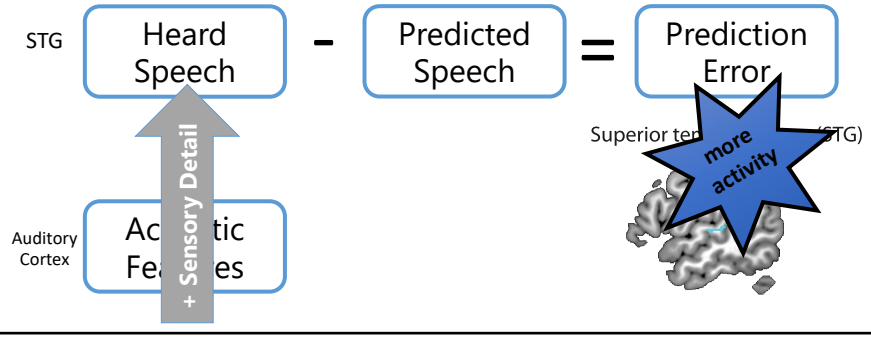


- Sohoglu, Peelle, Carlyon & Davis (2012, *J. Neuroscience*)

28

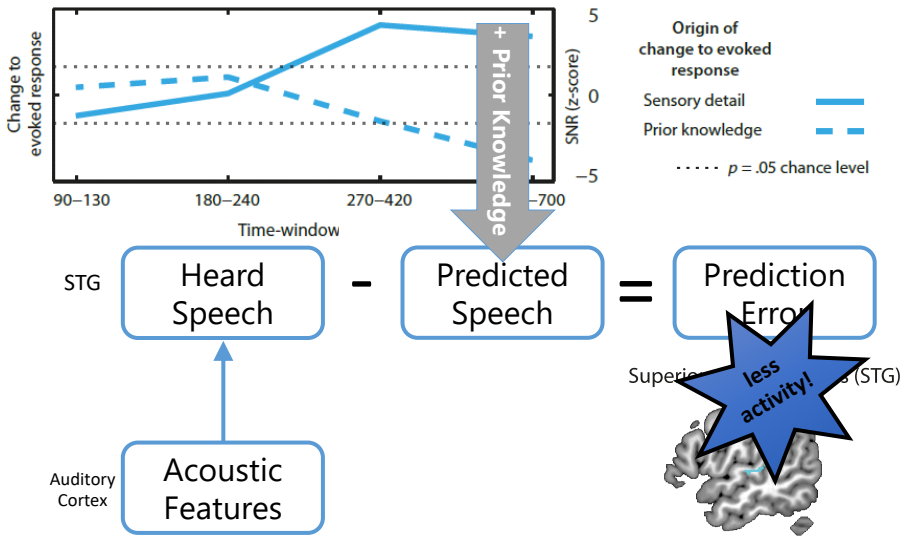
Predictive coding model of speech perception





30

Predictive coding model of speech perception

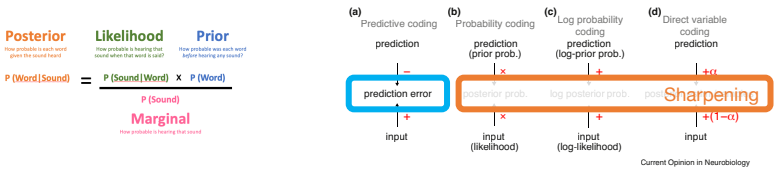


31

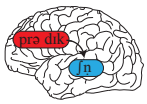
Processing Spoken Words

Localising vs explaining spoken word recognition

Bayesian inference in speech perception



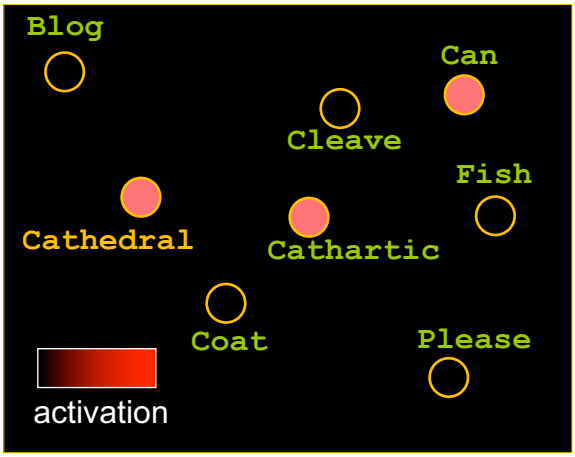
Predictive computations for word recognition

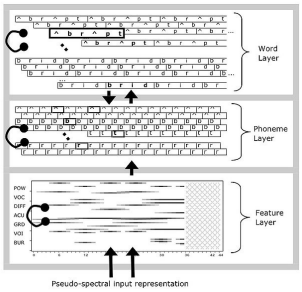


52

Recognising spoken words
(Cohort & TRACE Models)

"ca



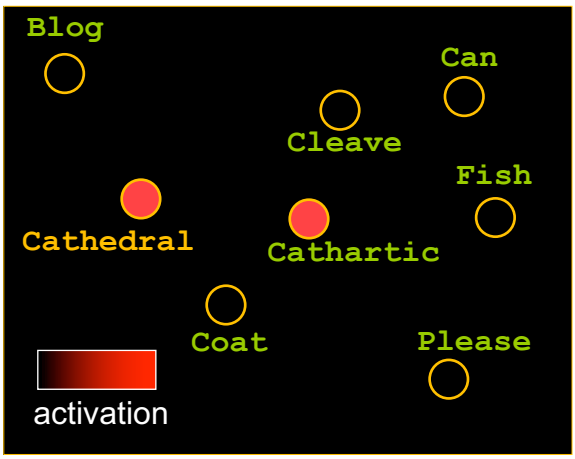


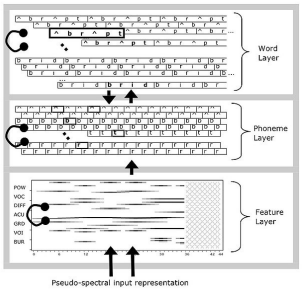
TRACE:
McClelland & Elman (1986, *Cog Psych*)

55

Recognising spoken words
(Cohort & TRACE Models)

"cath





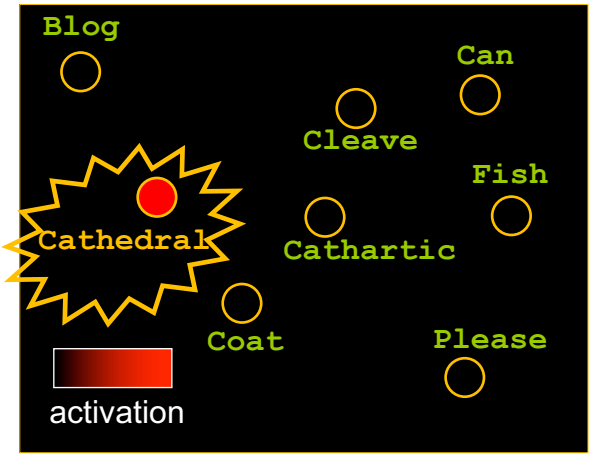
Marslen-Wilson & Tyler (1980, *Phil Trans B*)

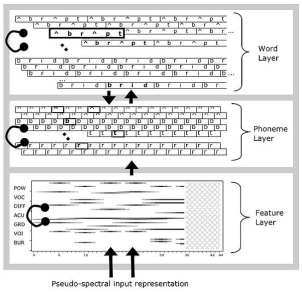
TRACE:
McClelland & Elman (1986, *Cog Psych*)

56

Recognising spoken words
(Cohort & TRACE Models)

"cathe



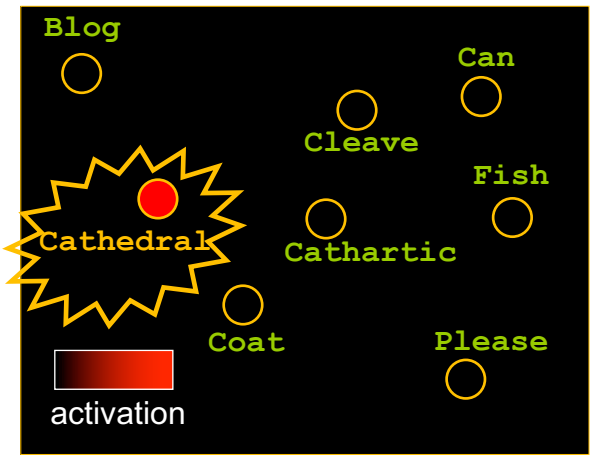


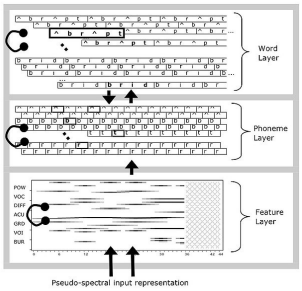
TRACE:
McClelland & Elman (1986, *Cog Psych*)

57

Recognising spoken words
(Cohort & TRACE Models)

"cathedral"



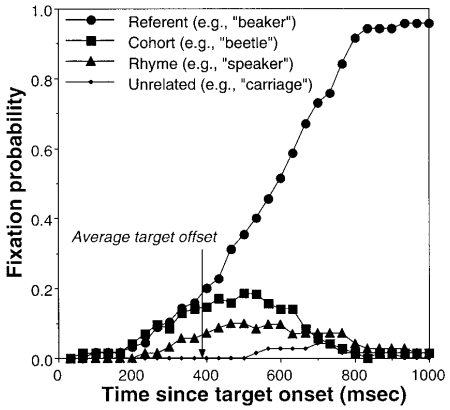


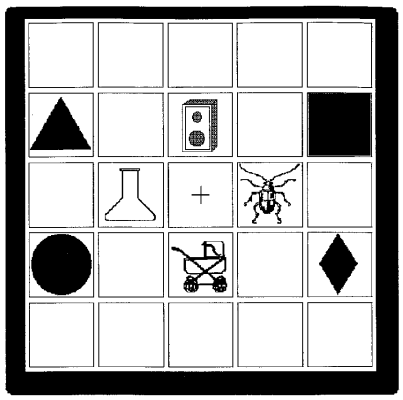
Marslen-Wilson & Tyler (1980, *Phil Trans B*)

TRACE:
McClelland & Elman (1986, *Cog Psych*)

59

Recognising spoken words

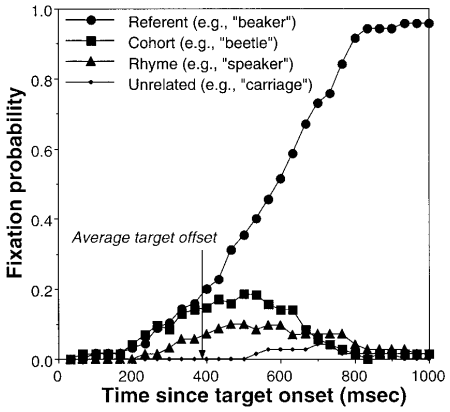


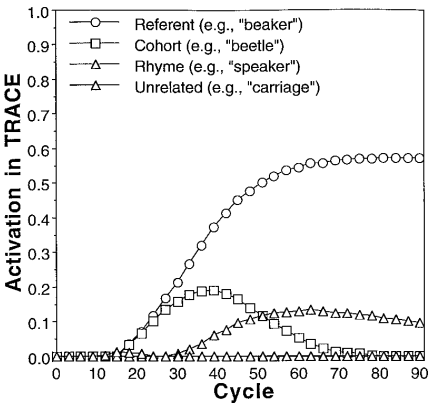


Allopenna, Magnuson & Tanenhaus (1998, JML)

60

Recognising spoken words

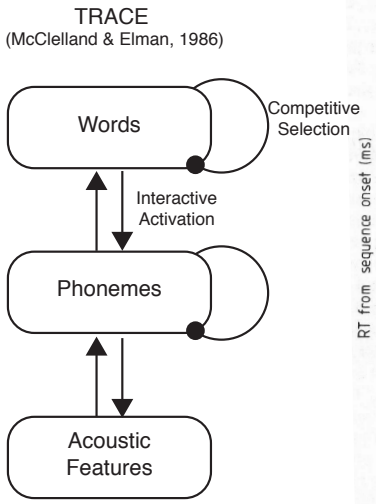


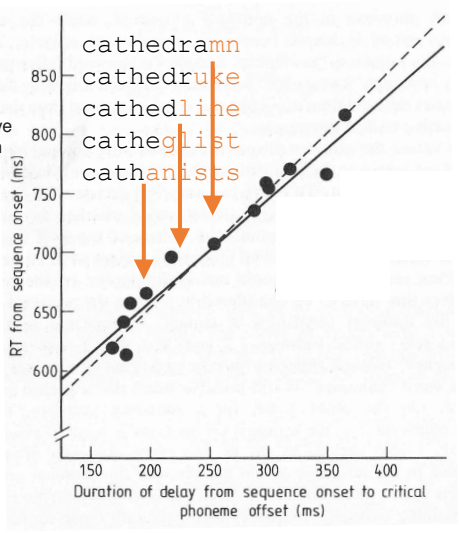


Allopenna Magnuson & Tanenhaus (1998, JML)

61

Competitve vs Predictive Selection

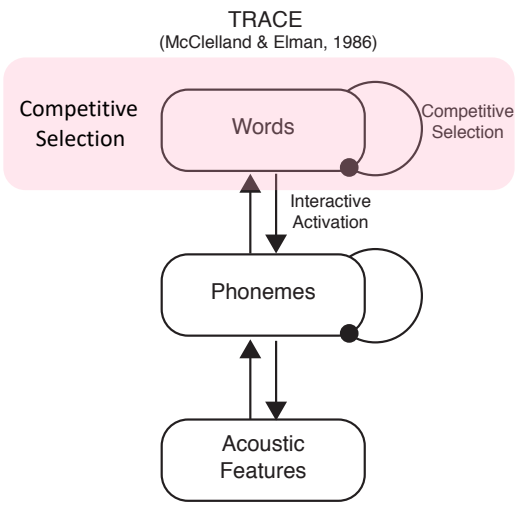


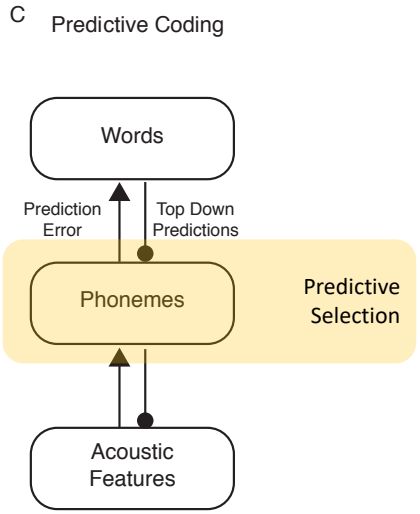


Marslen-Wilson (1984, Attention & Performance X)
Marslen-Wilson (1987, Cognition)

62

Sharpening vs Predictive Coding

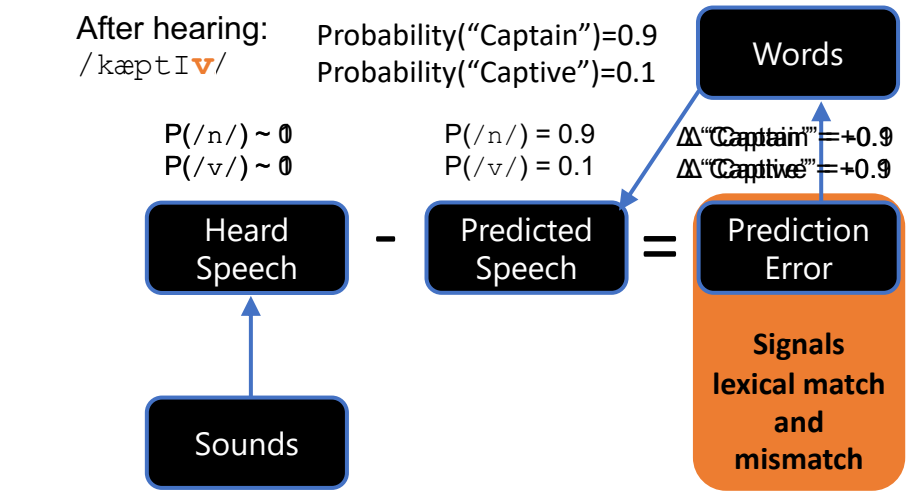




Davis & Sohoglu (2020 *Cog Neurosci 6, MIT Press)*
https://psyarxiv.com/qc4u6/

64

Predictive coding and word recognition



Gagnepain, Henson & Davis (2012, *Current Biology*)

67

Speech predictions change with learning

Hygiene
Hygiene
Hygiene ⇨ Expected Sounds

Hijack
Hijous ⇨ Heard Sounds

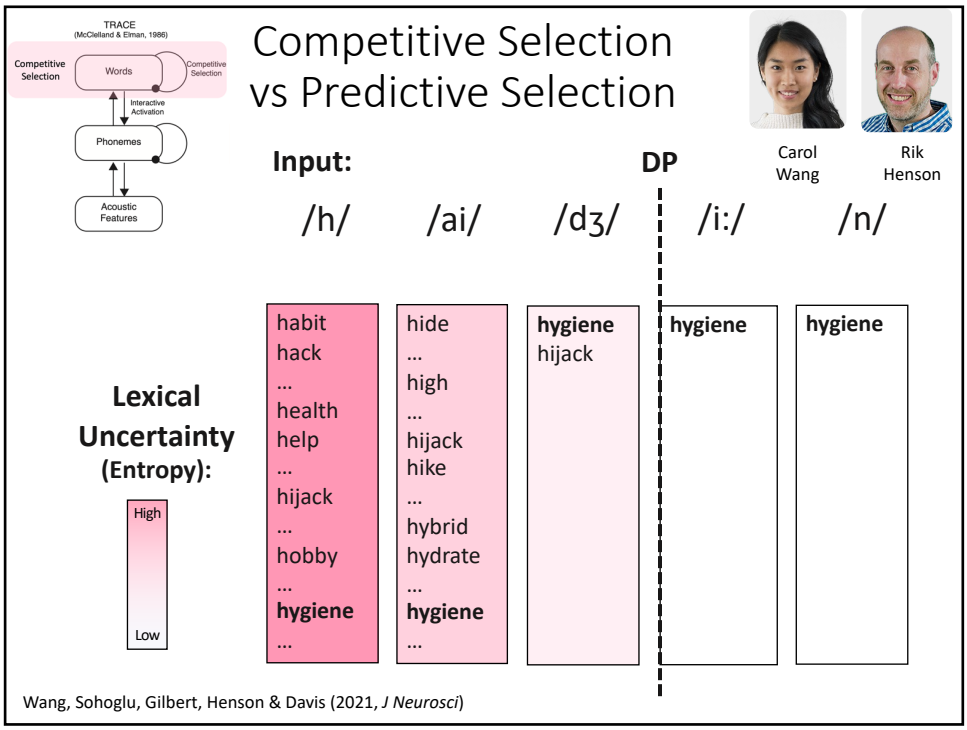
Prediction Error

Delays word recognition (Monsell & Hirsh, 1998)

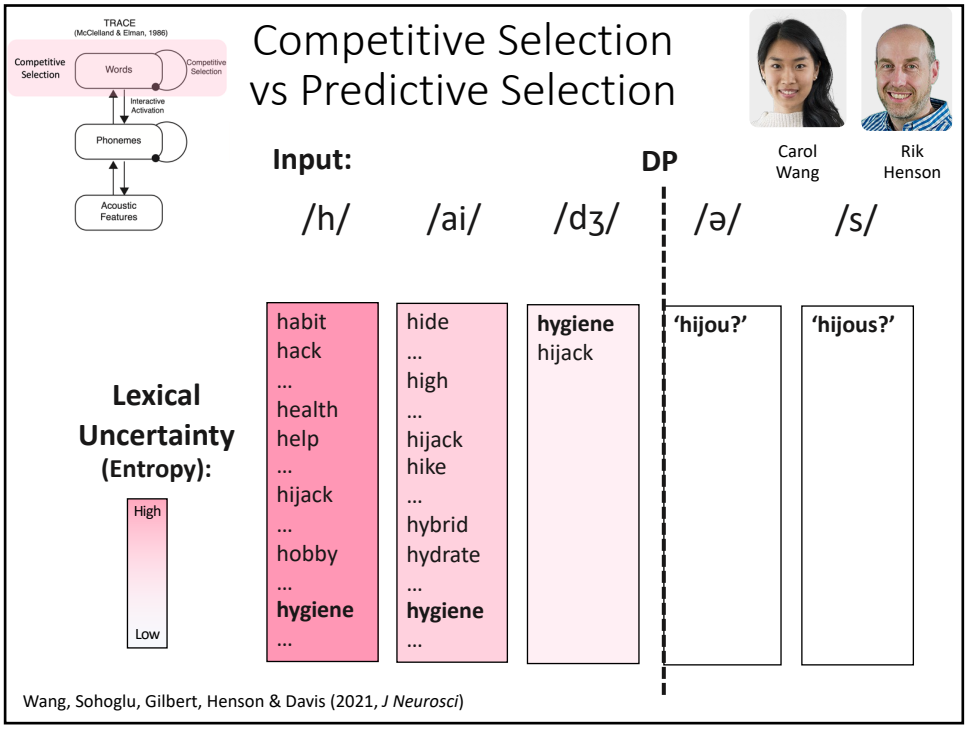
Predictive Selection (cf. Davis & Sohoglu, 2020)

Localise lexical prediction error to STG regions

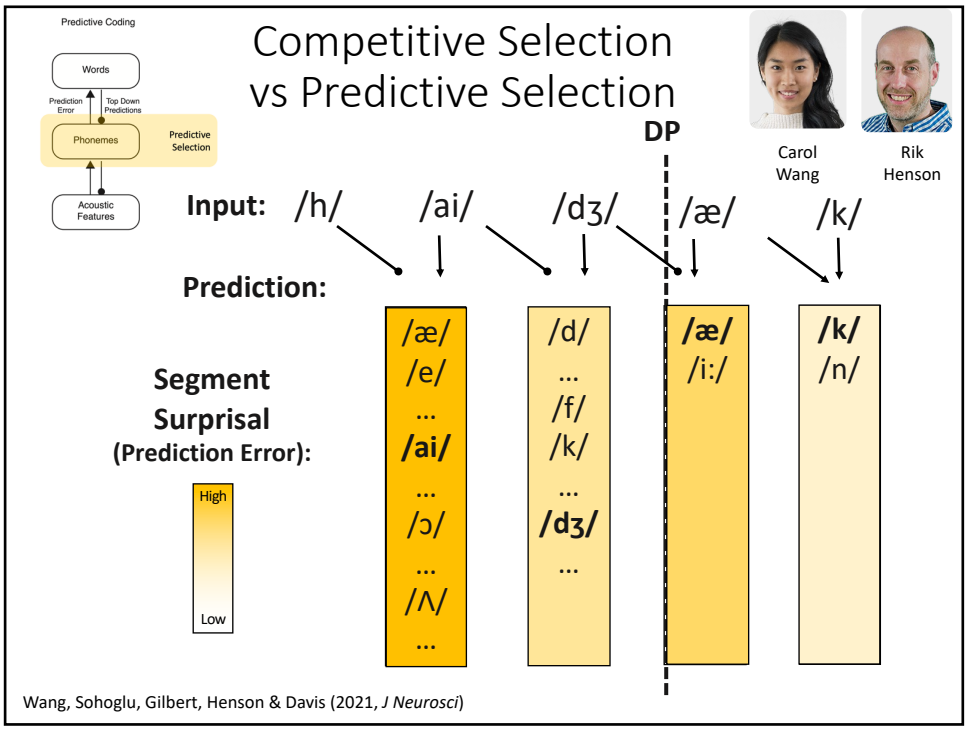
68


TRACE
(McClelland & Elman, 1986)
Competitive Selection
Words
Competitive Selection
Interactive Activation
Phonemes
Acoustic Features
Competitive Selection vs Predictive Selection
Carol Wang
Rik Henson
Input:
DP
/h/ /ai/ /dʒ/ /i:/ /n/
habit hack … health help … hijack … hobby … hygiene …
hide … high … hijack hike … hybrid hydrate … hygiene …
hygiene hijack
hygiene
hygiene
Lexical Uncertainty (Entropy):
High
Low
Wang, Sohoglu, Gilbert, Henson & Davis (2021, J Neurosci)

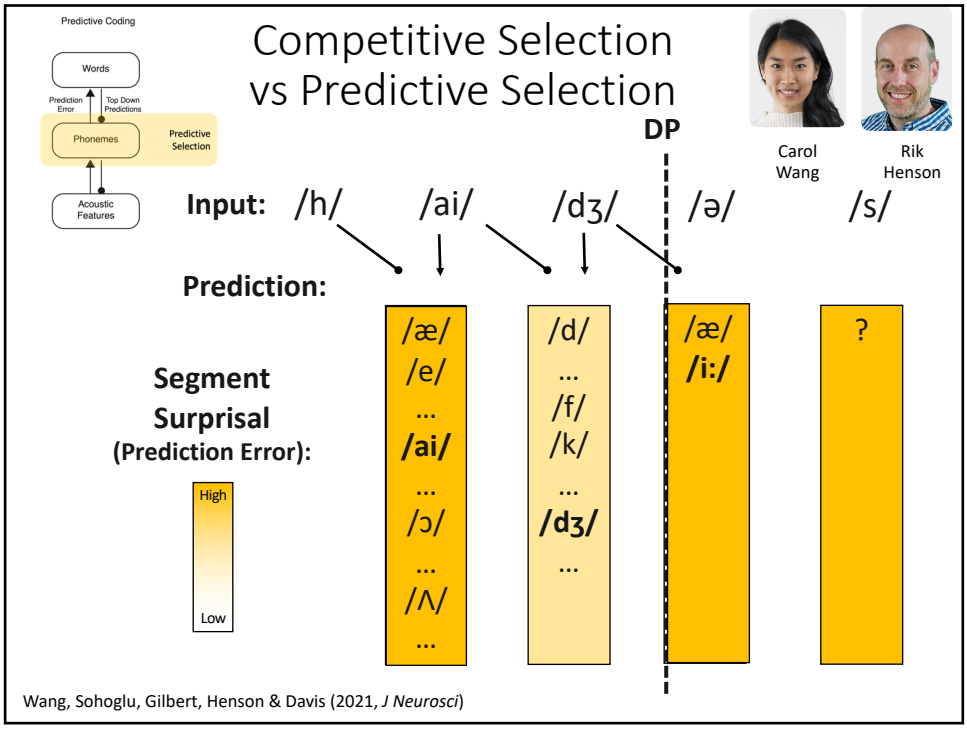
69


TRACE
(McClelland & Elman, 1986)
Competitive Selection
Words
Competitive Selection
Interactive Activation
Phonemes
Acoustic Features
Competitive Selection vs Predictive Selection
Carol Wang
Rik Henson
Input:
DP
/h/ /ai/ /dʒ/ /ə/ /s/
habit hack … health help … hijack … hobby … hygiene …
hide … high … hijack hike … hybrid hydrate … hygiene …
hygiene hijack
'hijou?'
'hijous?'
Lexical Uncertainty (Entropy):
High
Low
Wang, Sohoglu, Gilbert, Henson & Davis (2021, J Neurosci)

70


Predictive Coding
Words
Prediction Error
Top Down Predictions
Phonemes
Predictive Selection
Acoustic Features
Competitive Selection
vs Predictive Selection
DP
Carol Wang
Rik Henson
Input: /h/ /ai/ /dʒ/ /æ/ /k/
Prediction:
Segment Surprisal (Prediction Error):
High
Low
/æ/ /e/ ... /ai/ ... /ɔ/ ... /ʌ/ ...
/d/ ... /f/ /k/ ... /dʒ/ ...
/æ/ /i:/
/k/ /n/
Wang, Sohoglu, Gilbert, Henson & Davis (2021, J Neurosci)

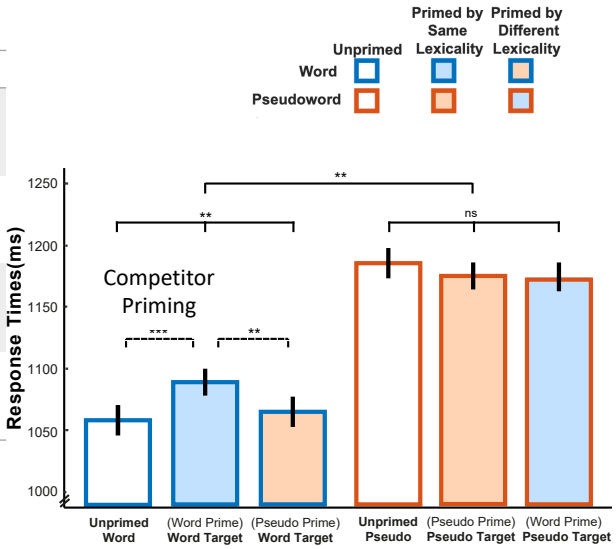
72


Predictive Coding
Words
Prediction Error
Top Down Predictions
Phonemes
Predictive Selection
Acoustic Features
Competitive Selection
vs Predictive Selection
DP
Carol Wang
Rik Henson
Input: /h/ /ai/ /dʒ/ /ə/ /s/
Prediction:
Segment Surprisal (Prediction Error):
High
Low
/æ/ /e/ ... /ai/ ... /ɔ/ ... /ʌ/ ...
/d/ ... /f/ /k/ ... /dʒ/ ...
/æ/ /i:/
?
Wang, Sohoglu, Gilbert, Henson & Davis (2021, J Neurosci)

73

Competitor Priming for Word Pairs (Behaviour)

Pairs separated by 20-80 trials (1-4 mins)	
Unprimed Word hygiene	(Word Primed) Word hijack
Unprimed Pseudo letto	(Pseudo Primed) Peeudo lettan
Unprimed Pseudo basef	(Pseudo primed) Word Basic
Unprimed Word boycott	(Word primed) Pseudo boyten

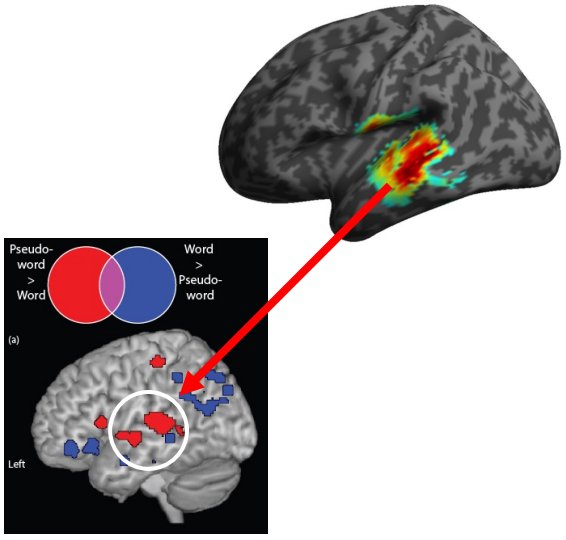


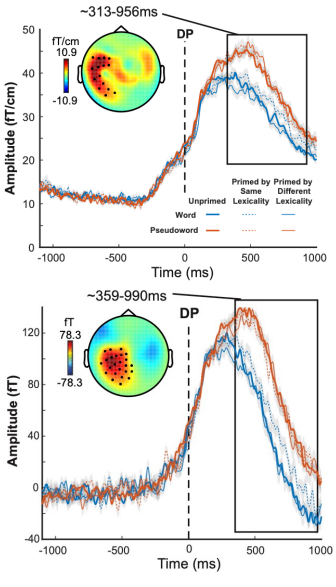
Wang, Sohoglu, Gilbert, Henson & Davis (2021, *J Neurosci*)

74

Competitor Priming for Word Pairs overlaps with Pseudo > Word (MEG)

STG Source of Lexicality Effect



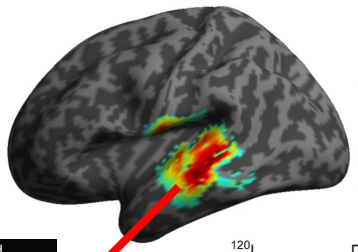


Wang, Sohoglu, Gilbert, Henson & Davis (2021, *J Neurosci*)

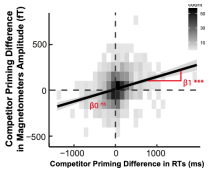
75

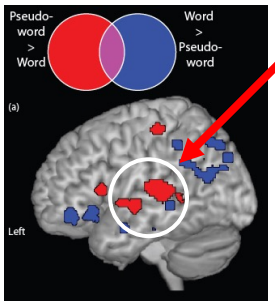
Competitor Priming for Word Pairs overlaps with Pseudo > Word (MEG)

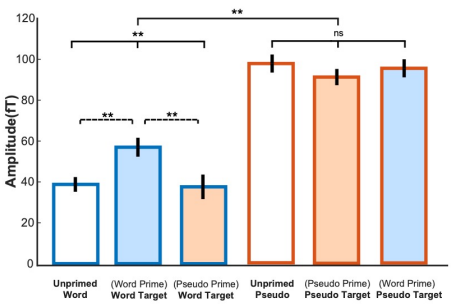
STG Source of Lexicality Effect



Trial-wise correlation of neural and behavioural priming





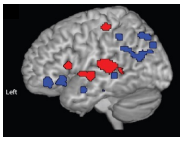


Wang, Sohoglu, Gilbert, Henson & Davis (2021, *J Neurosci*)

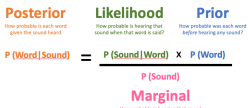
76

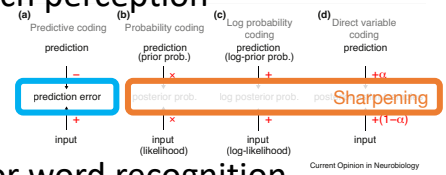
Processing Spoken Words

Localising vs explaining spoken word recognition



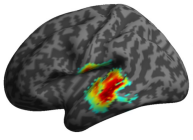
Bayesian inference in speech perception





Predictive computations for word recognition

Heard Speech - Predicted Speech = Prediction Error



79

Thank you!


Jo Taylor (UCL)


Ed Sohoglu (Sussex)


Carol Wang (ex-CBU)


Gareth Gaskell (York)


Pierre Gagnepain (Caen, FR)


Rik Henson (MRC-CBU)

UNIVERSITY OF CAMBRIDGE

UKRI Medical Research Council

@MattDavis@fediverse.science

Localising vs explaining spoken word recognition:

C.J. Price (2000, *J Anatomy*): Mapping box-and-arrow models of word recognition onto brain systems
J.L. McClelland & D.E. Rumelhart (1981, *Psych Review*): Interactive activation model of visual word recognition
J.S.H. Taylor, K. Rastle & M. H. Davis (2013, *Psych Bulletin*): Engagement and effort principles link word recognition models to brain activity
M.H. Davis & M.G. Gaskell (2009, *Phil Trans Roy Soc B*): Review and imaging meta-analysis of spoken word recognition & learning

Bayesian inference in speech perception:

D. Norris & J.M. McQueen (2008, *Psych Review*) Shortlist B, A Bayesian model of spoken word recognition
J.L. McClelland & J. L. Elman (1986, *Cognitive Psych)*: TRACE model of speech perception
R. V. Shannon et al (1995, *Science*): Introduced noise-vocoded speech
E. Sohoglu et al (2012, *J Neuroscience*): Opposite effects of prior knowledge & clarity on MEG responses suggesting predictive processing

Predictive computations for word recognition:

W.D. Marslen-Wilson & L.K. Tyler (1980, *Phil Trans Roy Soc B*): Cohort model of spoken word recognition
P. D. Allopenna, J.S. Magnuson, M.K. Tanenhaus (1998, *J Memory & Language*): Competitor dynamics during spoken word recognition
M.H. Davis & E. Sohoglu (2020, *Cognitive Neuroscience 6*): Three functions of prediction error for Bayesian inference in speech perception
Y. C. Wang et al (2021, *J Neuroscience*): Prediction error computations during competitor priming shown by behaviour linked to STG responses

80

Additional References

- RSA and decoding methods can further test sharpening vs Prediction error theories of spoken word recognition

fMRI: Blank & Davis (2016, PLoS Biology) https://doi.org/10.1371/journal.pbio.1002577
Blank et al (2018, J Neuroscience) https://doi.org/10.1523/JNEUROSCI.3258-17.2018

MEG: Sohoglu & Davis (2020, eLife) https://doi.org/10.7554/eLife.58077
Sohoglu, Beckers & Davis (in press, Nature Comms) https://www.biorxiv.org/content/10.1101/2023.10.03.560649v3

- Probabilistic predictors of neural activity for speech

Gwilliams, L.E., Davis, M.H. (2022) Extracting language content from speech sounds: The information theoretic approach. in Holt, Peelle et al (eds) Auditory Cognitive Neuroscience of Speech Perception: Springer Handbook of Auditory Research, Vol 74. Springer

https://doi.org/10.1007/978-3-030-81542-4_5

https://lauragwilliams.github.io/d/m/Gwilliams_Davis_2021.pdf

81