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

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.”

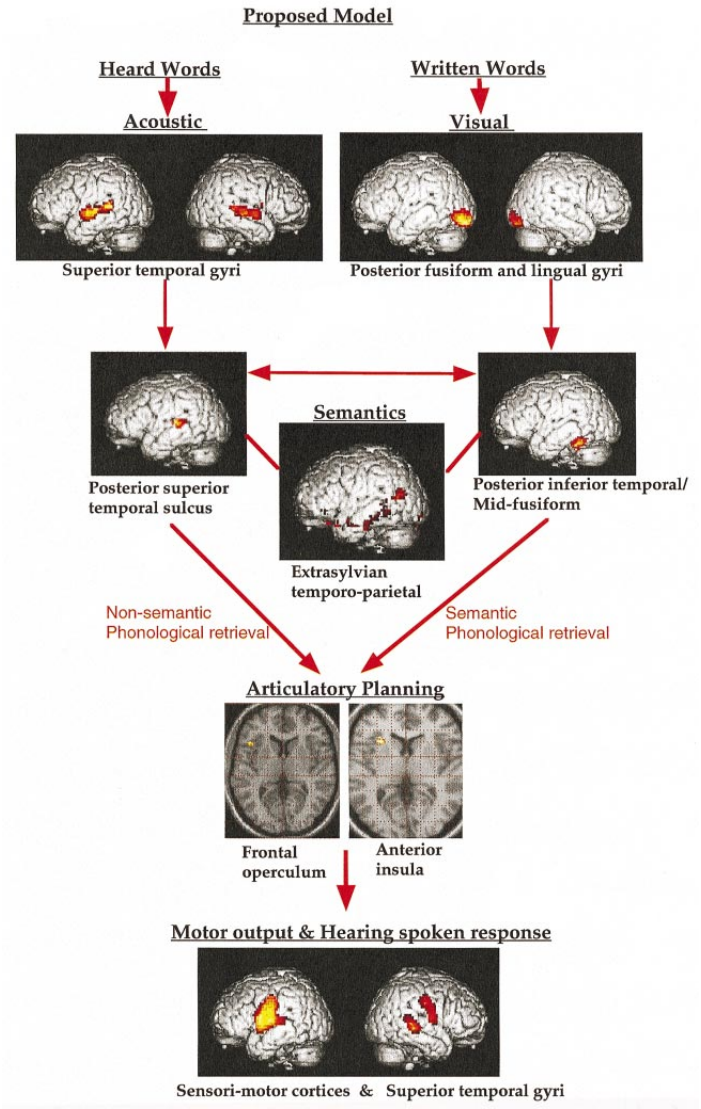
Processing Spoken Words

Localising vs explaining spoken word recognition

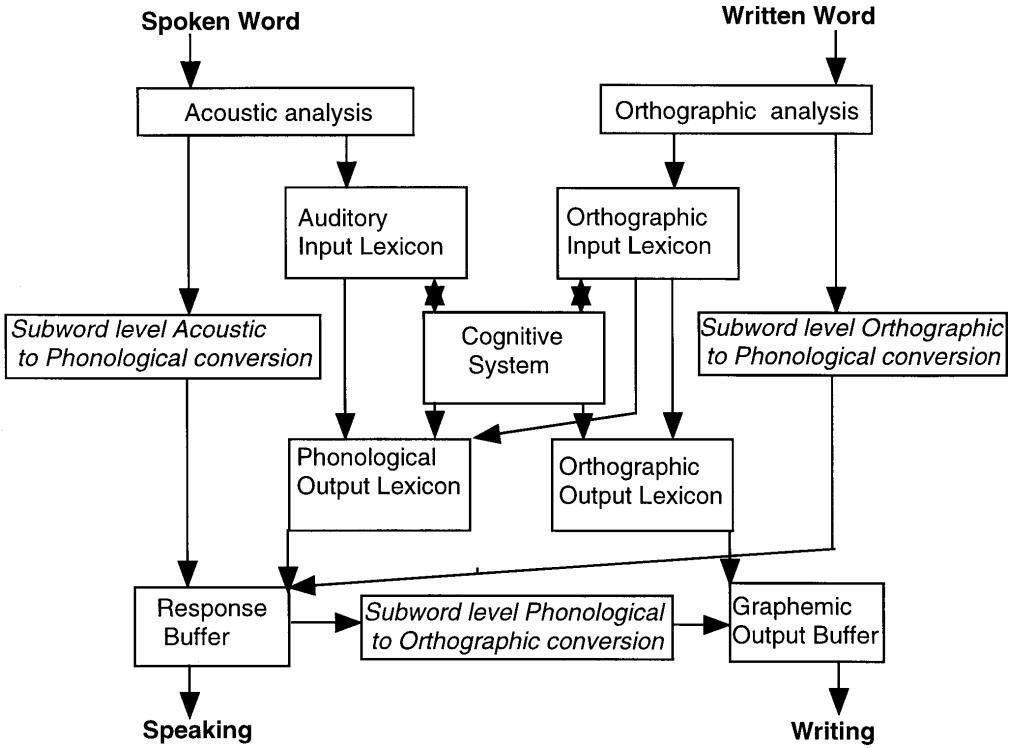
Bayesian inference in speech perception

Predictive computations for word recognition

Box & Arrow Models of Word Recognition



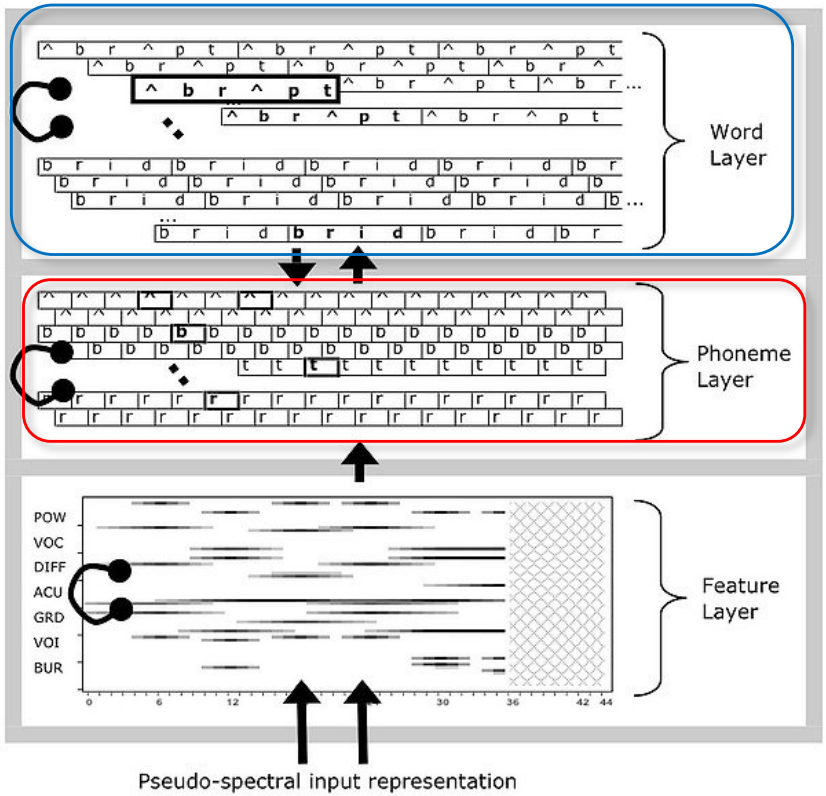
Price (2000, *J Anatomy*)



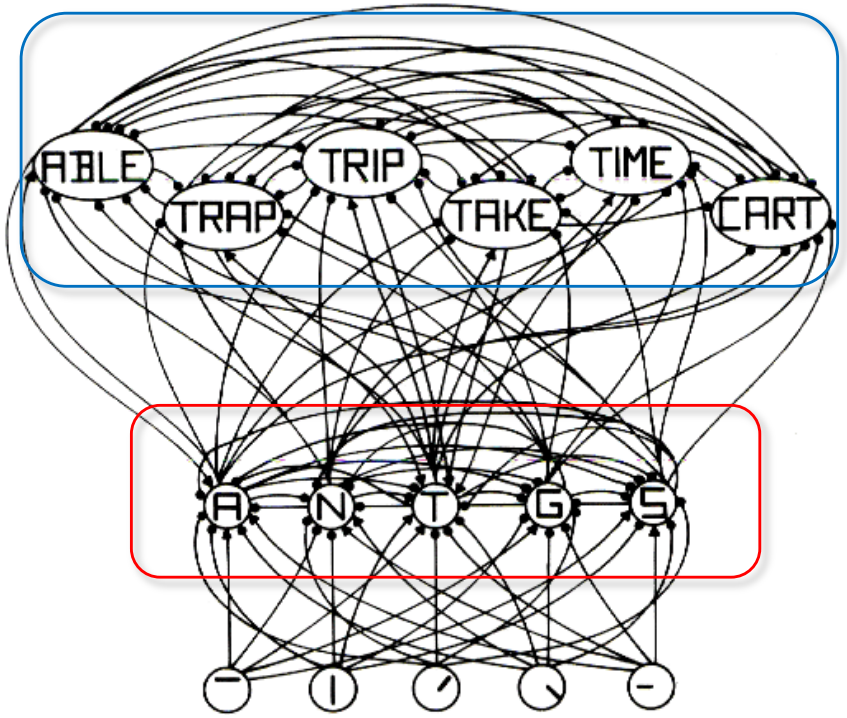
Morton (1969) Logogen Model
Patterson & Shewell (1987)

Computational accounts of Lexical Processing

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

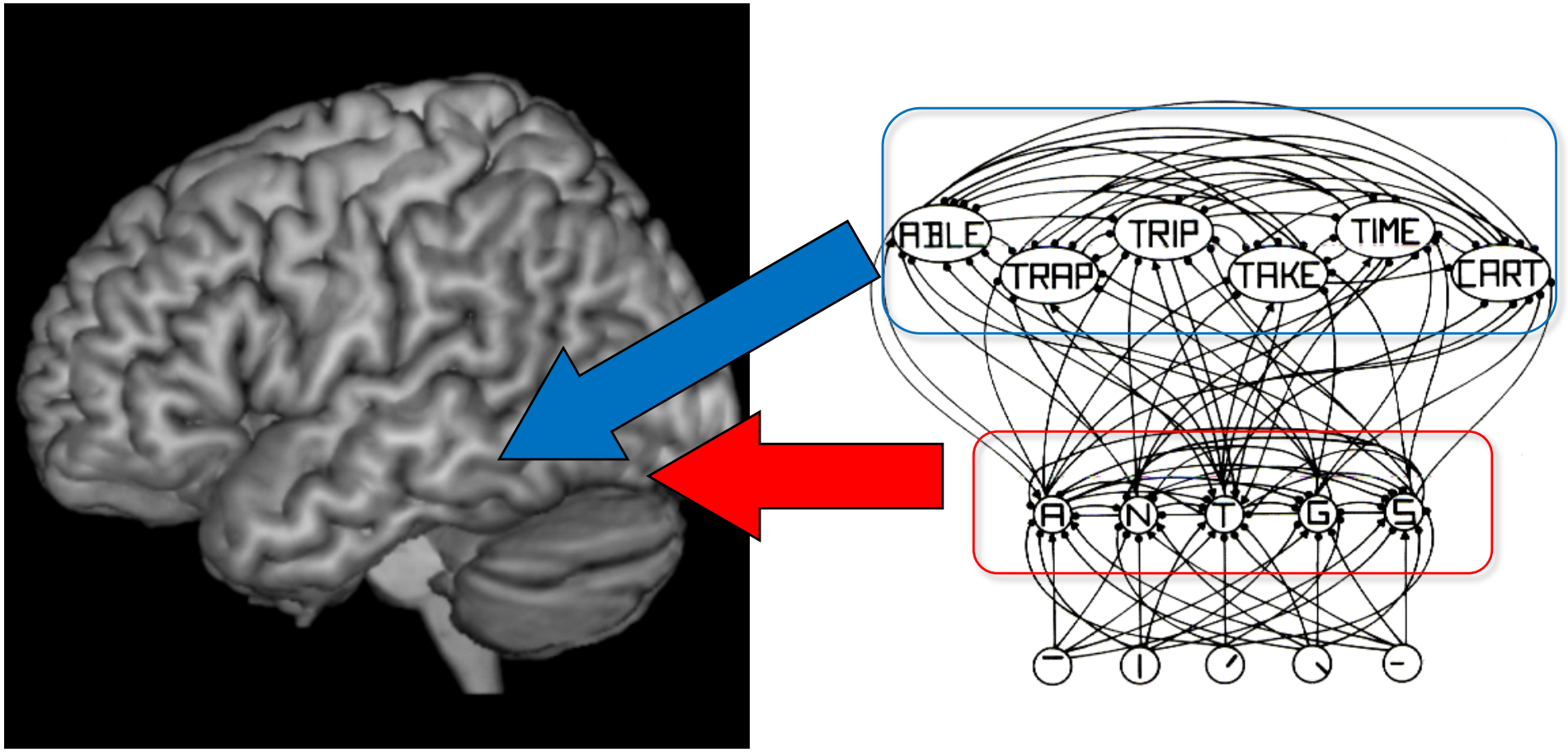


Interaction Activation Model of letter perception (McClelland & Rumelhart, 1981)



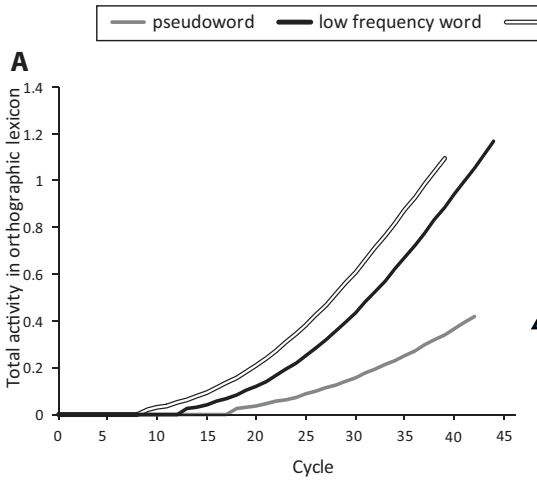
Mapping Computational Accounts onto the Brain

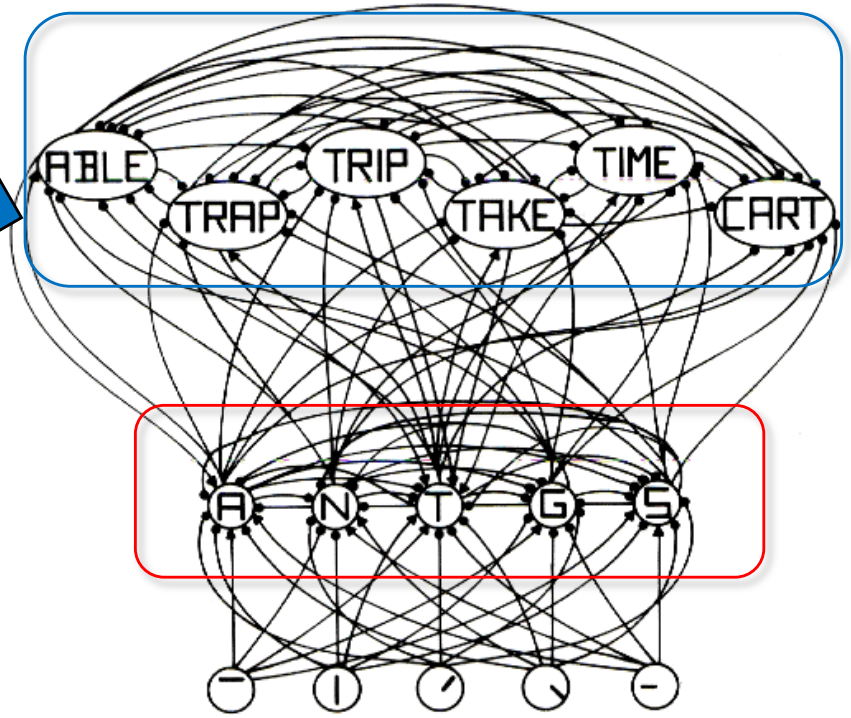
Interactive Activation Model
of Letter Perception
(McClelland & Rumelhart, 1981)



Mapping Computational Accounts onto the Brain

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



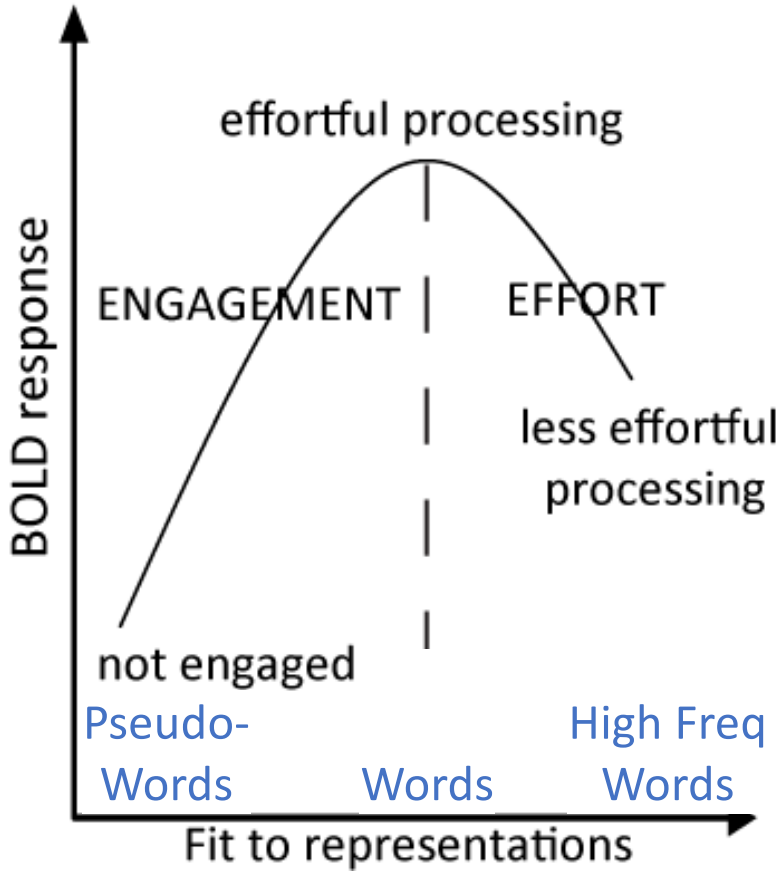


Taylor, Rastle, & Davis (2013)
Psychological Bulletin

Linking Computational Accounts to the Brain

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

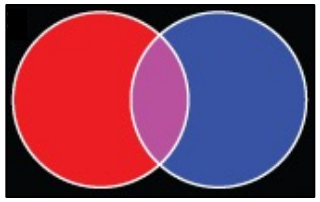
2. **Effort:**
   Stimuli that are a good fit to representations lead to less effort during neural processing
   (e.g. low > high frequency words)



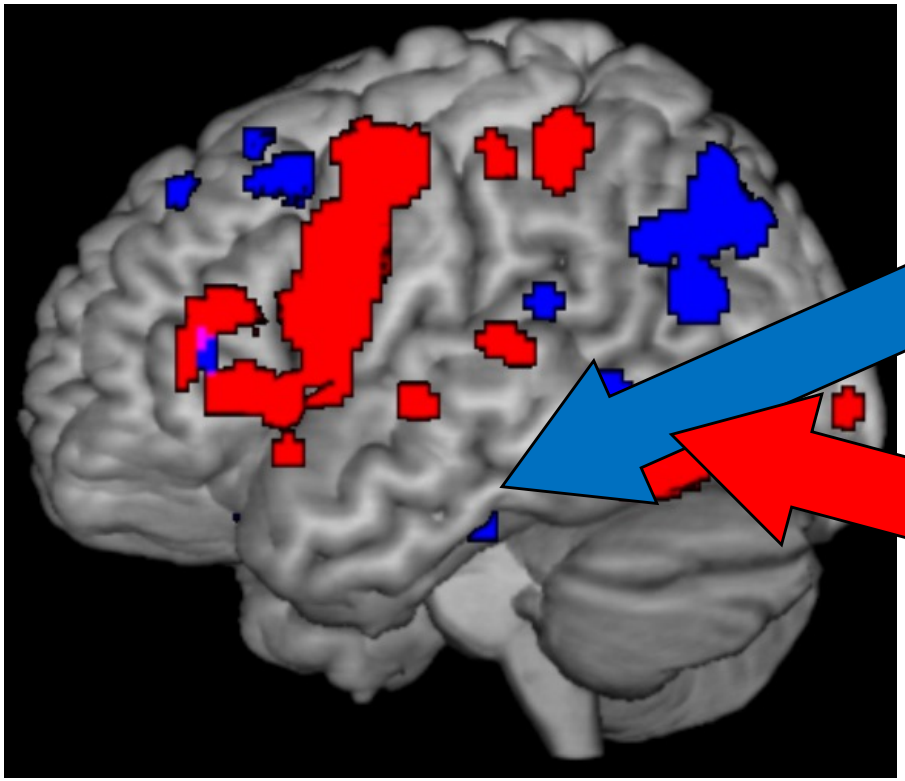
Taylor, Rastle, & Davis (2013)
Psychological Bulletin

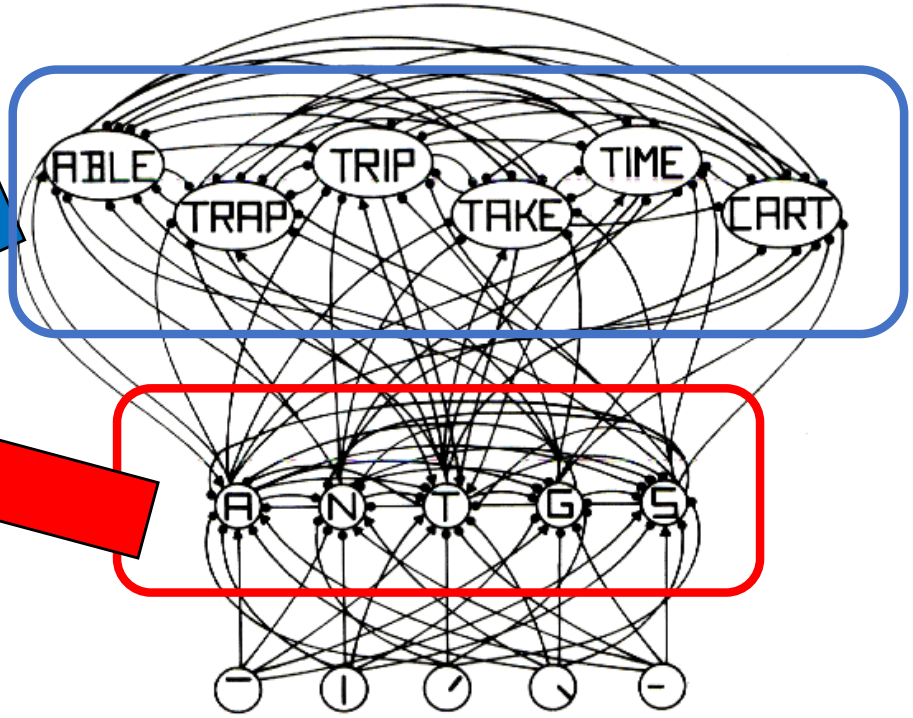
fMRI Meta-analysis: Written Words vs Pseudowords

pseudowords > words



words > pseudowords

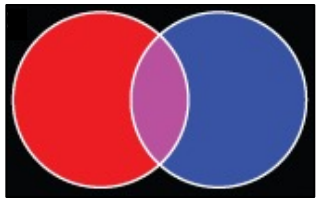




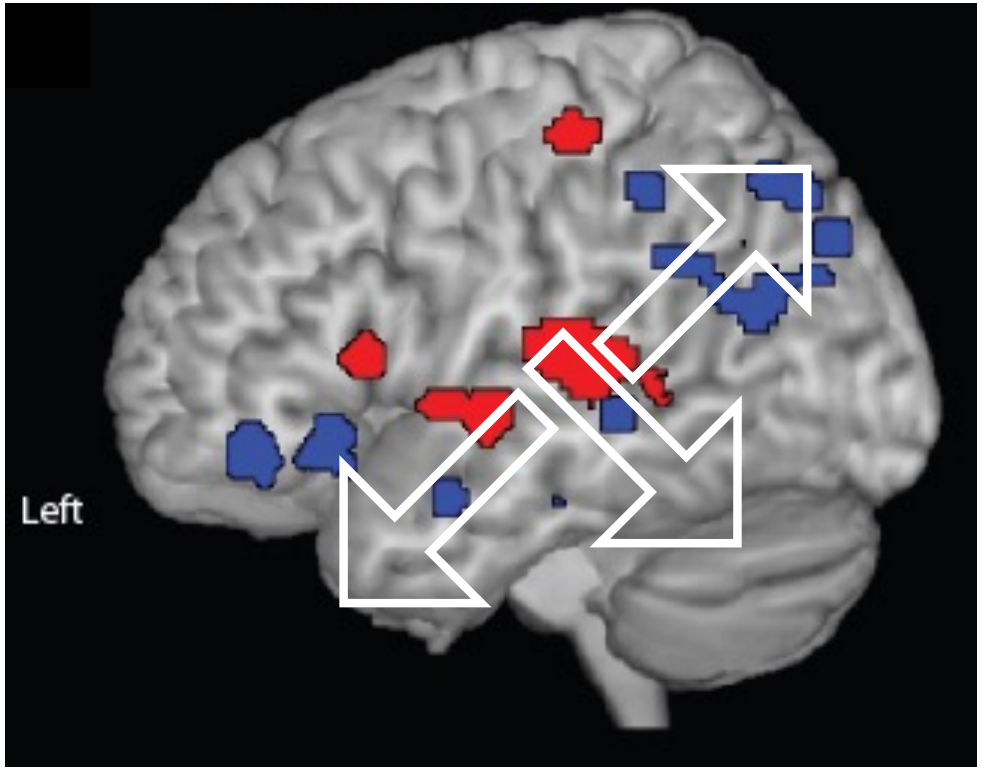
Taylor, Rastle, & Davis (2013)
Psychological Bulletin

fMRI Meta-analysis: Spoken Words vs Pseudowords

pseudowords > words

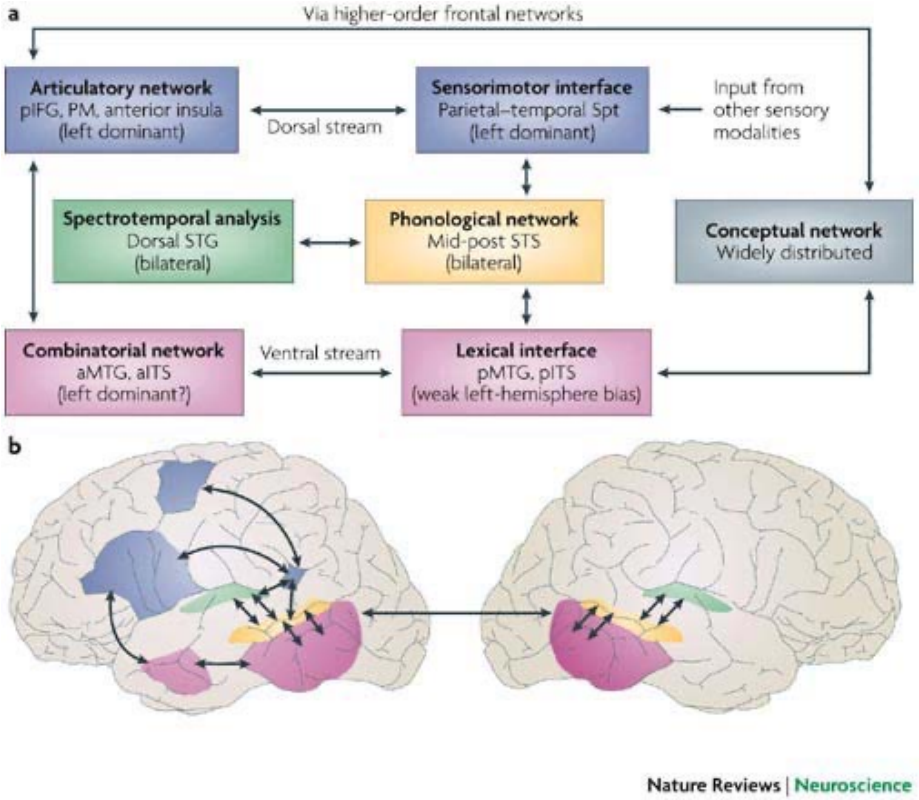


words > pseudowords



Davis & Gaskell (2013) Phil Trans Roy Soc B

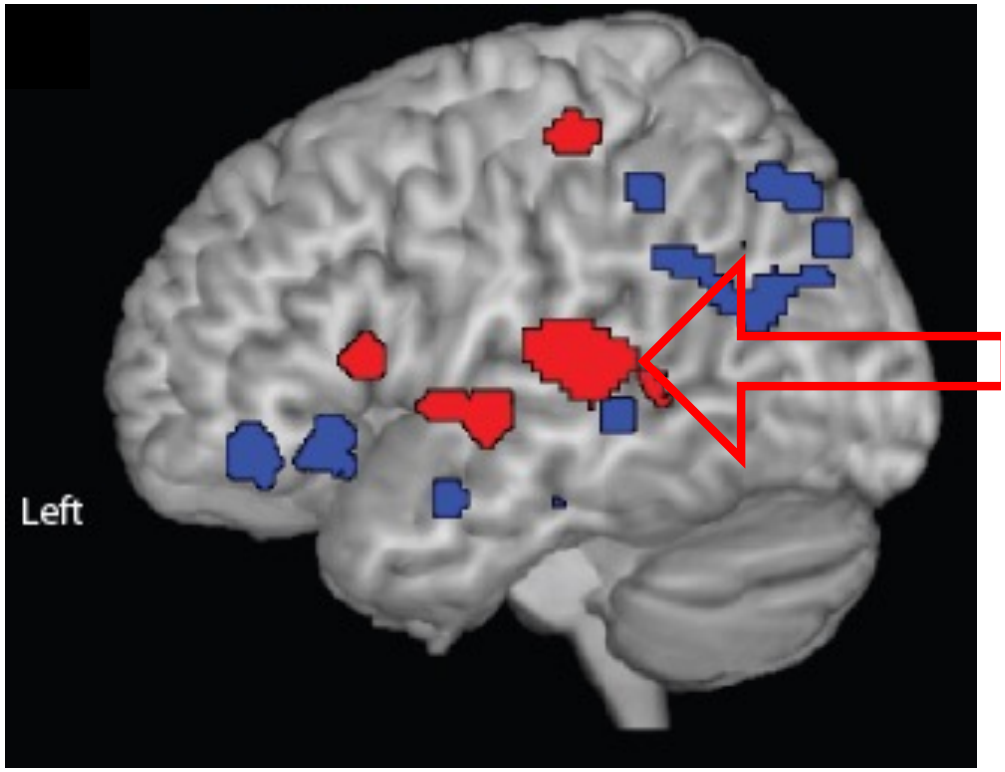
Dorsal vs Ventral Pathways

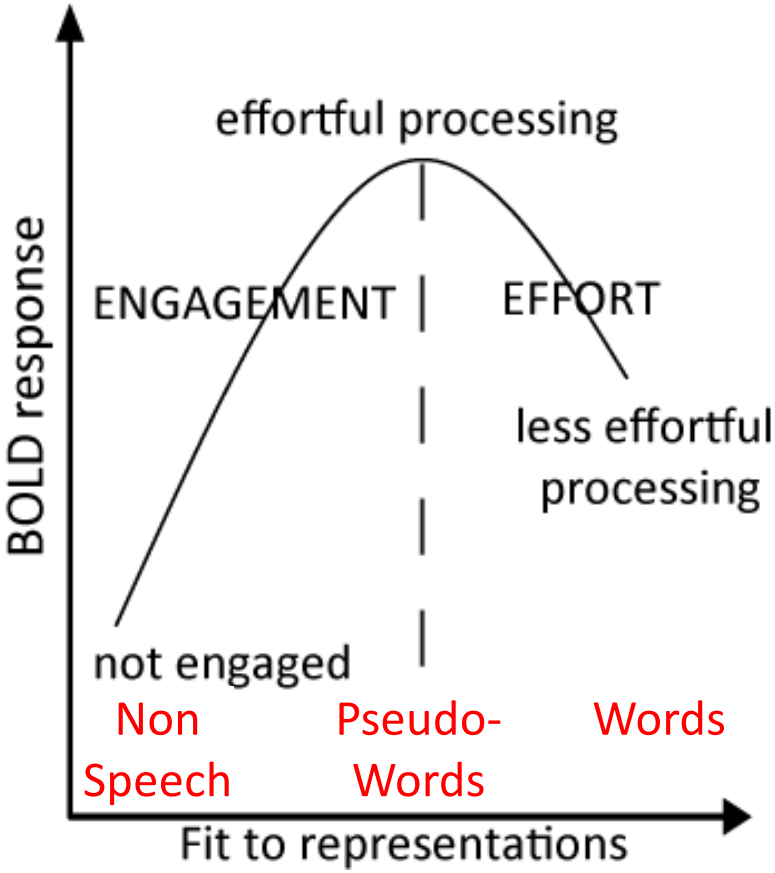


Hickok & Poeppel
(2007, *Nature Reviews Neuroscience*)

Processing Spoken Words

Localising vs explaining spoken word recognition





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

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?

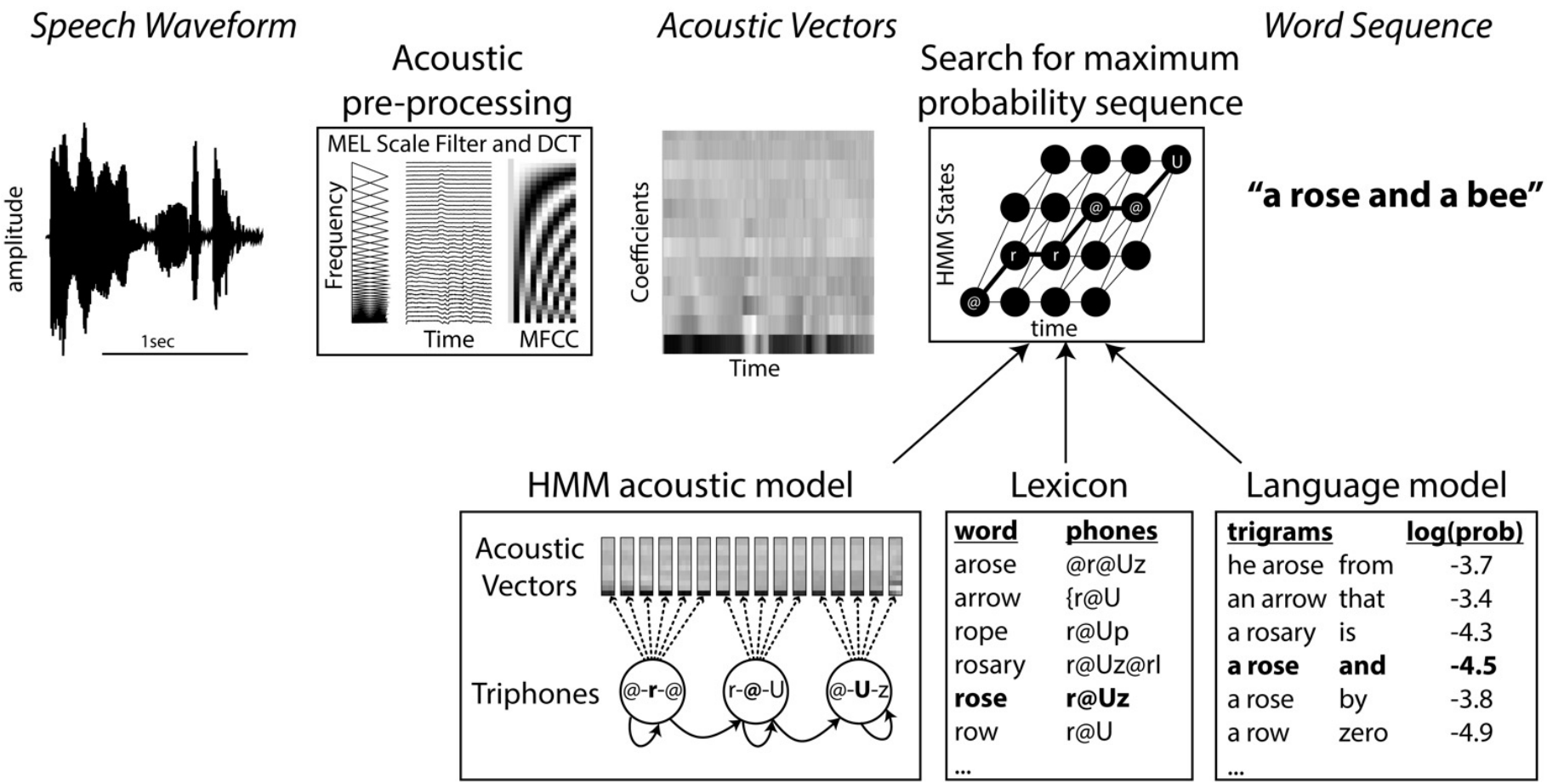
$$P(\text{Word}|\text{Sound}) = \frac{P(\text{Sound}|\text{Word}) \times P(\text{Word})}{P(\text{Sound})}$$

Marginal
How probable is hearing that sound

Thomas Bayes
1701-1761

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

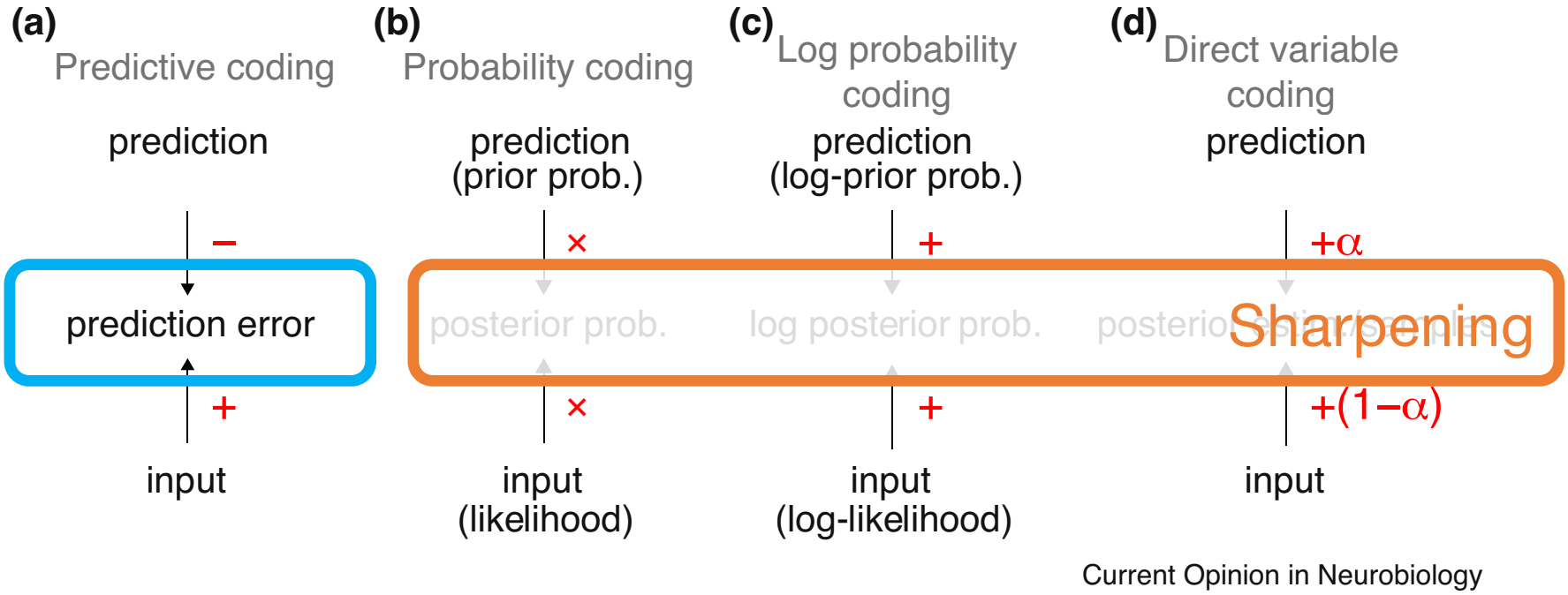
Speech perception by machines



Traditional ASR System

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

Neural Implementations of Bayesian Inference



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

Vocoded Speech

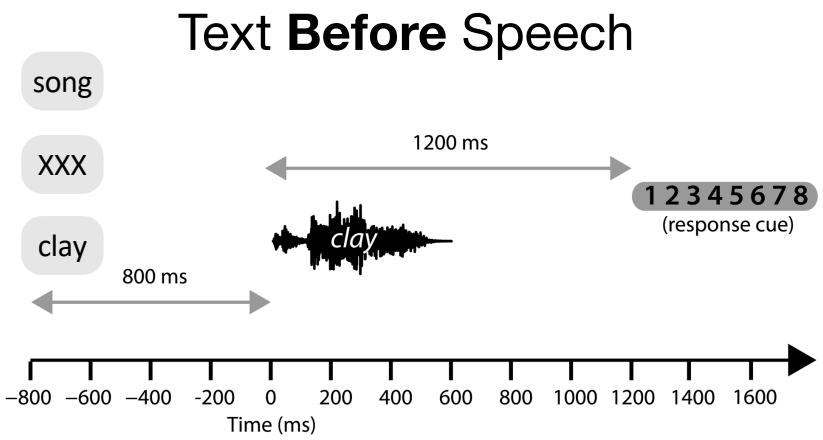
(Simulation of a Cochlear Implant)

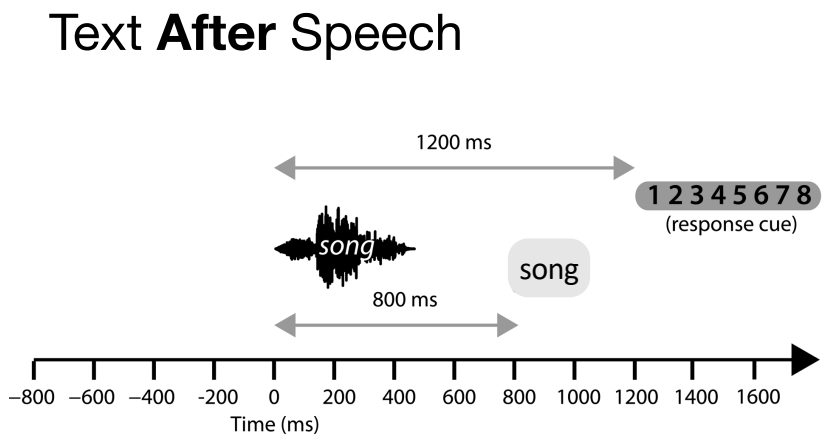
Shannon, Zeng, Kamath, Wygonski & Ekelid (1995, *Science*)

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

Prior knowledge enhances speech clarity

Ed Sohoglu



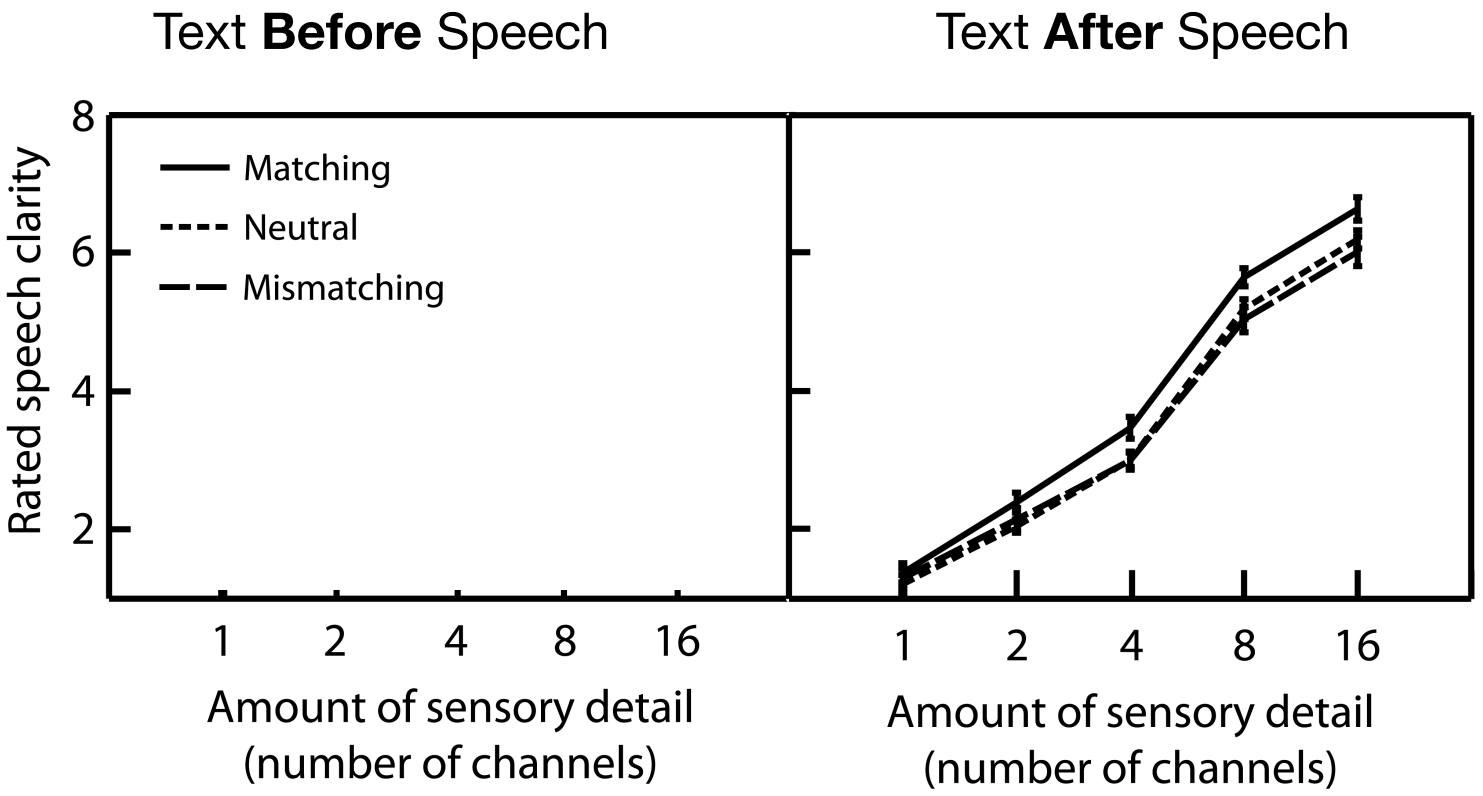


Rate clarity of 1/2/4/8/16-channel vocoded words
Paired with matching/neutral/mismatching text

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

Prior knowledge enhances speech clarity

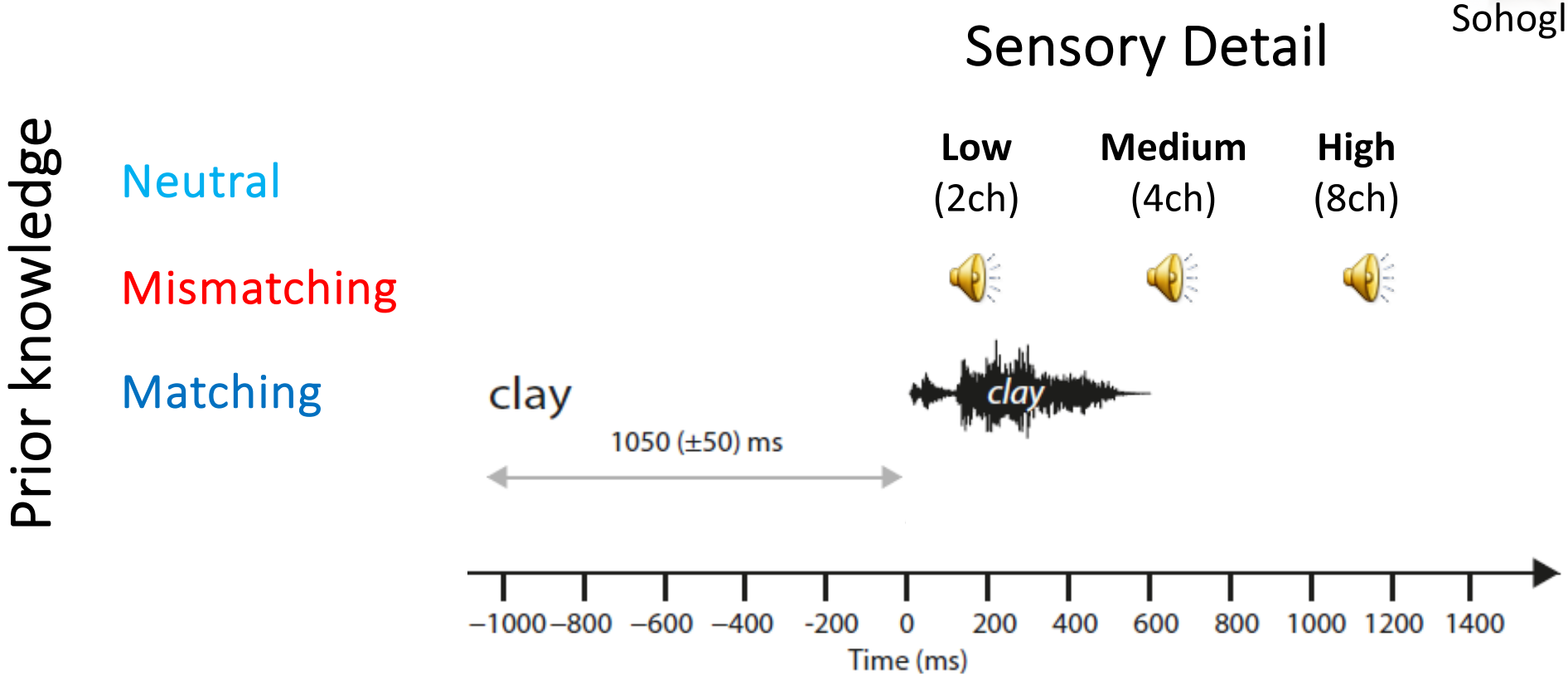
Ed Sohoglu



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

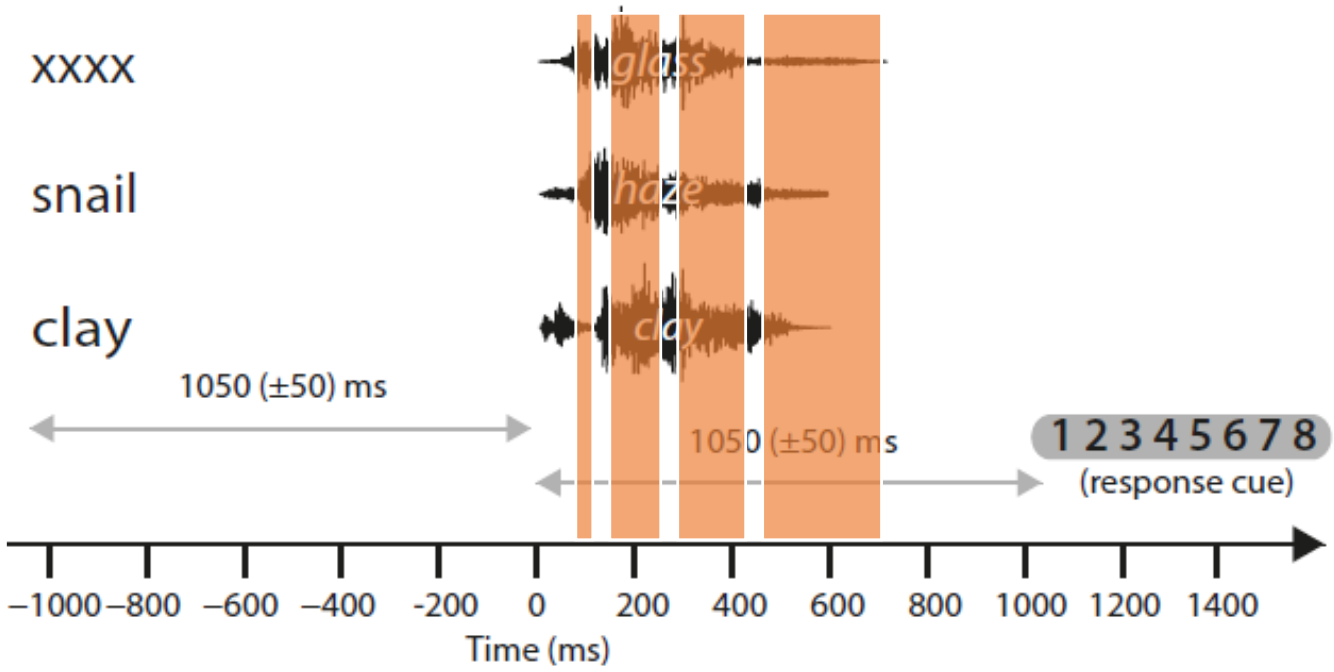
Prior knowledge and perception of speech

Ed Sohoglu



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

Prior knowledge and perception of speech

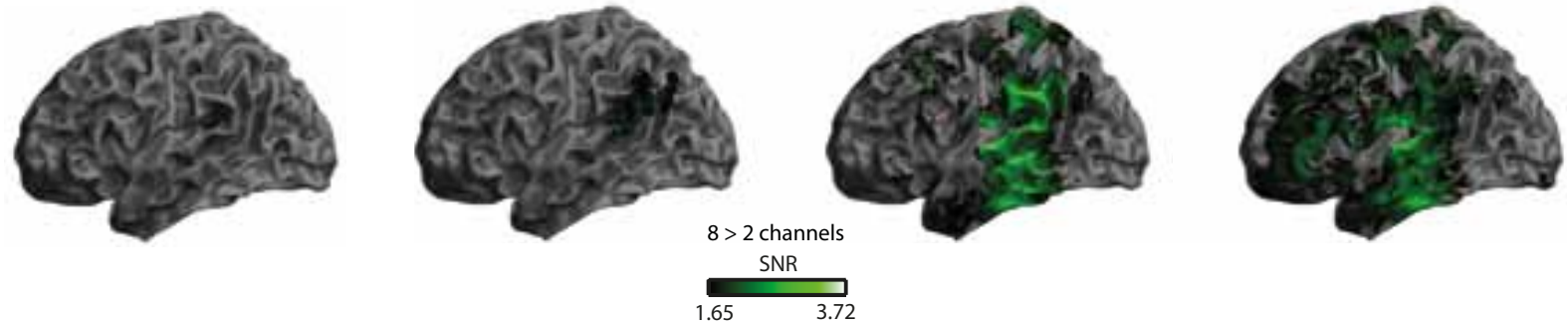


- Sohoglu, Peelle, Carlyon & Davis (2012, *J. Neuroscience*)

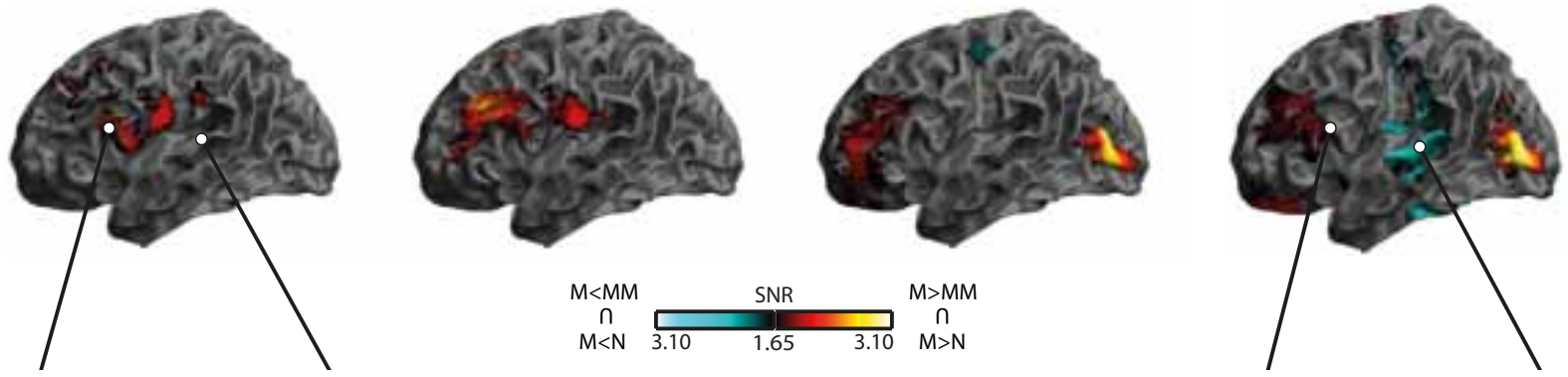
Prior knowledge and perception of speech

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

8 channel > 2 channel



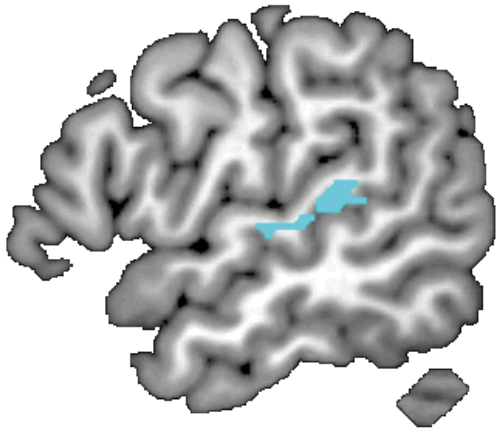
Match > Mismatch|Neutral **Match < Mismatch|Neutral**

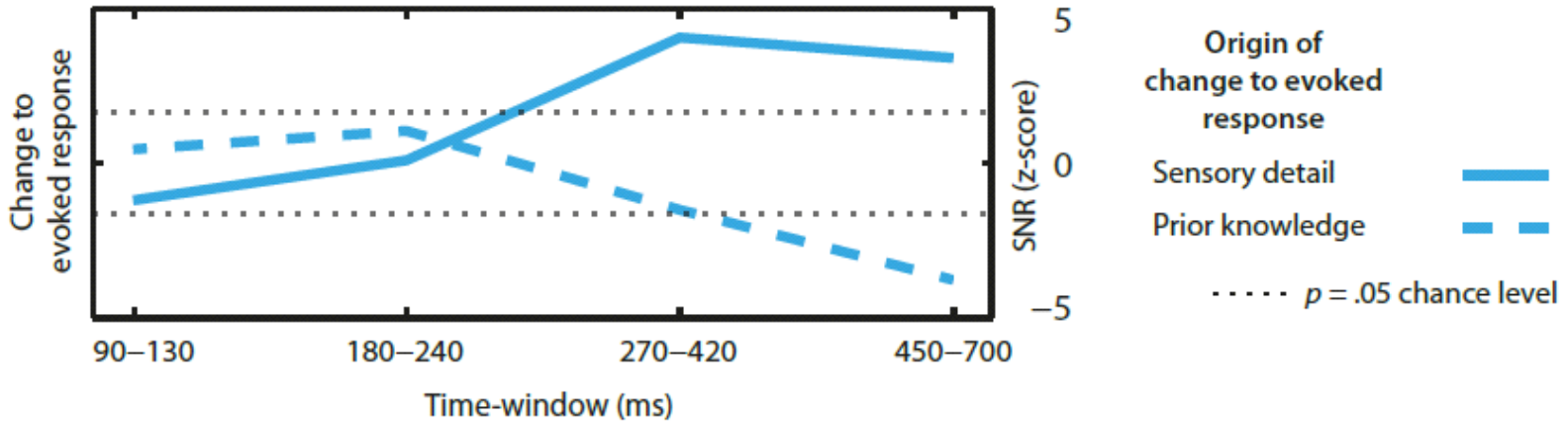


- Sohoglu, Peelle, Carlyon & Davis (2012, *J. Neuroscience*)

Prior knowledge and perception of speech

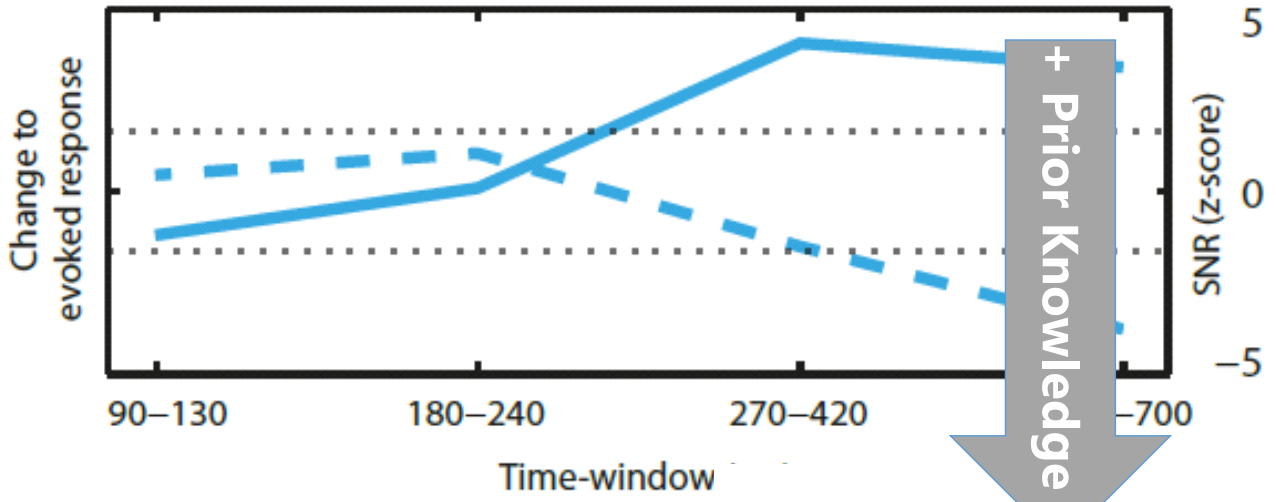
Superior temporal gyrus (STG)





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

Predictive coding model of speech perception



Processing Spoken Words

Localising vs explaining spoken word recognition

Bayesian inference in speech perception

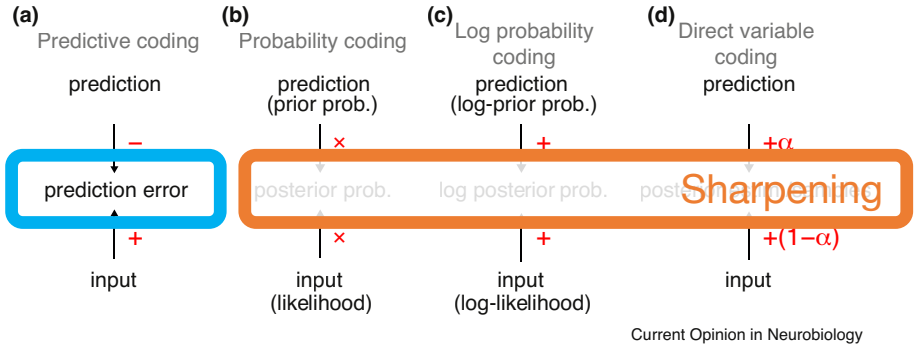
Posterior — How probable is each word given the sound heard

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

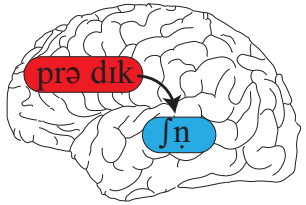
Prior — How probable was each word *before* hearing any sound?

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

Marginal — How probable is hearing that sound

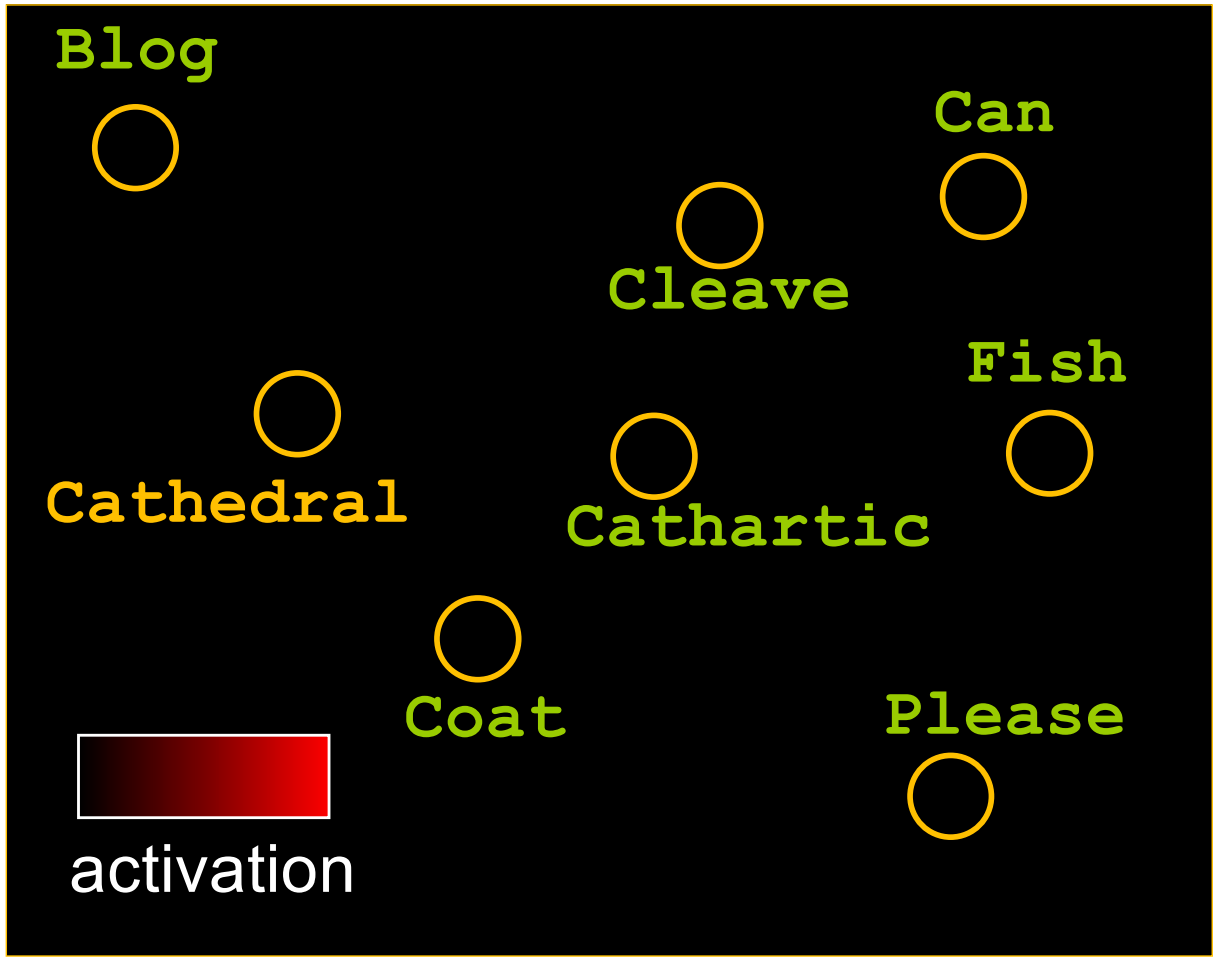


Predictive computations for word recognition

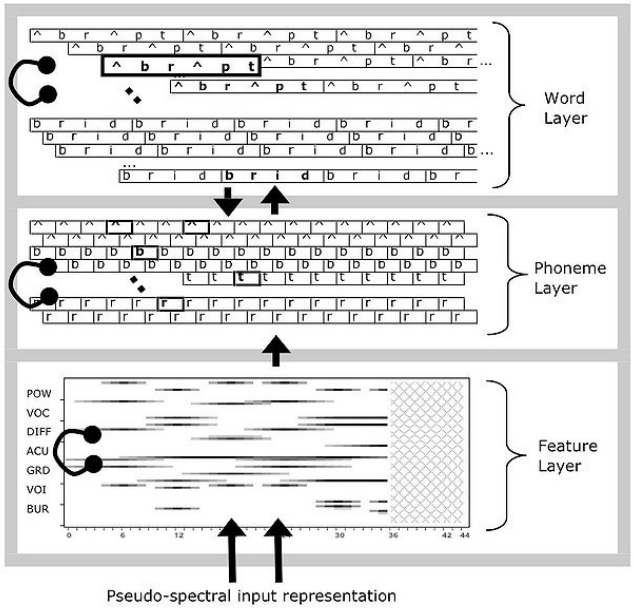


Recognising spoken words
(Cohort & TRACE Models)

“……”



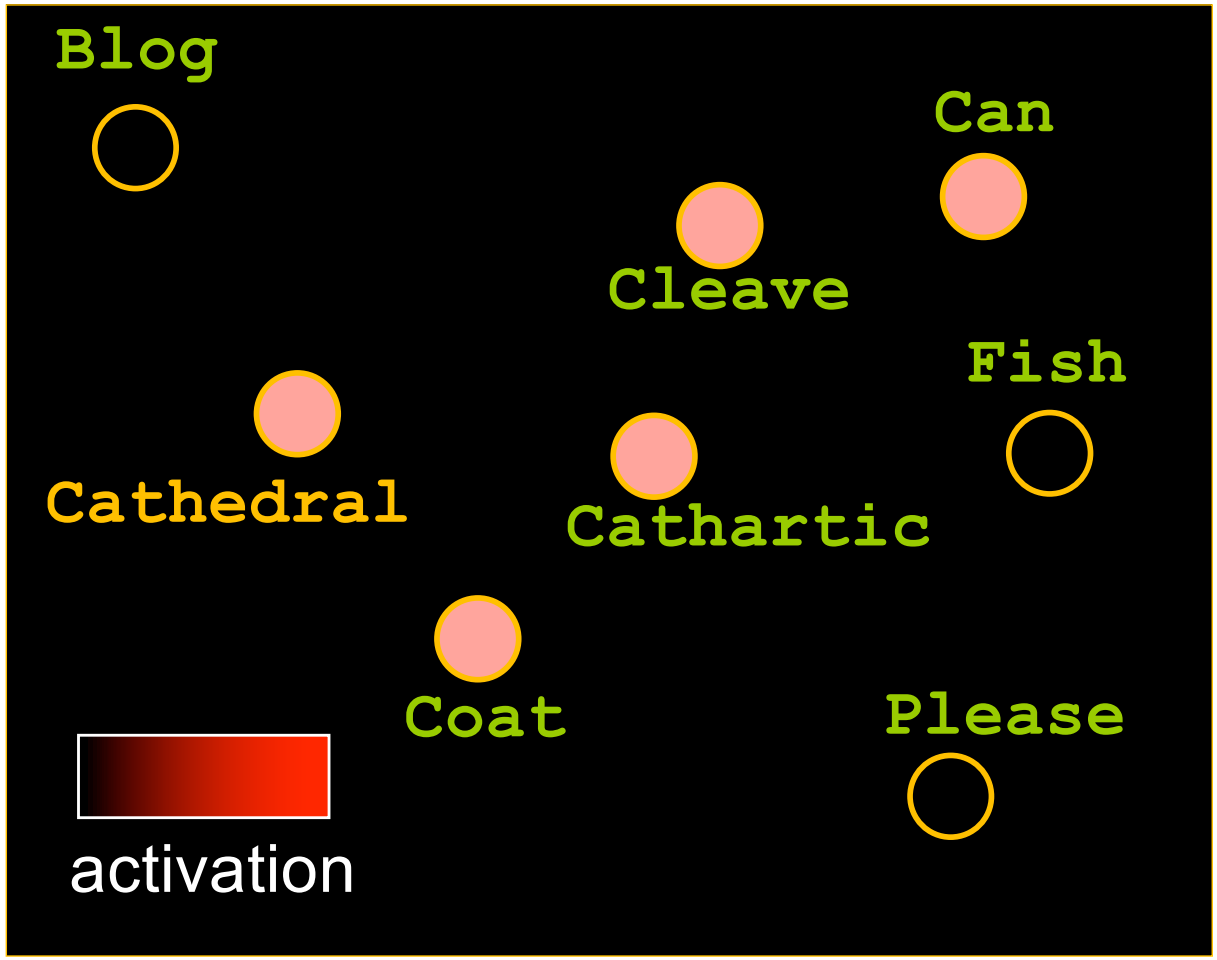
Marslen-Wilson & Tyler (1980, *Phil Trans B*)

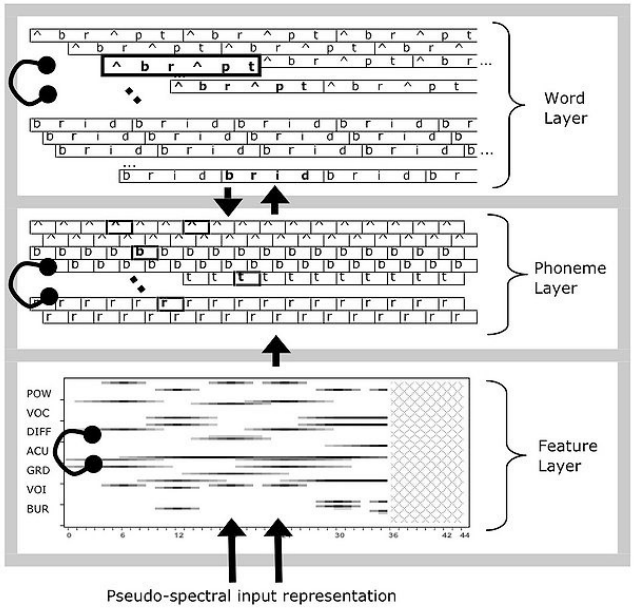


TRACE:
McClelland & Elman (1986, *Cog Psych*)

Recognising spoken words
(Cohort & TRACE Models)

“c

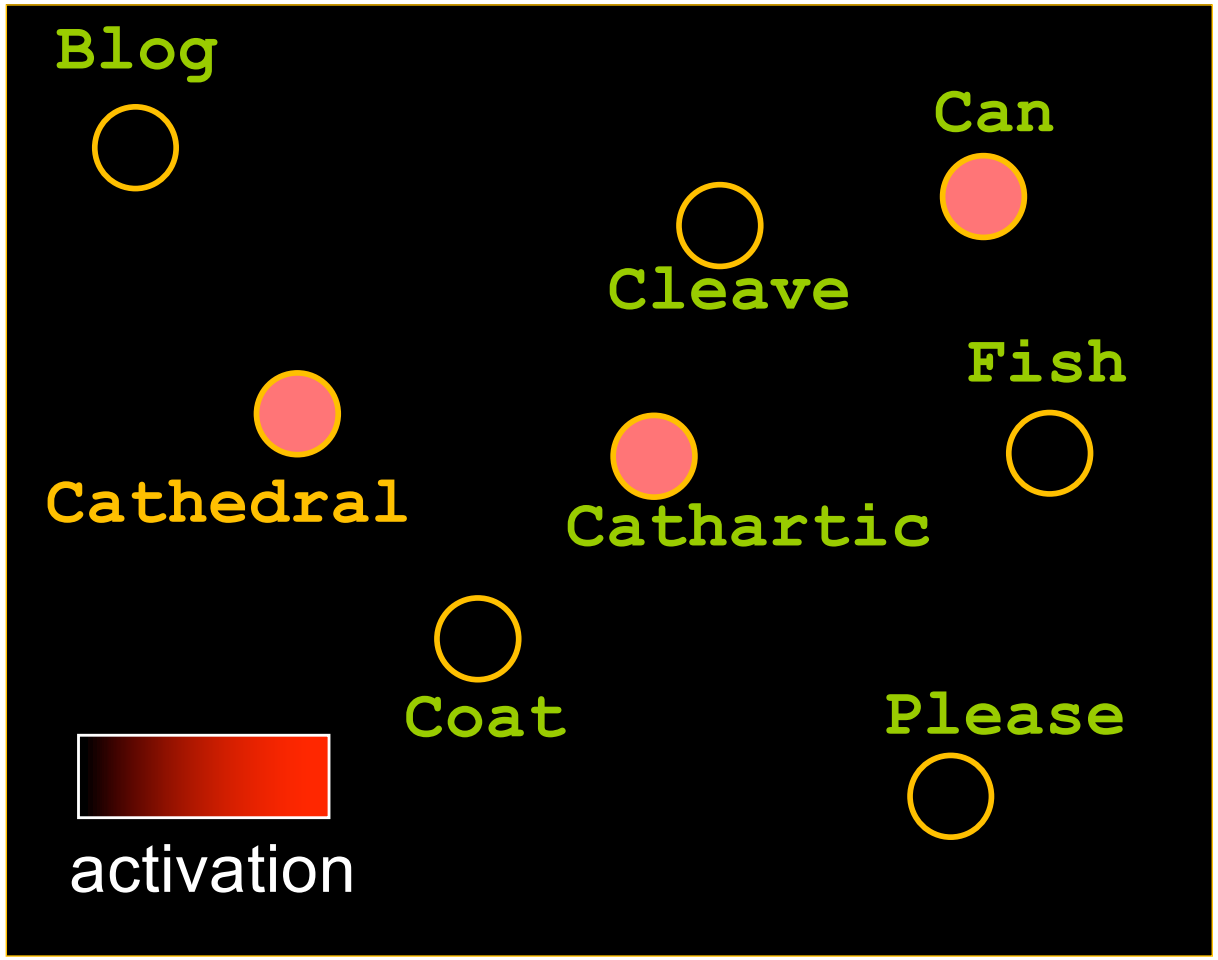


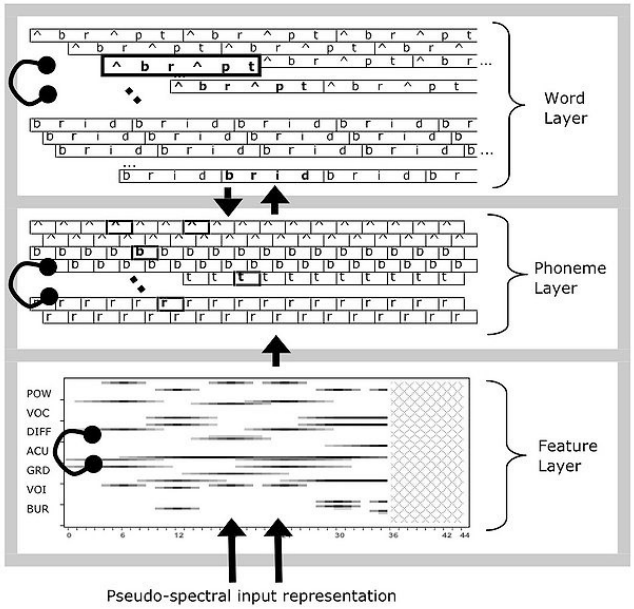


TRACE:
McClelland & Elman (1986, *Cog Psych*)

Recognising spoken words
(Cohort & TRACE Models)

“ca

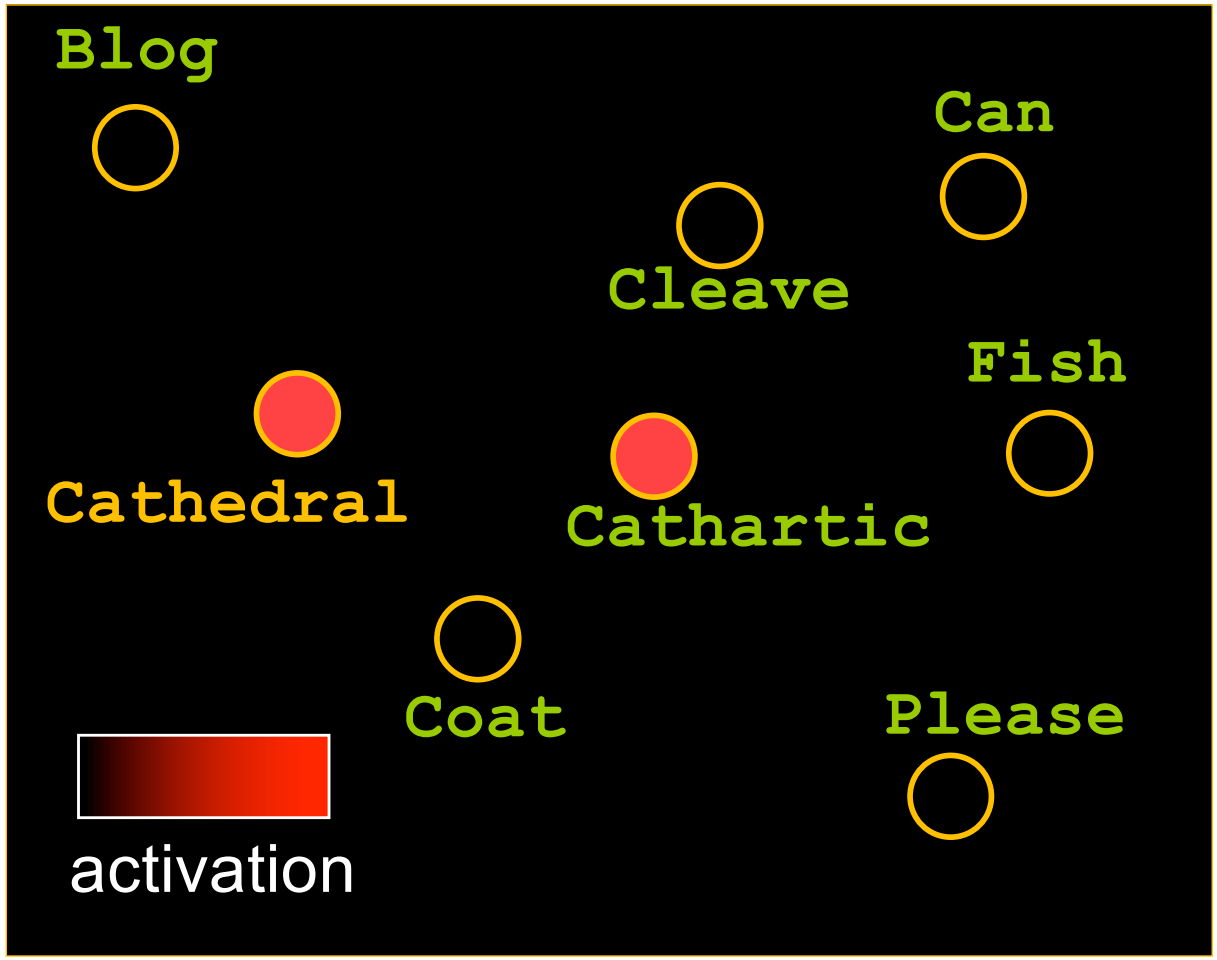




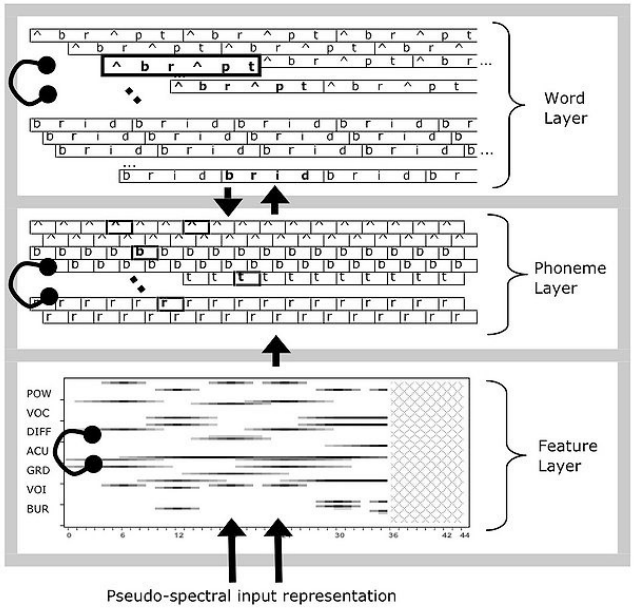
TRACE:
McClelland & Elman (1986, *Cog Psych*)

Recognising spoken words
(Cohort & TRACE Models)

“cath



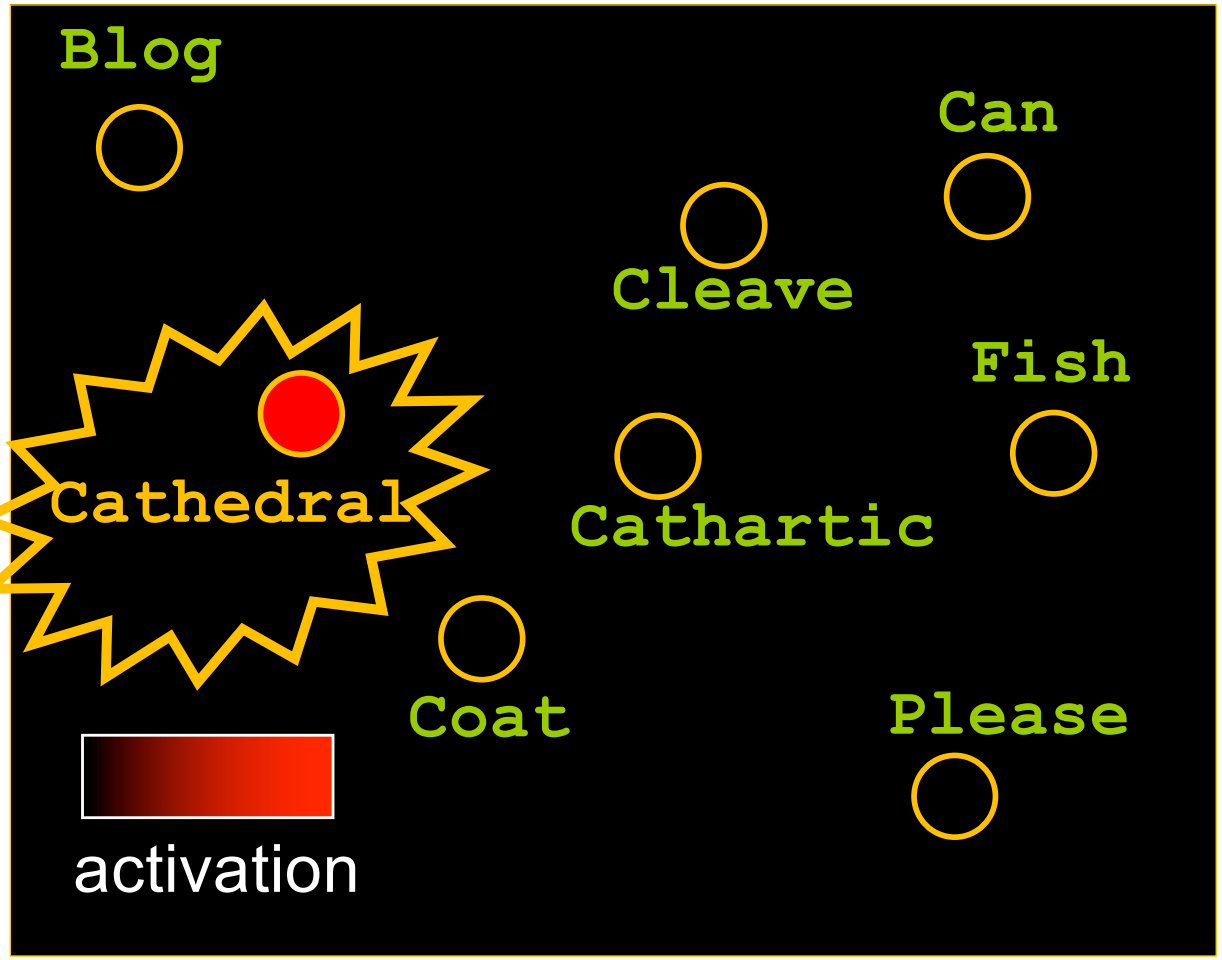
Marslen-Wilson & Tyler (1980, *Phil Trans B*)

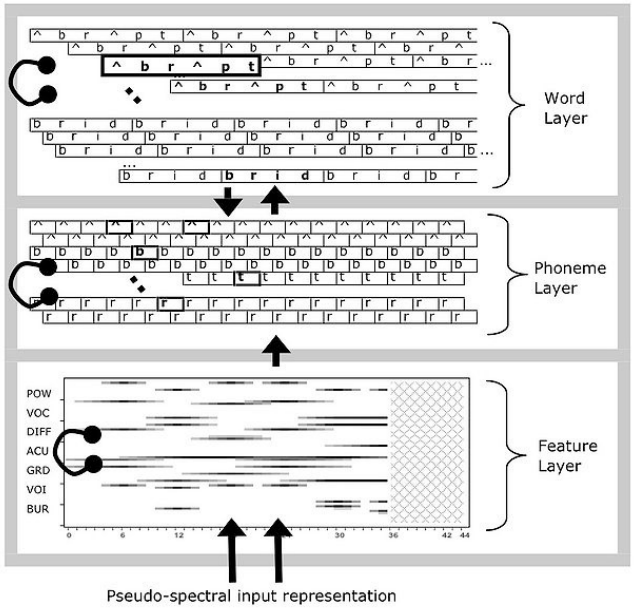


TRACE:
McClelland & Elman (1986, *Cog Psych*)

Recognising spoken words
(Cohort & TRACE Models)

“cathe

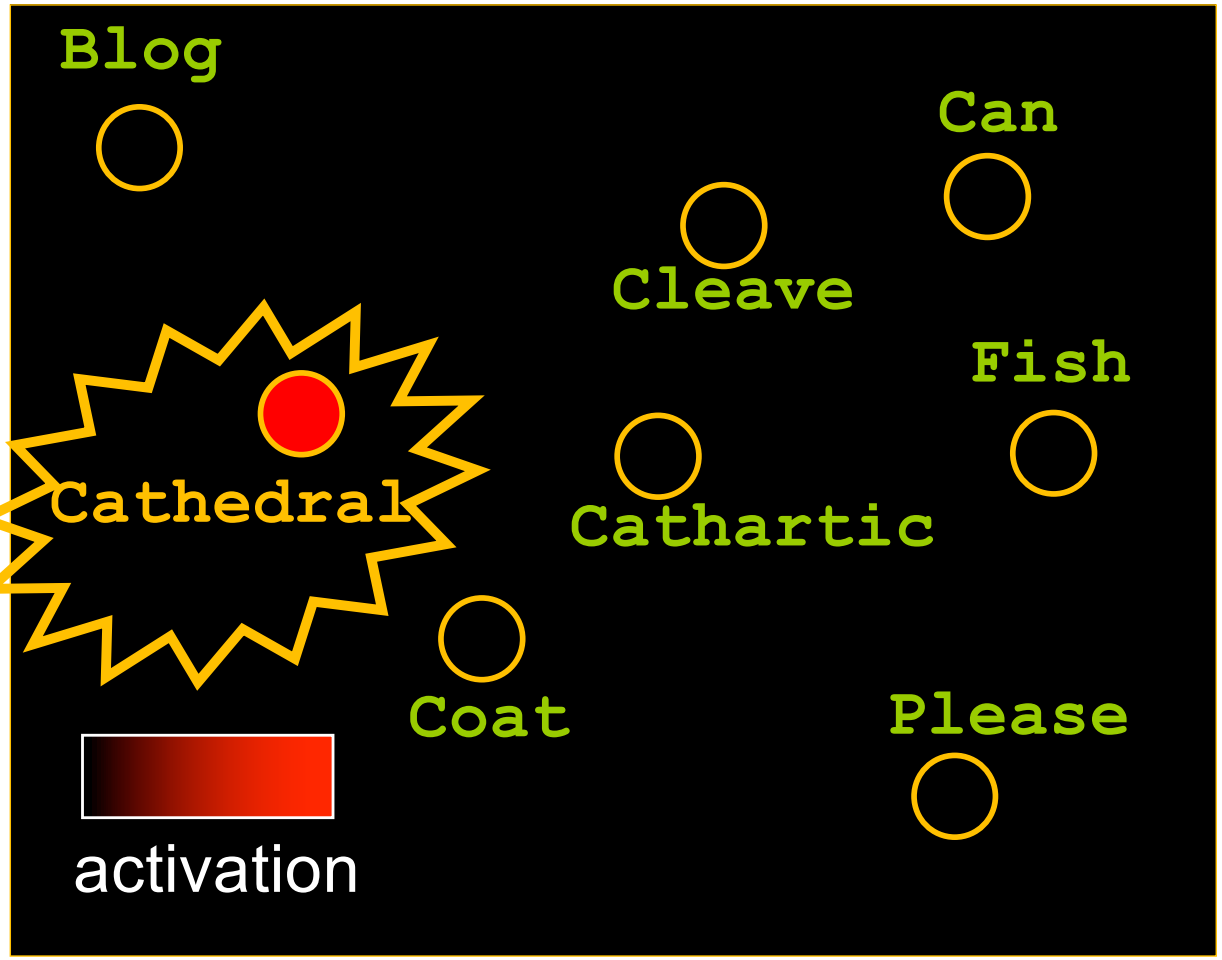




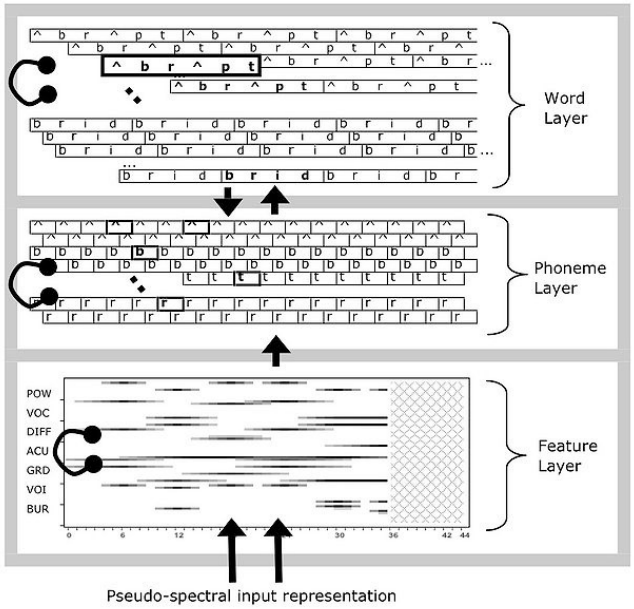
TRACE:
McClelland & Elman (1986, *Cog Psych*)

Recognising spoken words
(Cohort & TRACE Models)

“cathedr



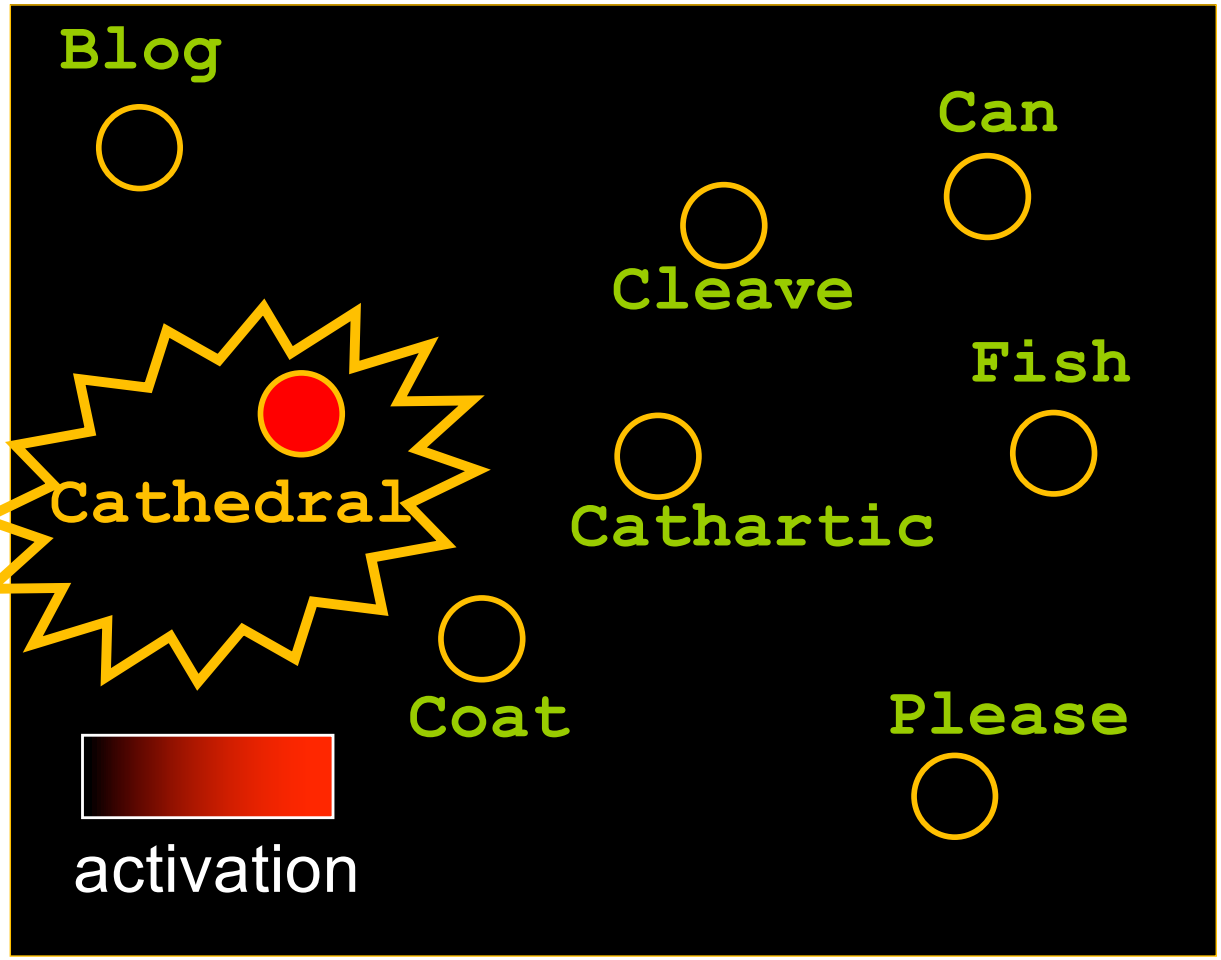
Marslen-Wilson & Tyler (1980, *Phil Trans B*)



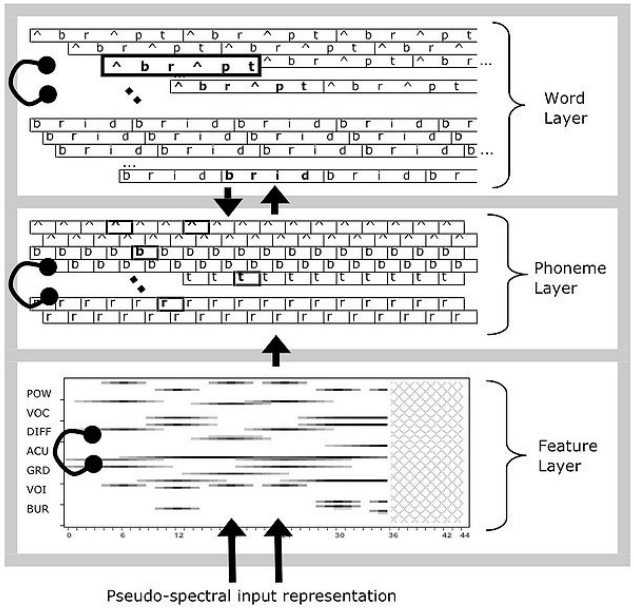
TRACE:
McClelland & Elman (1986, *Cog Psych*)

Recognising spoken words
(Cohort & TRACE Models)

“cathedral”

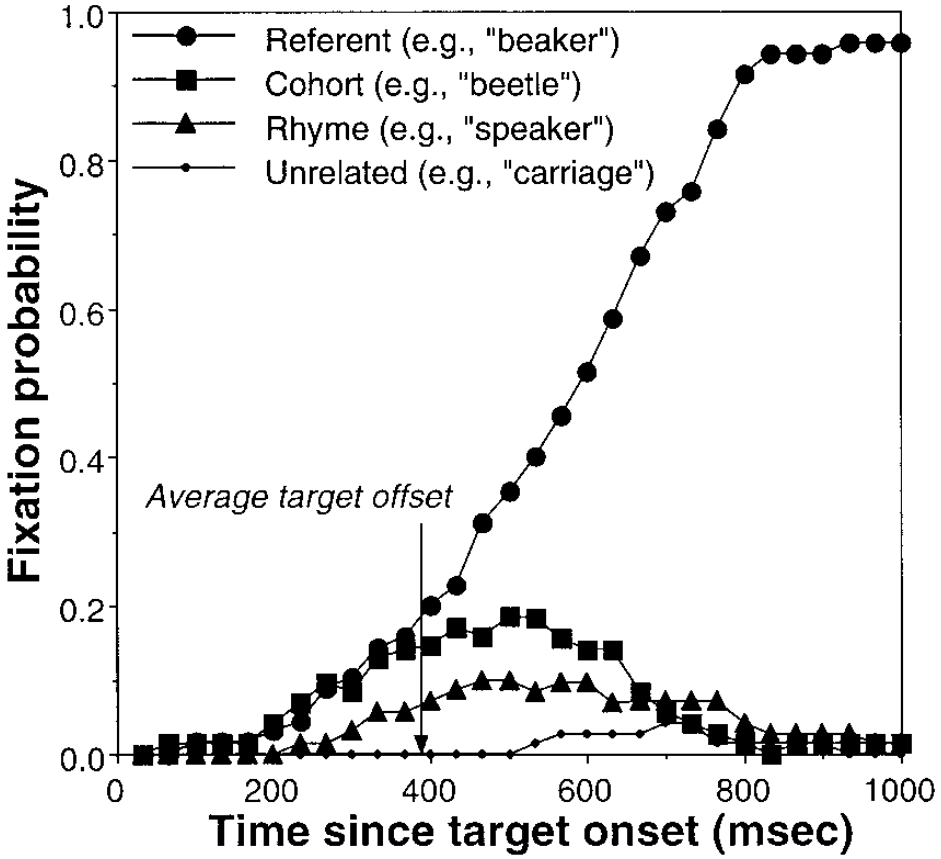


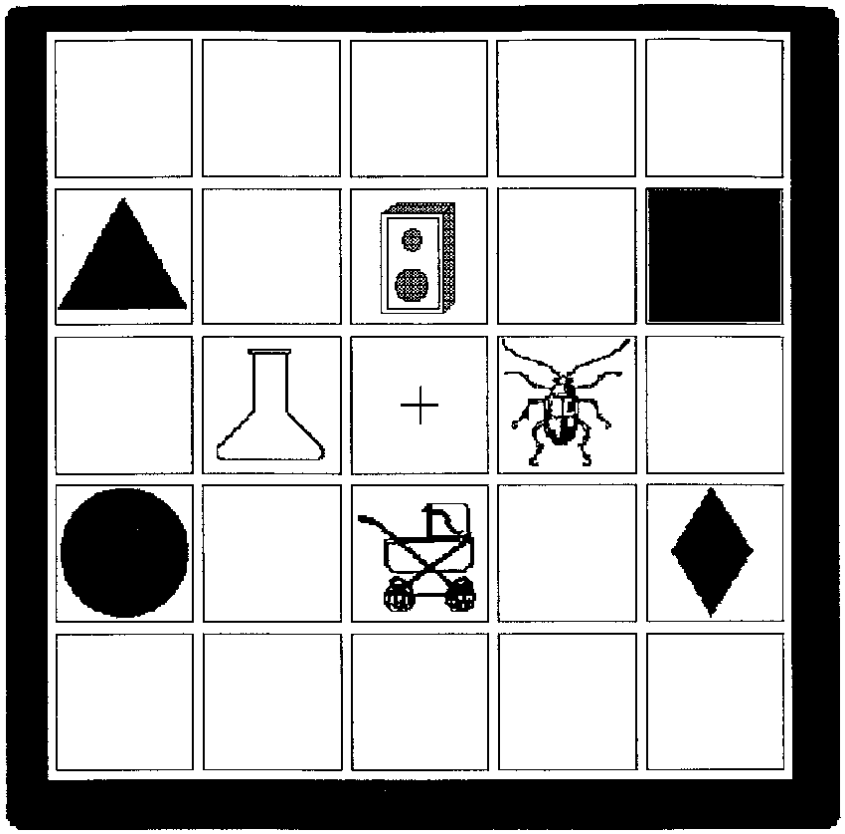
Marslen-Wilson & Tyler (1980, *Phil Trans B*)



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

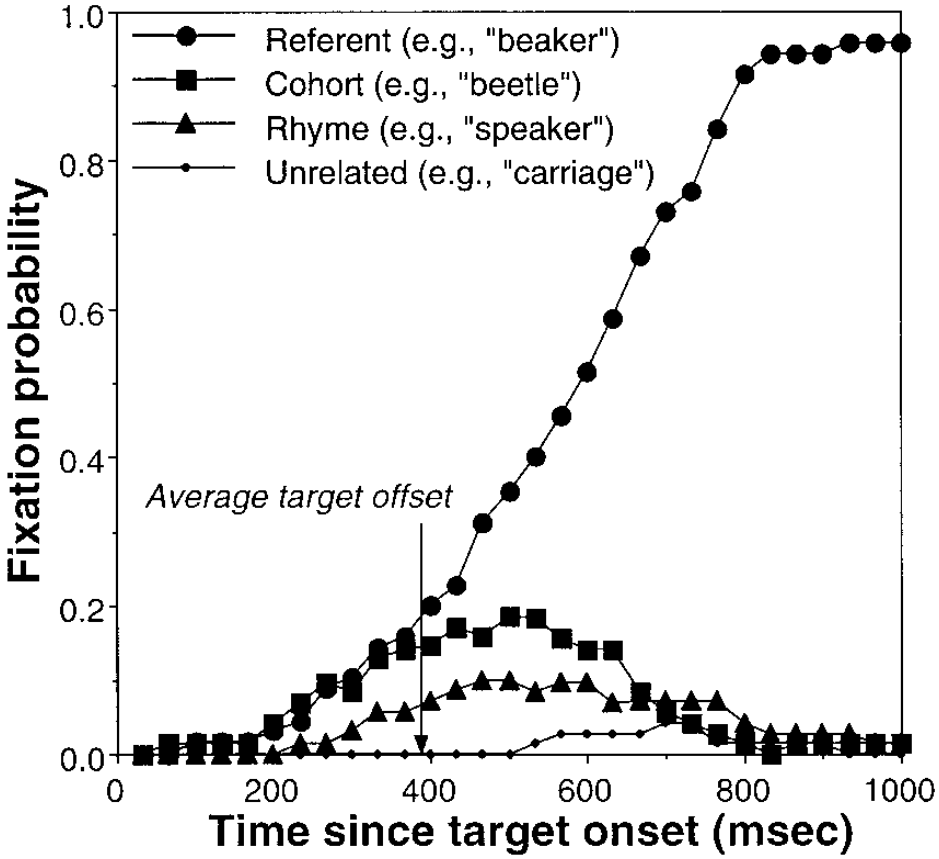
Recognising spoken words

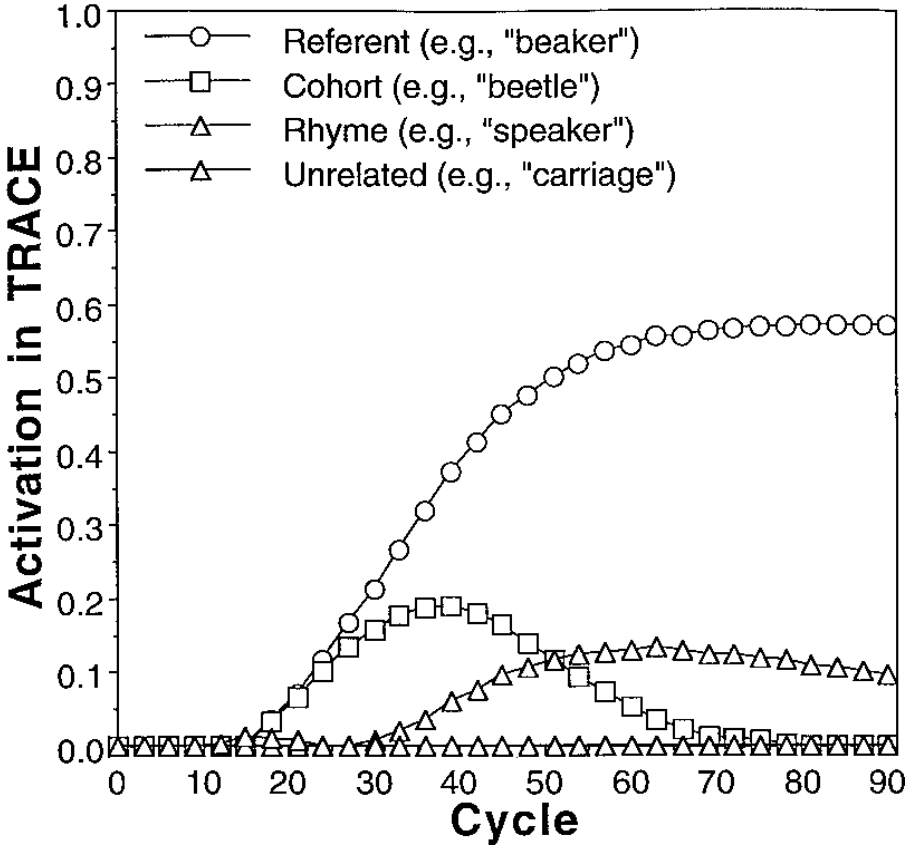




Allopenna, Magnuson & Tanenhaus (1998, JML)

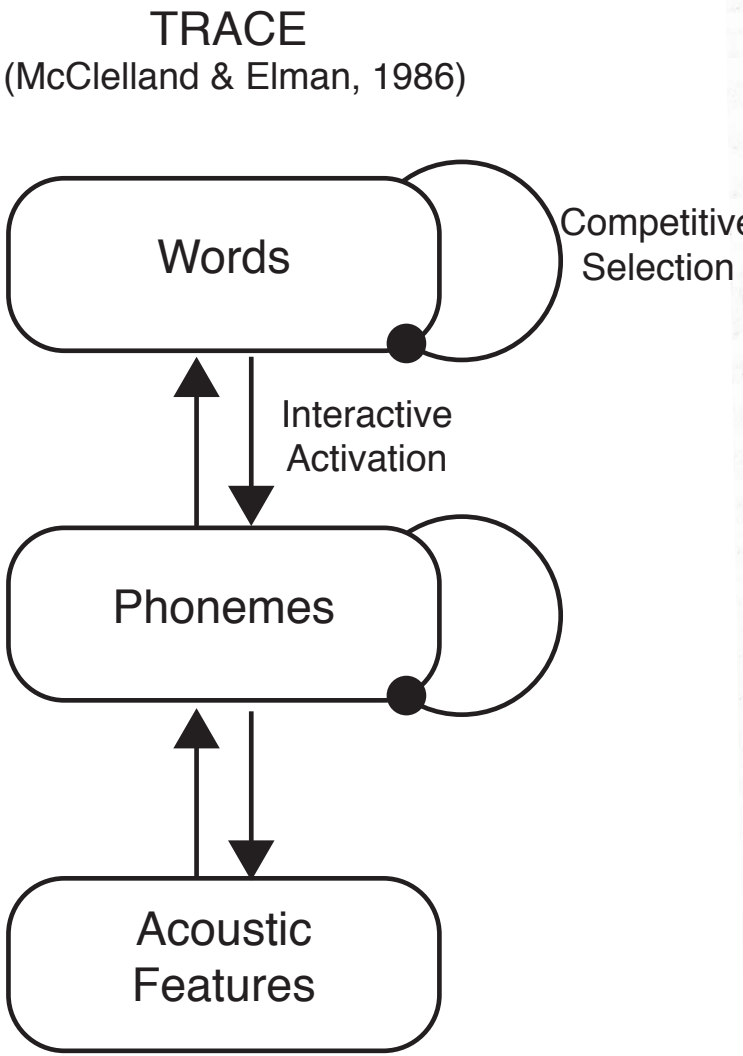
Recognising spoken words

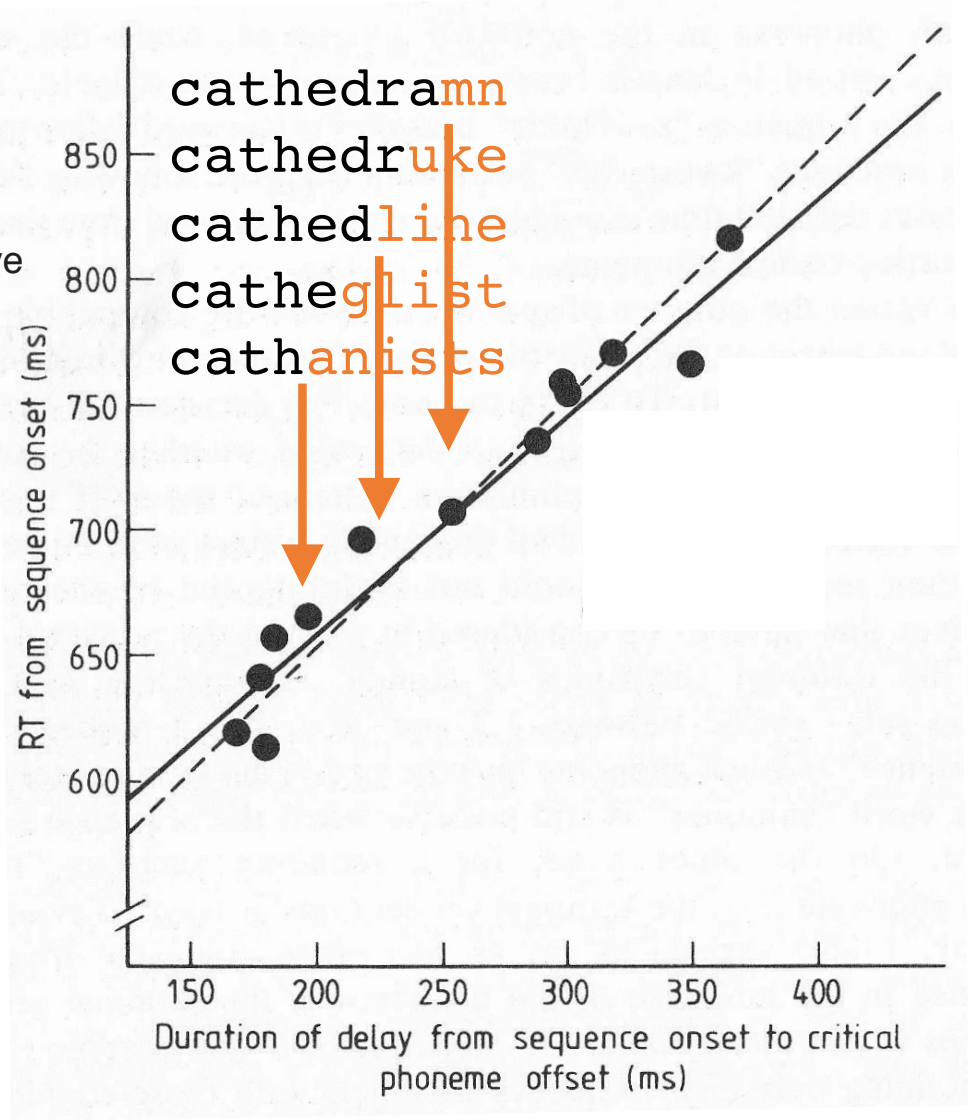




Allopenna Magnuson & Tanenhaus (1998, JML)

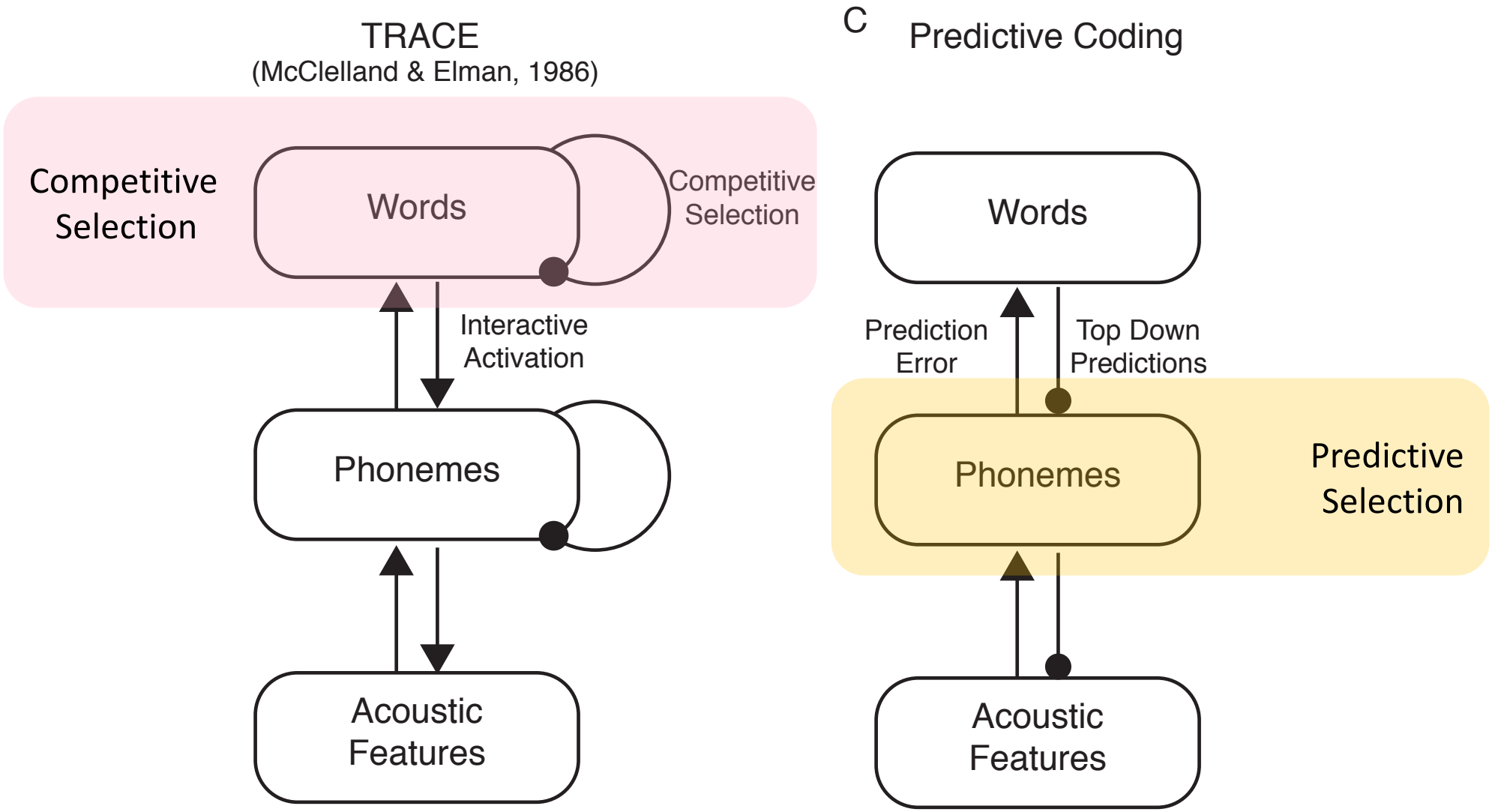
Competitve vs Predictive Selection





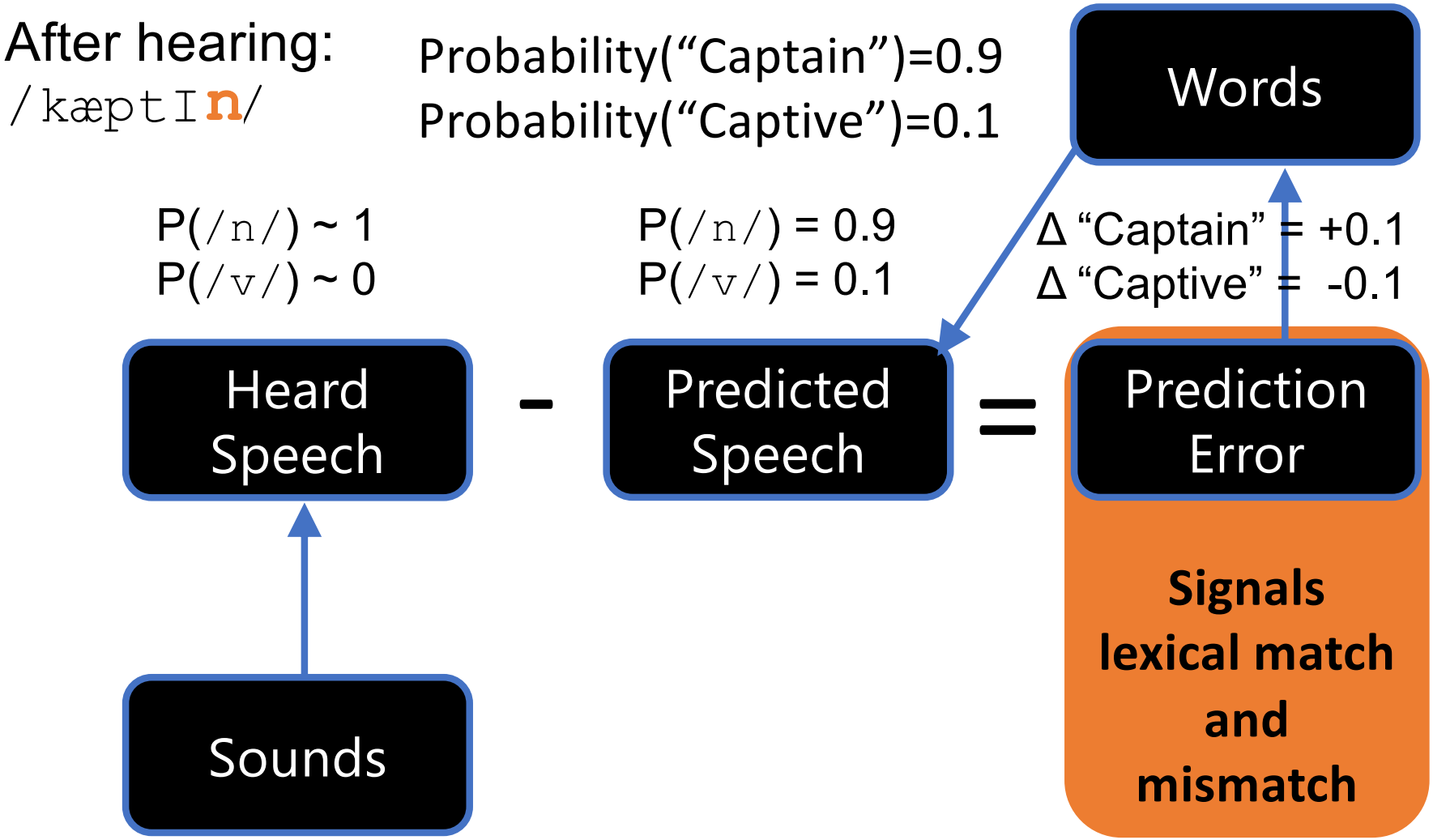
Marslen-Wilson (1984, Attention & Performance X)
Marslen-Wilson (1987, Cognition)

Sharpening vs Predictive Coding



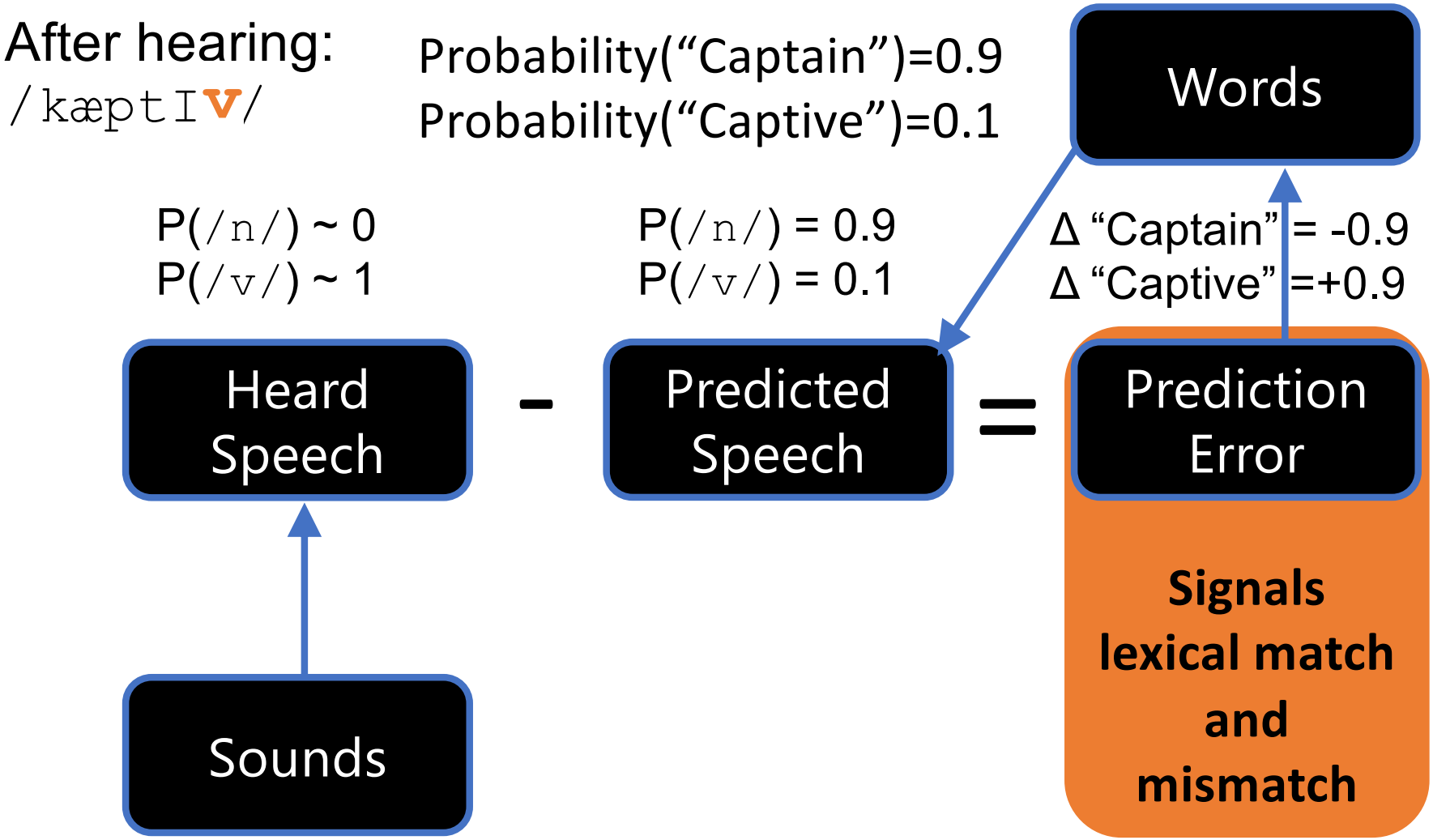
Davis & Sohoglu (2020 *Cog Neurosci 6, MIT Press)*
https://psyarxiv.com/qc4u6/

Predictive coding and word recognition



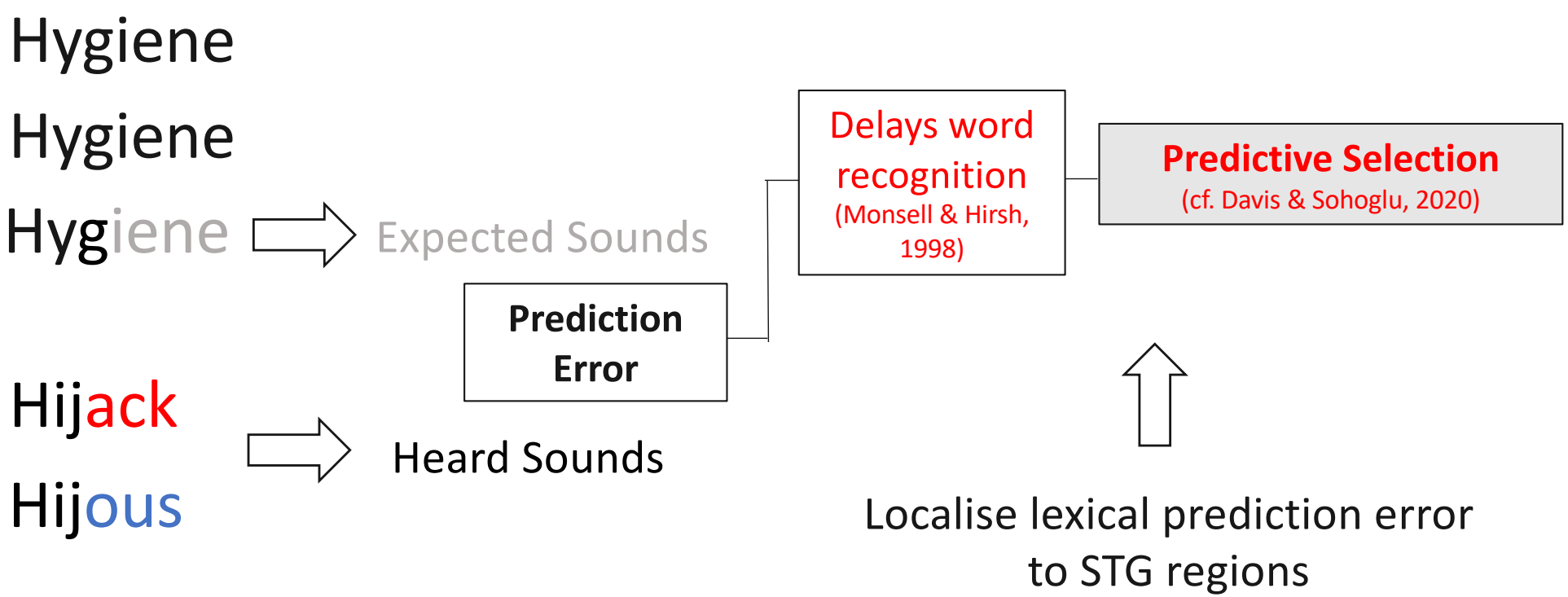
Gagnepain, Henson & Davis (2012, *Current Biology*)

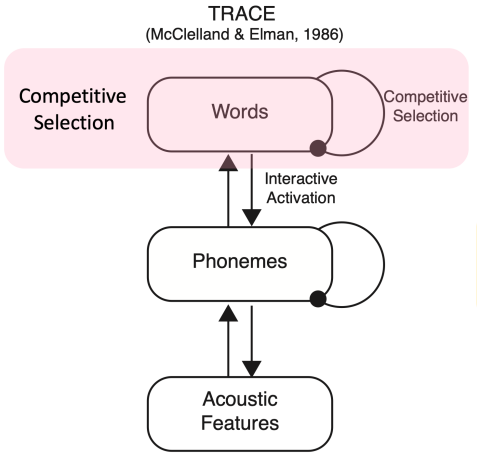
Predictive coding and word recognition



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

Speech predictions change with learning





Competitive Selection vs Predictive Selection

Carol Wang

Rik Henson

Input: /h/ /ai/ /dʒ/ **DP** /ə/ /s/

Lexical Uncertainty (Entropy):

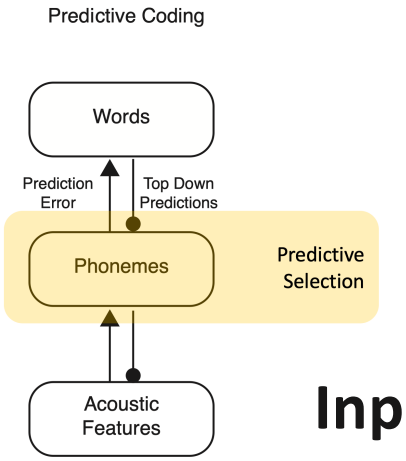
High

Low

habit hack … health help … hijack … hobby … **hygiene** …	hide … high … hijack hike … hybrid hydrate … **hygiene** …	**hygiene** hijack	**'hijou?'**	**'hijous?'**

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

Competitive Selection vs Predictive Selection



Carol Wang

Rik Henson

DP

Input: /h/ /ai/ /dʒ/ /i:/ /n/

Prediction:

Segment Surprisal (Prediction Error):

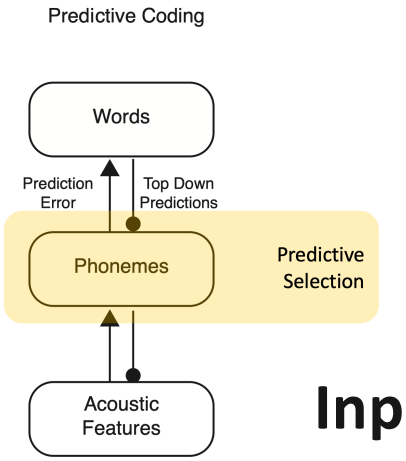
High

Low

/æ/	/d/	/æ/	**/n/**
/e/	…	**/i:/**	
…	/f/		
/ai/	/k/		
…	…		
/ɔ/	**/dʒ/**		
…	…		
/ʌ/			
…			

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

Competitive Selection vs Predictive Selection



Carol Wang

Rik Henson

DP

Input: /h/ /ai/ /dʒ/ /æ/ /k/

Prediction:

Segment Surprisal (Prediction Error):

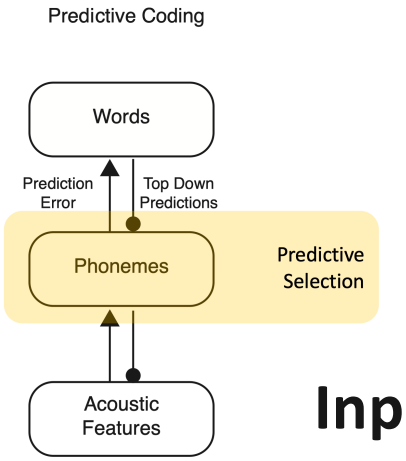
High

Low

/æ/	/d/	**/æ/**	**/k/**
/e/	…	/i:/	/n/
…	/f/		
/ai/	/k/		
…	…		
/ɔ/	**/dʒ/**		
…	…		
/ʌ/			
…			

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

Competitive Selection vs Predictive Selection



Carol Wang

Rik Henson

DP

Input: /h/ /ai/ /dʒ/ /ə/ /s/

Prediction:

Segment Surprisal (Prediction Error):

High

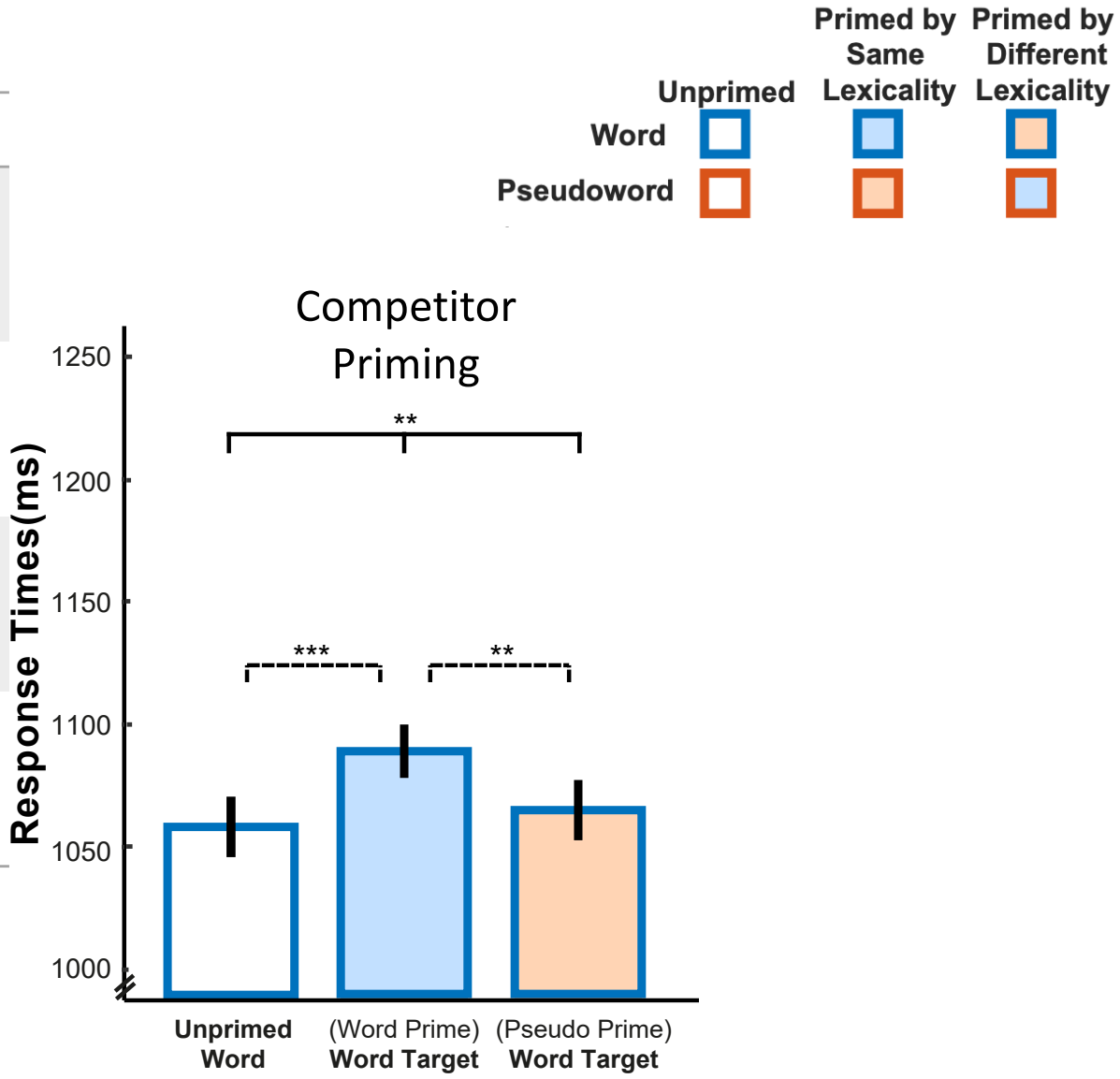
Low

/æ/	/d/	/æ/	?
/e/	…	**/i:/**	
…	/f/		
/ai/	/k/		
…	…		
/ɔ/	**/dʒ/**		
…	…		
/ʌ/			
…			

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

Competitor Priming for Word Pairs (Behaviour)

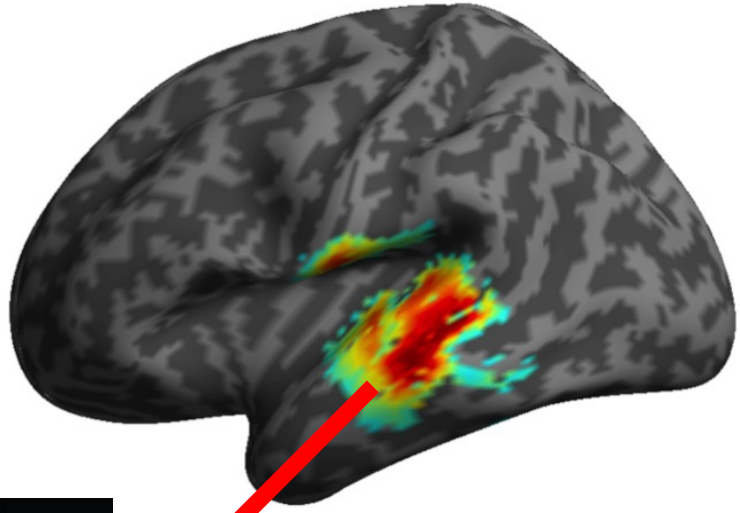
Pairs separated by 20-80 trials (1-4 mins)	
Unprimed **Word** hygiene	(Word Primed) **Word** hijack
Unprimed **Pseudo** letto	(Pseudo Primed) **Peeudo** lettan
Unprimed **Pseudo** basef	(Pseudo primed) **Word** Basic
Unprimed **Word** boycott	(Word primed) **Pseudo** boyten

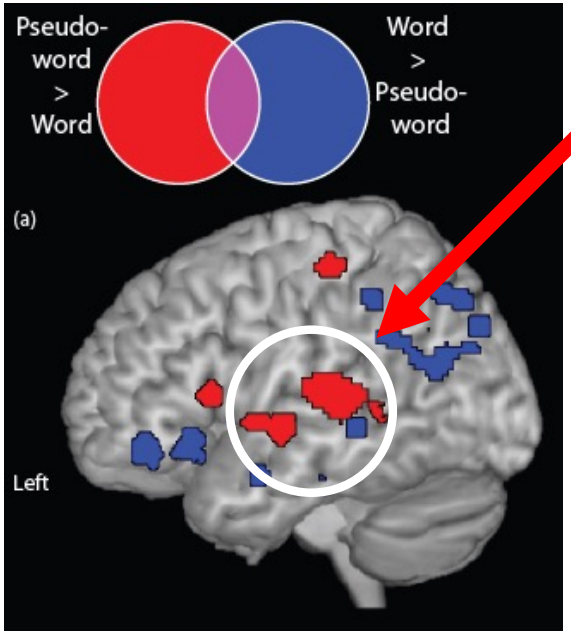


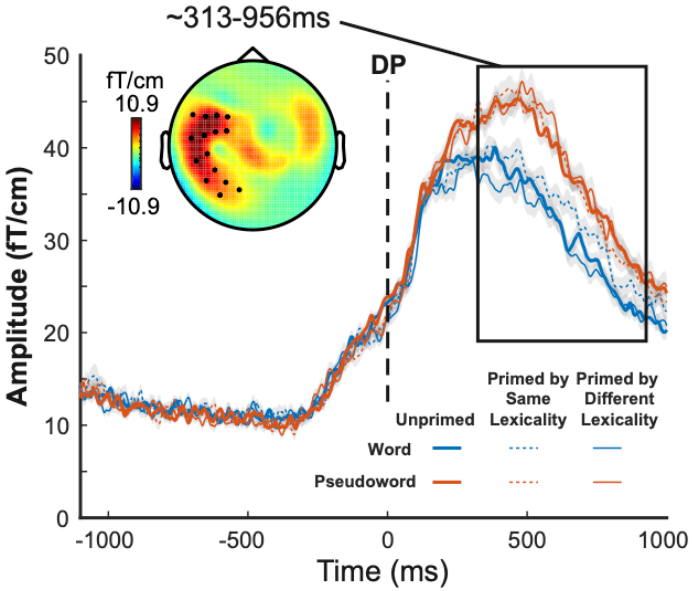
Wang, Sohoglu, Gilbert, Henson & Davis (2021, *J Neurosci*)

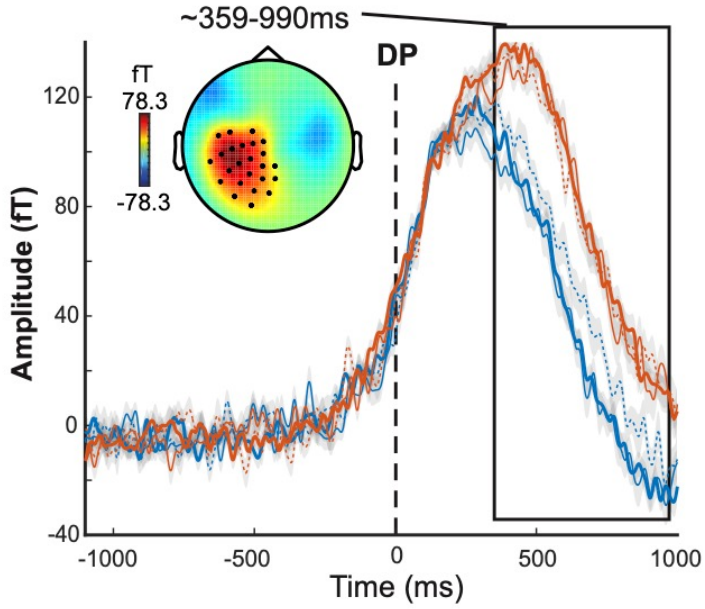
Competitor Priming for Word Pairs overlaps with Pseudo > Word (MEG)

STG Source of Lexicality Effect





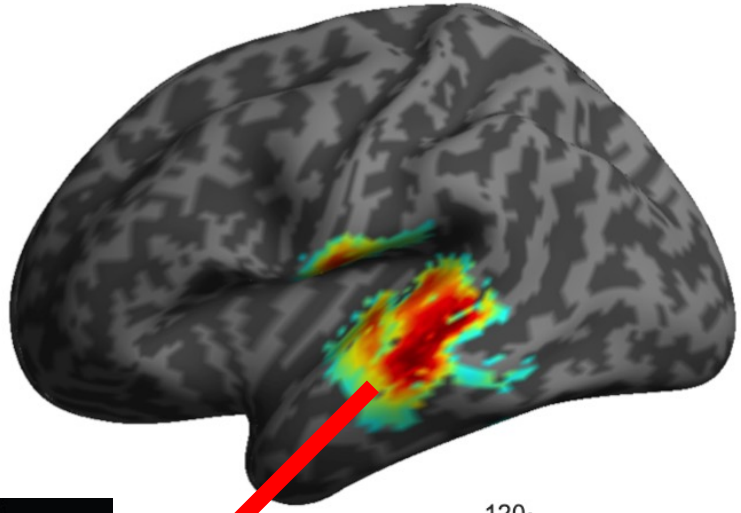




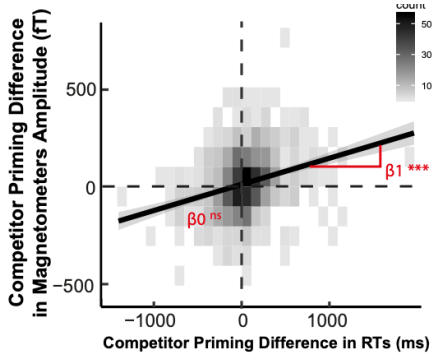
Wang, Sohoglu, Gilbert, Henson & Davis (2021, *J Neurosci*)

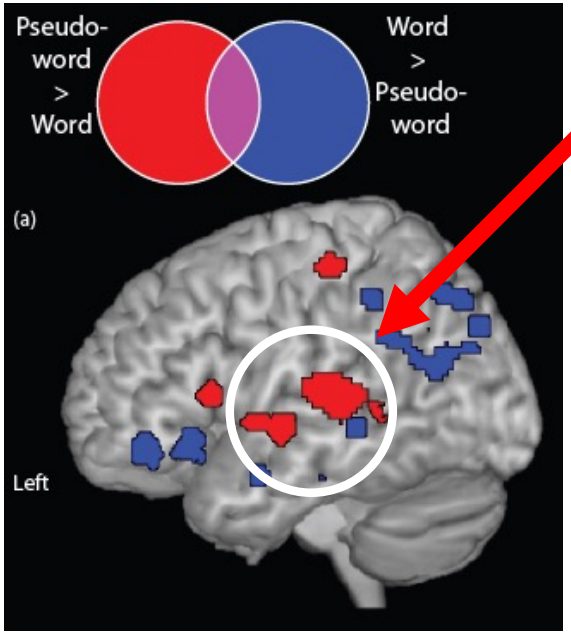
Competitor Priming for Word Pairs overlaps with Pseudo > Word (MEG)

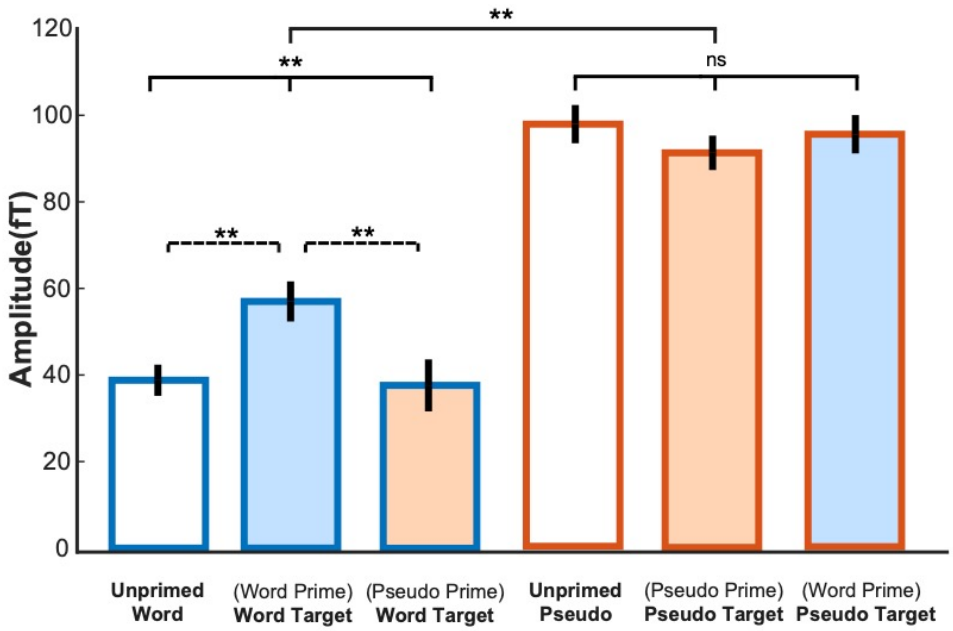
STG Source of Lexicality Effect



Trial-wise correlation of neural and behavioural priming



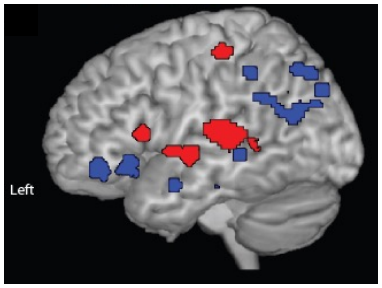




Wang, Sohoglu, Gilbert, Henson & Davis (2021, *J Neurosci*)

Processing Spoken Words

Localising vs explaining spoken word recognition



Bayesian inference in speech perception

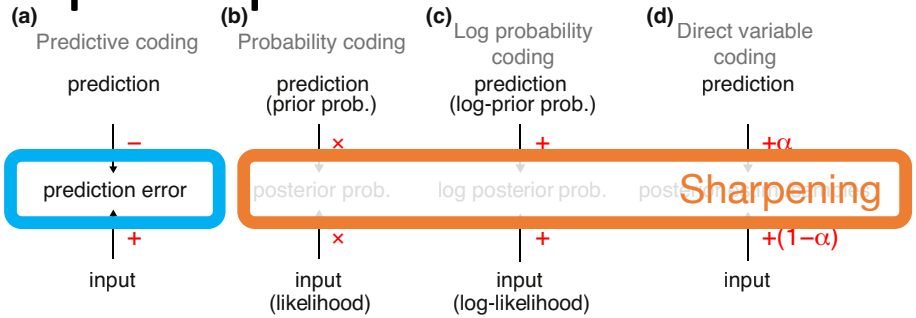
Posterior — How probable is each word given the sound heard

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

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

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

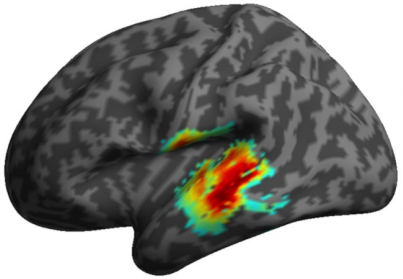
Marginal — How probable is hearing that sound



Current Opinion in Neurobiology

Predictive computations for word recognition

Heard Speech − Predicted Speech = Prediction Error



References from question period:

- Word learning and overnight consolidation:

  Behaviour: Dumay & Gaskell (2007, Psych Science) https://doi.org/10.1111/j.1467-9280.2007.01845.x
  fMRI: Davis et al (2009, JoCN) . https://doi.org/10.1162/jocn.2009.21059
  MEG: Gagnepain et al (2012, Current Biology) https://doi.org/10.1016/j.cub.2012.02.015
  Theory / Review: Davis & Gaskell (2009, Phil Trans Roy Soc B) https://doi.org/10.1098/rstb.2009.0111

- Responses to pseudowords of different lengths:

Behaviour: Marslen-Wilson (1984, Attention & Performance) https://www.researchgate.net/publication/260320948
EEG: O'Rourke & Holcomb (2002, Biological Psychology) https://doi.org/10.1016/s0301-0511(02)00045-5
fMRI: Zhuang et al (2014, Cerebral Cortex) https://doi.org/10.1093/cercor/bhs366

- RSA & Decoding tests TRACE/Sharpening vs Prediction error

  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