CS 182

Leon Barrett and Will Chang

Thanks to Eva Mok and Joe Makin

Q: What did the Hollywood film director say after he finished making a movie about myelin?

A: "That's a wrap!"

http://faculty.washington.edu/chudler/jokes.html

Announcements

- a2 is out, due next Tuesday 11:59pm
  - play with **`fannExplorer`**
  - you can run it on the server, but if you want to do a lot of playing, you might want to download it; search for FANN

Where we stand

- Last Week
  - Neurons
  - Neural development
- This Week
  - Idealized neurons and connectionist units
  - Spreading Activation, triangle nodes
- Coming up
  - Backprop (review your calculus!)

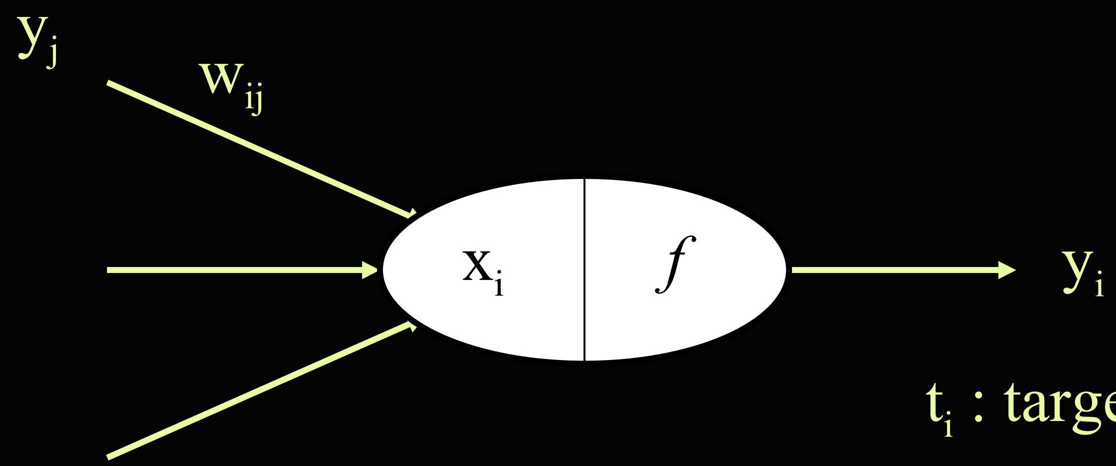
Serial and Parallel

- What is the difference between serial and parallel computation? Can you give an example of this contrast?
  - Serial computation takes time for every piece of the computation
  - Parallel computations occur at the same time
  - For example, face recognition is parallel – don't have to search through all possible people
  - Listing family members is serial – can't think of them all at once

Neuron models

- What is a McCullough-Pitts neuron?
  - Calculate “activation” by weighted sum of inputs
  - Use a function to map activation to output
- What sorts of output functions does it use?
  - Linear – why is this bad?
  - Threshold
  - Sigmoid – smoothed threshold

The McCullough-Pitts Neuron



t_i : target

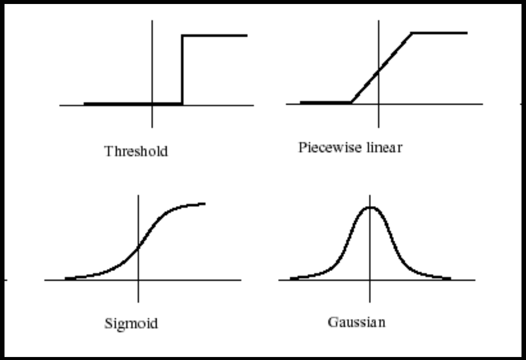
$$x_i = \sum_j w_{ij}\, y_j$$

$$y_i = f(x_i)$$

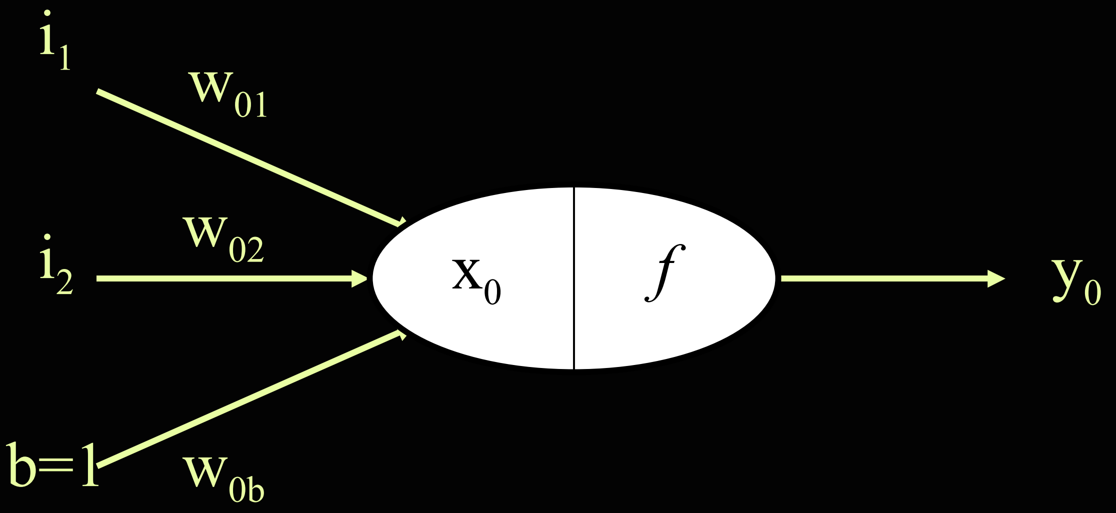
y_j: output from unit j

W_{ij}: weight on connection from j to i

x_i: weighted sum of input to unit i



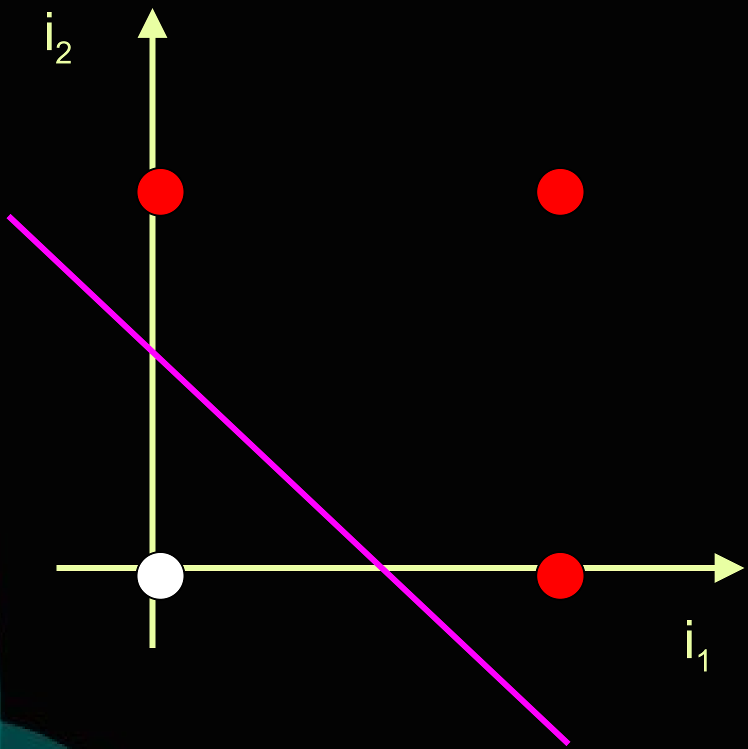
Let's try an example: the OR function



i_1	i_2	y_0
0	0	0
0	1	1
1	0	1
1	1	1

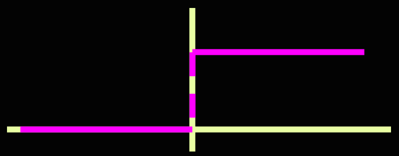
- Assume you have a threshold function centered at the origin
- What should you set w_{01}, w_{02} and w_{0b} to be so that you can get the right answers for y_0?

Many answers would work



$$y = f\,(w_{01}i_1 + w_{02}i_2 + w_{0b}b)$$

recall the threshold function



the separation happens when

$$w_{01}i_1 + w_{02}i_2 + w_{0b}b = 0$$

move things around and you get

$$i_2 = -\,(w_{01}/w_{02})i_1 - (w_{0b}b/w_{02})$$

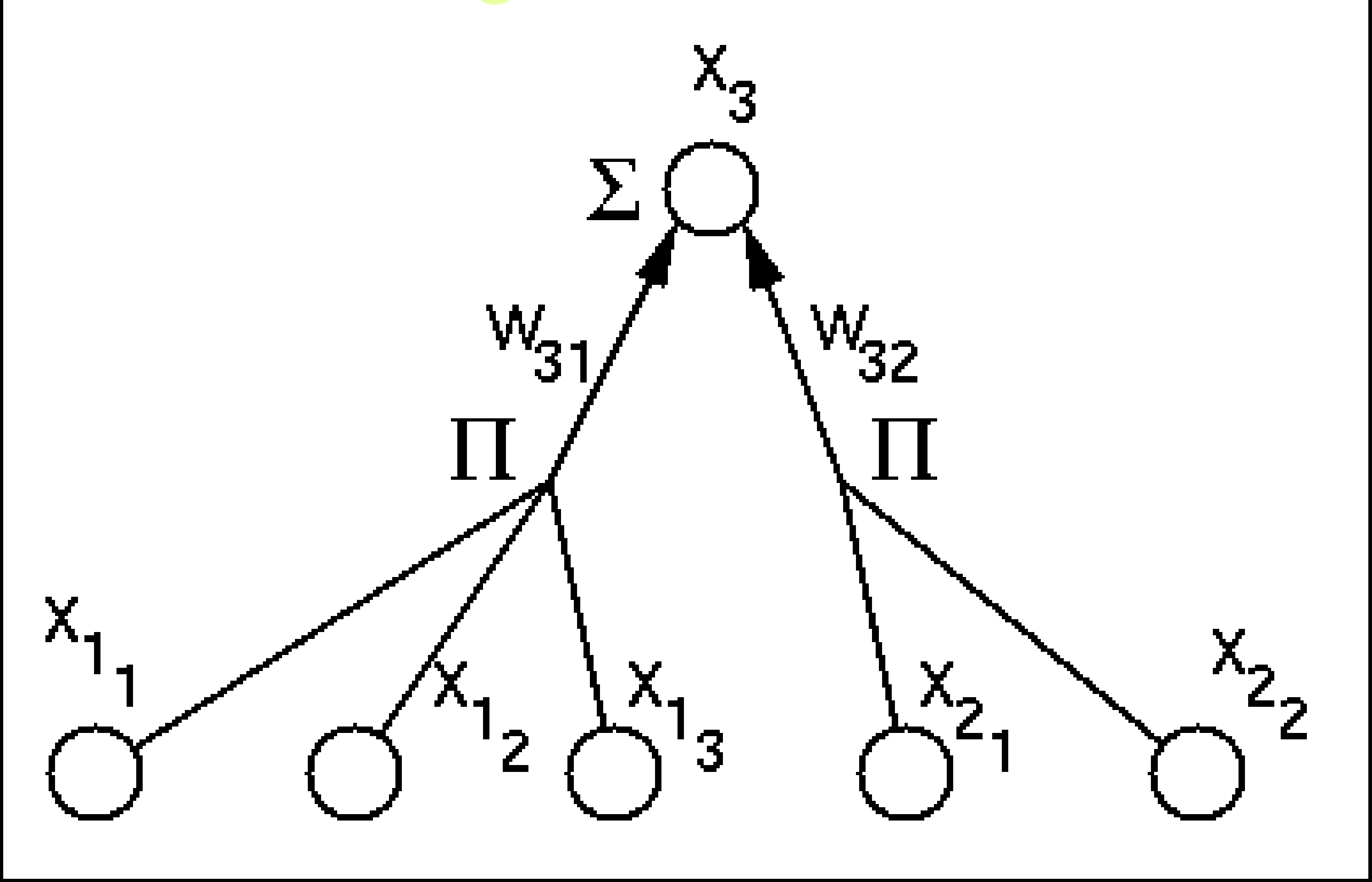
Neuron models

- If neurons spike, what biological feature do non-integer values correspond to?
  - Firing rate
- What is unbiological about the McCullough-Pitts model?
  - Time is ignored
  - Weights are unlimited
  - Hugely simplifies computations
    - Some people propose that real neurons do complicated quantum-mechanical things

Neuron models

- How does a Sigma-Pi neuron differ from a McCullough-Pitts neuron?
  - It multiplies sets of inputs

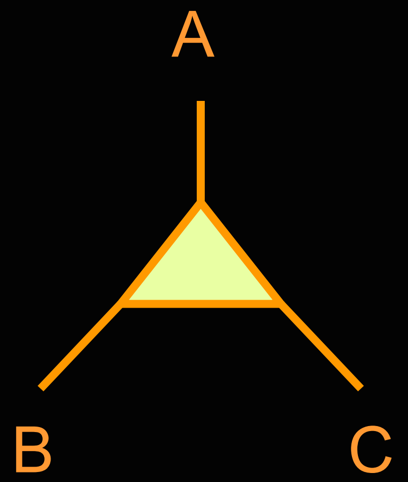
Sigma-Pi units

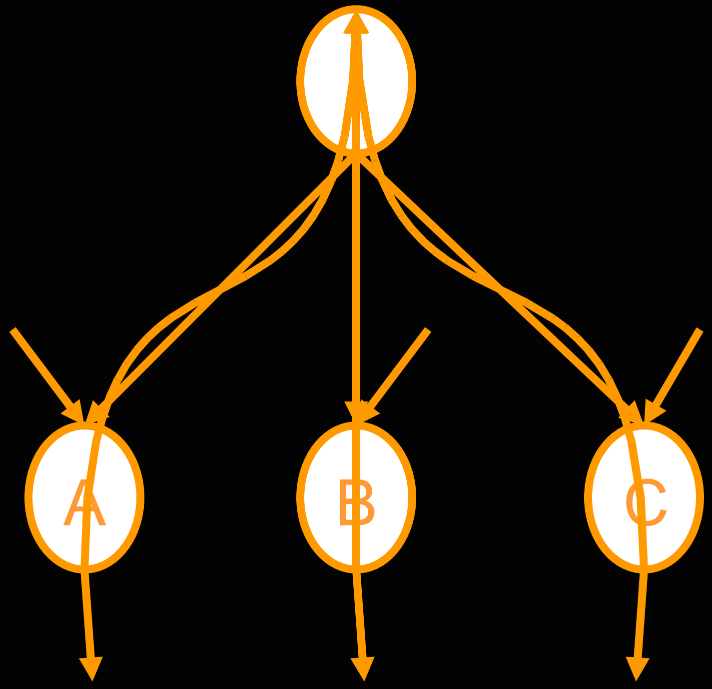


Sigma-Pi Unit

$$h_j = \sum_i w_{ji} \prod_k x_{i_k}$$

How we can model the triangle node with McCullough-Pitts Neurons?





Necker-Cube like illusion

Stroop effect

- What is the Stroop effect?
  - takes longer to say what color a word is printed in if it names a different color
  - suggests interaction of form and meaning (as opposed to an encapsulated ‘language module’)
- Red Blue Green Orange Red Yellow
- Green Yellow Red Blue Yellow Orange

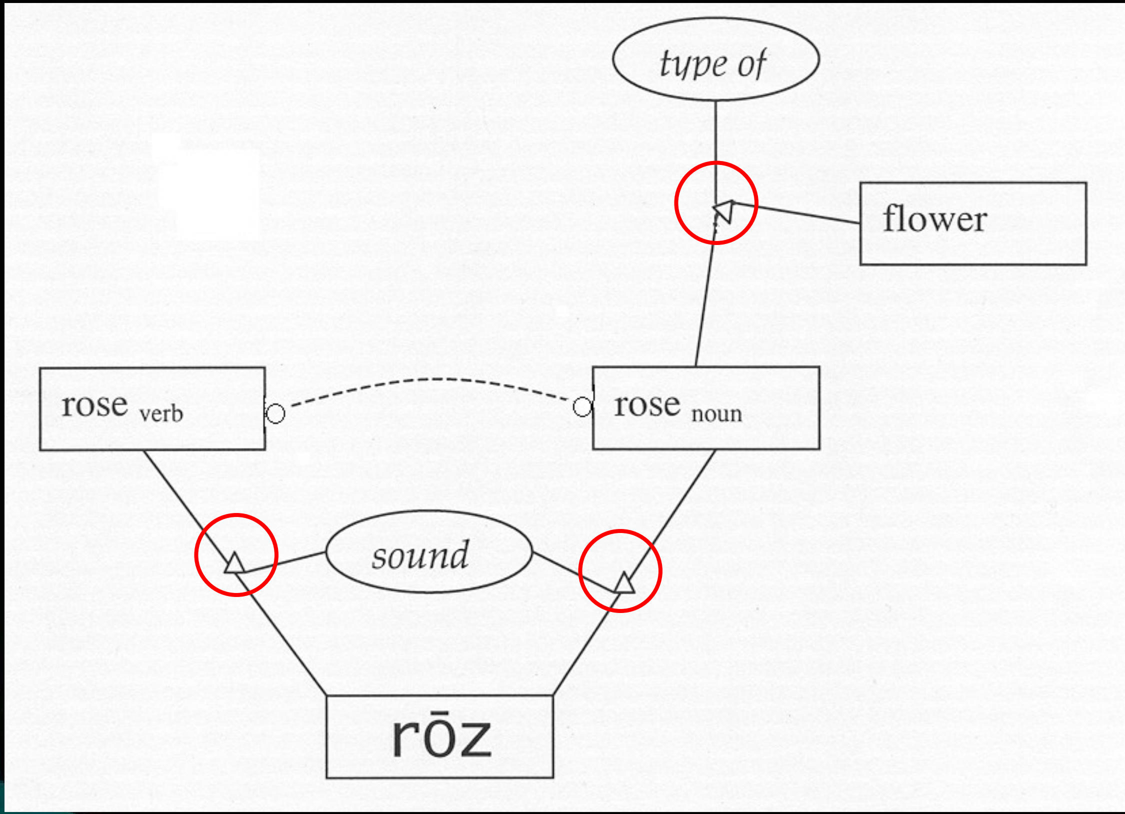
‘Word superiority effect’

- What is the word superiority effect?
  - it’s easier to remember letters if they are seen in the context of a word
  - militates against ‘bottom-up’ model, where word recognition is built up from letters
  - suggestion: there are top-down and bottom-up processes which interact

Priming and Related Concepts

- What is priming, and how do we measure it?
  - Faster completion of a task
  - Present stimulus while subject performing task
  - Different stimulus timing causes different effects
- What is the effect of suggesting an alternate meaning of a word, like saying “They all rose”?
  - If early, primes
  - If late, decreases priming

“They all rose”



triangle nodes:

when two of the neurons fire, the third also fires

model of spreading activation

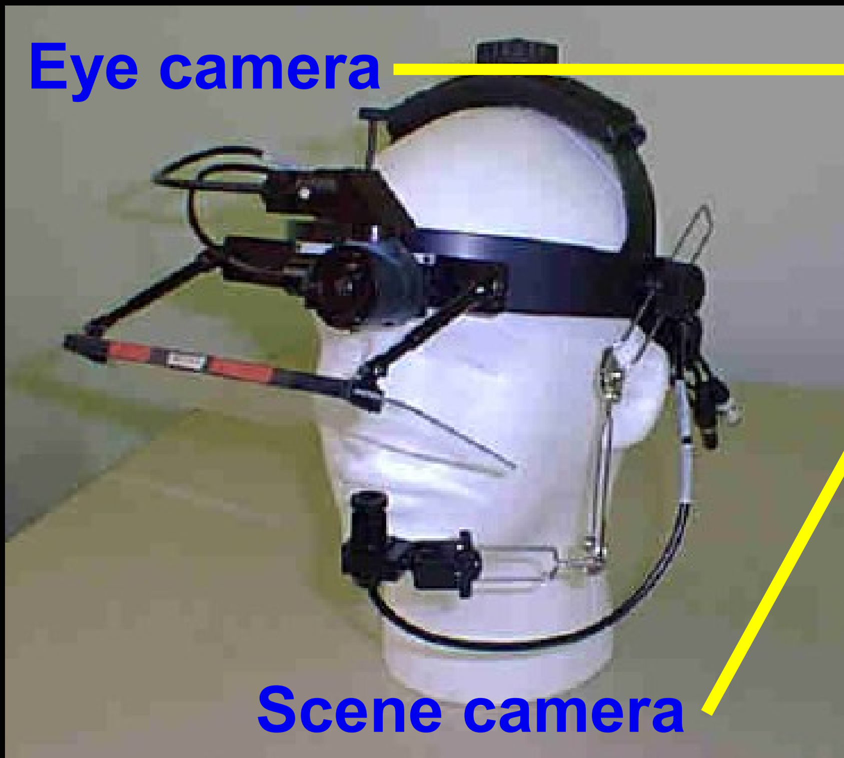
Body part understanding task

- Tests priming and interference when showing images of actions
- It turns out that viewing actions affects recognizing action words
- Can you explain the results (next slide)?

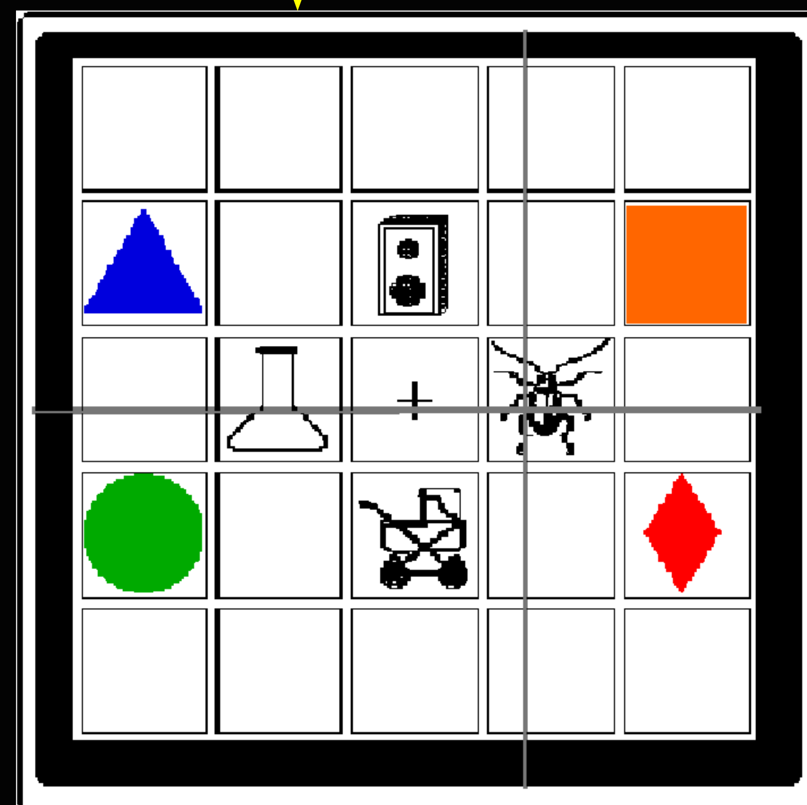
Preliminary Behavior Results

	Same Action	Other Effector	Same Effector
40 Native Speakers	788	804	871
Eliminate RT > 2 sec.	767	785	825

Allopenna, Magnuson & Tanenhaus (1998)



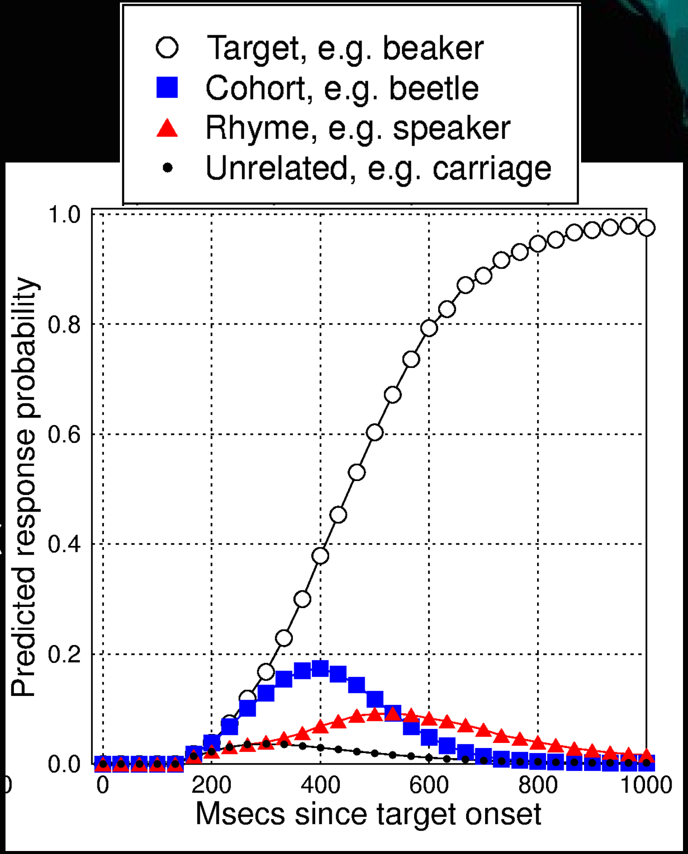
Eye tracking computer



‘Pick up the beaker’

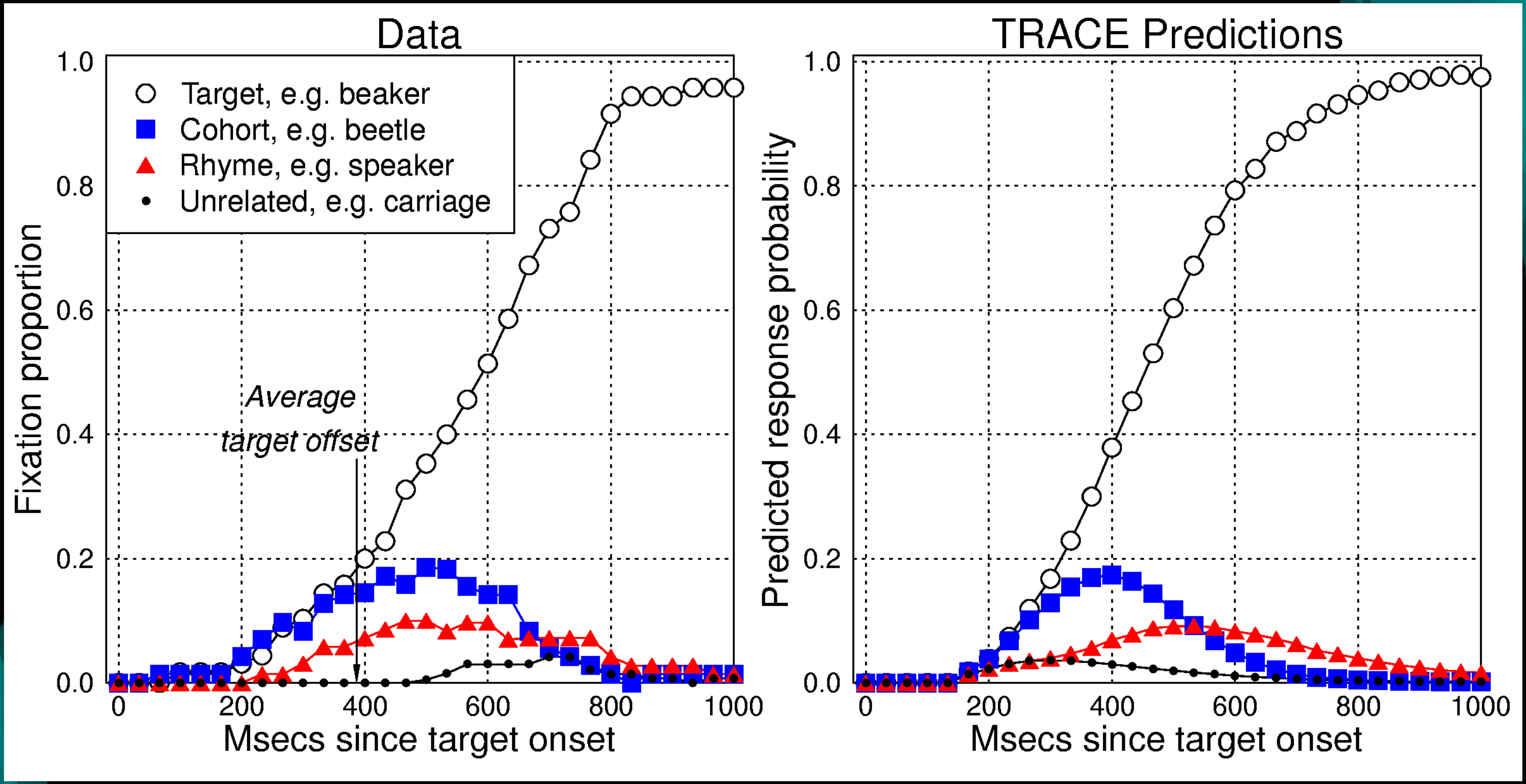
Do rhymes compete?

- Cohort (Marlsen-Wilson): onset similarity is primary because of the incremental nature of speech
  (serial/staged; Shortlist/Merge)
  - Cat activates cap, cast, cattle, camera, etc.
  - Rhymes won't compete
- NAM (Neighborhood Activation Model; Luce): global similarity is primary
  - Cat activates bat, rat, cot, cast, etc
  - Rhymes among set of strong competitors
- TRACE (McClelland & Elman): global similarity constrained by incremental nature of speech
  - Cohorts and rhymes compete, but with different time course



TRACE predictions

Allopenna et al. Results



Study 1 Conclusions

- As predicted by interactive models, cohorts and rhymes are activated, with different time courses
- Eye movement paradigm
  - More sensitive than conventional paradigms
  - More naturalistic
  - Simultaneous measures of multiple items
  - Transparently linkable to computational model
- Time locked to speech at a fine grain

Theoretical conclusions

- Natural contexts provide strong constraints *that are used*
- When those constraints are extremely predictive, they are integrated as quickly as we can measure
- Suggests rapid, continuous interaction among
  - Linguistic levels
  - Nonlinguistic context
- Even for processes assumed to be low-level and automatic
- Constrains processing theories, also has implications for, e.g., learnability