

Learning to Recognise Relevant Ambient Speech: a computational model

July 7, 2007

Gregory A. Caza
MSc Candidate, Cognitive Science
Advisor: Dr. Alistair Knott
University of Otago, Dunedin

Uncorrelated Streams

- Infants exposed to potentially overwhelming verbal information
- Parallel streams with low correlation
 - verbal/auditory information
 - sensorimotor/visual information

| | | | | | | |
|-------|---------|--------|-------|-------|--------|-------|
| DOG | DOG | MOTHER | CAT | FOX | DOG | BABY |
| “run” | “sleep” | “baby” | “cat” | “cat” | “jump” | “eat” |

Uncorrelated Streams

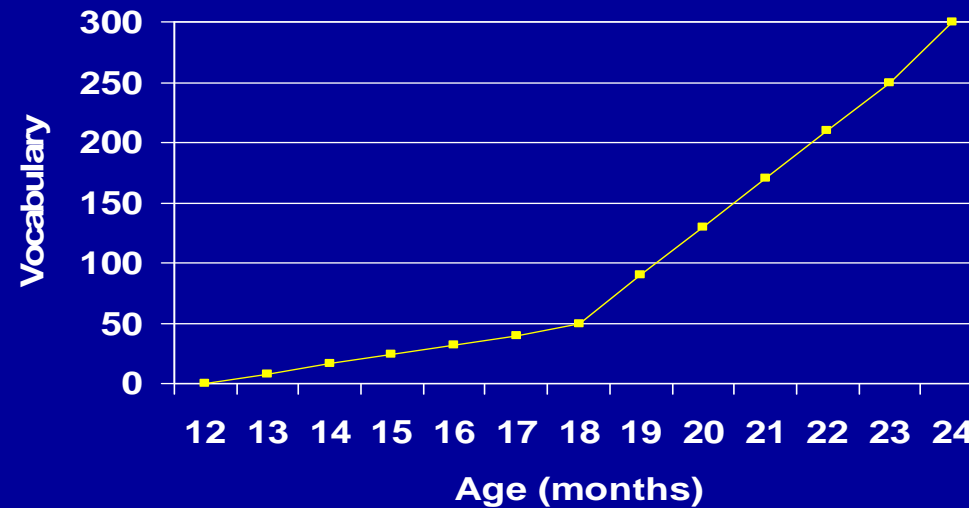
- Infants exposed to potentially overwhelming verbal information
- Parallel streams with low correlation
 - verbal/auditory information
 - sensorimotor information

| | | | | | | |
|-------|---------|--------|-------|-------|--------|-------|
| DOG | DOG | MOTHER | CAT | FOX | DOG | BABY |
| “run” | “sleep” | “baby” | “cat” | “cat” | “jump” | “eat” |

- How are words learned?

Word Learning

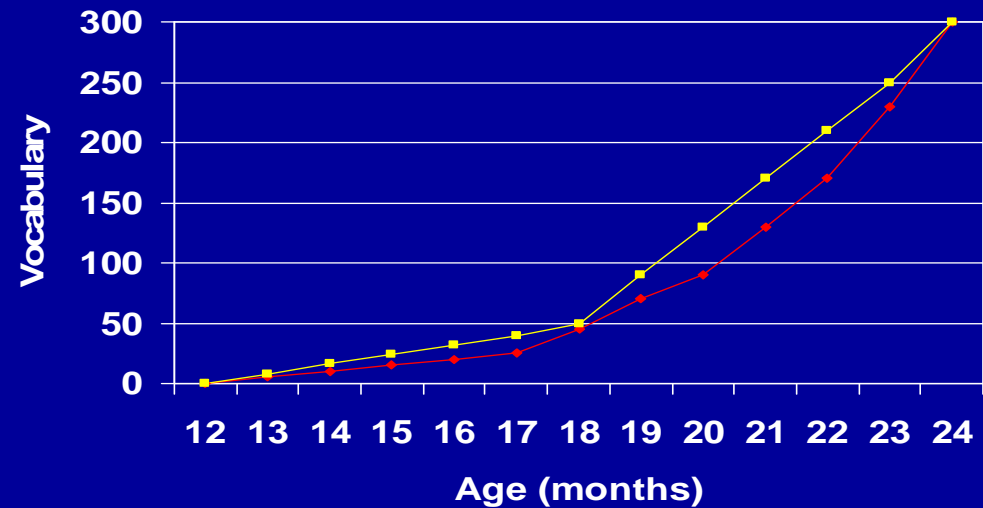
- Is there a vocabulary spurt?



adapted from O'Grady
(2004)

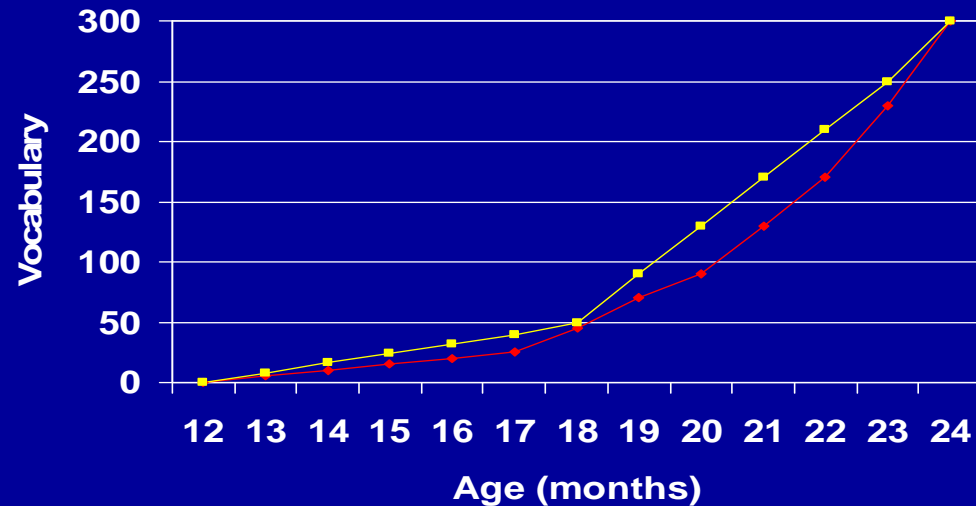
Word Learning

- Is there a vocabulary spurt?



Word Learning

- Is there a vocabulary spurt?



- Possibly only true for 18% of children¹
- Start slow, get faster
- Which cognitive skills facilitate word learning?

¹Ganger and Brent, 2004

Communicative Situations

- 9- to 12-month-olds share, follow, and direct attention
- Word learning requires understanding the intentions of utterances
- Crucial skills of intention-reading are developed around the first birthday

Tomasello, 2003

Joint Attention

- Triadic interaction between infant and caregiver
- Robust gaze following possible by 18 months¹
- 19- to 20-month-olds give referential cues priority² when learning object labels
- Assignment may be dependent on clear referential intent³

¹Triesch et al., 2006

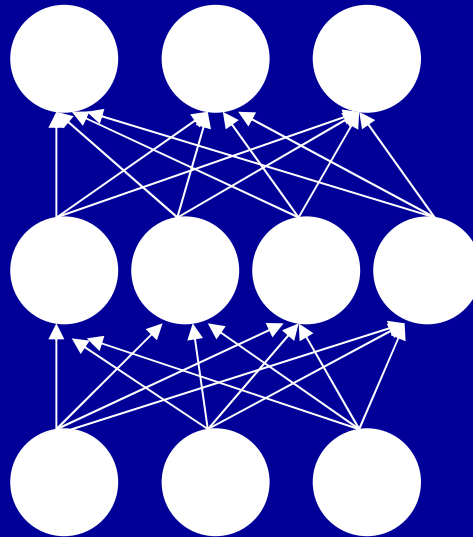
²Baldwin, 1993

³Baldwin et al., 1996

Model of Simple Word Learning

- Simple, back-propagation network

Output: Current Word?



Input: SM Item

Model of Simple Word Learning

- Possibly useful for bootstrapping
- Does not model rapid word learning
- Learning is cluttered with noise and SLOW

Model of Simple Word Learning

- Possibly useful for bootstrapping
 - Does not model rapid word learning
 - Learning is cluttered with noise and SLOW
-
- **How can the model be improved?**

Modelling Joint Attention

| | | | | | | |
|--|--|--------|---------------------------------|--|--|--|
| | | MOTHER | <i>follow Mother's gaze</i> | | | |
| | | | ??? | | | |

Modelling Joint Attention

| | | | | | | |
|--|--|--------|---------------------------------|--|--|--|
| | | MOTHER | <i>follow Mother's gaze</i> | | | |
| | | | ??? | | | |

| | | | | | | |
|--|--|--------|---------------------------------|--|--|--|
| | | MOTHER | <i>follow Mother's gaze</i> | | | |
| | | TALK | | | | |
| | | | <i>hear correct label</i> | | | |

Modelling Joint Attention

| | | | | | | |
|--|--|--------|---------------------------------|--|--|--|
| | | MOTHER | <i>follow Mother's gaze</i> | | | |
| | | | ??? | | | |

| | | | | | | |
|--|--|--------|---------------------------------|--|--|--|
| | | MOTHER | <i>follow Mother's gaze</i> | | | |
| | | TALK | | | | |
| | | | <i>hear correct label</i> | | | |

| | | | | | | |
|-------|---------|--------|-------|-------|--------|-------|
| DOG | DOG | MOTHER | CAT | FOX | DOG | BABY |
| | JUMP | TALK | | | RUN | SLEEP |
| "run" | "sleep" | | "cat" | "cat" | "jump" | "eat" |

Modelling Joint Attention

| | | | | | | |
|--|--|--------|---------------------------------|--|--|--|
| | | MOTHER | <i>follow Mother's gaze</i> | | | |
| | | | ??? | | | |

| | | | | | | |
|--|--|--------|---------------------------------|--|--|--|
| | | MOTHER | <i>follow Mother's gaze</i> | | | |
| | | TALK | | | | |
| | | | <i>hear correct label</i> | | | |

| | | | | | | |
|-------|---------|--------|-------|-------|--------|-------|
| DOG | DOG | MOTHER | CAT | FOX | DOG | BABY |
| | JUMP | TALK | | | RUN | SLEEP |
| "run" | "sleep" | | "cat" | "cat" | "jump" | "eat" |

Recognising Relevant Speech

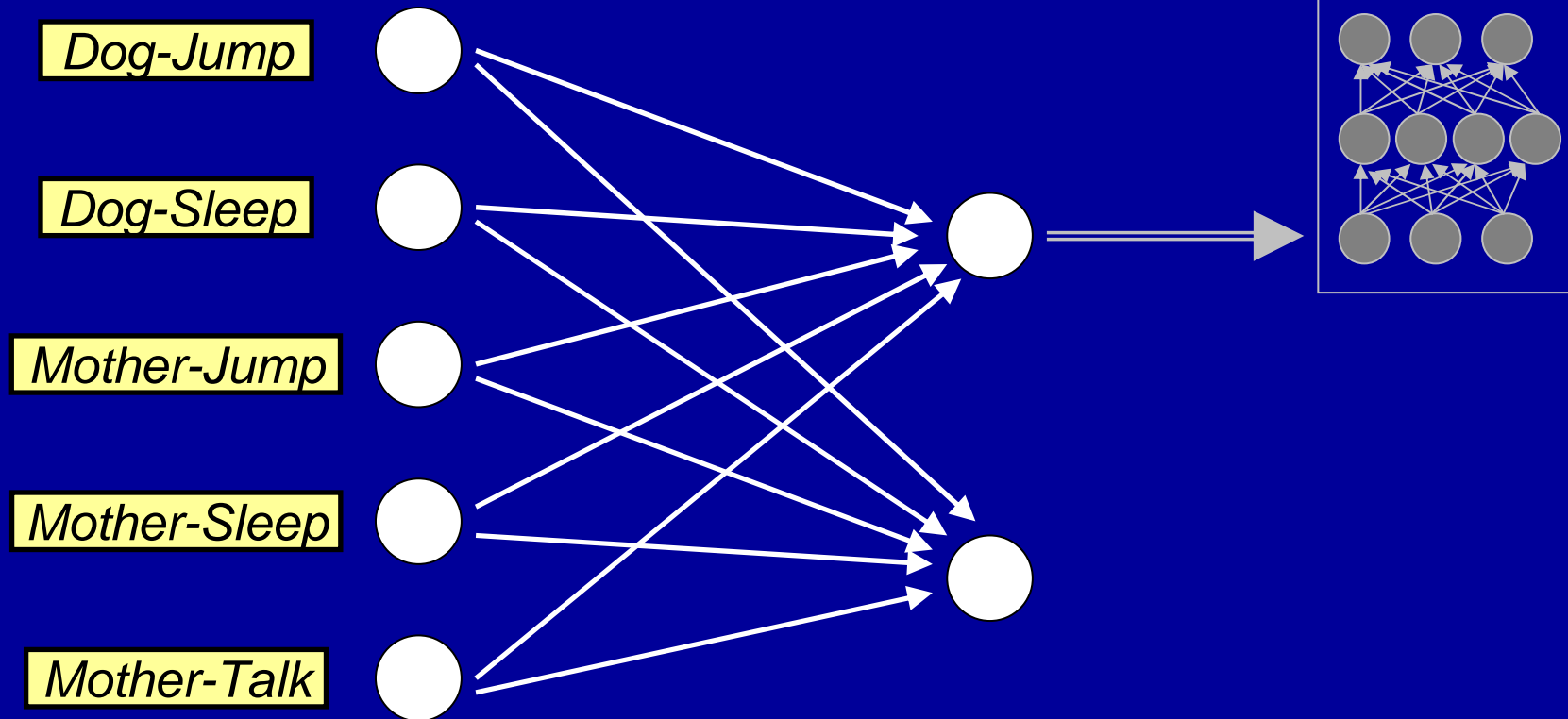
- Noise can be filtered by gating the input to single word learning
- Language learning is only enabled when the crucial trigger is recognised

MOTHER-TALK => follow gaze

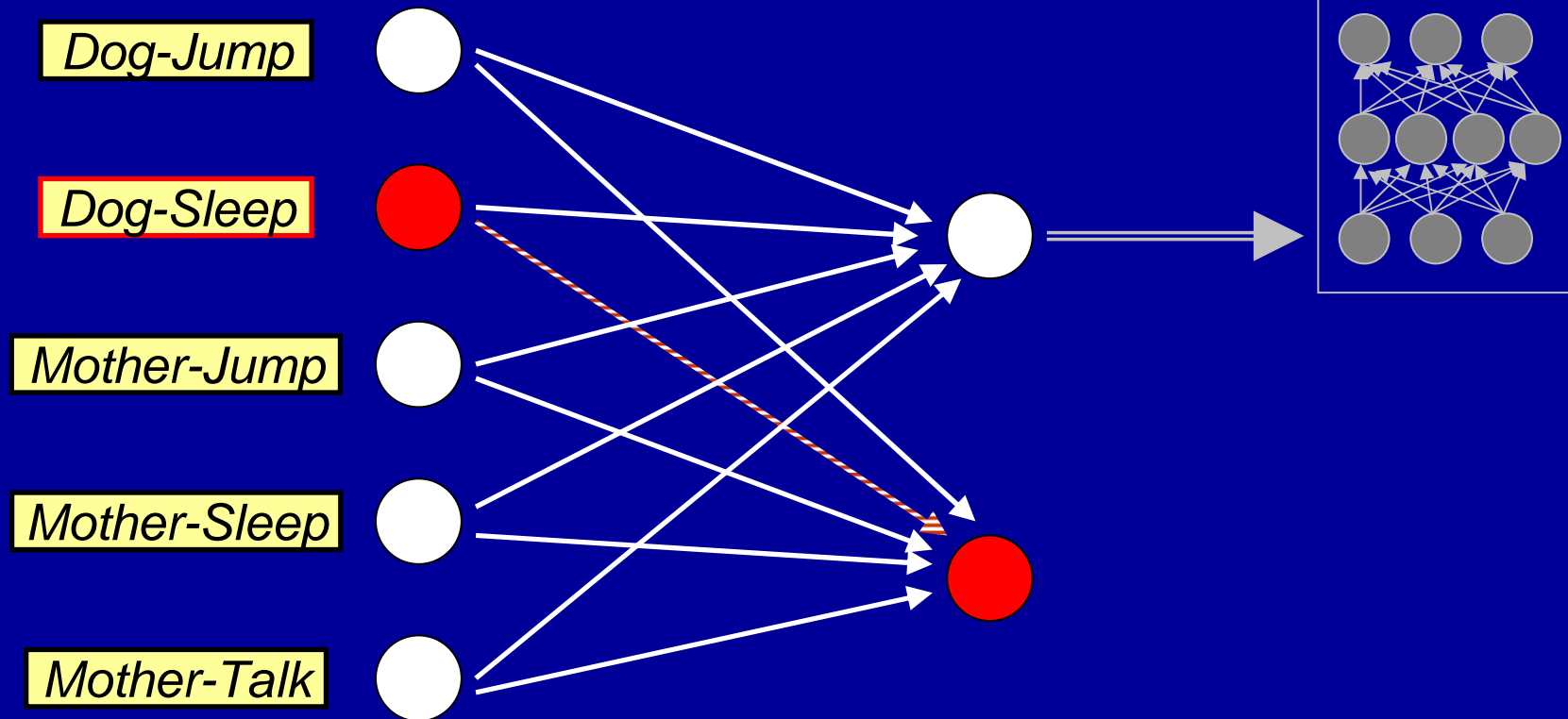
- Learning is reinforcement-based¹
- Feedback from word-learning network controls reward/punishment

¹Knecht et al., 2004

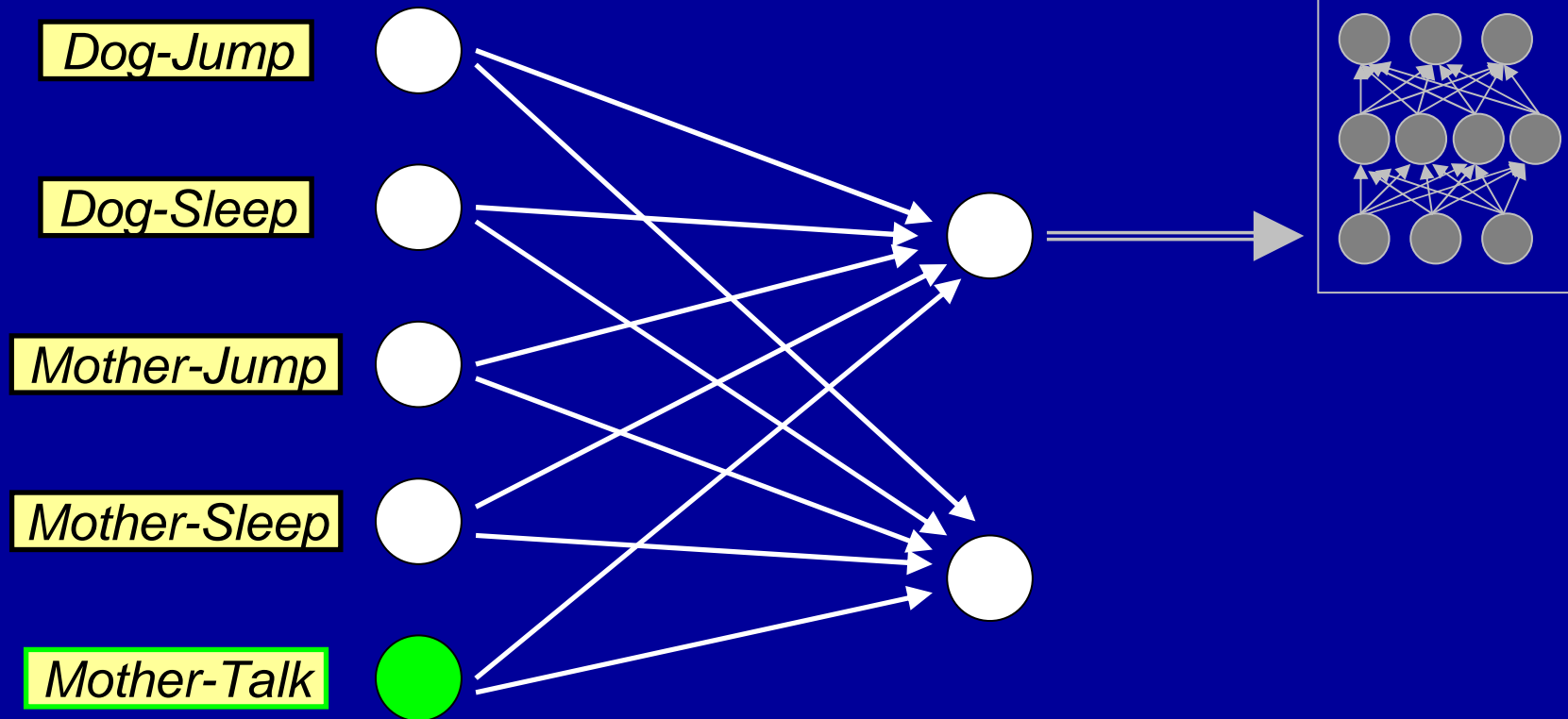
Network



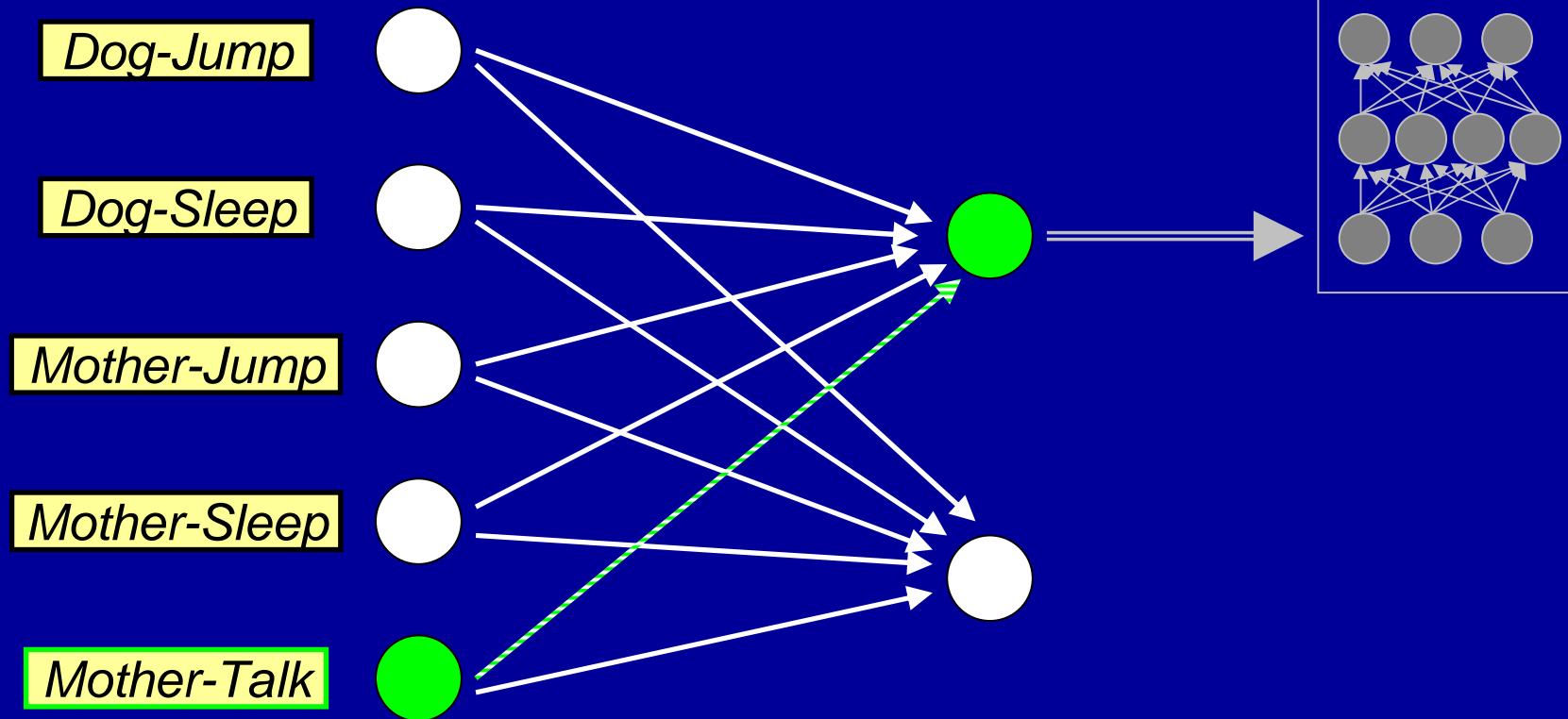
Network



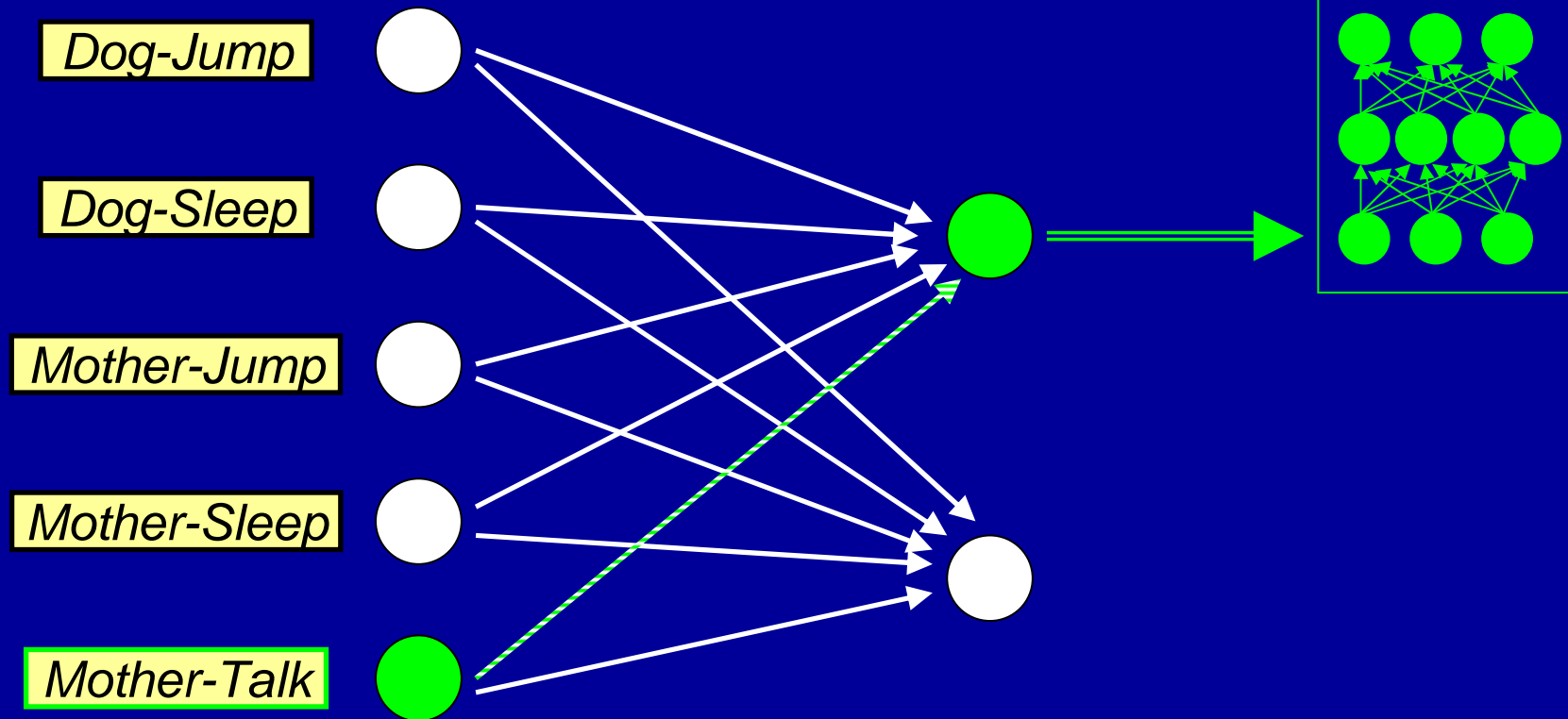
Network



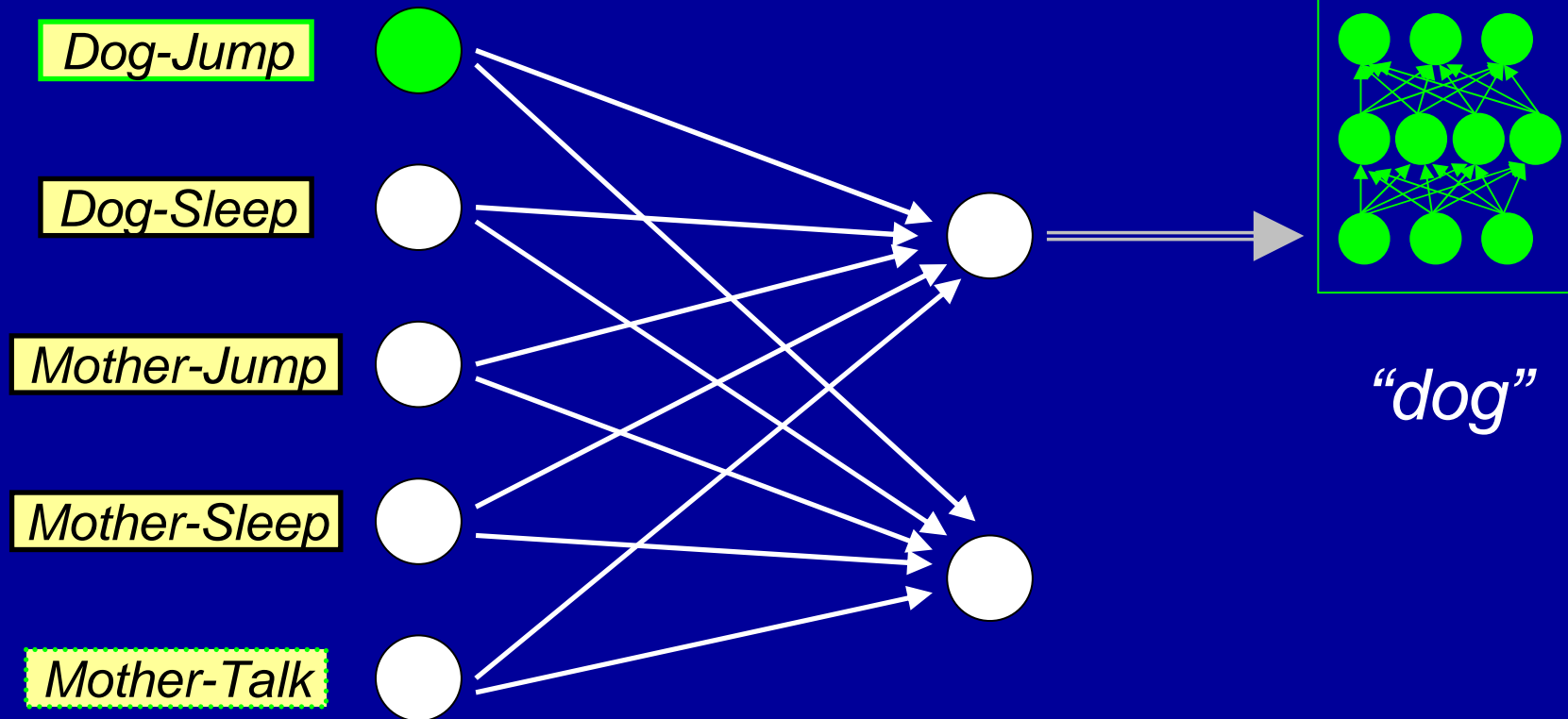
Network



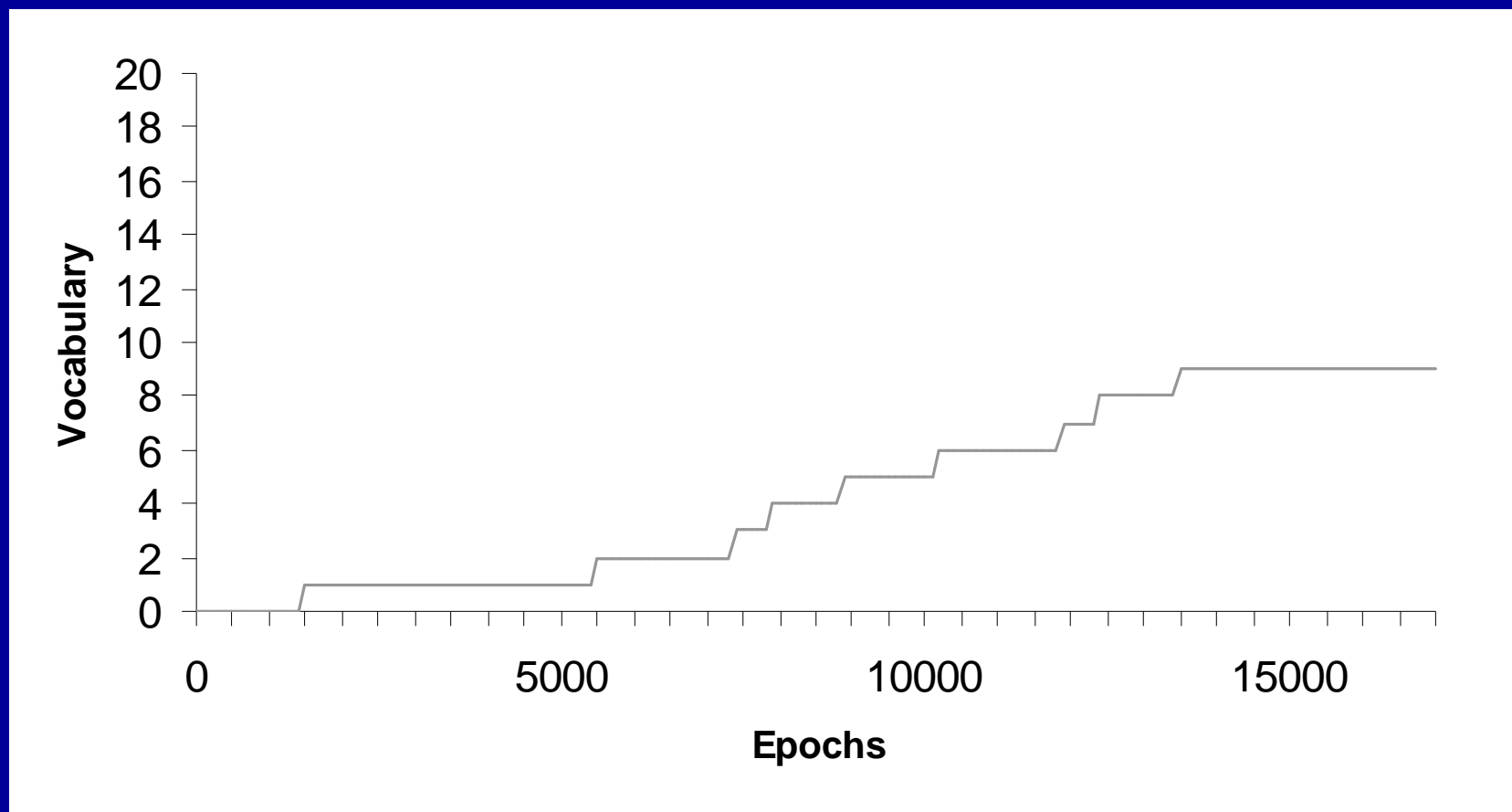
Network



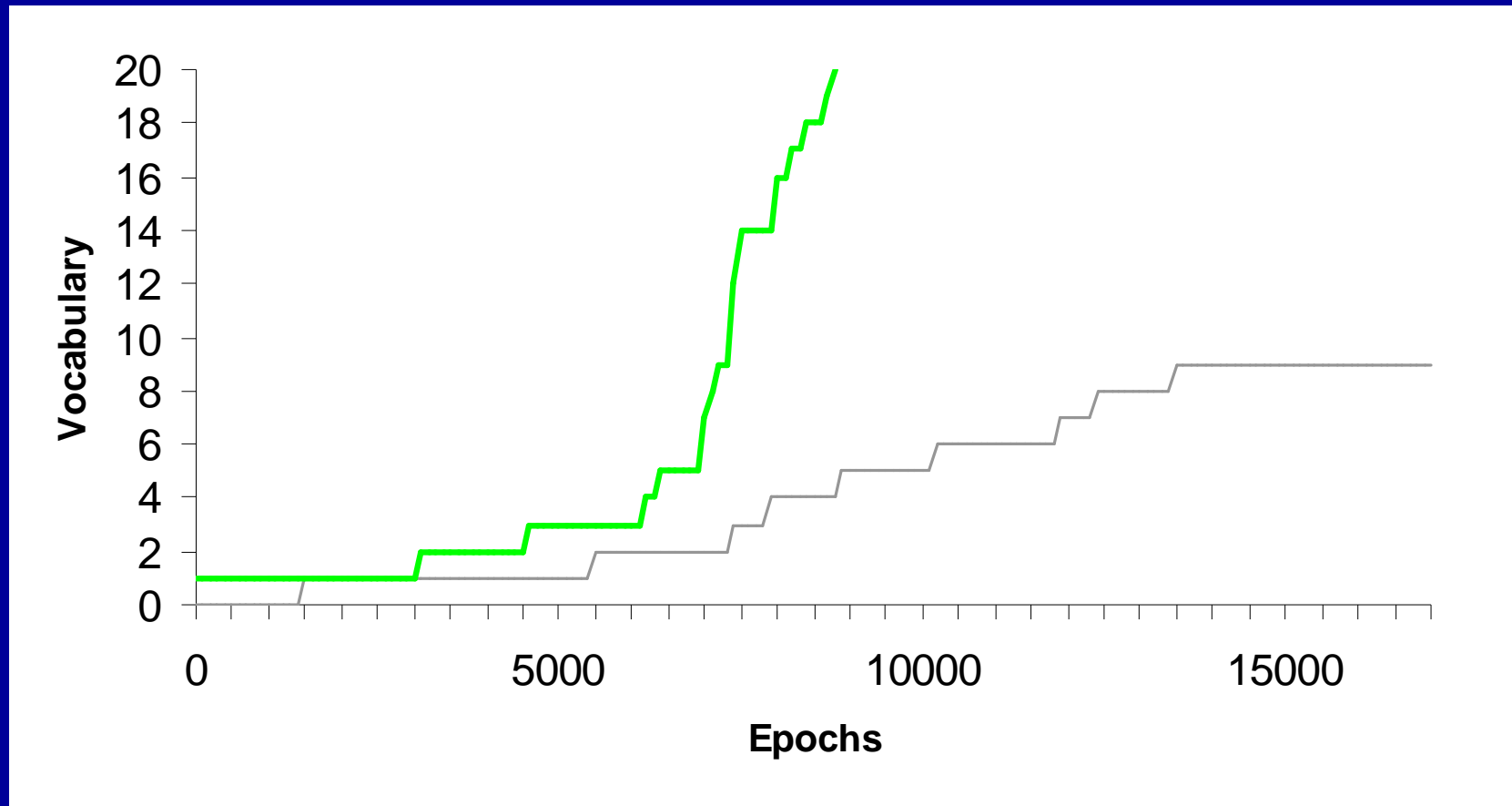
Network



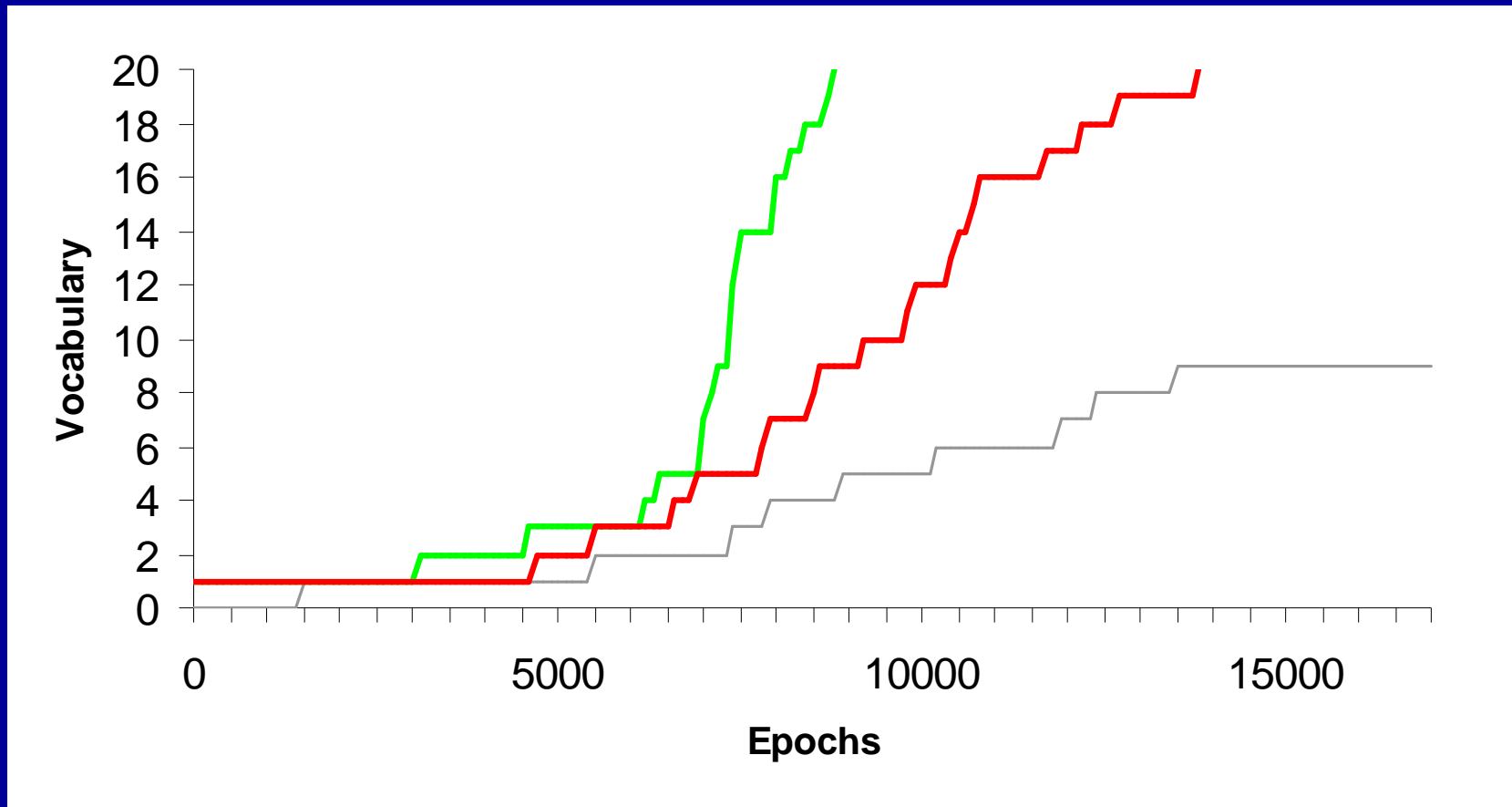
Results



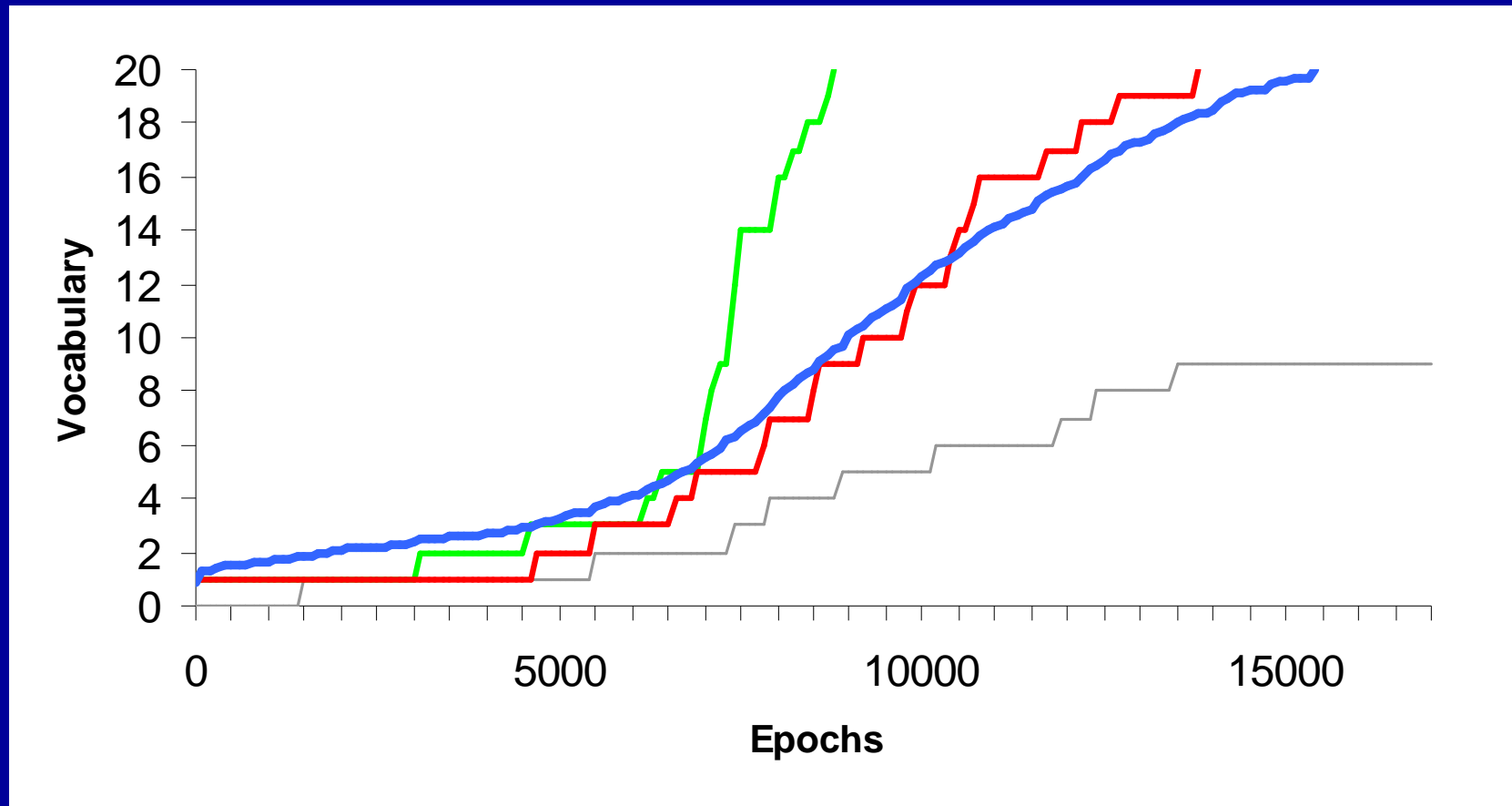
Results



Results



Results



Future Work

- Work with a larger vocabulary
- Experiment with different critical periods
- Upgrade reinforcement scheme to full-fledged temporal difference learning
- Dynamically grow network as vocabulary increases

References

- Baldwin, D. A. (1993). Early Referential Understanding: Infants' Ability to Recognize Referential Acts for What They Are. *Developmental Psychology*, **29**(5), 832-843
- Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., and Tidball, G. (1996). Infants' Reliance on a Social Criterion for Establishing Word-Object Relations. *Child Development*, **67**, 3135-3153.
- Ganger, J and Brent, M. R. (2004). Reexamining the Vocabulary Spurt. *Developmental Psychology*, **40**(4), 621-632.
- Knecht, S., Breitenstein, C., Bushuven, S., Wailke, S., Kamping, S., Flöel, A., Zwisterlood, P, and Ringelstein, E. B. (2004). Levodopa: Faster and Better Word Learning in Normal Humans. *Annals of Neurology*, **56**(1), 20-26.
- O'Grady, W. (2005). *How Children Learn Language*. Cambridge, UK, Cambridge University Press.
- Tomasello, M. (2003). *Constructing a Language: a usage-based theory of language acquisition*. Cambridge, MA, USA, Harvard University Press.
- Triesch, J., Teuscher, C., Deak, G. O., and Carlson, E. (2006). Gaze following: why (not) learn it? *Developmental Science*, **9**(2), 125-157

Acknowledgements

Dr. Alistair Knott

Edwin van der Ham, University of Twente

Joost van Oijen, University of Twente