COSC421: Neural Models of Language

Lecture 7: How infants learn words

Alistair Knott

Dept. of Computer Science, University of Otago

Recap

- Lecture 1: the visual system
- Lecture 2: execution of reach-to-grasp actions
- Lecture 3: perception of reach-to-grasp actions
- Lecture 4: the ‘who’ pathway, and sequential structure in the reach-to-grasp system
- Lecture 5: WM for episodes and phonological material
- Lecture 6: language networks in the brain

Recap

Lecture 6 again:

- How phonological representations are encoded in the brain
  - The idea of a mirror system for phonemes
  - The idea of phonological input and output buffers
- How words are encoded in the brain
  - As phonological units
  - As semantic units
  - As associations between phonological and semantic units
- Grammatical processing in the brain
  - Broca's aphasia and Wernicke's aphasia

Outline of today's lecture

1. Development of phonological word representations
2. Development of object and action concepts
3. Development of knowledge of word meanings

Outline of today's lecture

Development of phonological word representations

There's good evidence that infants can form representations of words *as commonly-occurring phonological sequences* by around 8 months (Saffran *et al.*, 1996).

- In a training phase, an infant is played a stream of phonemes, where some phoneme sequences occur more often than chance. E.g. *ga bi ro to ba di ga bi ro pa la ku to ba di*

Development of phonological word representations

There's good evidence that infants can form representations of words *as commonly-occurring phonological sequences* by around 8 months (Saffran *et al.*, 1996).

- In a training phase, an infant is played a stream of phonemes, where some phoneme sequences occur more often than chance. E.g. *ga bi ro to ba di ga bi ro pa la ku to ba di*

Development of phonological word representations

There's good evidence that infants can form representations of words *as commonly-occurring phonological sequences* by around 8 months (Saffran *et al.*, 1996).

- In a training phase, an infant is played a stream of phonemes, where some phoneme sequences occur more often than chance. E.g. *ga bi ro to ba di ga bi ro pa la ku to ba di*

Development of phonological word representations

There's good evidence that infants can form representations of words *as commonly-occurring phonological sequences* by around 8 months (Saffran *et al.*, 1996).

- In a training phase, an infant is played a stream of phonemes, where some phoneme sequences occur more often than chance. E.g. *ga bi ro to ba di ga bi ro pa la ku to ba di*
- In a testing phase, infants were presented with a mixture of high-probability sequences (*to ba di*) and low-probability sequences (*di ga bi*).

Development of phonological word representations

There's good evidence that infants can form representations of words *as commonly-occurring phonological sequences* by around 8 months (Saffran *et al.*, 1996).

- In a training phase, an infant is played a stream of phonemes, where some phoneme sequences occur more often than chance. E.g. *ga bi ro to ba di ga bi ro pa la ku to ba di*
- In a testing phase, infants were presented with a mixture of high-probability sequences (*to ba di*) and low-probability sequences (*di ga bi*).
- They showed longer listening times for the low-probablity sequences, suggesting that they were perceived as novel.

Development of phonological word representations

There's good evidence that infants can form representations of words *as commonly-occurring phonological sequences* by around 8 months (Saffran *et al.*, 1996).

- In a training phase, an infant is played a stream of phonemes, where some phoneme sequences occur more often than chance. E.g. *ga bi ro to ba di ga bi ro pa la ku to ba di*
- In a testing phase, infants were presented with a mixture of high-probability sequences (*to ba di*) and low-probability sequences (*di ga bi*).
- They showed longer listening times for the low-probablity sequences, suggesting that they were perceived as novel.

By 9 months, infants prefer to listen to nonwords which conform to the phonological rules of the exposure language (Jusczyk *et al.*, 1994)

Development of phonological output representations

Infants produce sounds from birth.

- Children *imitate* the speech sounds they hear from very early. E.g. vowels are imitated from around 1 month old (Kuhl and Meltzoff, 1996).
- So the mirror system for phonemes starts to be learned very early.

Infants start to use **syllables** ('babble') around 6 months.

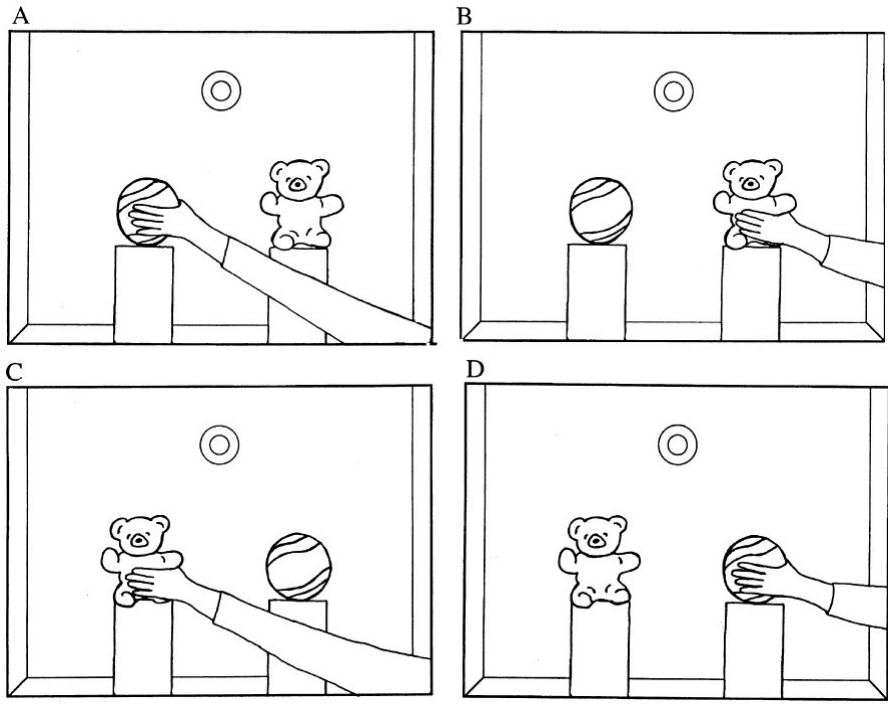
- To begin with, babbling is repeating a single syllable (*bababa*).
- By around 8 months, infants can produce short sequences of varied syllables (e.g. *baga*).
- Infants begin to *imitate* syllables at the time they start to babble (see e.g. Oller, 2000).

Outline of today's lecture

Development of object and action concepts

By 8 months:

- Infants can perform reach-to-grasp actions (Dimitrijevi and Bjelakovi, 2004).
- Infants are sensitive to the intended target of a reach-to-grasp action (Woodward, 1998).



Development of object and action concepts

By 8 months:

- Infants have a good understanding of the support/contact relations involved in grasping (Leslie, 1984).
- Infants have a good awareness of the spatiotemporal continuity of objects (Spelke *et al.*, 1994).
- Infants can recognise common objects in a variety of poses (Ruff, 1978).

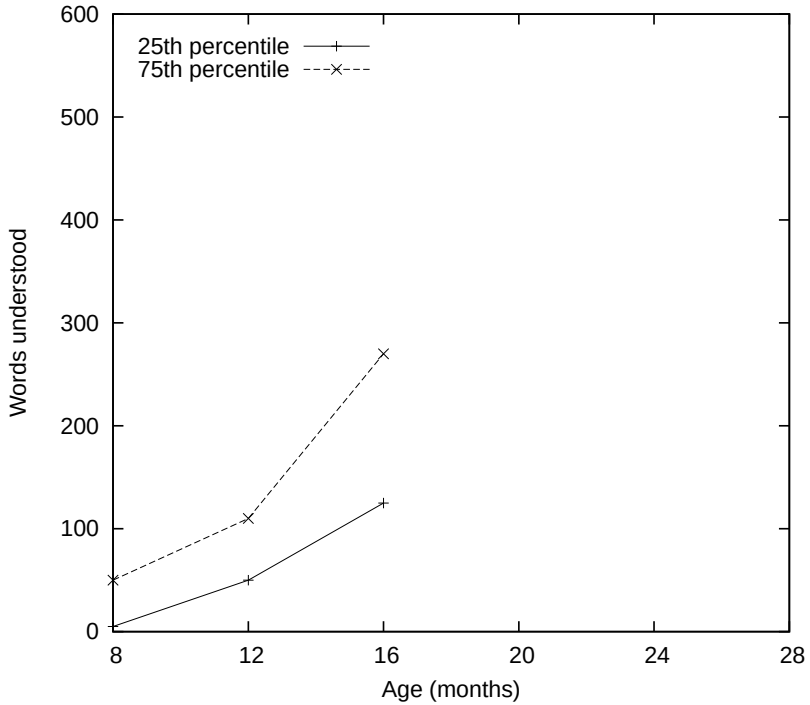
So at 8 months:

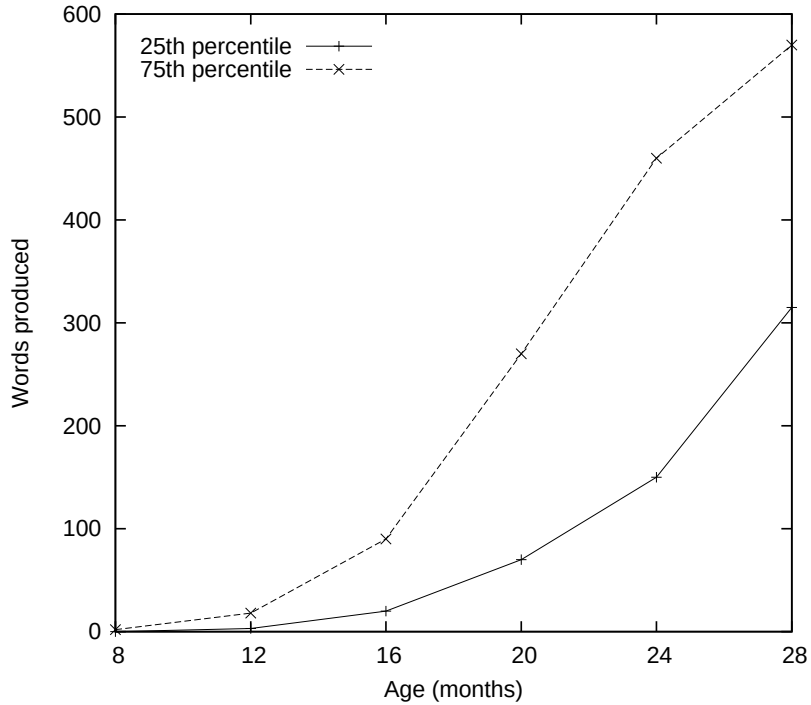
- Infants have well-developed phonological word representations;
- Infants (apparently) have well-enough developed SM concepts to serve as semantic representations of concrete words.

Outline of today's lecture

Development of knowledge of word meanings

Some summary data from a well-normed study of US infants (Fenson *et al.*, 1994):





- 8–12 months: infants acquire a small **passive vocabulary**.
- Active vocabulary really takes off at around 12 months.

Basic associationist model of word meaning learning

The basic model:

- An infant perceives an object or action, and evokes a semantic representation.
- An adult says the word which denotes this object/action, and the infant evokes a phonological representation.
- Associations are strengthened in the infant between the semantic and phonological representations.

This simple idea is partly right. But there are a few problems.

Noisy learning situations

The stream of words infants hear is often **very decorrelated** from the stream of semantic representations they evoke.

Correspondences between word forms and word meanings take the form of weak *statistical tendencies*, reflecting the fact that parents talk about things the infant is attending to ‘more than chance’.

Cross-situational learning (Siskind, 1996)

Key idea: in any given **situation**, only a subset of his semantic concepts will be activated.

- The infant explores the situation, evoking a stream of concepts.
- The situation also contains **mature speakers**, whose utterances often describe the current situation.
- Assume the infant is processing these utterances phonologically, and activating a stream of word forms.
- *Any word form the infant hears in this situation is more likely than chance to denote the semantic concept he is currently evoking*.

So if we treat the set of word-concept pairs as training data, we'll eventually learn word meanings.

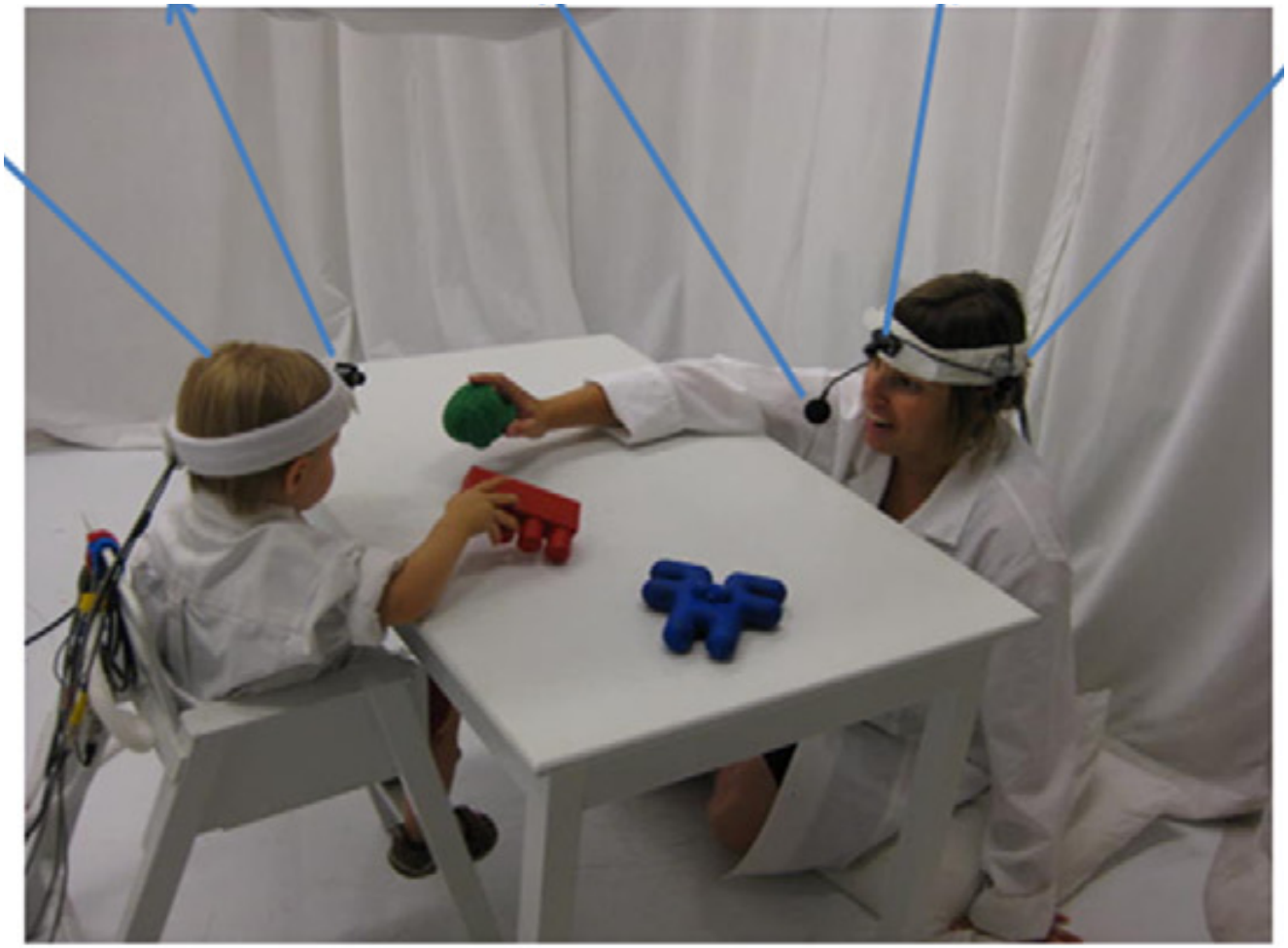
Cross-situational learning (Siskind, 1996)

Cross-situational learning works. But it's *slow*.
(Because situation-based correlations are only very weak.)

There are various improvements we can make.

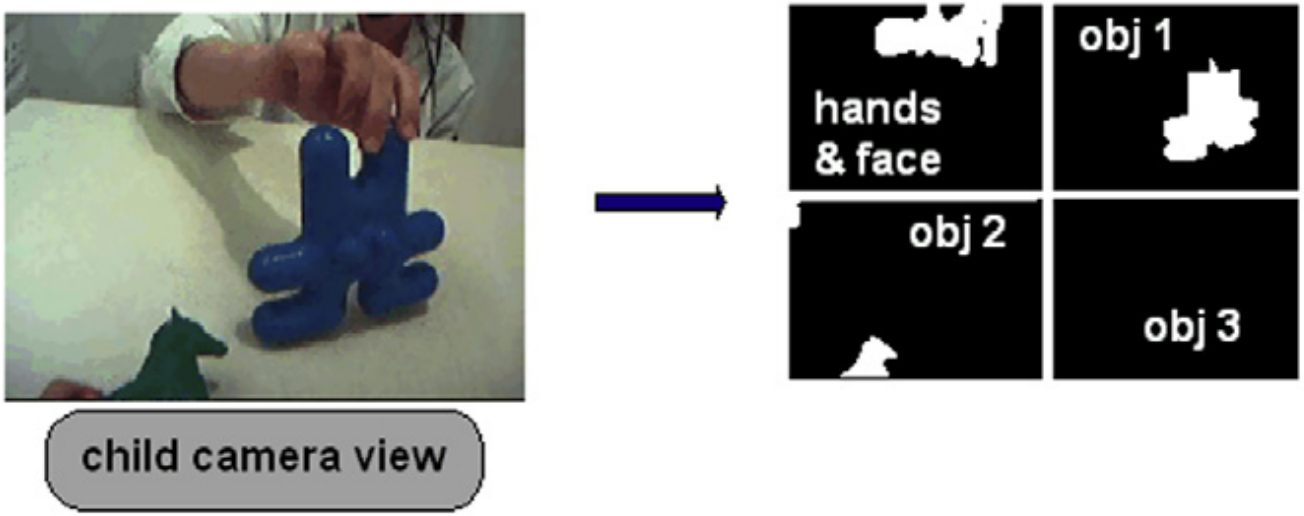
Improvements to cross-situational learning (1)

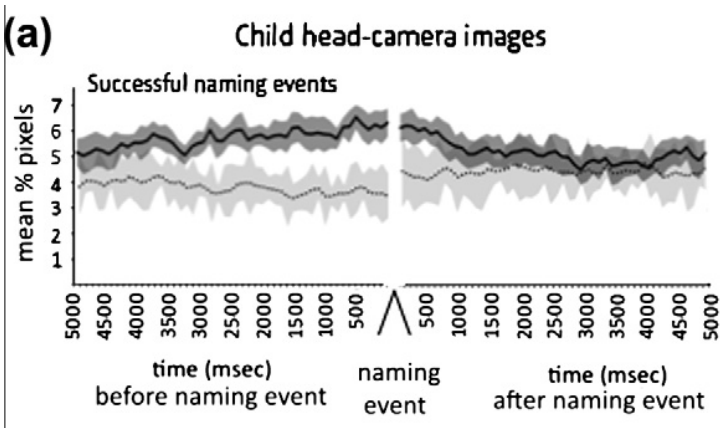
One improvement is *for adults to focus the child's attention on an object being named*. (Yu and Smith, 2012)

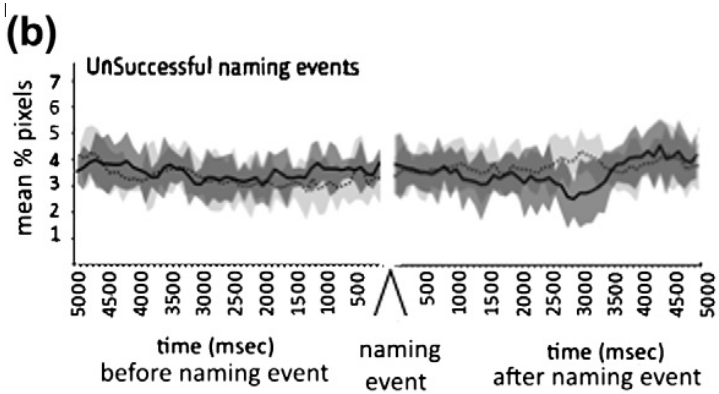


Improvements to cross-situational learning (1)

Yu and Smith: in the situations where an infant successfully learns a word from a naming event, the named object 'fills the field of view'.







Improvements to cross-situational learning

But: infants learn language even when they're not being expressly 'taught' by an adult.

In some cultures, adults don't talk much to children at all!//(Pinker, 1994)

In such cases, infants have to find ways of dealing with the noisy training data *themselves*.

Improvements to cross-situational learning (2)

A problem with cross-sit. learning is the **sparseness** of good training data. It's very unlikely that the child hears a word at *exactly* the time he activates the associated concept.

- The infant can alleviate this problem if she associates the current concept not just with the current word, but with *all the words in the phonological buffer*.
- Baddeley *et al.* (1998): children's phonological short-term memory predicts with their word-learning abilities. (E.g. at ages 4-5)

Refinements to cross-situational learning (3)

Maybe children aren't doing simple associative learning. Maybe they have a better way of identifying good word-learning opportunities.

Note:

- Nonhuman animals can do associative learning. But they can't learn word meanings like human infants do.
- If human infants use simple associative learning to acquire word meanings, it's surprising that they don't learn *faster* from around 8 months.

Most current models of word learning assume there are mechanisms which go beyond simple associative learning.

Social/pragmatic models of word meaning learning

An influential model by Michael Tomasello (2003) holds that infants need to acquire a set of **social/pragmatic skills** before they can learn words in any numbers.

There are two skills:

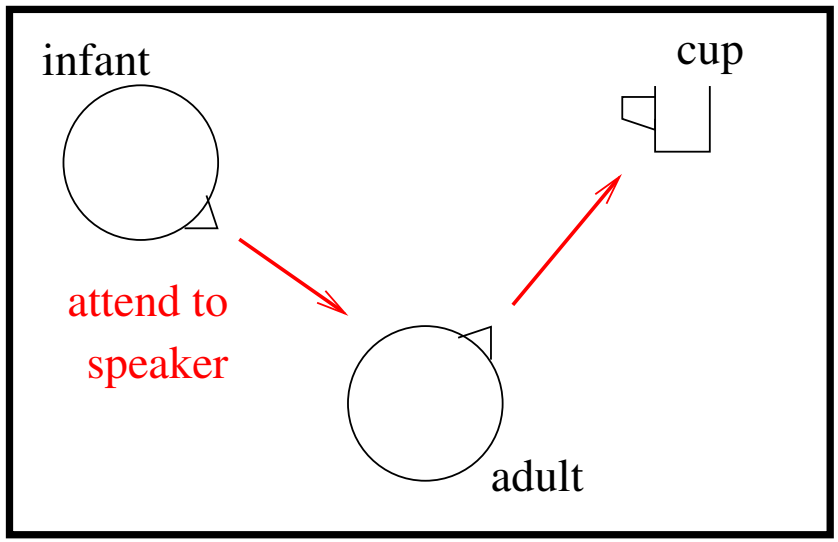
- **Joint attention**: the ability to follow an observed agent's eye gaze.
- The ability to recognise **communicative intentions**.

The role of joint attention in word meaning learning

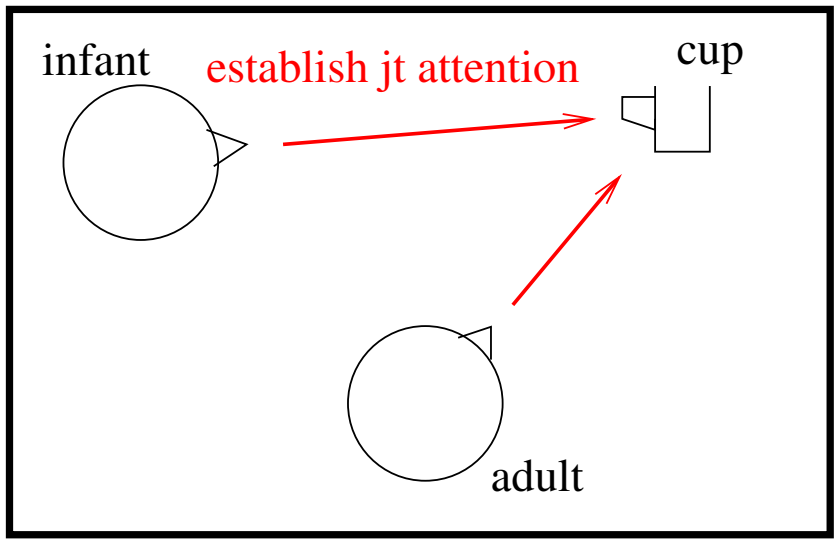
The joint attention model:

- The infant watches mature speakers talking, maybe to each other.
- While an observed speaker makes an utterance, the infant follows the speaker's gaze, and evokes semantic representations of *what the speaker is attending to*.
  (Assumption: speakers often talk about what they're attending to.)
- The infant associates *these* semantic reps with the speaker's words.

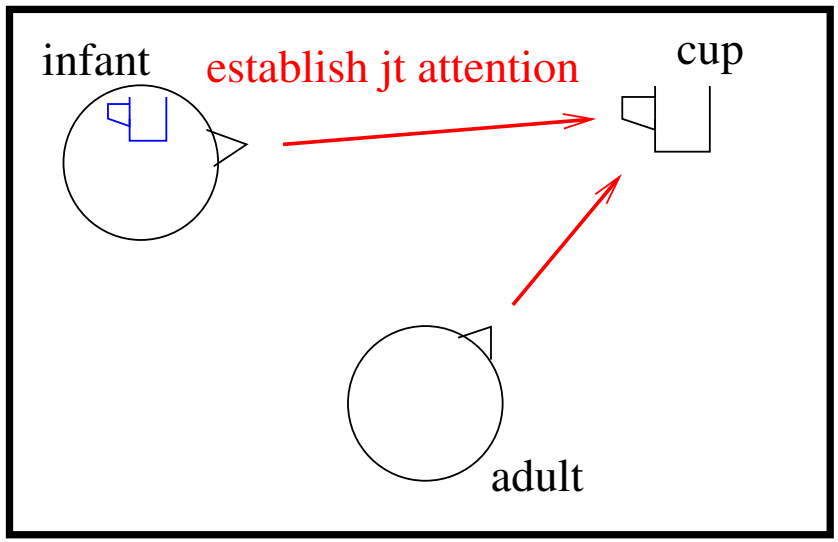
The joint attention word-learning scenario



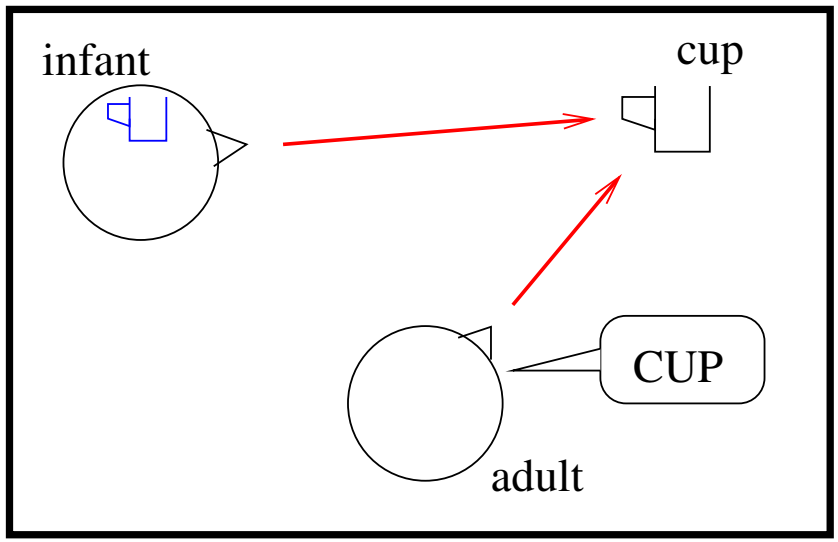
The joint attention word-learning scenario



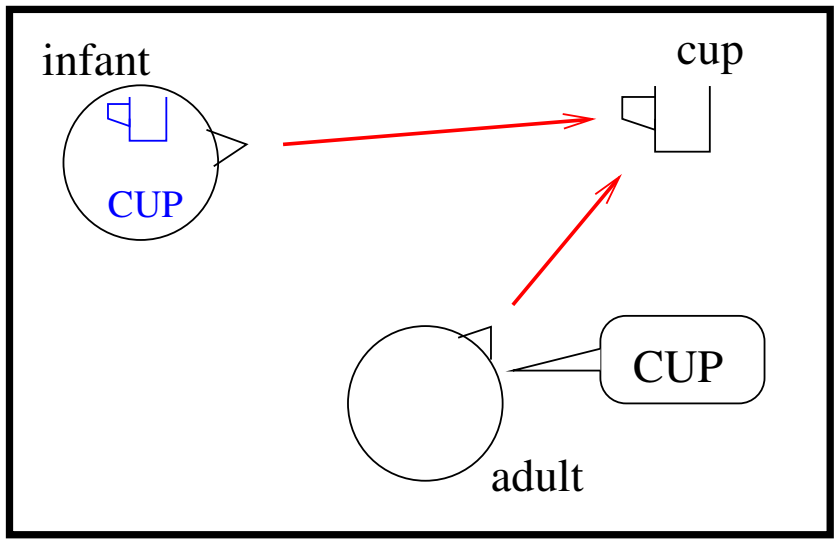
The joint attention word-learning scenario



The joint attention word-learning scenario



The joint attention word-learning scenario



Development of the joint attention capacity

Infants' ablility to establish joint attention develops gradually.

- In simple contexts (where the angle the infant must traverse to establish joint attention is small), the ability is reliable around 10 months (Scaife and Bruner, 1975).
- In more complex contexts (where the angle is larger), it doesn't become reliable until around 18 months (Butterworth and Jarrett, 1991).

Evidence for the role of joint attention in word learning

Infants' ability to establish joint attention at 12 and 18 months is predictive of their language ability at 24 months (Mundy *et al.*, 2007).

Baldwin (1991, 1993):

- Infants will follow the gaze of an observed speaker to determine the referent of a new word the speaker is using.

Evidence for the role of joint attention in word learning

Infants' ability to establish joint attention at 12 and 18 months is predictive of their language ability at 24 months (Mundy *et al.*, 2007).

Baldwin (1991, 1993):

- Infants will follow the gaze of an observed speaker to determine the referent of a new word the speaker is using.

Baldwin (1994): for 18-20-month-olds, simple correlation of words and evoked concepts isn't enough.

Evidence for the role of joint attention in word learning

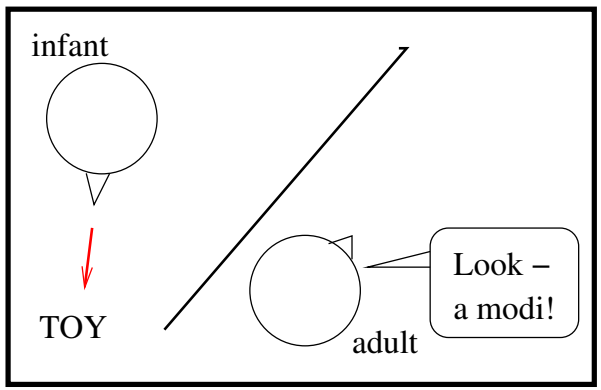
Infants' ability to establish joint attention at 12 and 18 months is predictive of their language ability at 24 months (Mundy *et al.*, 2007).

Baldwin (1991, 1993):

- Infants will follow the gaze of an observed speaker to determine the referent of a new word the speaker is using.

Baldwin (1994): for 18-20-month-olds, simple correlation of words and evoked concepts isn't enough.

- Bad learning:



Evidence for the role of joint attention in word learning

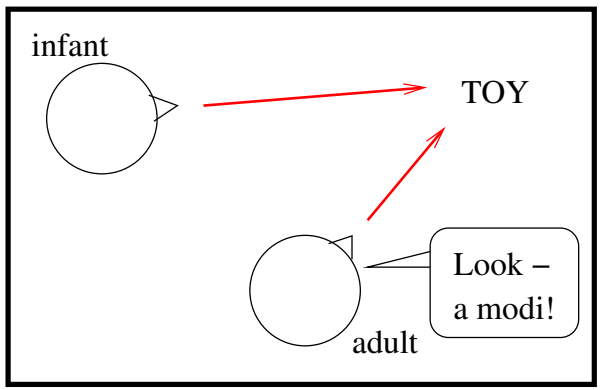
Infants' ability to establish joint attention at 12 and 18 months is predictive of their language ability at 24 months (Mundy *et al.*, 2007).

Baldwin (1991, 1993):

- Infants will follow the gaze of an observed speaker to determine the referent of a new word the speaker is using.

Baldwin (1994): for 18-20-month-olds, simple correlation of words and evoked concepts isn't enough.

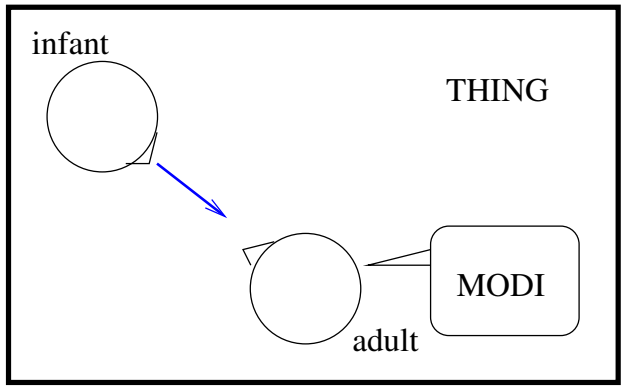
- Good learning:



Pragmatic understanding and word learning

Infants don't only need to learn what words mean.
They also need to learn *what communication is*.

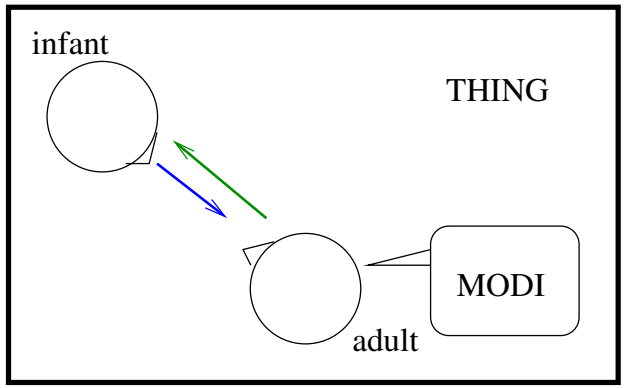
> *'Communication is where* **Agent1** *gets* **Agent2** *to attend to* **Thing**, *by performing a special type of action (an* **utterance**).'



Pragmatic understanding and word learning

Infants don't only need to learn what words mean.
They also need to learn *what communication is*.

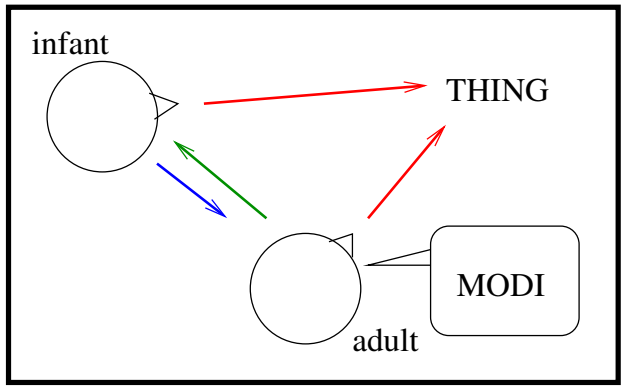
> *'Communication is where* **Agent1** *gets* **Agent2** *to attend to* **Thing**, *by performing a special type of action (an* **utterance**).'



Pragmatic understanding and word learning

Infants don't only need to learn what words mean.
They also need to learn *what communication is*.

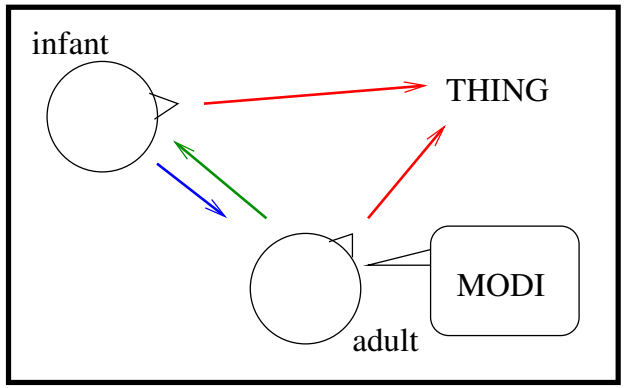
> *'Communication is where* **Agent1** *gets* **Agent2** *to attend to* **Thing**, *by performing a special type of action (an* **utterance**).'



Pragmatic understanding and word learning

Infants don't only need to learn what words mean.
They also need to learn *what communication is*.

> *'Communication is where* **Agent1** *gets* **Agent2** *to attend to* **Thing**, *by performing a special type of action (an* **utterance**).'

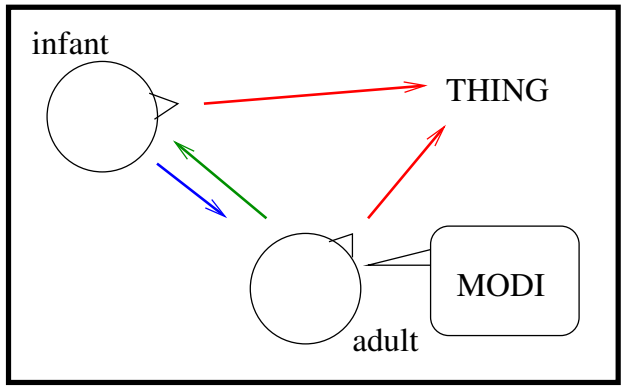


The infant has to learn that words have an attention-directing function.

Pragmatic understanding and word learning

Infants don't only need to learn what words mean.
They also need to learn *what communication is*.

> *'Communication is where* **Agent1** *gets* **Agent2** *to attend to* **Thing**, *by performing a special type of action (an* **utterance**).'

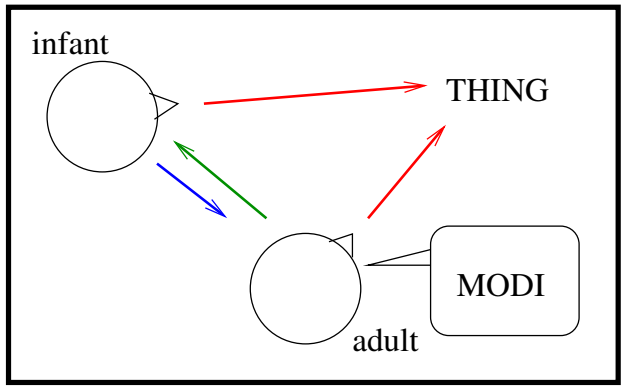


The infant has to learn that words have an attention-directing function.
Joint attention is important in learning this.

Pragmatic understanding and word learning

Infants don't only need to learn what words mean.
They also need to learn *what communication is*.

> *'Communication is where* **Agent1** *gets* **Agent2** *to attend to* **Thing**, *by performing a special type of action (an* **utterance**).'

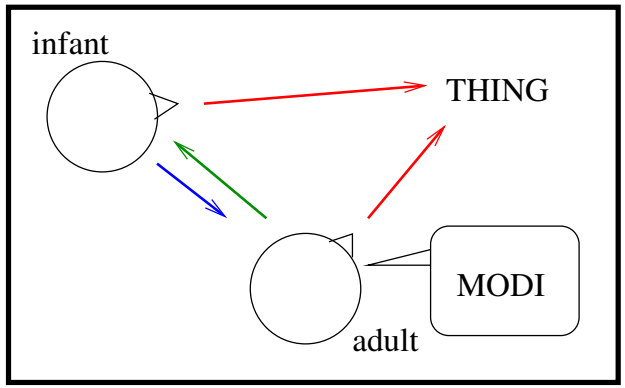


The infant has to learn that words have an attention-directing function.
Training: joint attention regularly co-occurs with observed utterances.

Pragmatic understanding and word learning

Infants don't only need to learn what words mean.
They also need to learn *what communication is*.

> *'Communication is where* **Agent1** *gets* **Agent2** *to attend to* **Thing**, *by performing a special type of action (an* **utterance**).'



The infant has to learn that words have an attention-directing function.
After training, utterances actively direct an infant's attention.

Developing the concept of communicative intentions

Infants begin following **pointing gestures** around 11–12 months (Butterworth, 2006; Carpenter *et al.*, 1998).

Behne *et al.* (2005): infants had the task of looking for a hidden object.

- The experimenter helped them by giving them nonverbal 'communicative gestures' (alternately looking at the infant and the hiding place, with and without pointing at the hiding place).
- These gestures began to be understood at 14 months.
- 'Non-communicative' looking / pointing was not helpful.

Apes and communicative intention recognition

We know that monkeys/great apes *can* recognise the intentions of other agents. And we know they *can* follow gaze.

Interestingly, great apes don't seem able to recognise *specifically communicative* intentions (Call *et al.*, 2004).

- E.g. They can't make use of 'communicative gestures' (looking or pointing) in a hide-and-seek game.

Tomasello: nonhuman primates can't represent communicative gestures. This is why they don't have language.

Questions about Tomasello's model of word learning

1. What's the *mechanism* via which communicative action recognition supports word meaning learning?

Questions about Tomasello's model of word learning

1. What's the *mechanism* via which communicative action recognition supports word meaning learning?

2. How do infants *represent* communicative actions?

   Mum **says** *[that P is the case].*

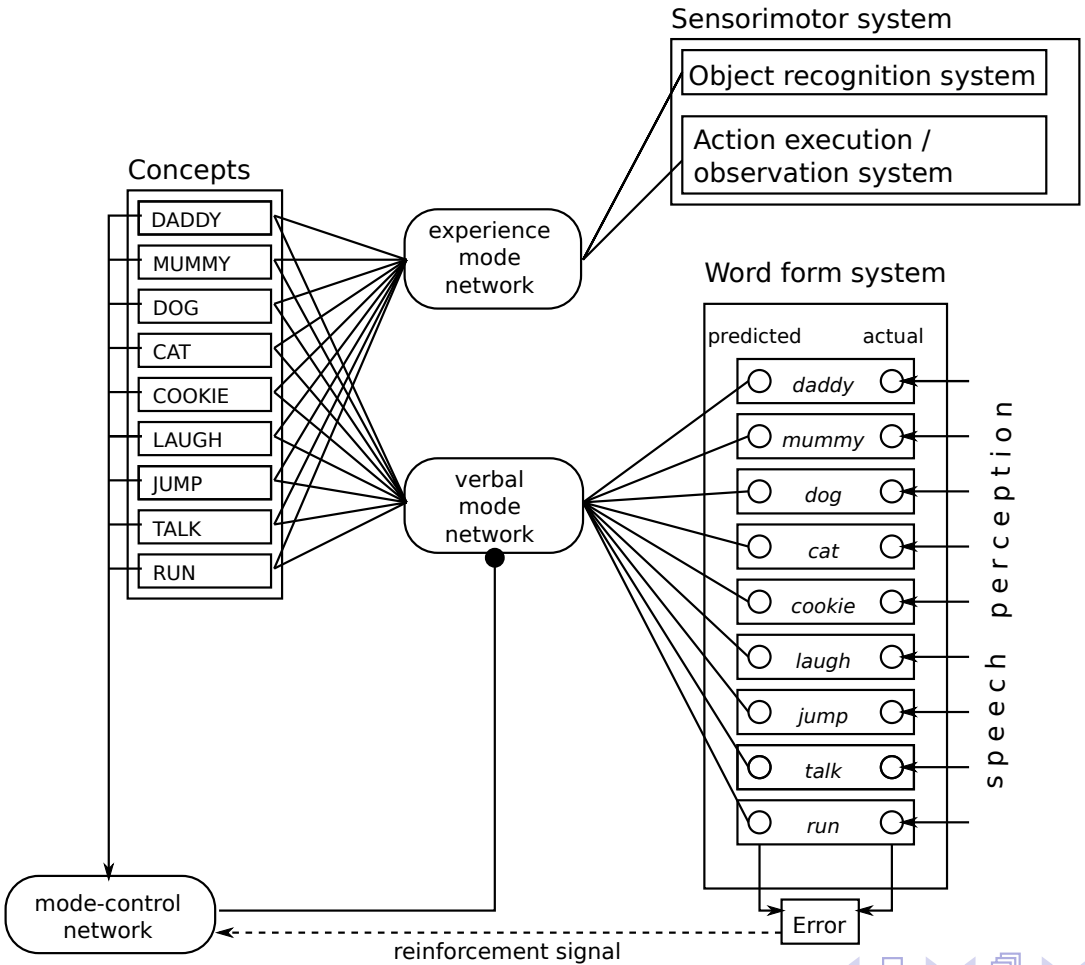
A neural network model of word learning

Caza & Knott, *Language and Cognitive Development* (2012)

A model of how the concept of communicative actions develops, and how this facilitates word meaning learning.

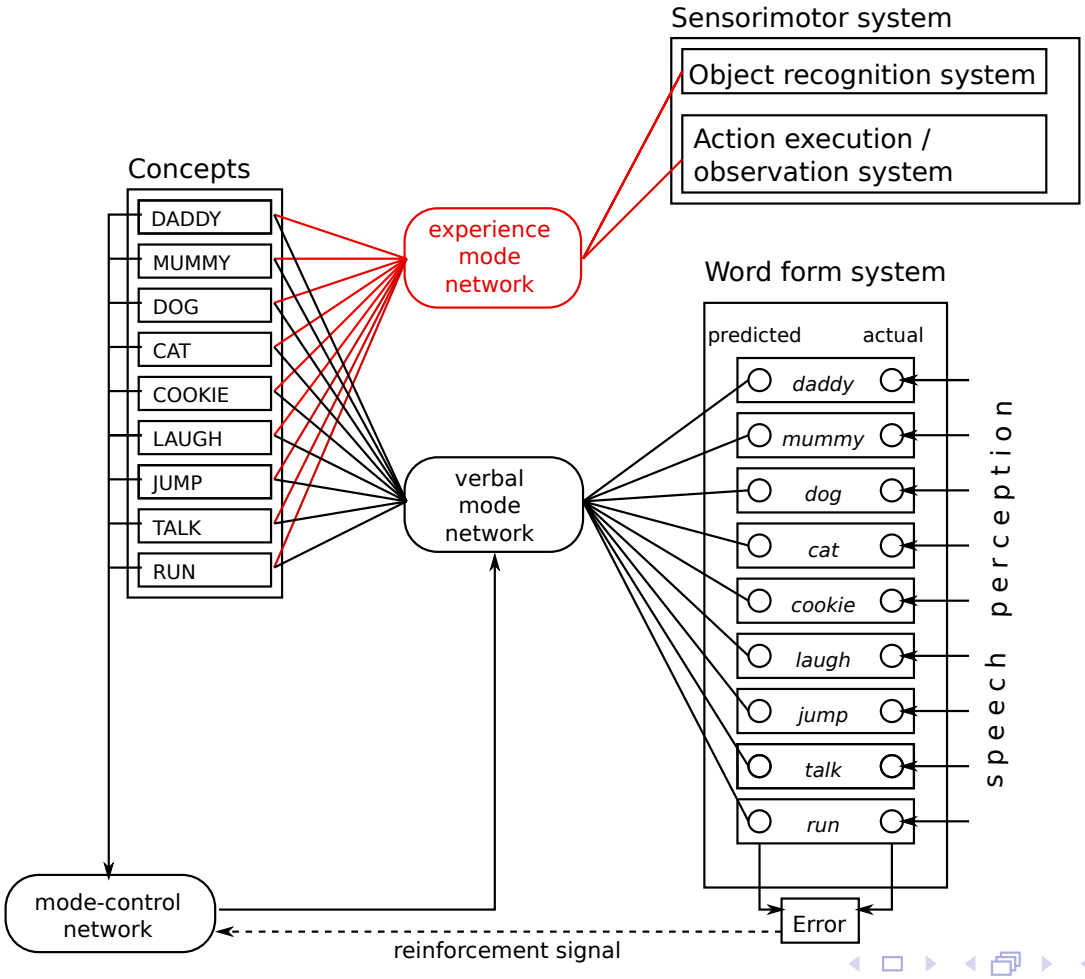
A neural network model of word learning

Caza & Knott, *Language and Cognitive Development* (2012)



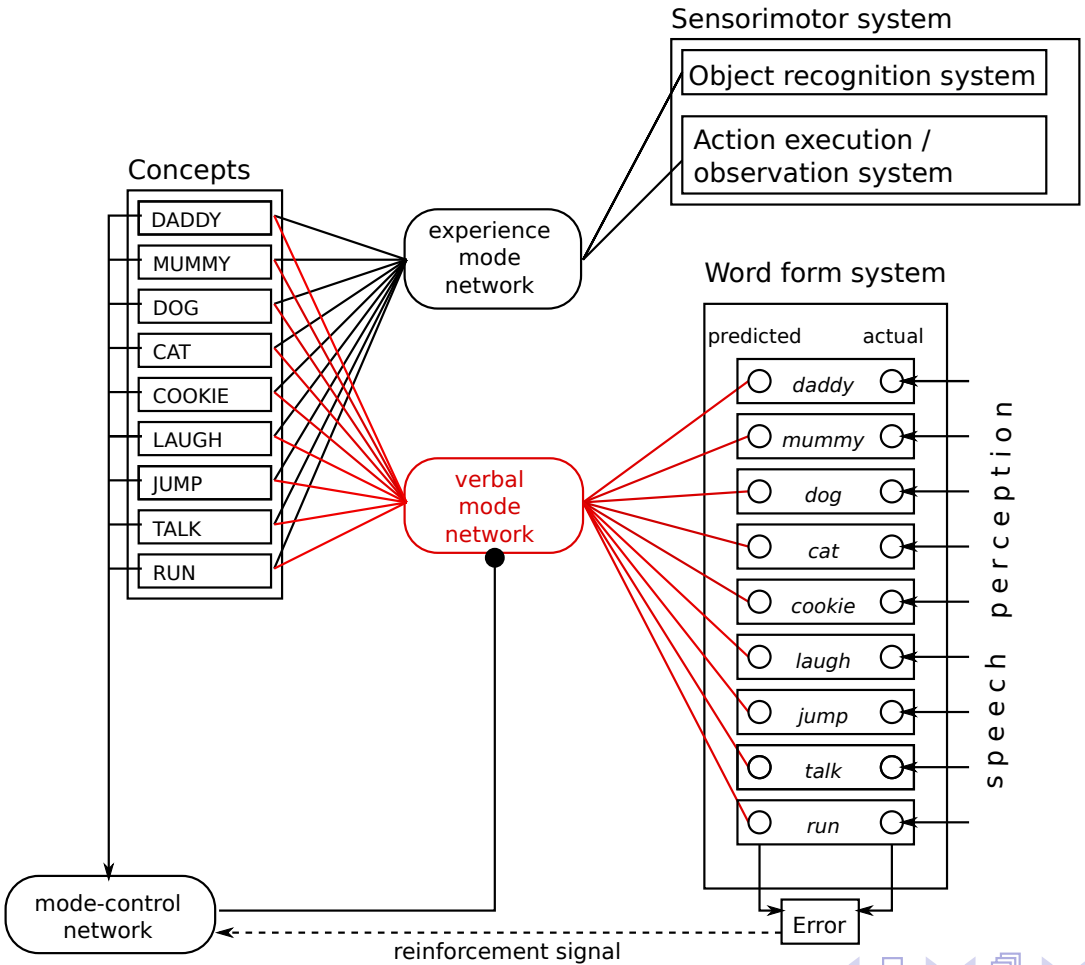
A neural network model of word learning

The **experience mode network** associates concepts with sensorimotor stimuli.



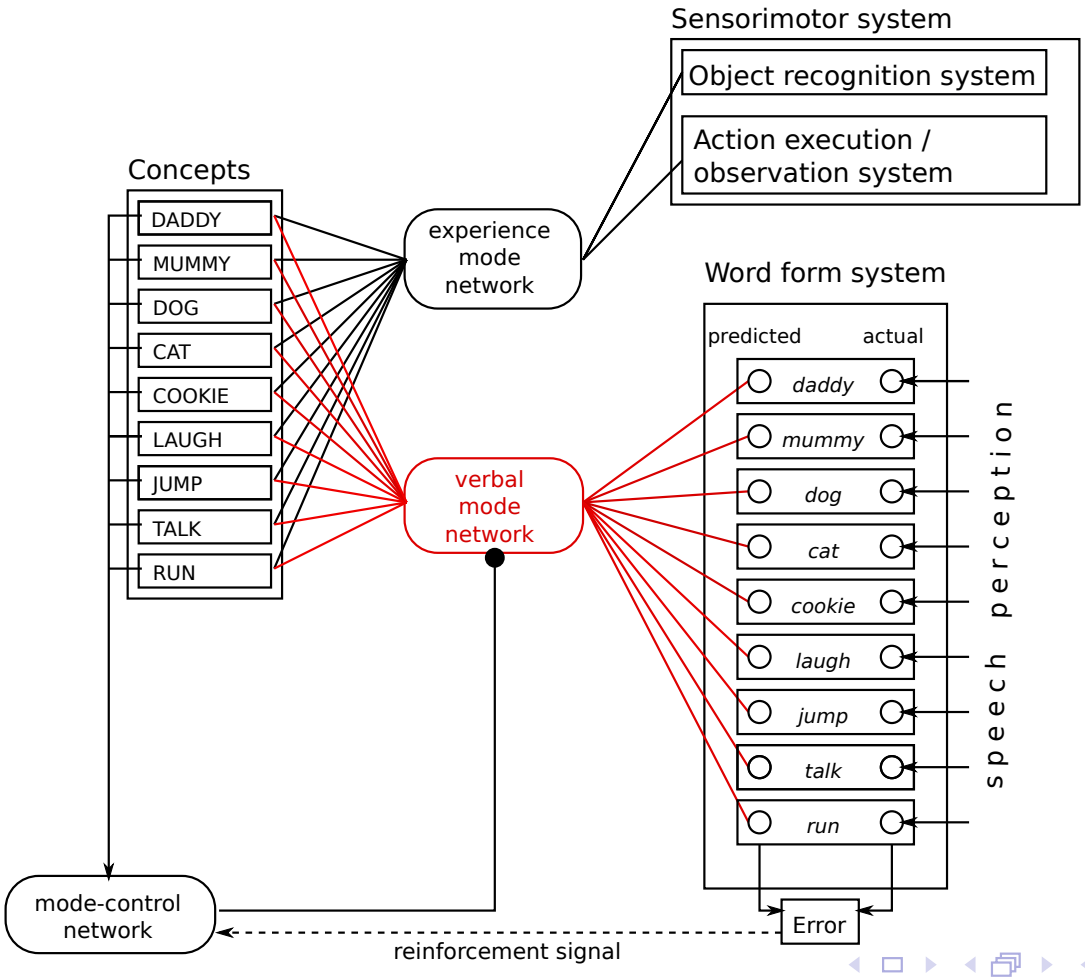
A neural network model of word learning

The **verbal mode network** associates concepts with word forms.



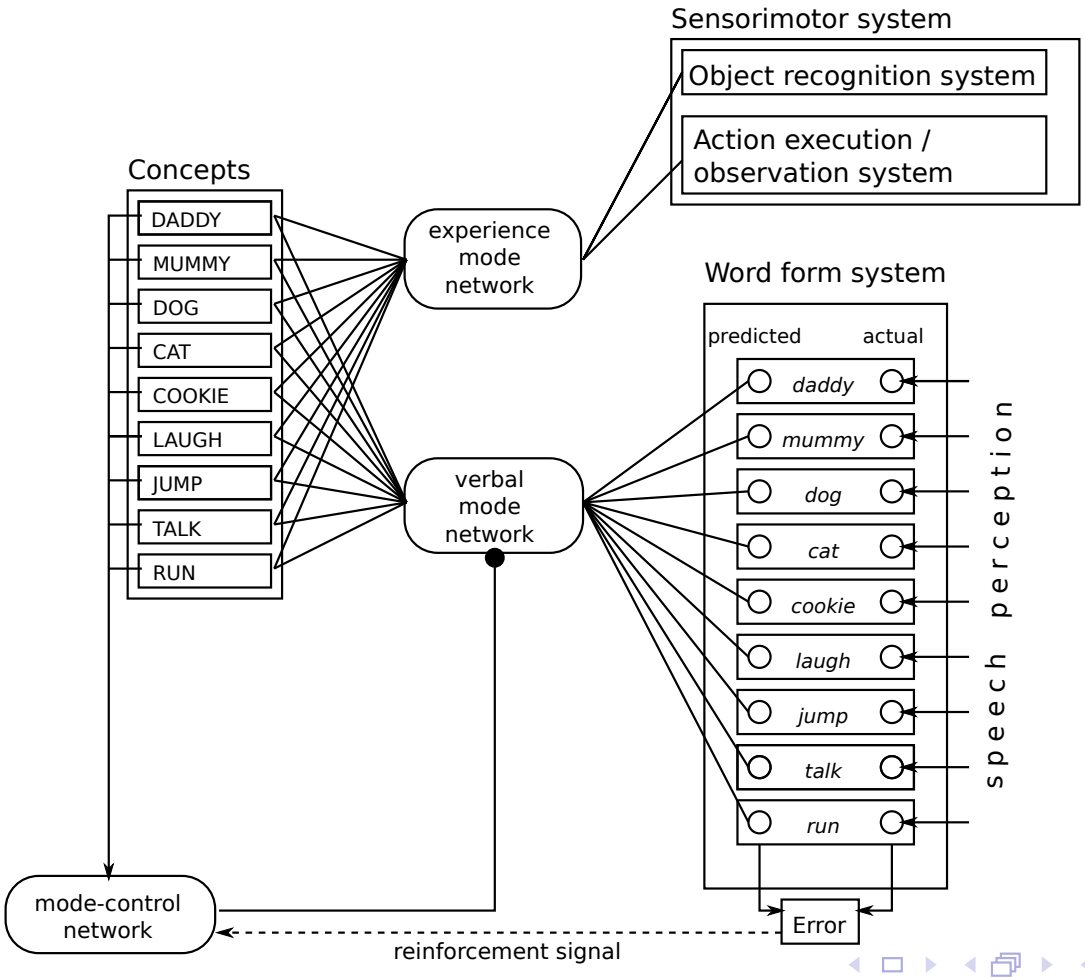
A neural network model of word learning

Verbal mode can be turned on and off.



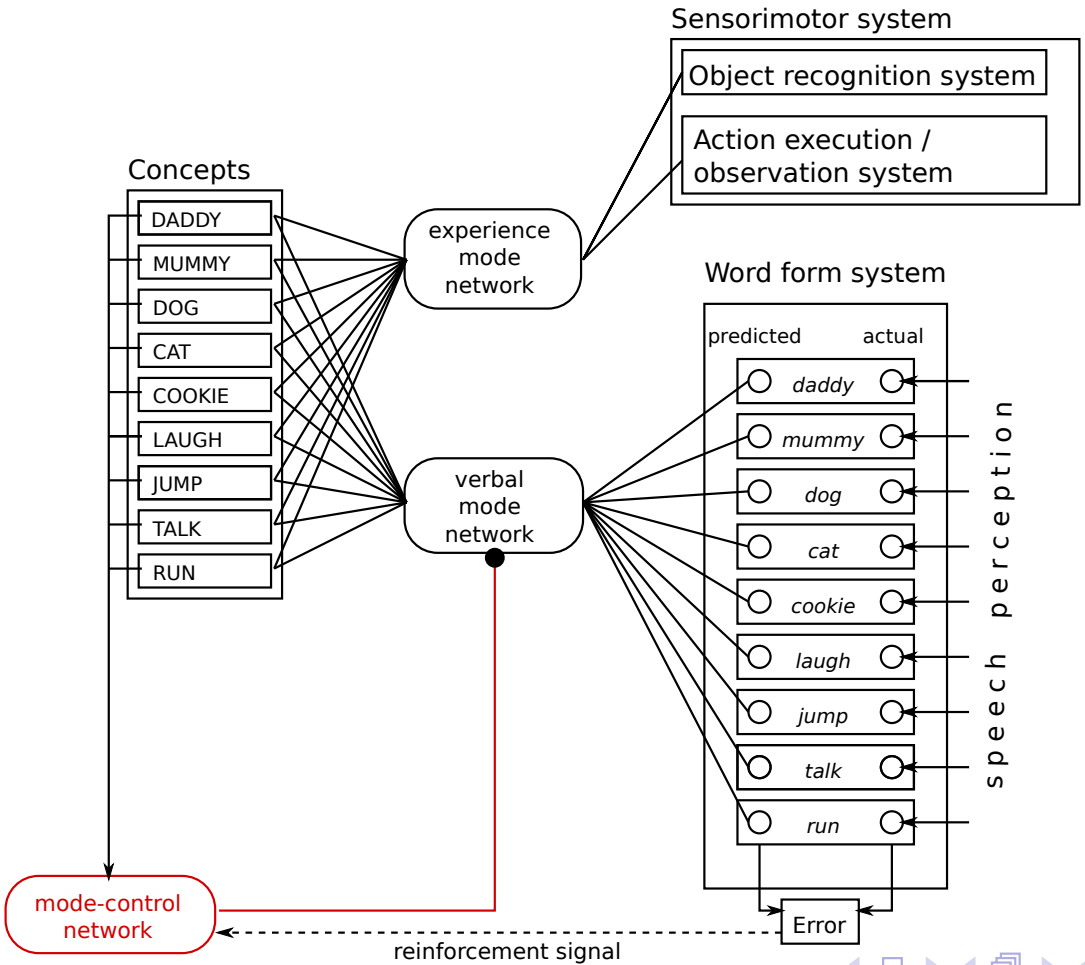
A neural network model of word learning

Verbal mode can be turned on and off.



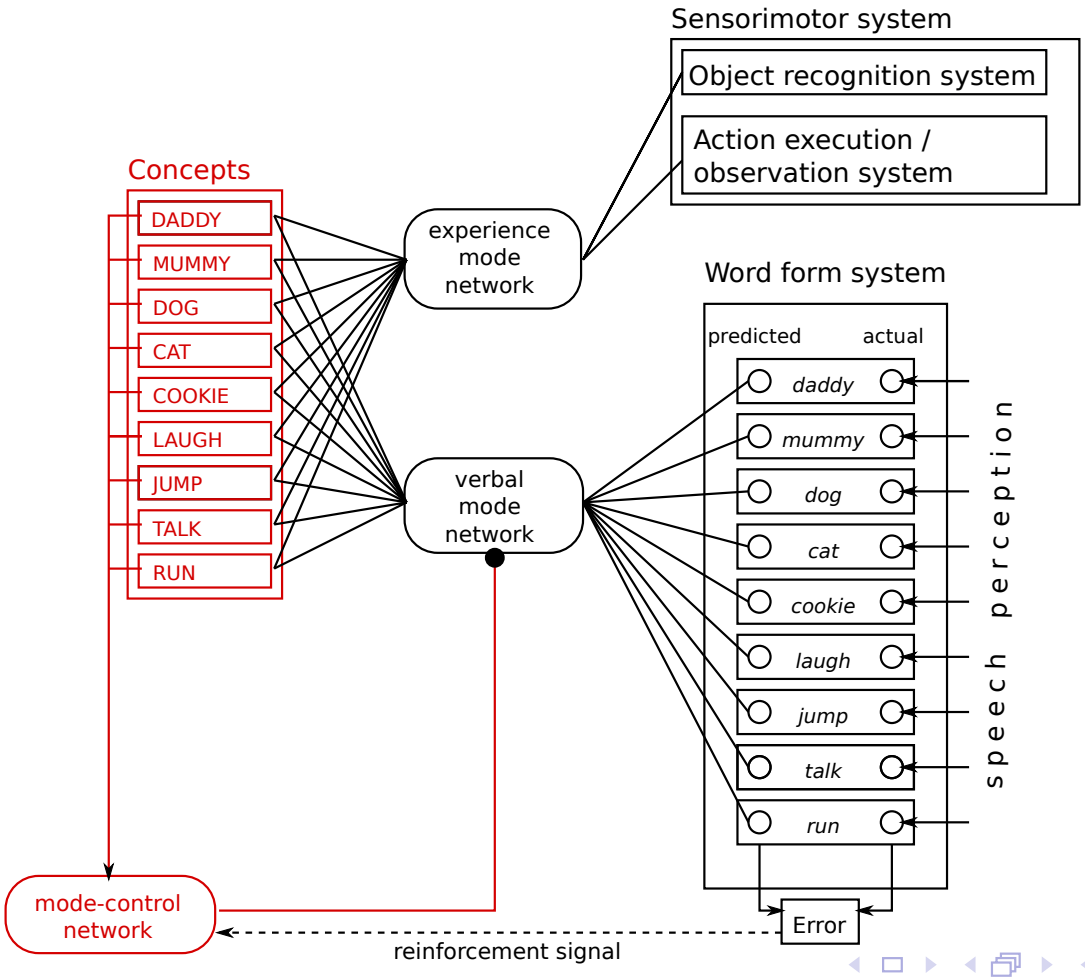
A neural network model of word learning

The **mode control network** learns when to turn verbal mode on.



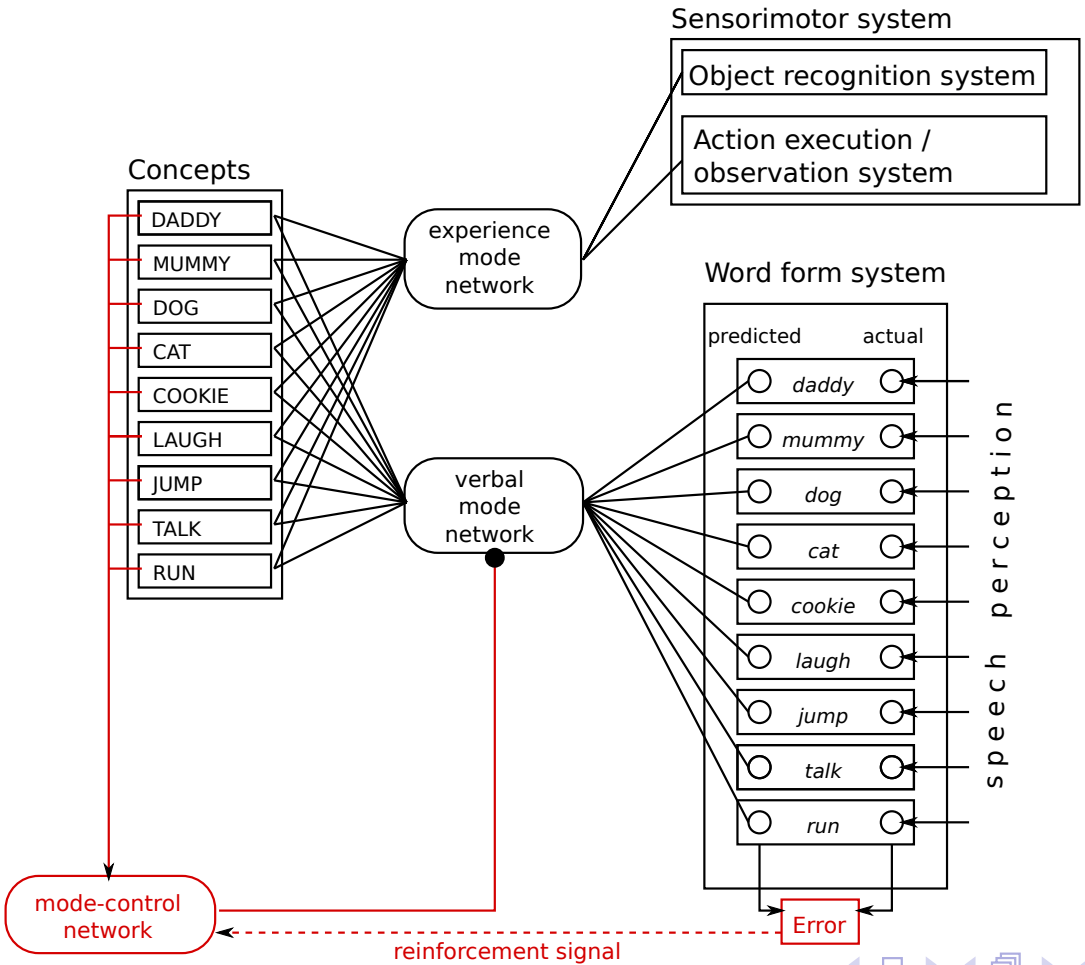
A neural network model of word learning

The mode control network takes concepts as input.



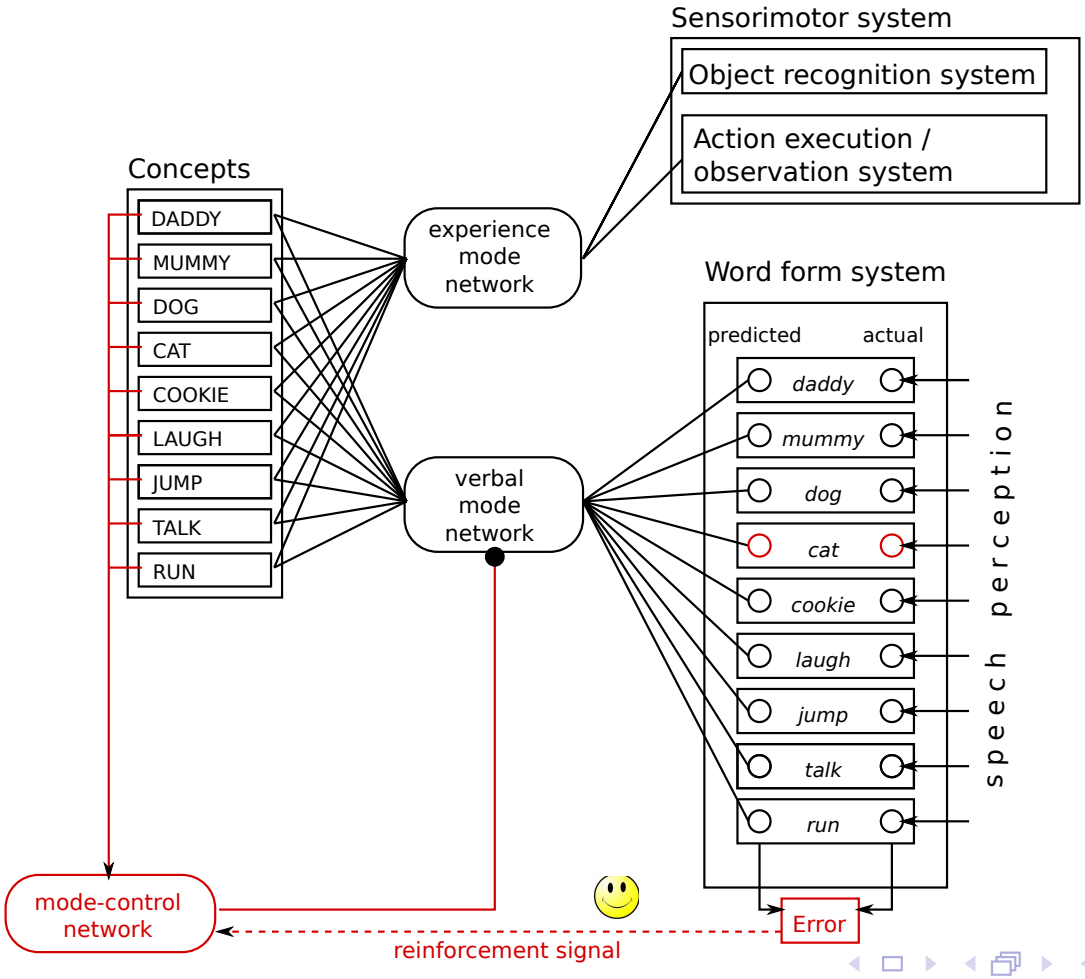
A neural network model of word learning

It learns when to turn verbal mode on through *reinforcement*.



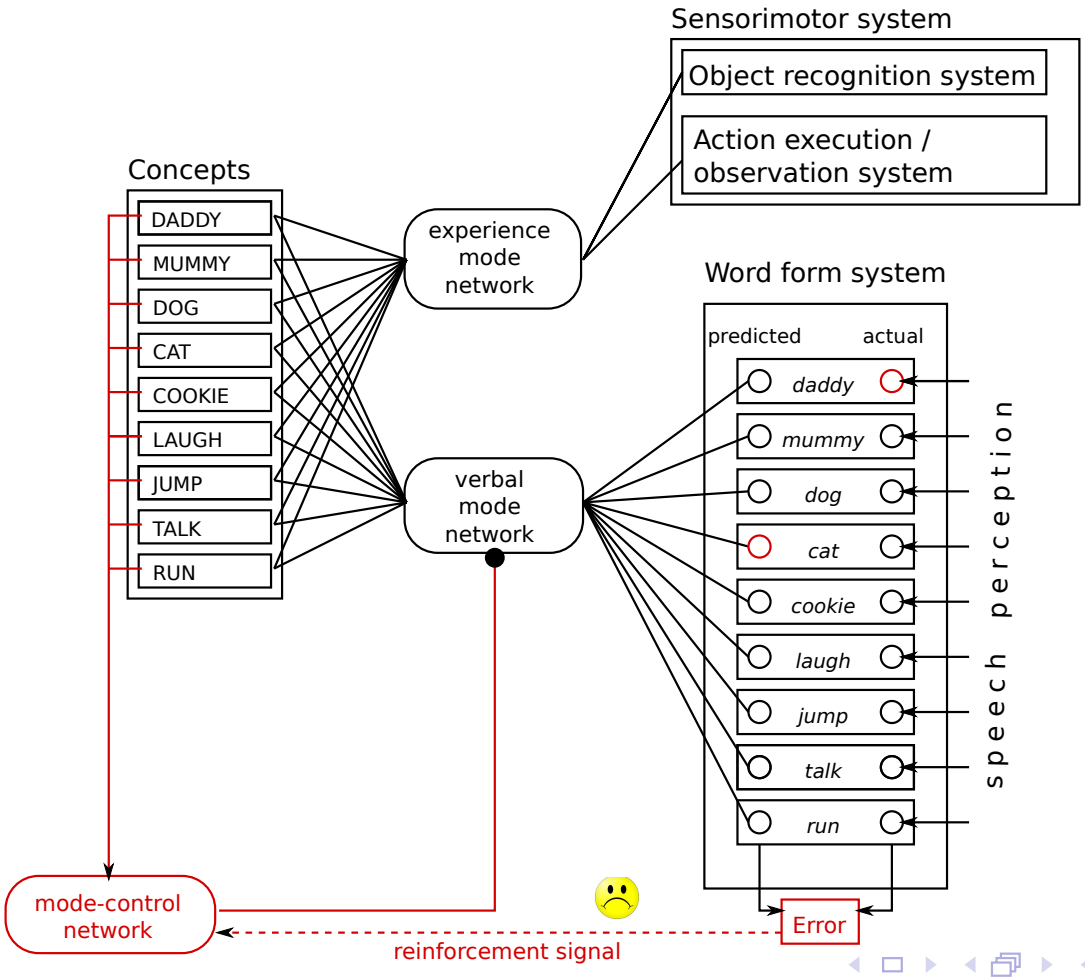
A neural network model of word learning

A reward is generated when the verbal mode network correctly predicts incoming word forms at the next time point.



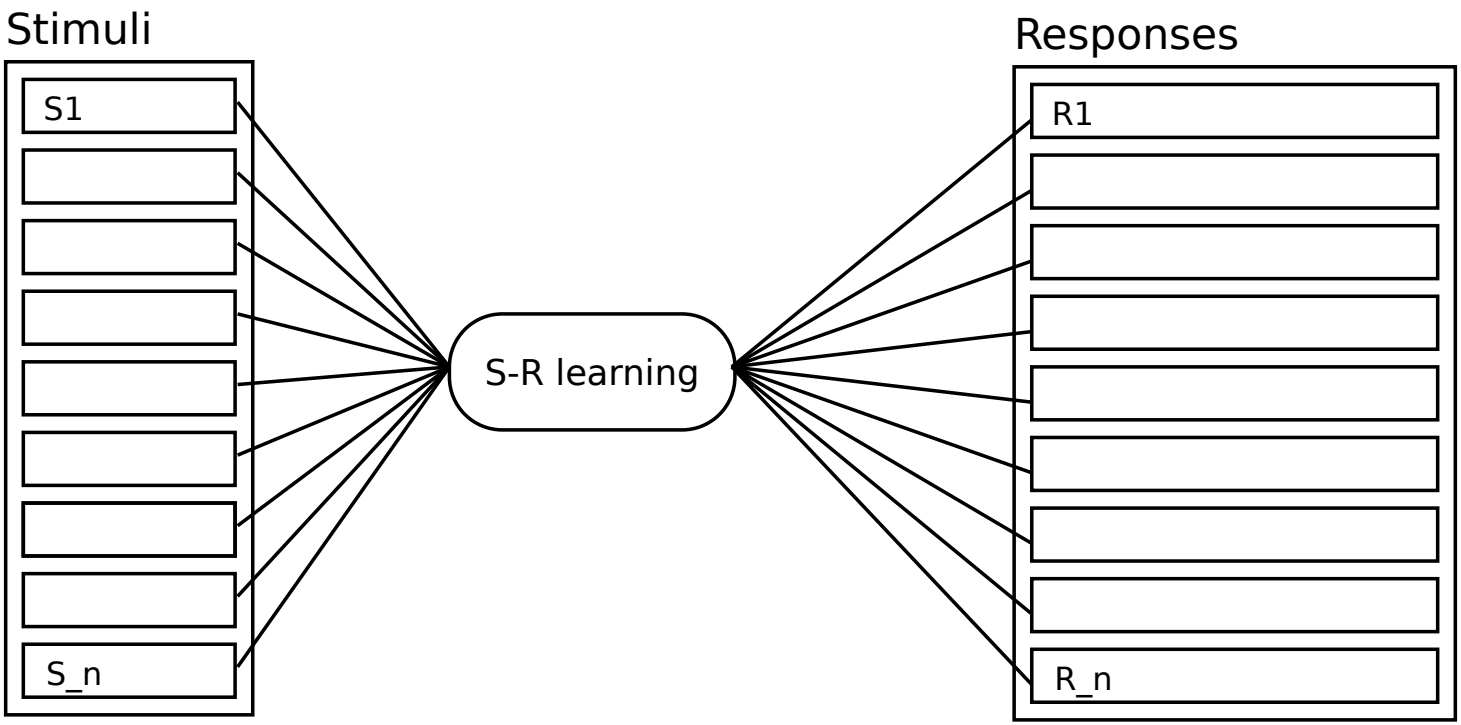
A neural network model of word learning

Otherwise a punishment is generated.



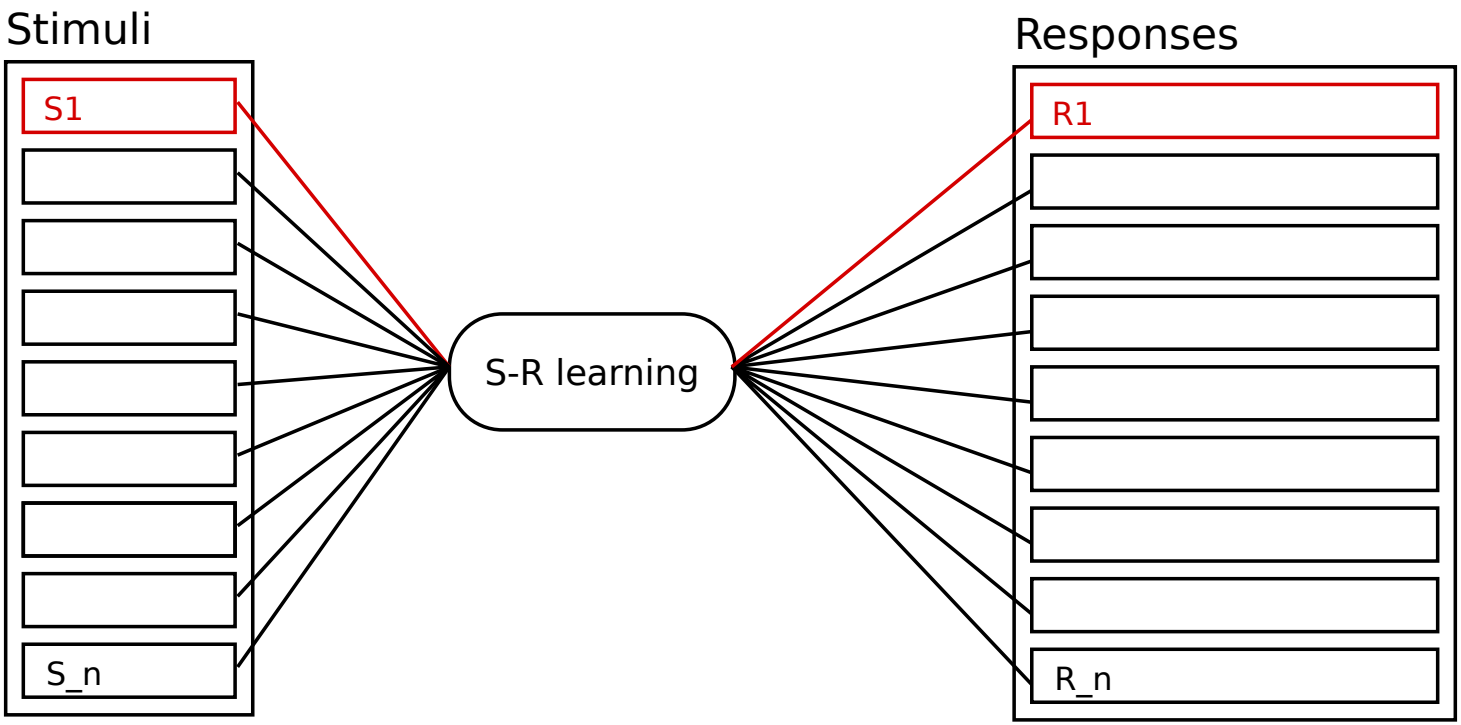
A neural network model of word learning

We think of the mode control network as part of the general network responsible for operant learning.



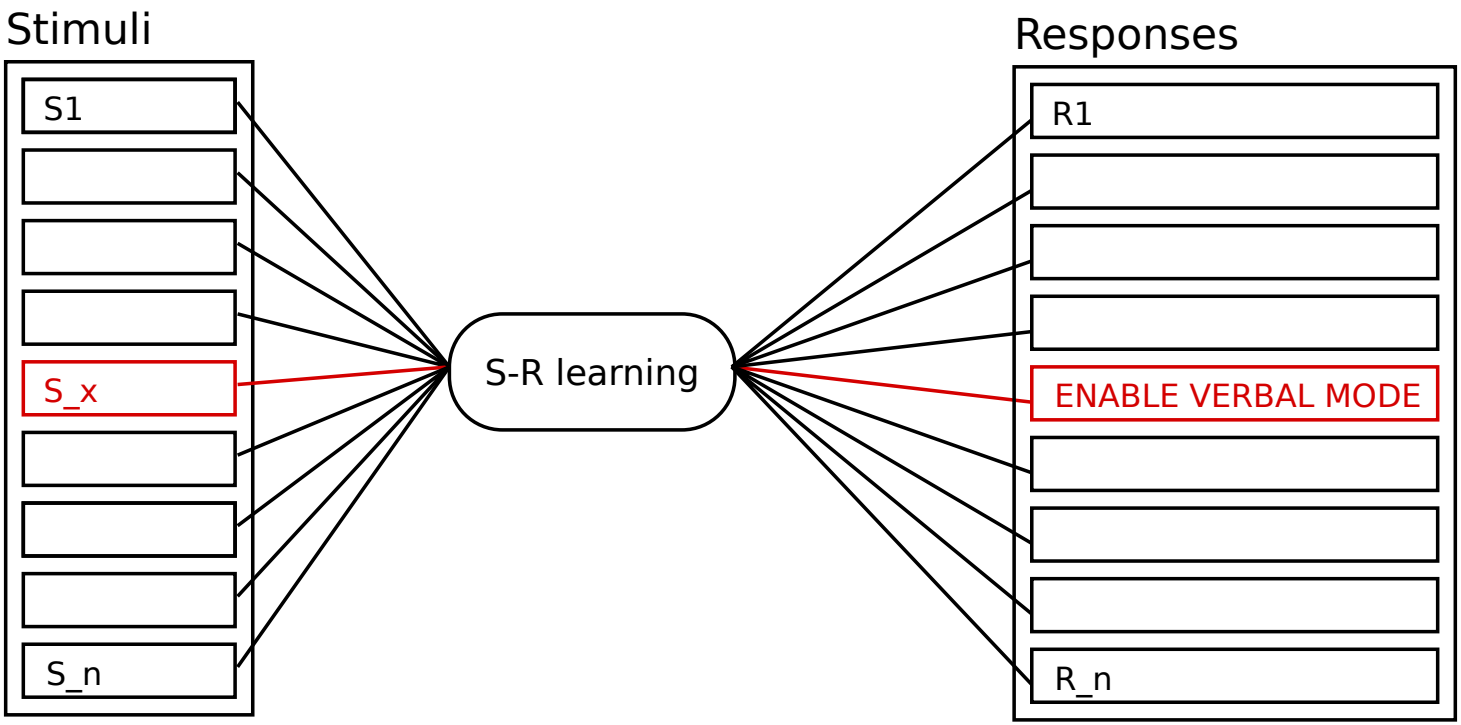
A neural network model of word learning

We think of the mode control network as part of the general network responsible for operant learning.



A neural network model of word learning

We think of the mode control network as part of the general network responsible for operant learning.



Training data

Time	t_1	t_2	t_3	t_4	t_5
Concepts	DADDY	MUMMY	DOG	MUMMY	DOG
	DRINK	PAT		TALK	CHASE
	BEER	DOG			KITTEN
Word forms	*minister...* *exports...*	*office...* *budget...*	*shopping...* *coffee...*	*dog... chasing... kitten...*	

Training data

Time	t_1	t_2	t_3	t_4	t_5
Concepts	**daddy**	MUMMY	DOG	MUMMY	DOG
	drink	PAT		TALK	CHASE
	beer	DOG			KITTEN
Word forms	*minister...* *exports...*	*office...* *budget...*	*shopping...* *coffee...*	*dog... chasing... kitten...*	

Most of the time, concepts and words are very decorrelated.

Training data

Time	t_1	t_2	t_3	t_4	t_5
Concepts	DADDY	**mummy**	**dog**	MUMMY	DOG
	DRINK	**pat**		TALK	CHASE
	BEER	**dog**			KITTEN
Word forms	*minister. . .* *exports. . .*	*office. . .* *budget. . .*	*shopping. . .* *coffee. . .*	*dog. . . chasing. . .* *kitten. . .*	

But if the infant has learned to establish joint attention with observed adults, there is some sequential structure to the stream of concepts.

Training data

Time	t_1	t_2	t_3	t_4	t_5
Concepts	DADDY	MUMMY	DOG	**mummy**	**dog**
	DRINK	PAT		**talk**	**chase**
	BEER	DOG			**kitten**
Word forms	*minister...* *exports...*	*office...* *budget...*	*shopping...* *coffee...*	**dog...chasing... kitten...**	

There are now some special moments where words and concepts become particularly well correlated.

Training data

Time	t_1	t_2	t_3	t_4	t_5
Concepts	DADDY	MUMMY	DOG	**mummy**	**dog**
	DRINK	PAT		**talk**	**chase**
	BEER	DOG			**kitten**
Word forms	*minister...* *exports...*	*office...* *budget...*	*shopping...* *coffee...*	**dog...chasing... kitten...**	

These are when the infant observes an agent *talking*.

- After noticing this, she establishes joint attention with the speaker.
- She is then (often) looking at the event the speaker is describing.

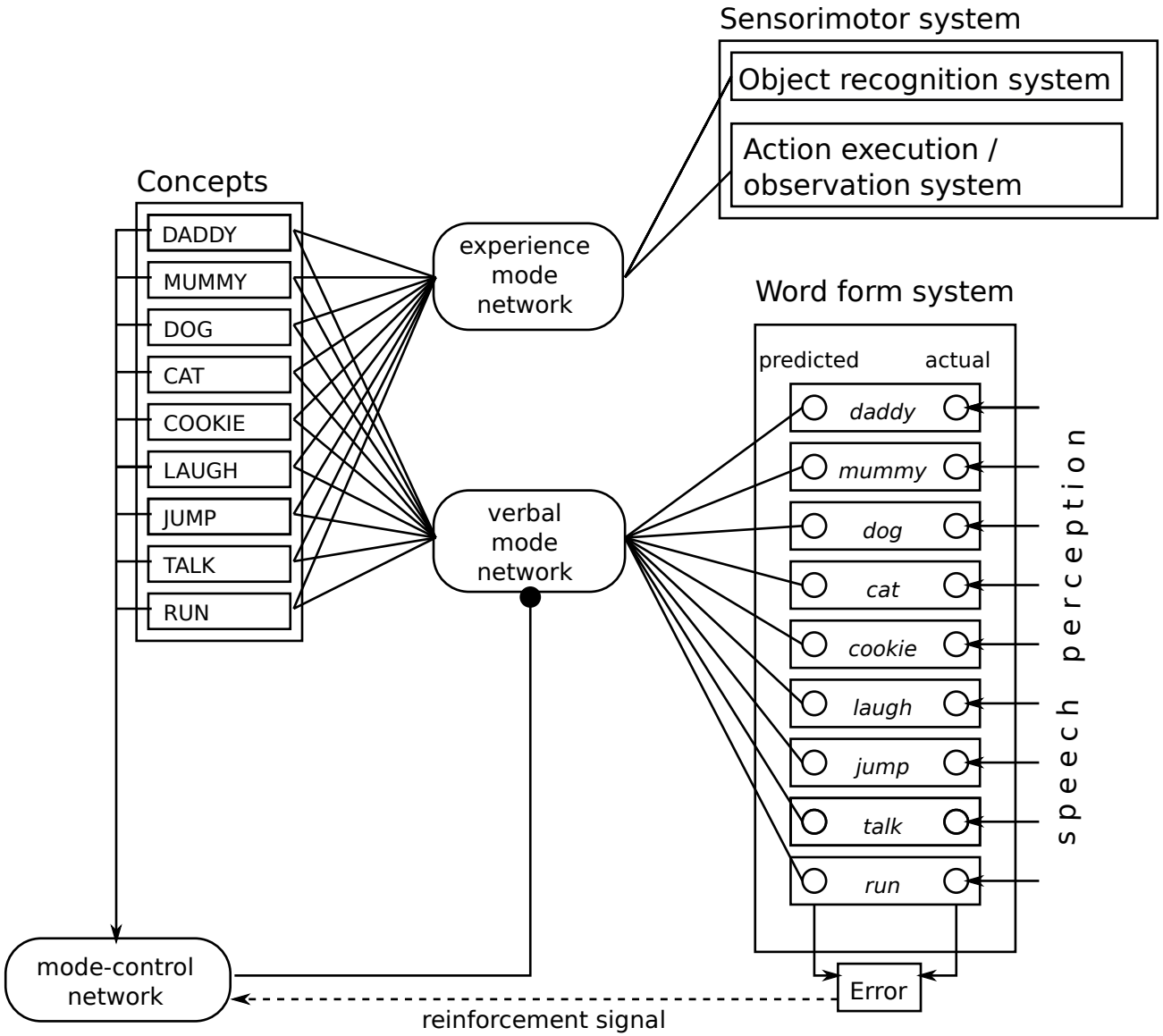
Training data

Time	t_1	t_2	t_3	t_4	t_5
Concepts	DADDY	MUMMY	DOG	**mummy**	**dog**
	DRINK	PAT		**talk**	**chase**
	BEER	DOG			**kitten**
Word forms	*minister...* *exports...*	*office...* *budget...*	*shopping...* *coffee...*	**dog...chasing... kitten...**	

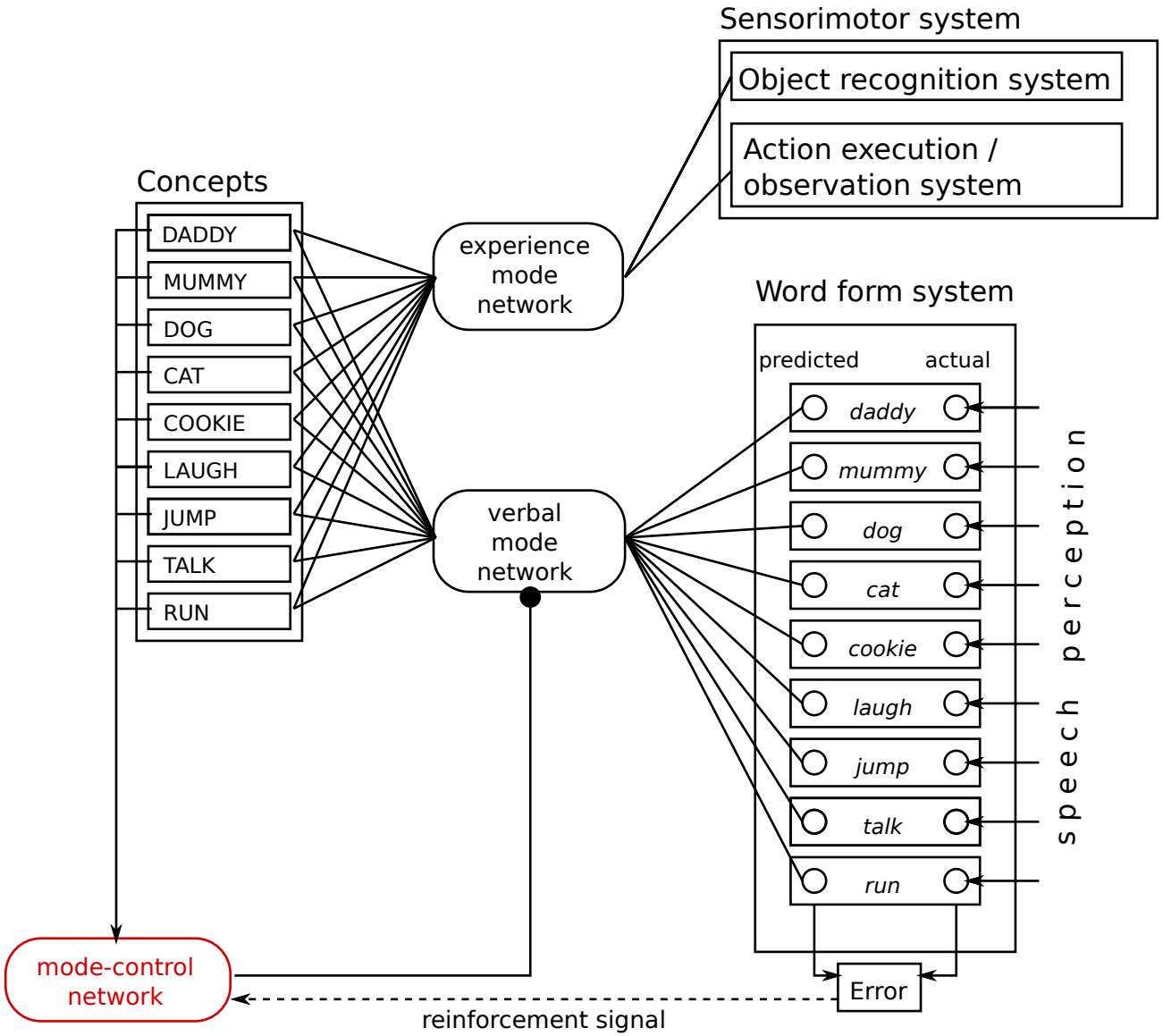
The infant can use TALK actions as a **cue** to engage verbal mode.

- After a TALK action, the verbal mode network's predictions about words are more likely to be correct.
- Through reinforcement learning, the infant learns to engage verbal mode after noticing a TALK action.

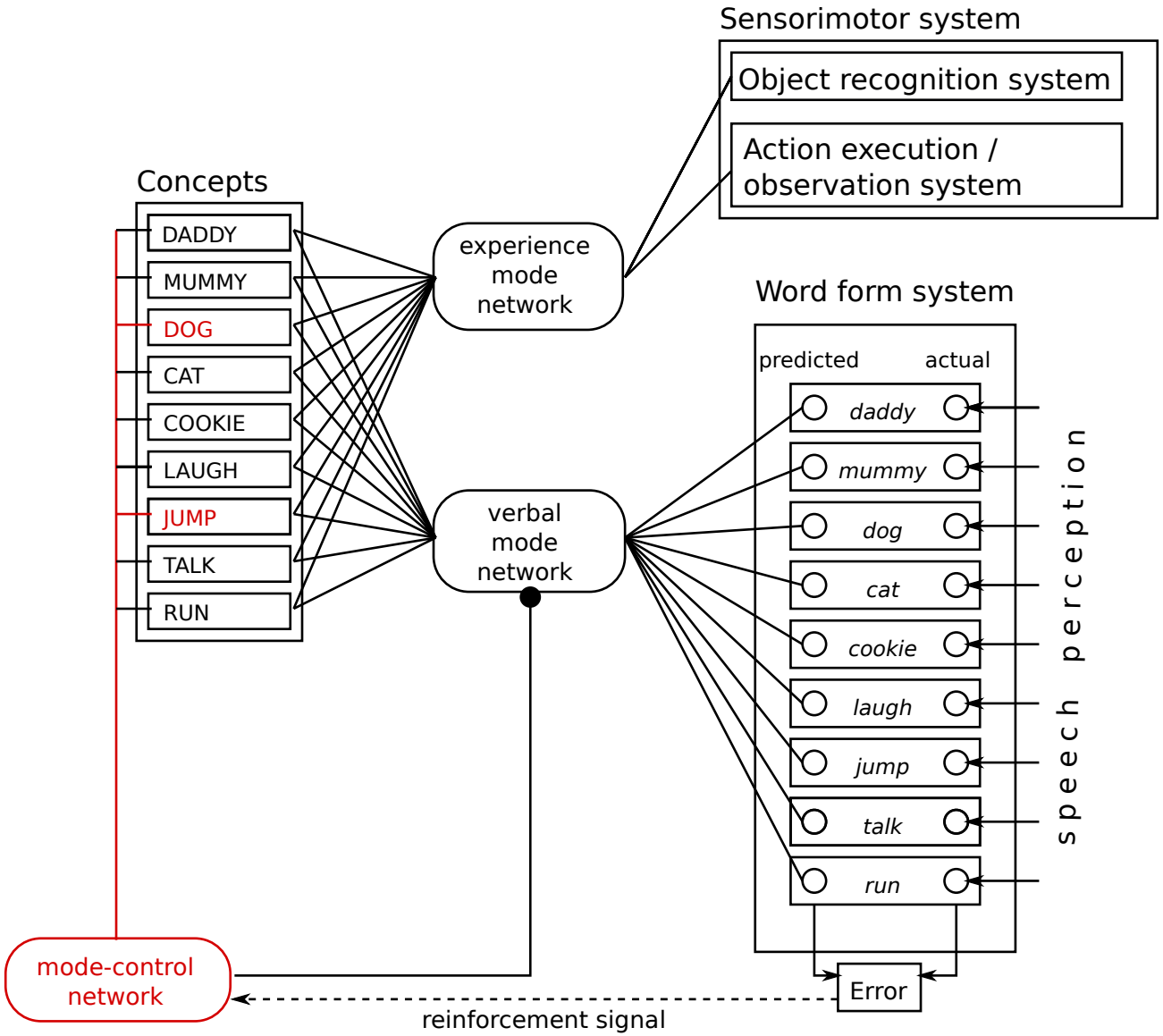
Talk actions as cues to engage verbal mode



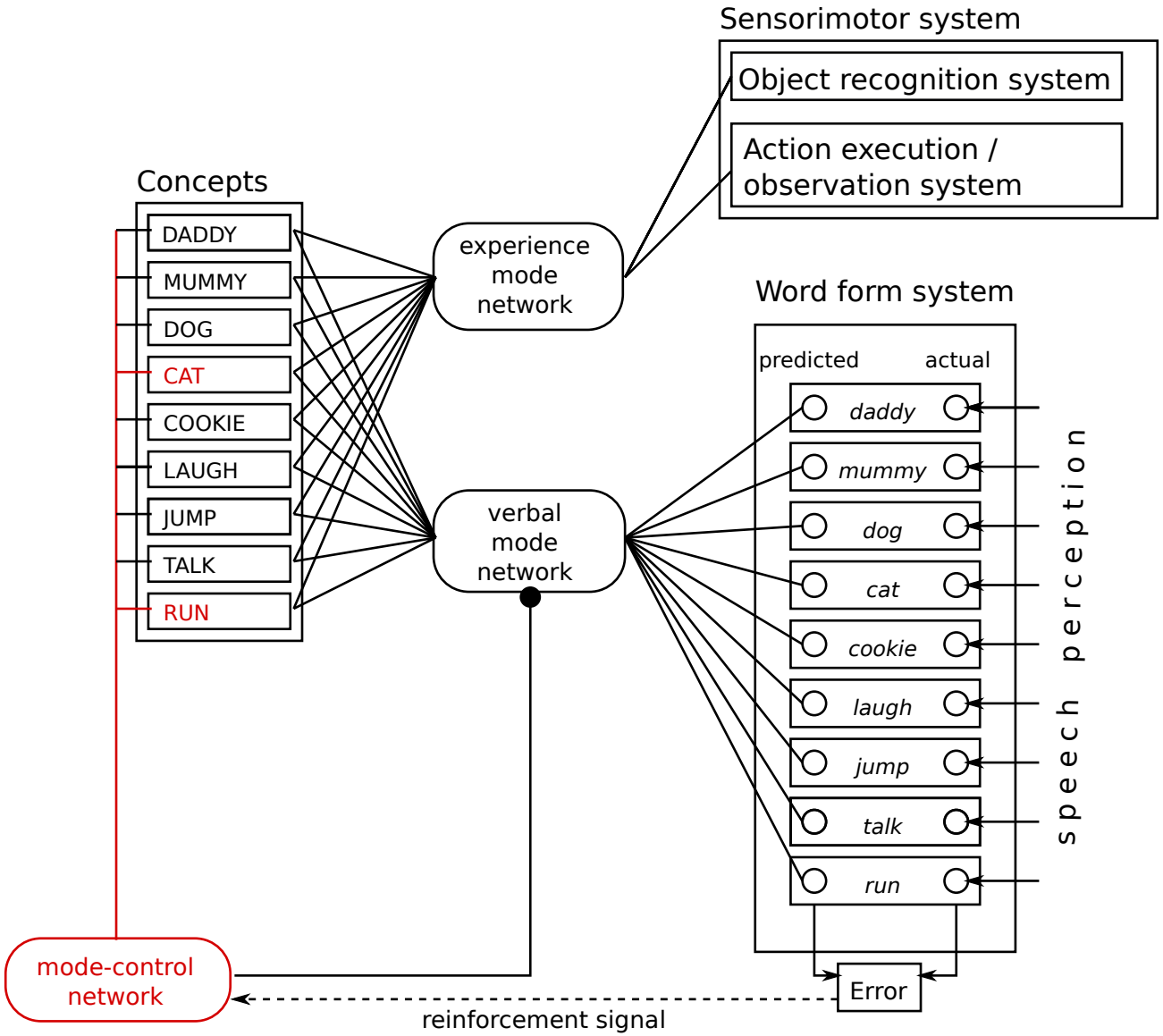
Talk actions as cues to engage verbal mode



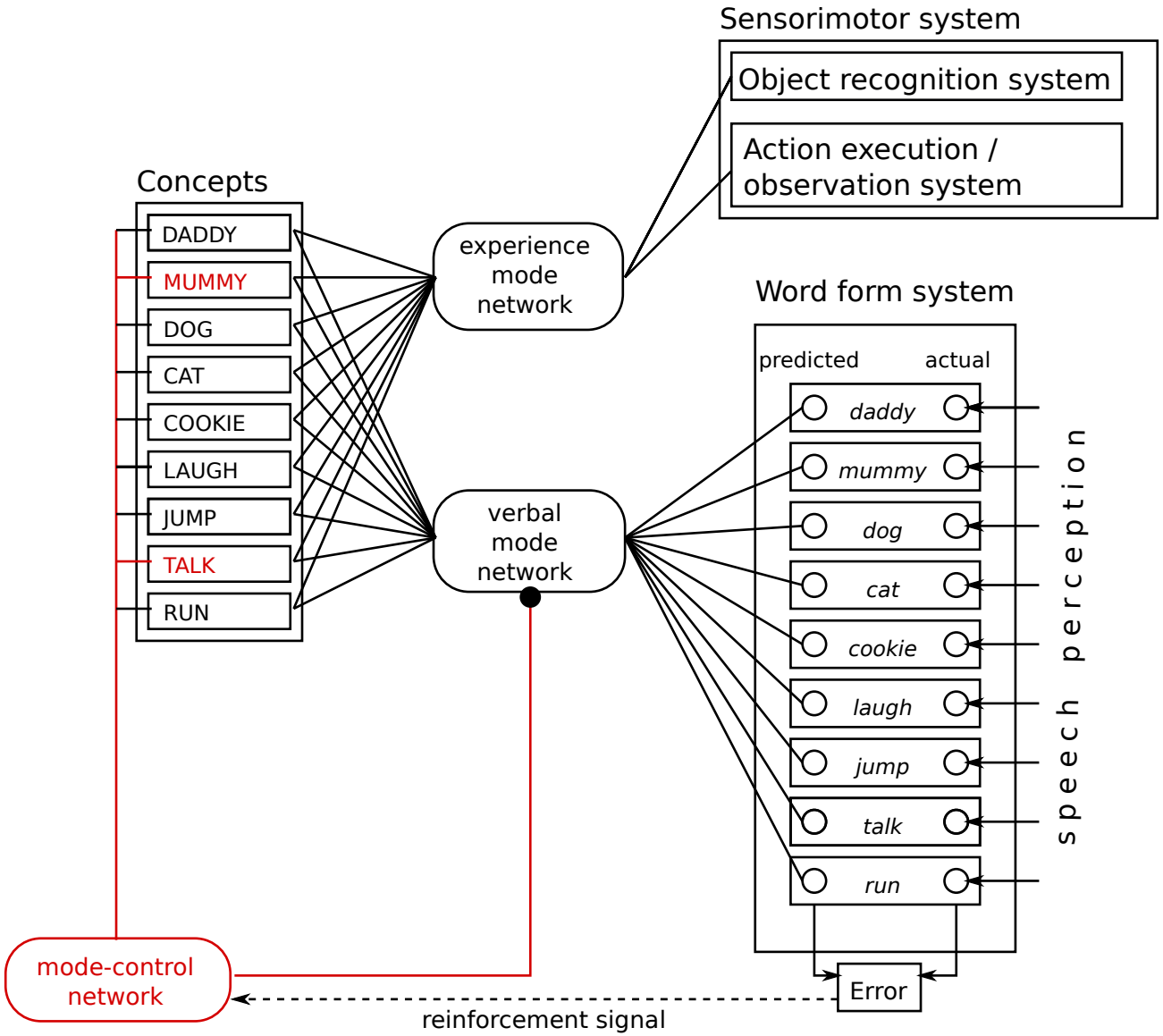
Talk actions as cues to engage verbal mode



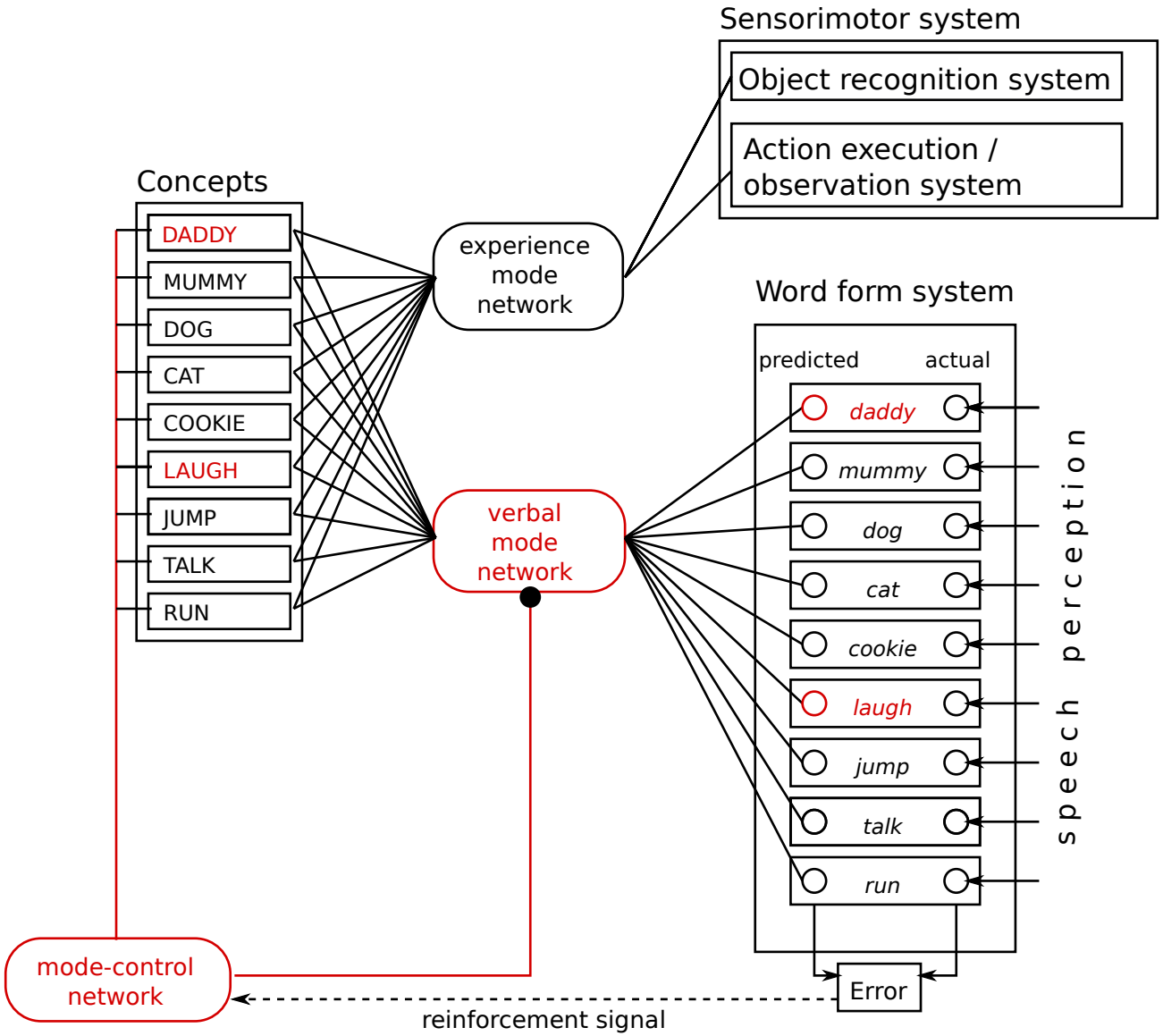
Talk actions as cues to engage verbal mode



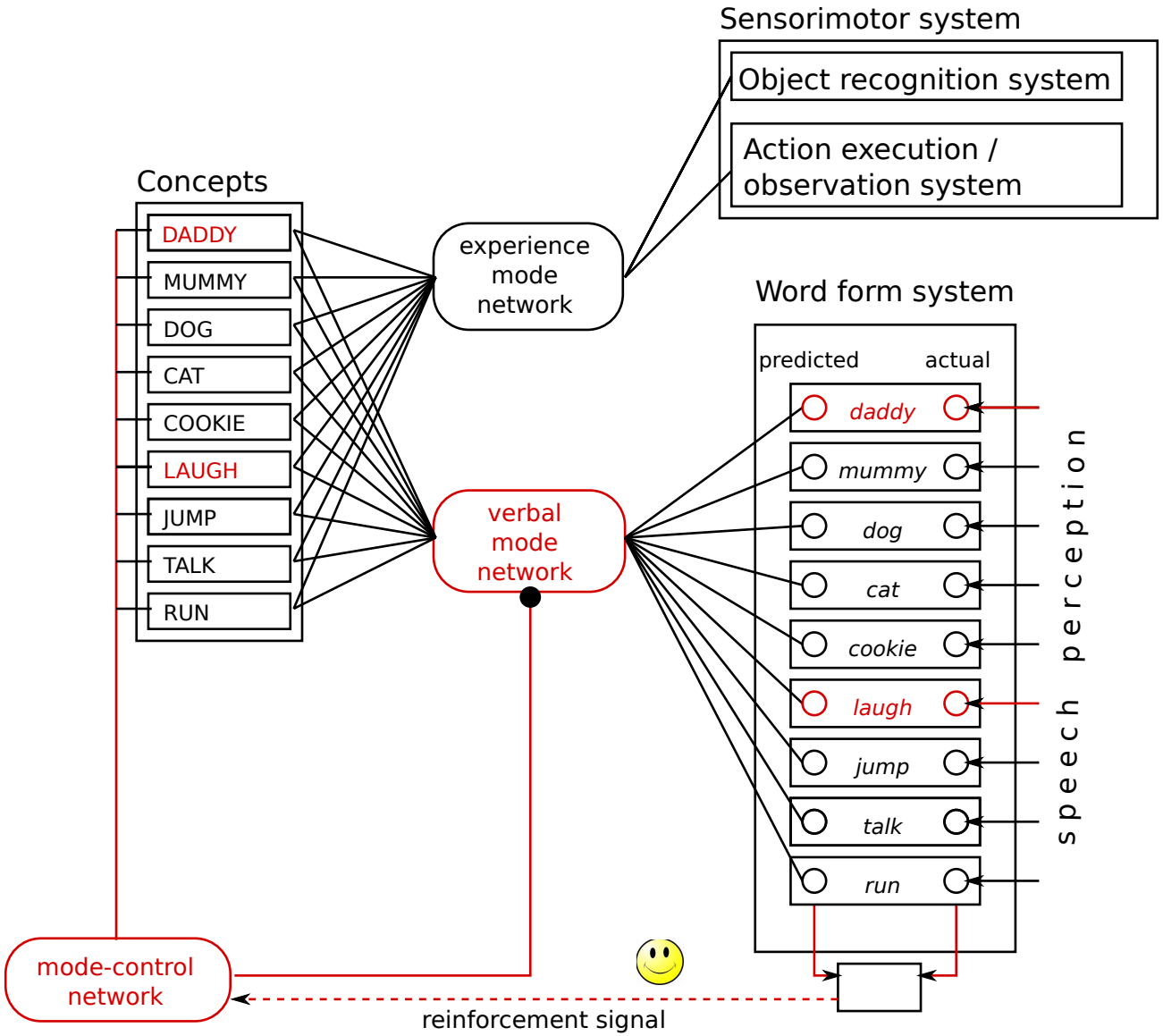
Talk actions as cues to engage verbal mode



Talk actions as cues to engage verbal mode

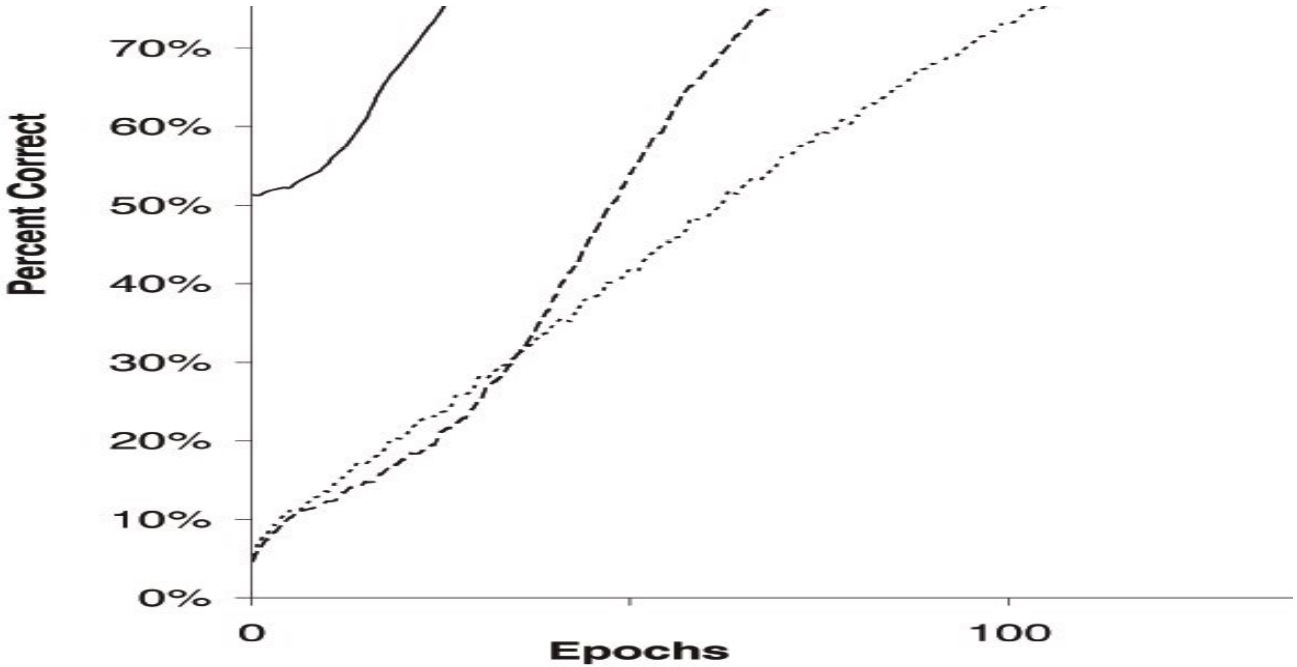


Talk actions as cues to engage verbal mode



Pragmatic bootstrapping

The verbal mode network's learning of word meanings happens *in parallel* with the mode control network's learning about TALK actions.



Two questions for Tomasello revisited

Q1. How do infants learn that linguistic actions are special?

Two questions for Tomasello revisited

Q1. How do infants learn that linguistic actions are special?
For Tomasello: how does learning this enable word learning?

Two questions for Tomasello revisited

Q1. How do infants learn that linguistic actions are special?
For Tomasello: how does learning this enable word learning?

A1. Through learning in the network just described.

Two questions for Tomasello revisited

Q1. How do infants learn that linguistic actions are special?
For Tomasello: how does learning this enable word learning?

A1. Through learning in the network just described.

2. How do infants *represent* communicative actions?

Mum **says** *[that P is the case].*

Propositional attitudes

The tricky thing about communicative actions is that they have **propositional content**.

John believes [that Mary loves Bill].
John desires [that Mary loves Bill].
John fears [that Mary loves Bill].
John **says** *[that Mary loves Bill].*

Two questions about propositional attitudes

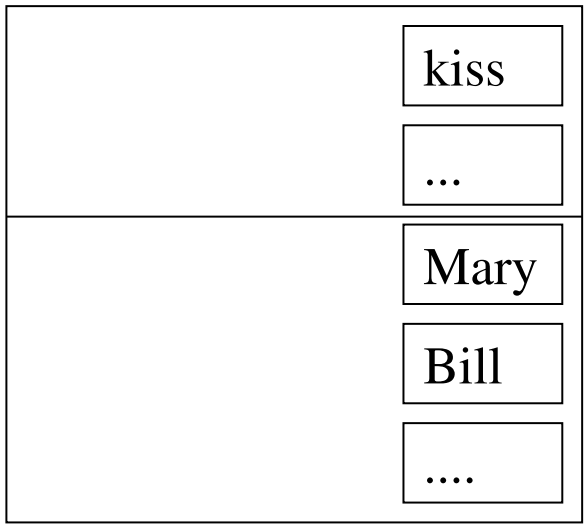
- How are propositional attitudes represented in the brain?
- How do infants learn to form these representations?

Difficulties in representing propositional attitudes

1. No-one knows how the brain encodes facts.

This even goes for ‘simple’ facts like *Mary kissed Bill*.

The simplest idea is that a fact is just a collection of active concepts.

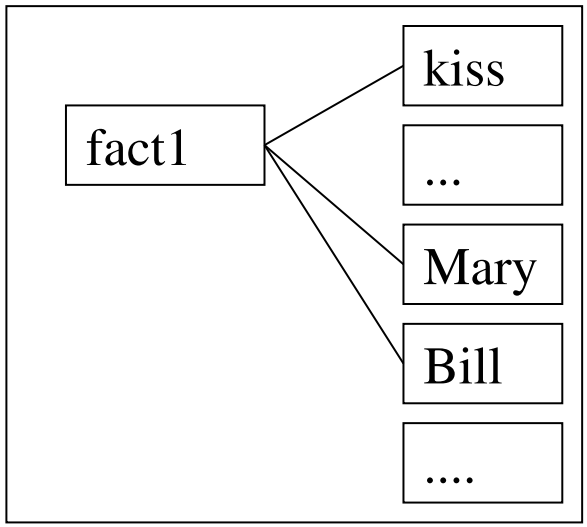


Difficulties in representing propositional attitudes

1. No-one knows how the brain encodes facts.

This even goes for ‘simple’ facts like *Mary kissed Bill*.

The simplest idea is that a fact is just a collection of active concepts.

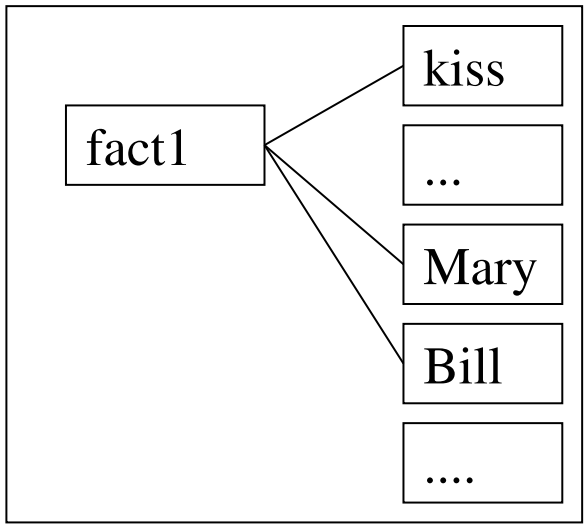


Difficulties in representing propositional attitudes

1. No-one knows how the brain encodes facts.

This even goes for ‘simple’ facts like *Mary kissed Bill*.

The simplest idea is that a fact is just a collection of active concepts.



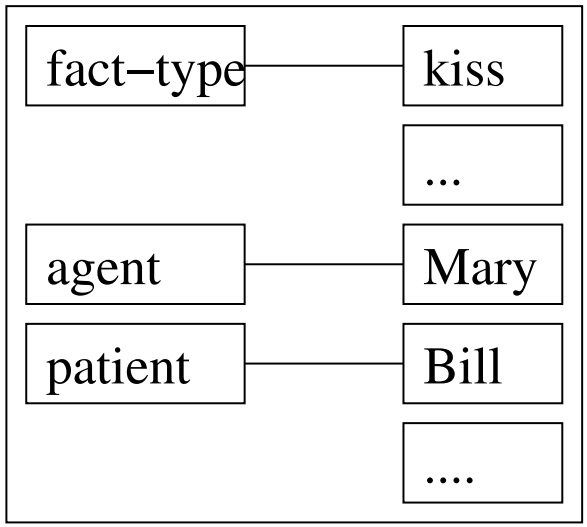
But the two individuals in *Mary kissed Bill* play different roles: there is the agent and the patient.

Difficulties in representing propositional attitudes

1. No-one knows how the brain encodes facts.

This even goes for ‘simple’ facts like *Mary kissed Bill*.

The simplest idea is that a fact is just a collection of active concepts.



Here’s a more elaborate representation, where the fact-denoting medium encodes *one fact at a time*.

Difficulties in representing propositional attitudes

2. Facts about propositional attitudes have recursive structure.

   Mary kissed Bill.
   John believes [Mary kissed Bill].
   Bill fears [John believes [Mary kissed Bill]].

Difficulties in representing propositional attitudes

2. Facts about propositional attitudes have recursive structure.

   Mary kissed Bill.
   John believes [Mary kissed Bill].
   Bill fears [John believes [Mary kissed Bill]].

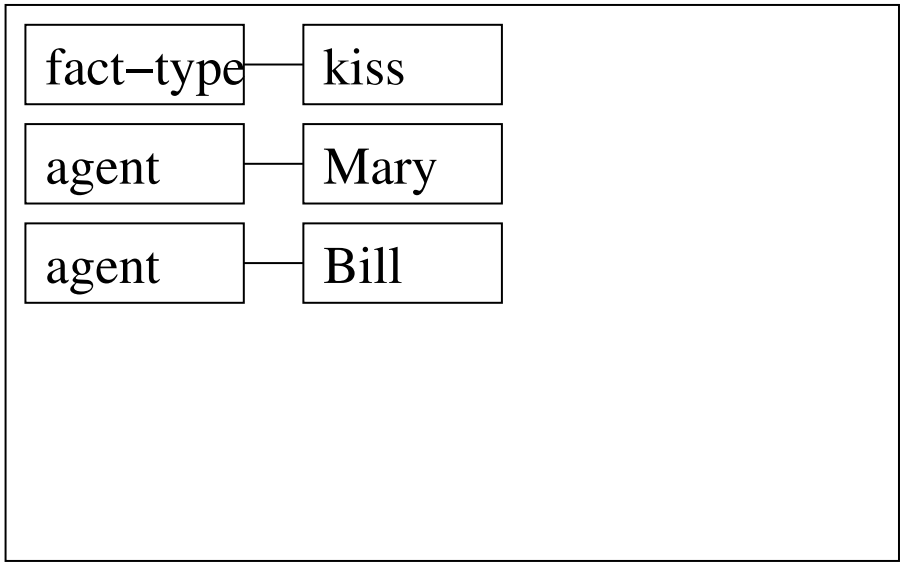
It is particularly hard to find a way of encoding nested propositions in a neural network.

Difficulties in representing propositional attitudes

2. Facts about propositional attitudes have recursive structure.

   Mary kissed Bill.
   John believes [Mary kissed Bill].
   Bill fears [John believes [Mary kissed Bill]].

It is particularly hard to find a way of encoding nested propositions in a neural network.

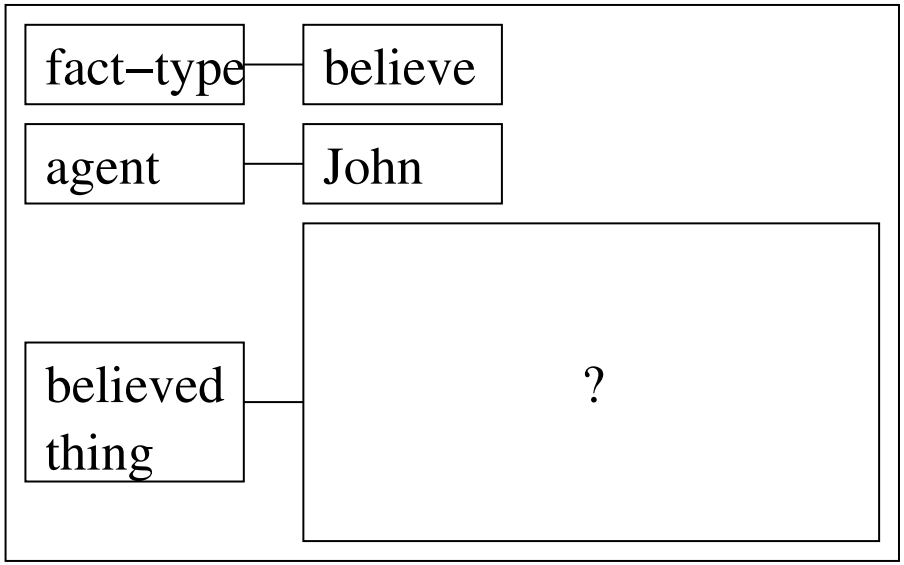


Difficulties in representing propositional attitudes

2. Facts about propositional attitudes have recursive structure.

   Mary kissed Bill.
   John believes [Mary kissed Bill].
   Bill fears [John believes [Mary kissed Bill]].

It is particularly hard to find a way of encoding nested propositions in a neural network.

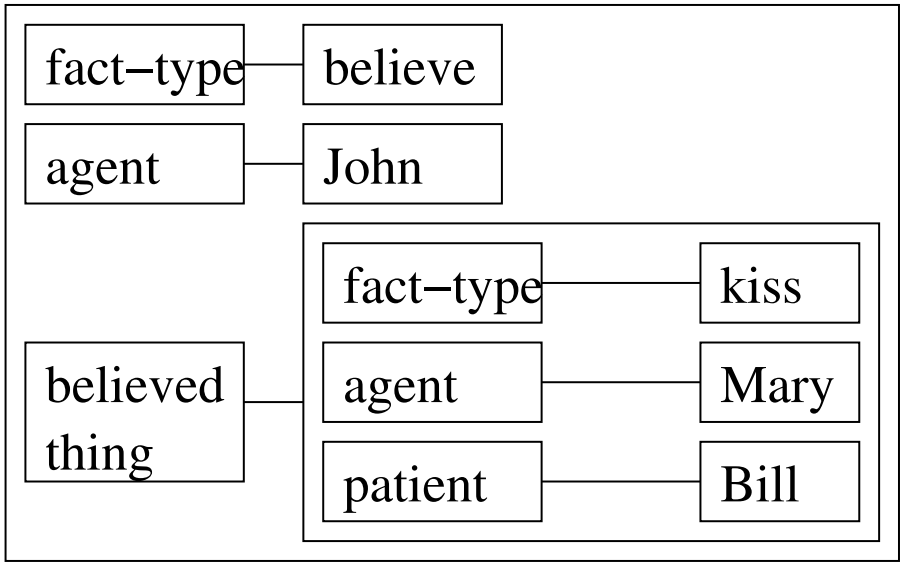


Difficulties in representing propositional attitudes

2. Facts about propositional attitudes have recursive structure.

   Mary kissed Bill.
   John believes [Mary kissed Bill].
   Bill fears [John believes [Mary kissed Bill]].

It is particularly hard to find a way of encoding nested propositions in a neural network.



Difficulties in representing propositional attitudes

3. The semantics of propositional attitudes is hard to capture.

A belief (/fear/utterance/. . .) is ‘about’ some proposition *P*. . .

Difficulties in representing propositional attitudes

3. The semantics of propositional attitudes is hard to capture.

Consider

John believes Mary loves Bill.

A belief (/fear/utterance/...) is 'about' some proposition *P*...

Difficulties in representing propositional attitudes

3. The semantics of propositional attitudes is hard to capture.

Consider

John believes Mary loves Bill.

- What's the relation between John's belief attitude and its content? A belief (/fear/utterance/...) is 'about' some proposition *P*...

Difficulties in representing propositional attitudes

3. The semantics of propositional attitudes is hard to capture.

Consider

John believes Mary loves Bill.

- What's the relation between John's belief attitude and its content? A belief (/fear/utterance/...) is 'about' some proposition *P*...
- Statements about beliefs are *intensional*: their truth depends on how the content of the belief is expressed. How should this be captured?

Talk actions as early propositional attitudes

If Tomasello is right, an infant's representations of communicative actions are early (proto-) propositional attitude representations.

John believes [that Mary loves Bill].
John desires [that Mary loves Bill].
John fears [that Mary loves Bill].
John **says** *[that Mary loves Bill].*

Talk actions as early propositional attitudes

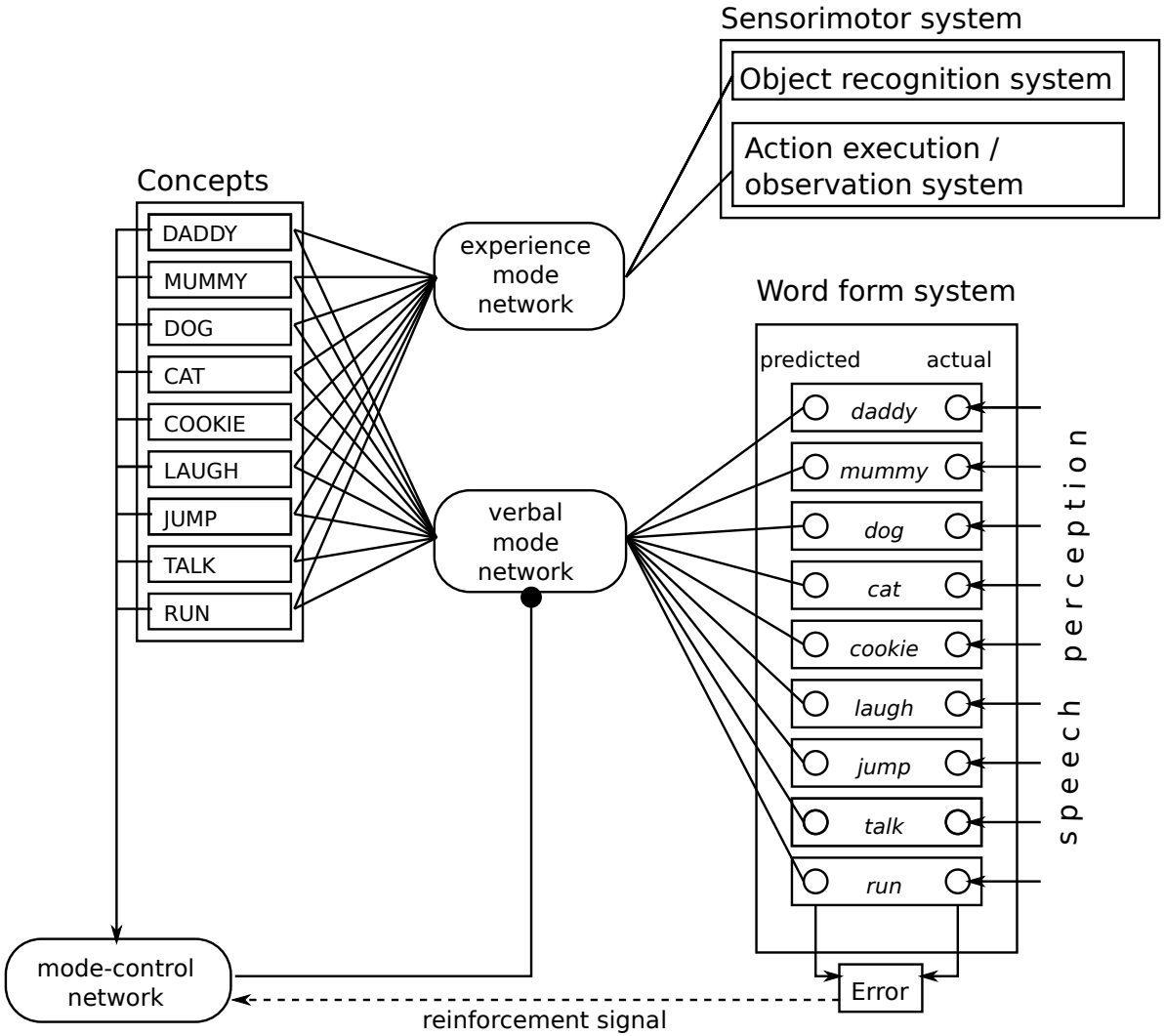
If Tomasello is right, an infant's representations of communicative actions are early (proto-) propositional attitude representations.

John believes [that Mary loves Bill].
John desires [that Mary loves Bill].
John fears [that Mary loves Bill].
John **says** *[that Mary loves Bill].*

I'll propose a model of these early propositional attitude representations.

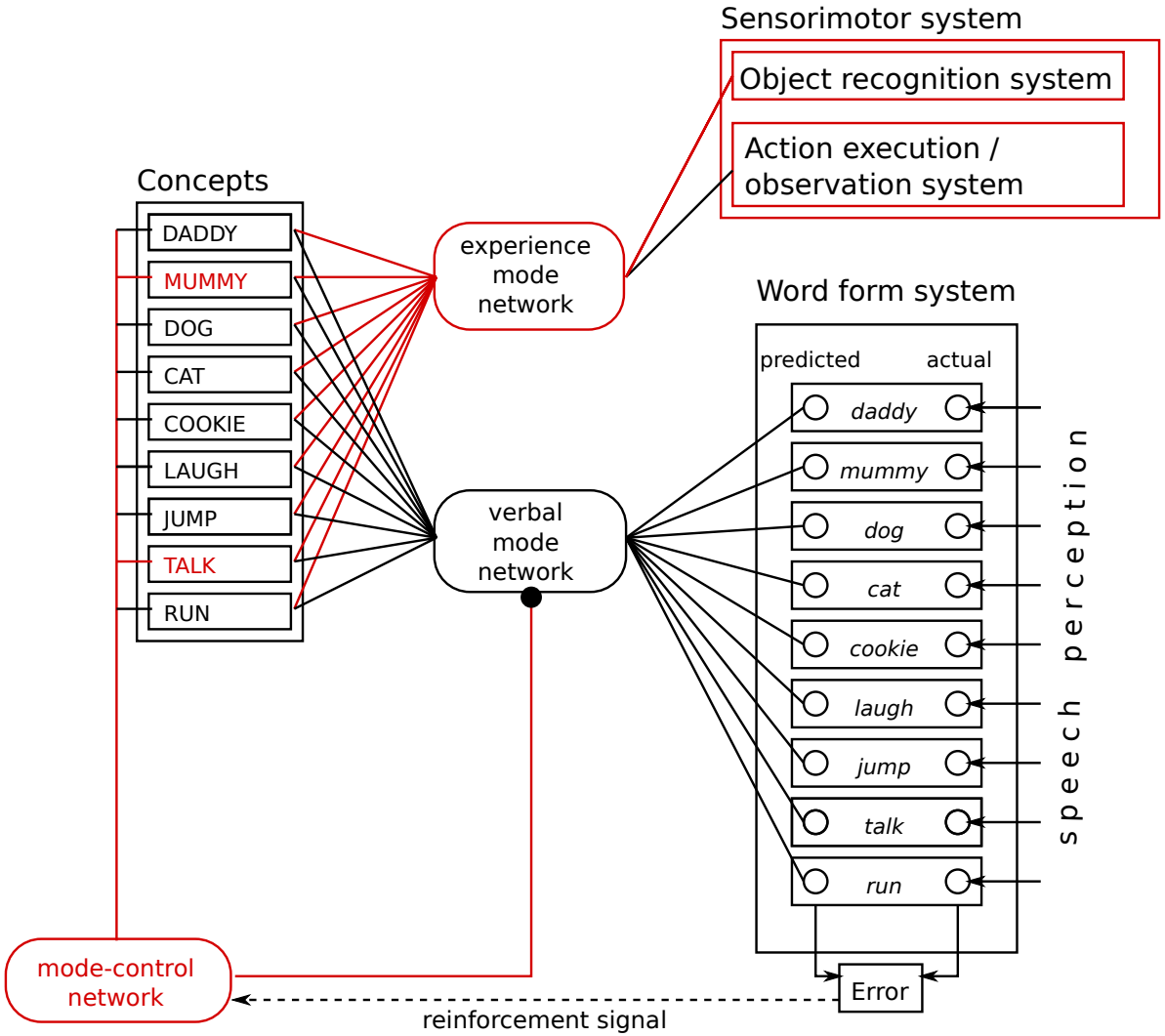
A mature network

Assume at a certain point in development. . .



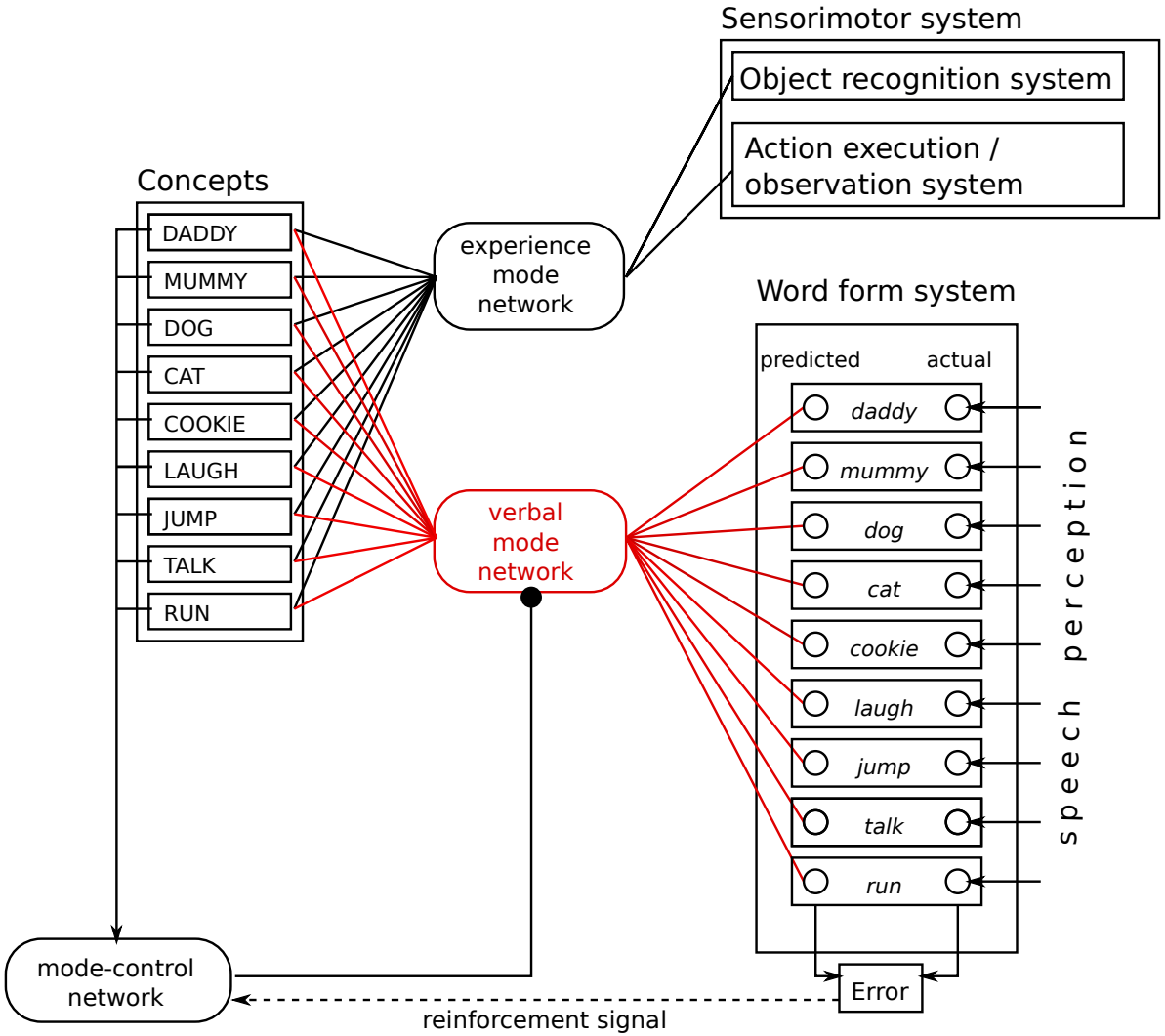
A mature network

1. The infant has learned to engage verbal mode after TALK actions.



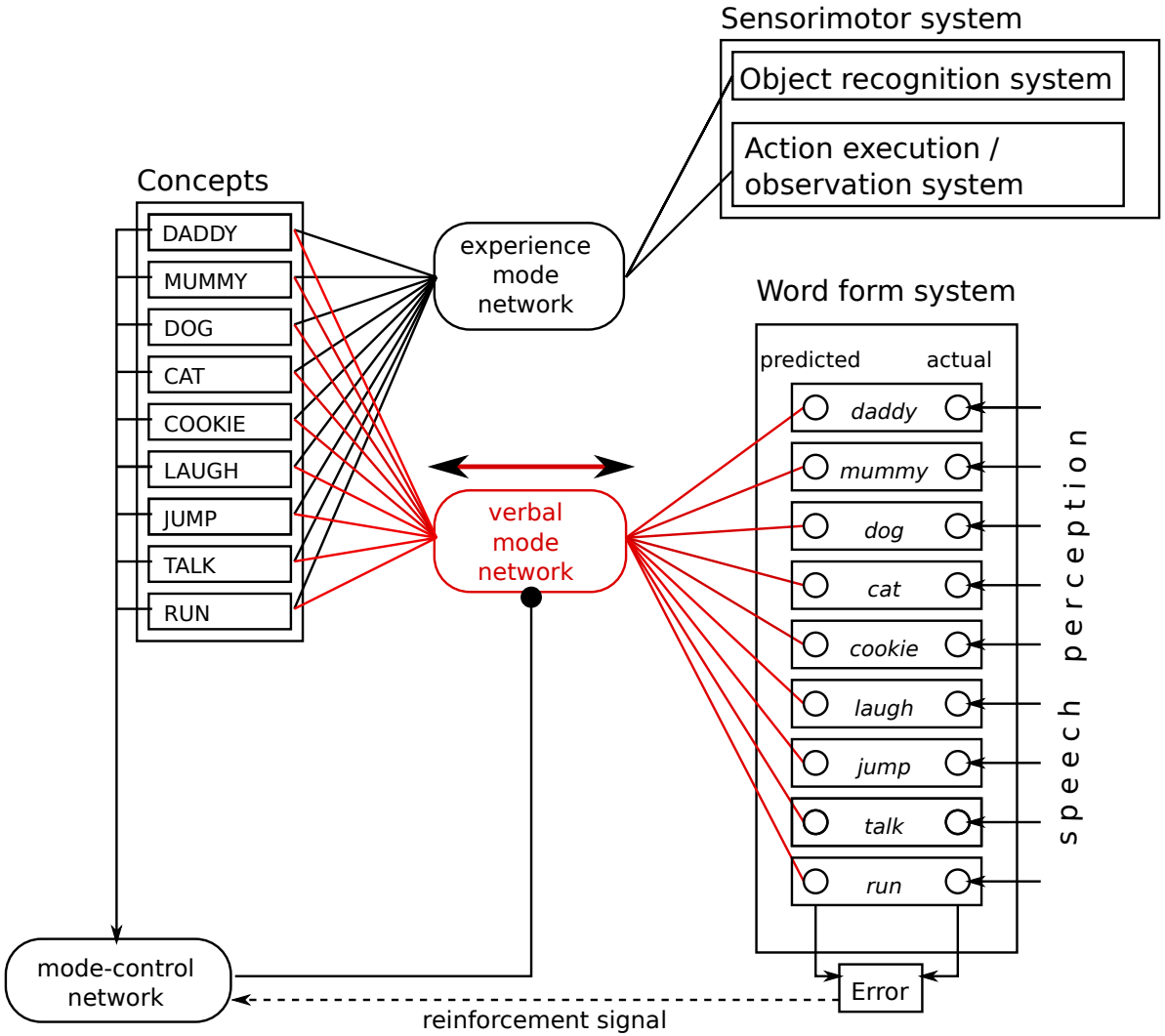
A mature network

2. Verbal mode *alternates* with experience mode.



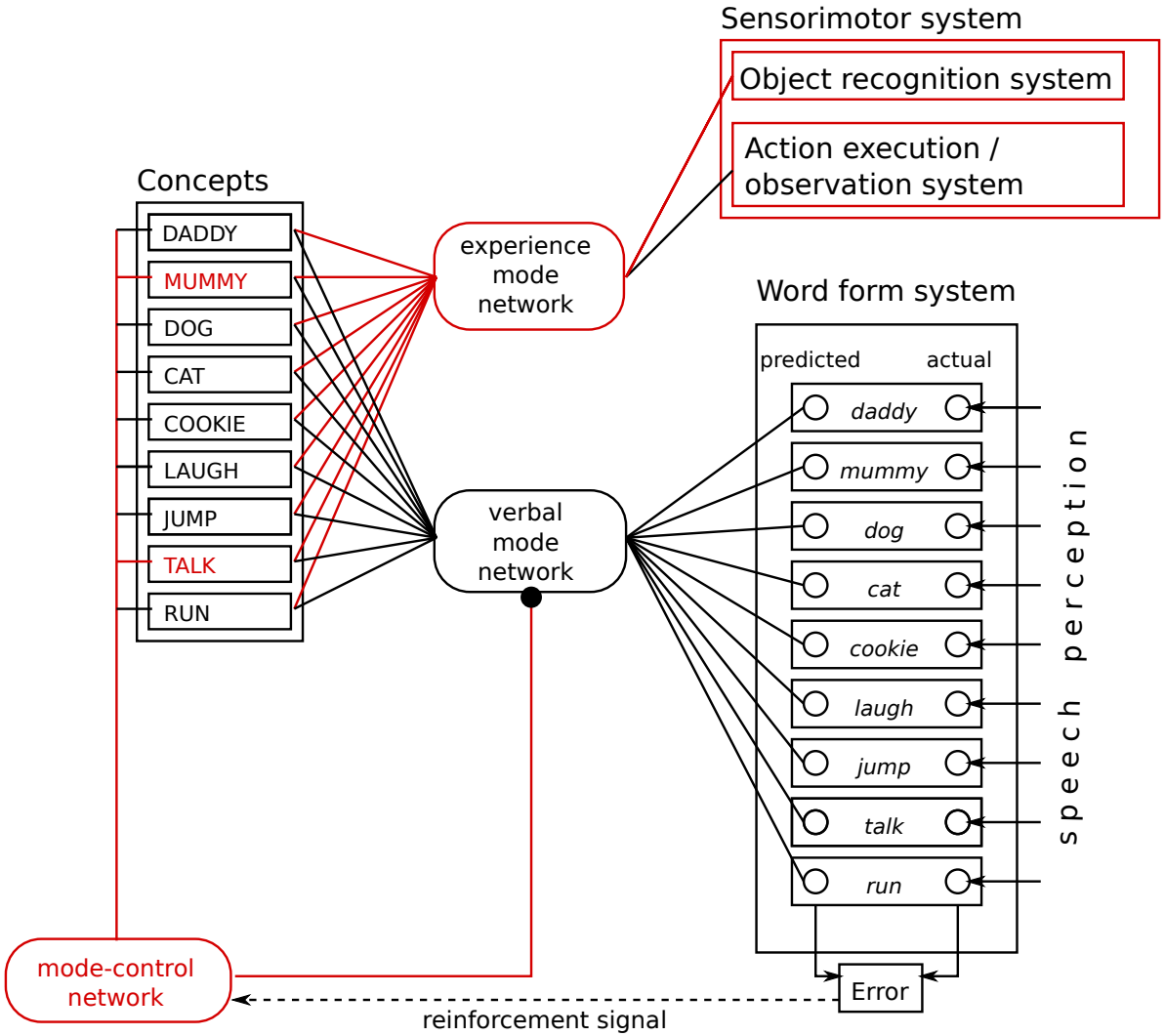
A mature network

3. The associations in verbal mode are *bidirectional*.



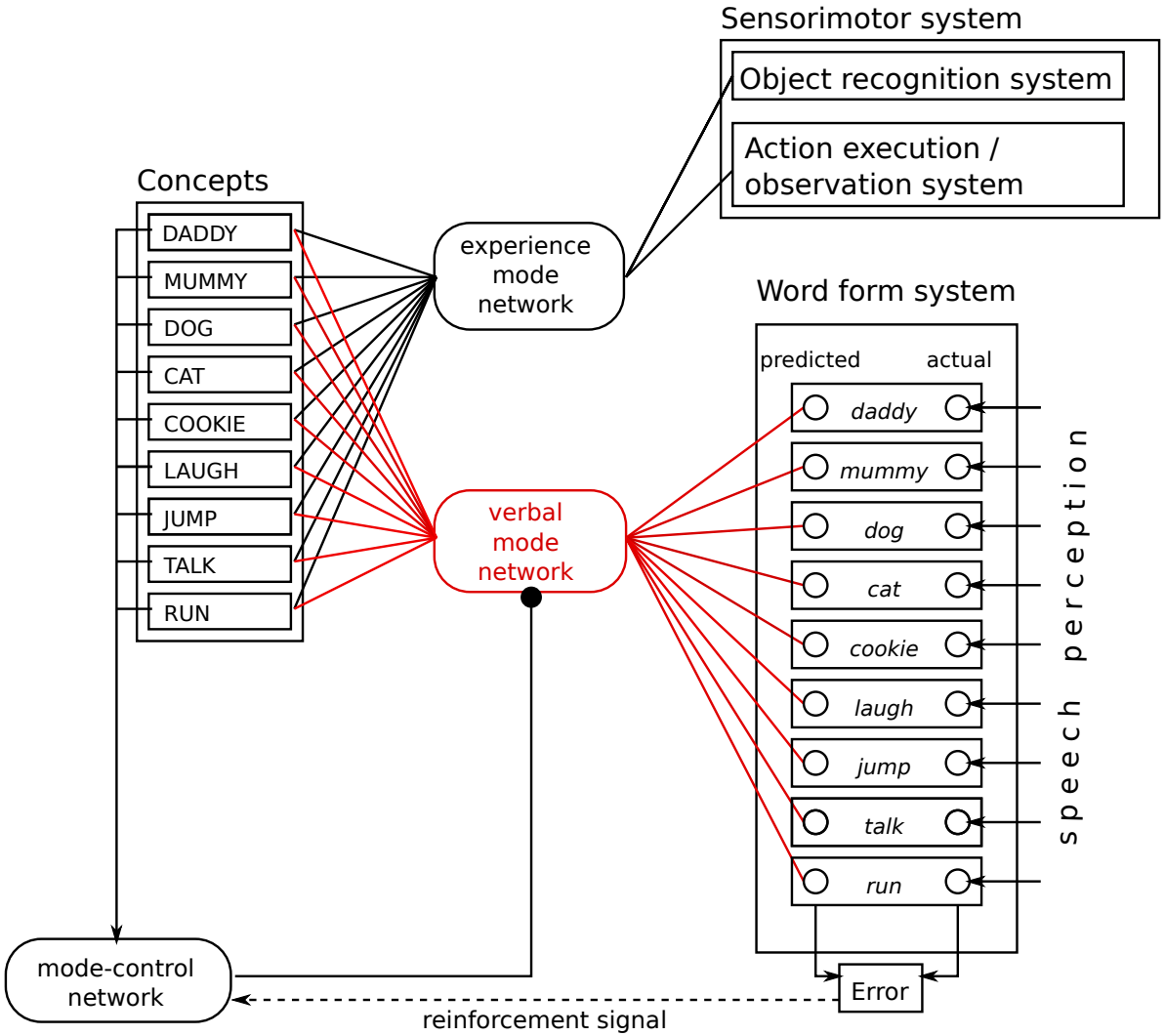
A mature network

Now, when the infant observes a TALK action. . .



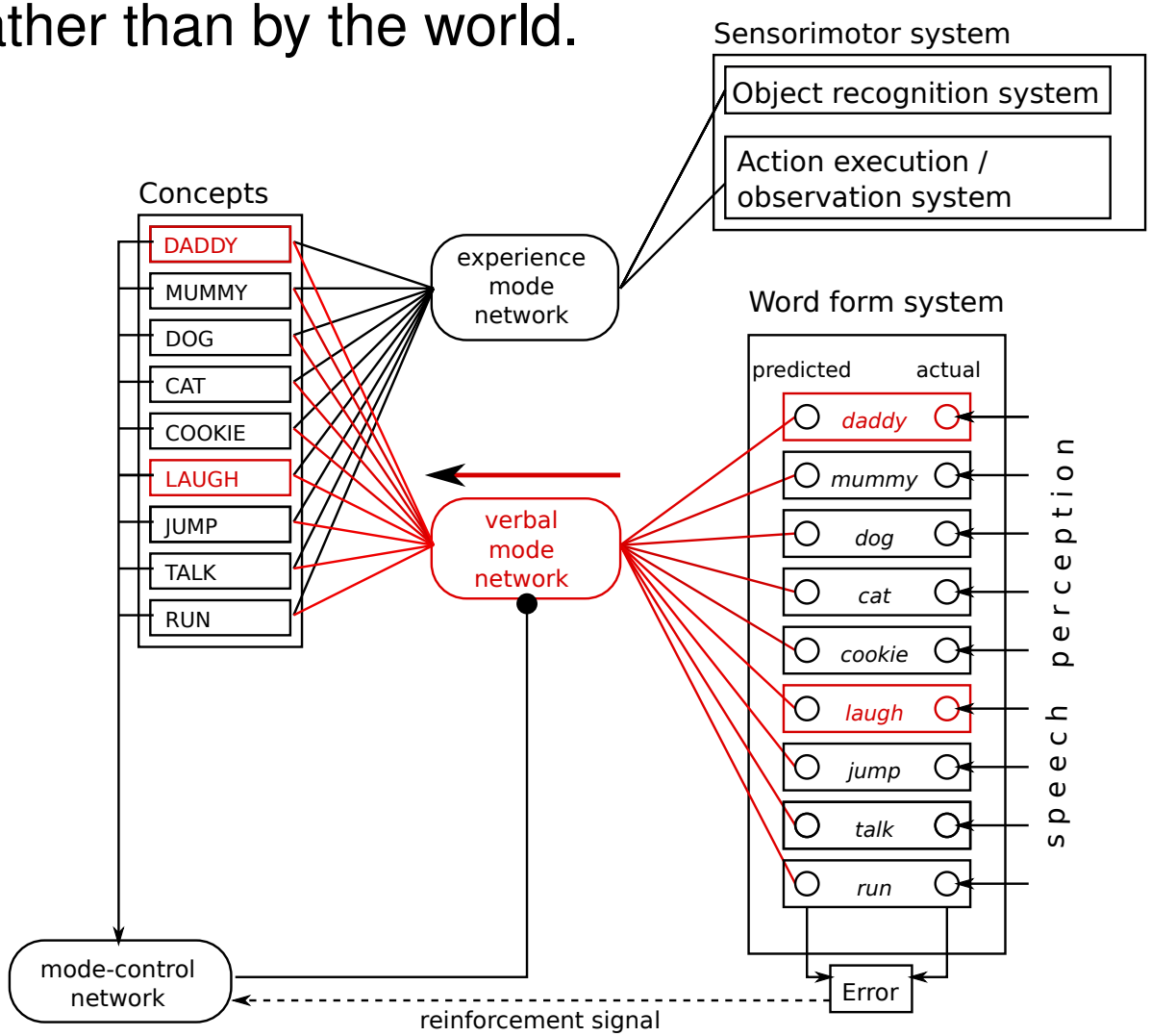
A mature network

. . . and enters verbal mode. . .

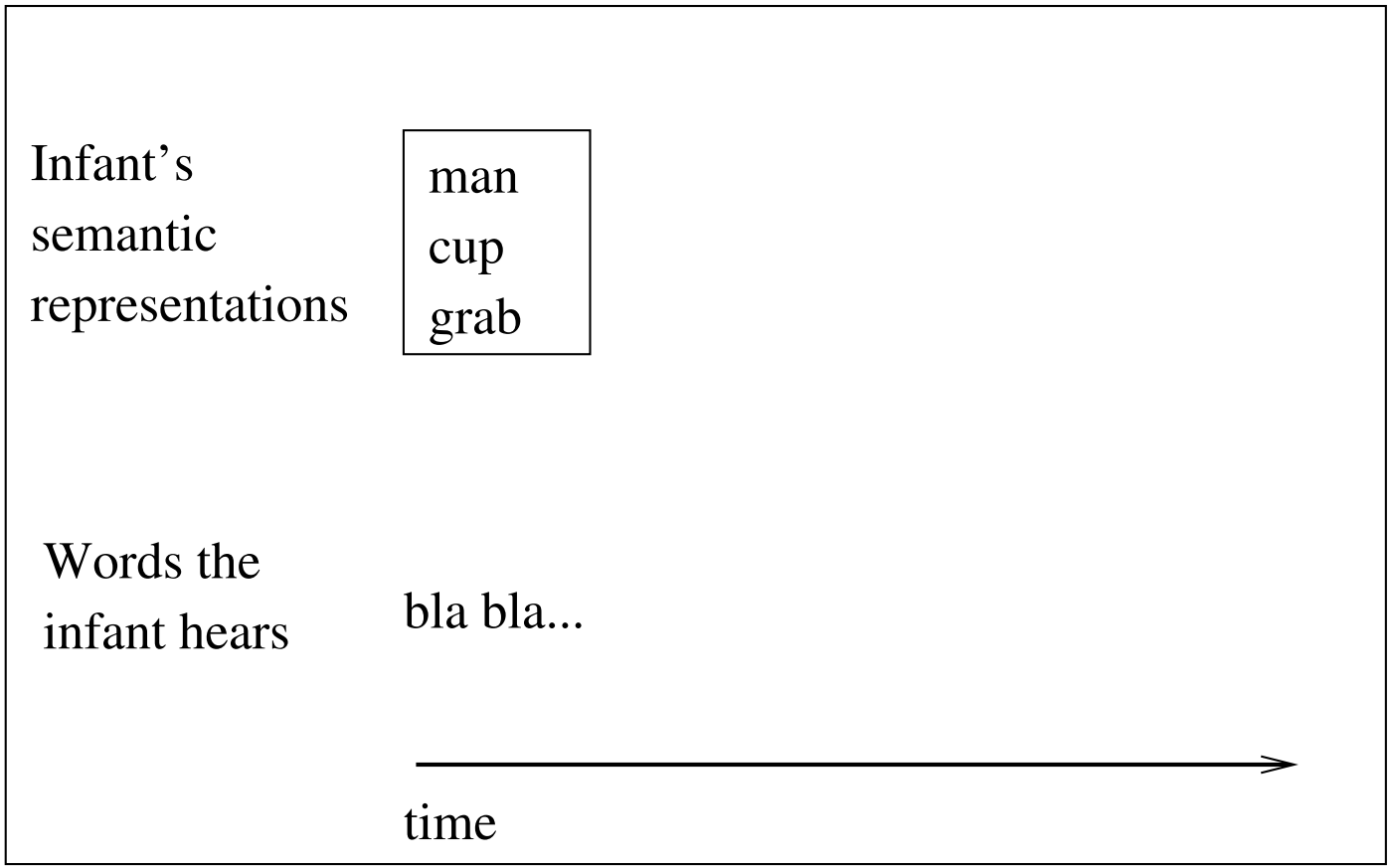


A mature network

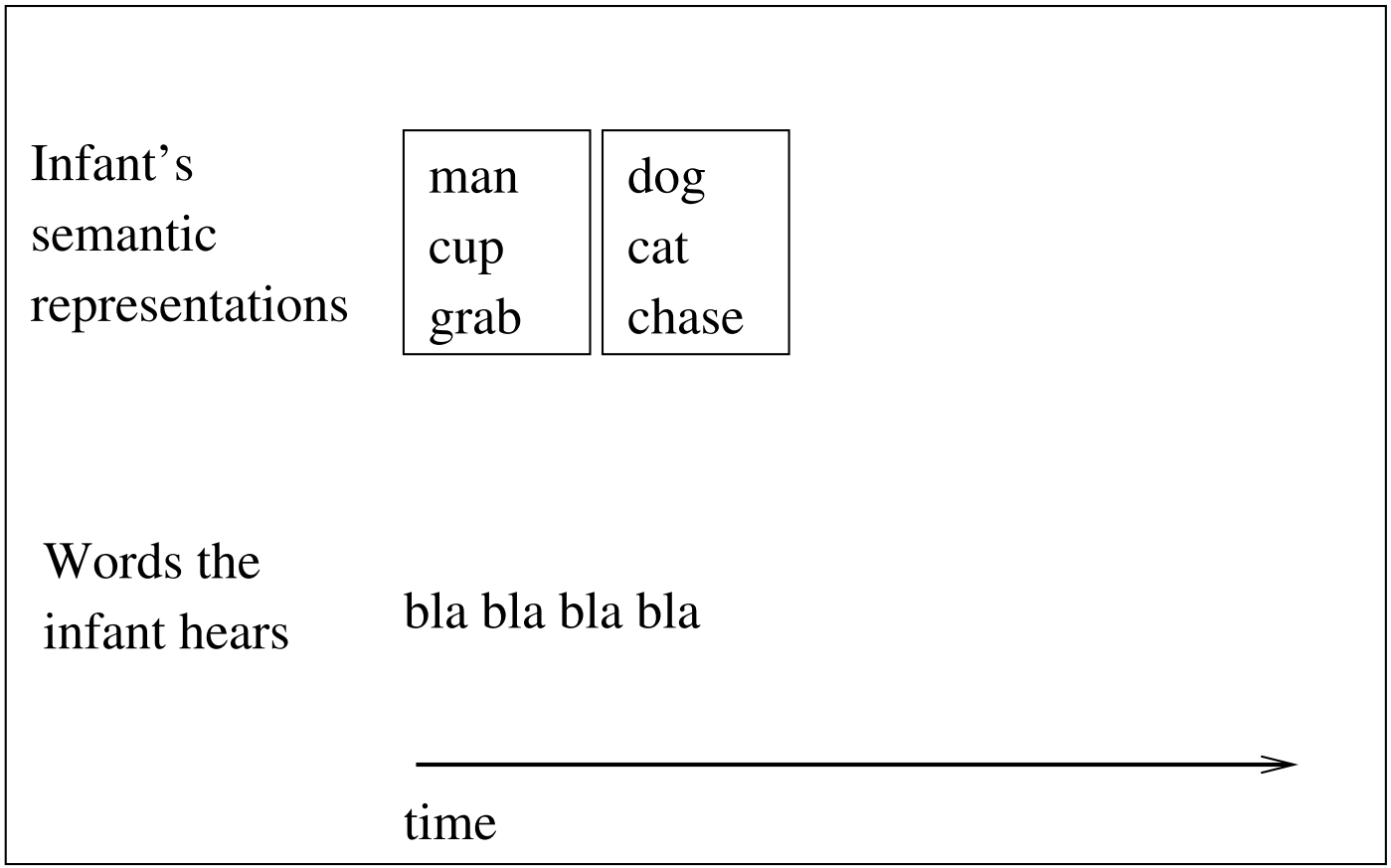
. . . her next semantic representations will be triggered by the words of the action, rather than by the world.



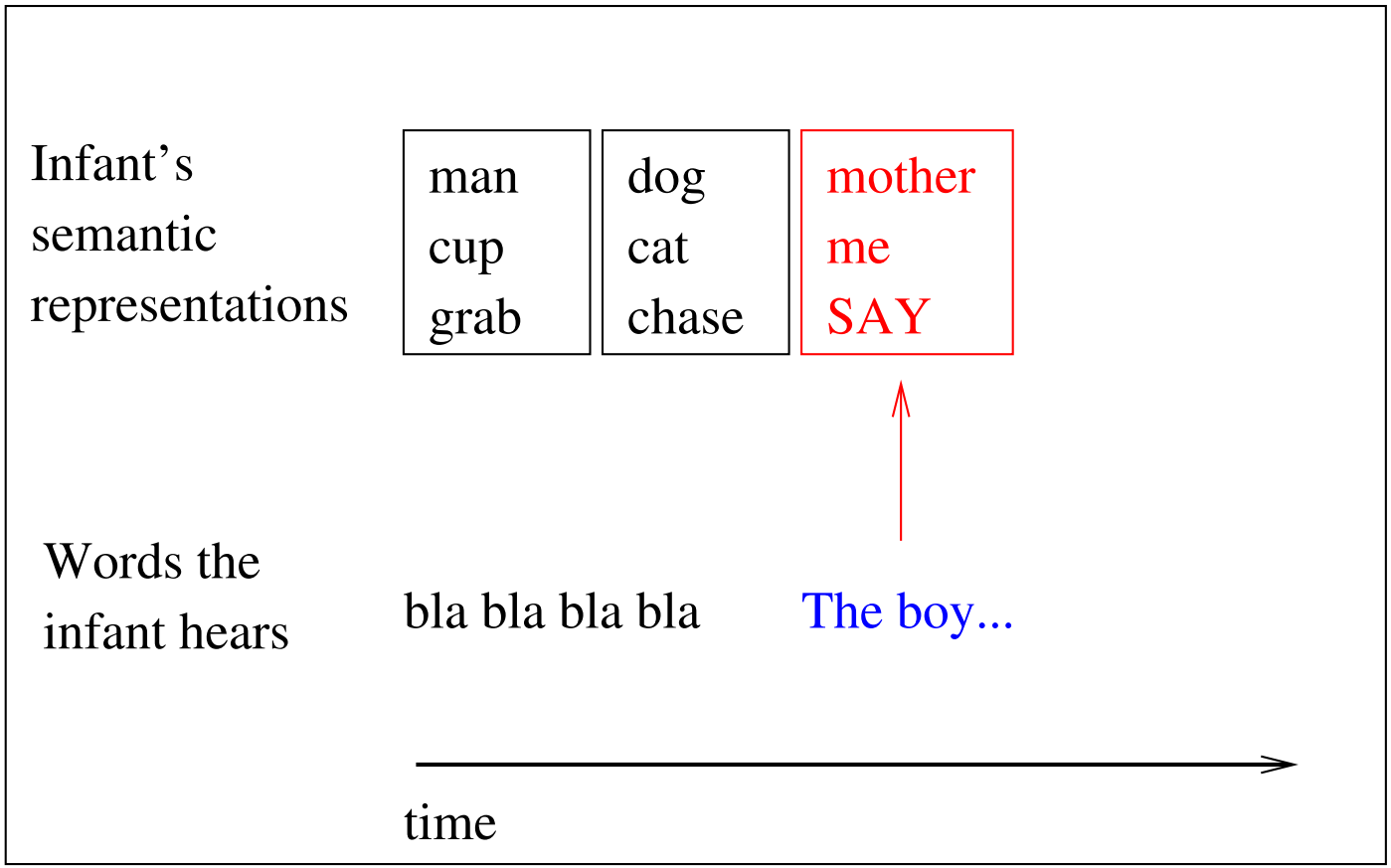
A model of communicative action representations



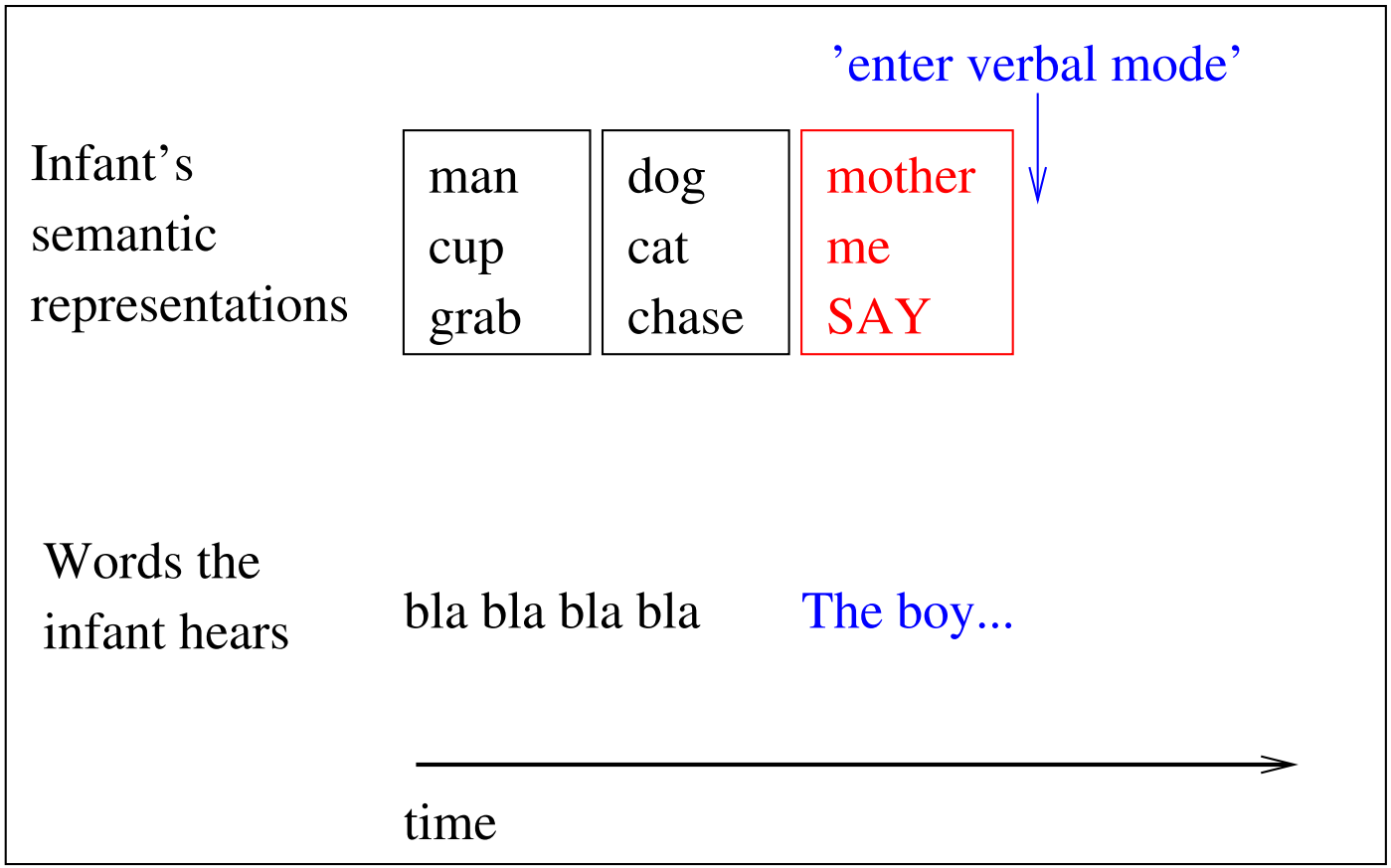
A model of communicative action representations



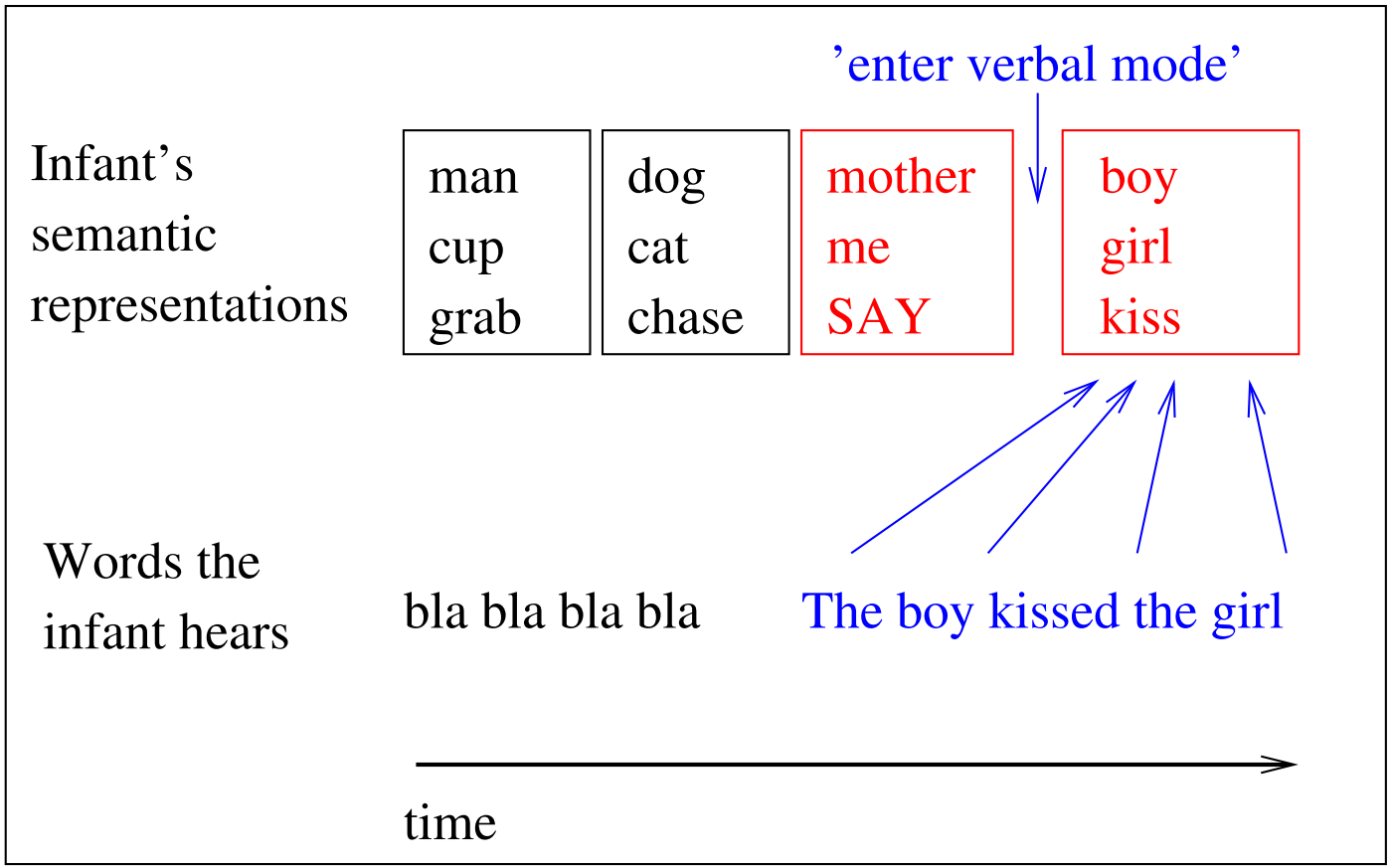
A model of communicative action representations



A model of communicative action representations

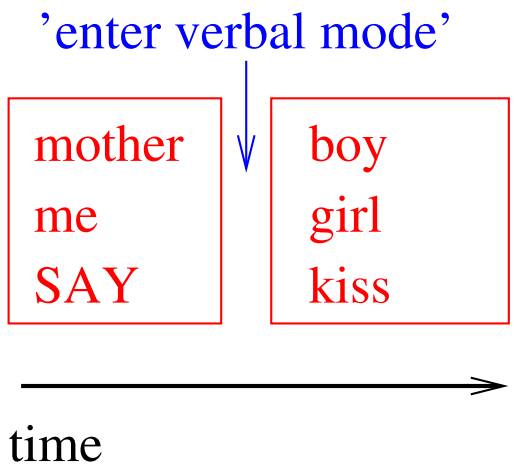


A model of communicative action representations



A model of communicative action representations

Proposal: we represent communicative actions using a pair of consecutive episode representations, interleaved by the operation of entering verbal mode.



Semantic representations as sequences

Recall: I've already argued that neural representations of concrete episodes in the world are structured as *sequences*.

- Experiencing an episode in the world requires a canonically structured sequence of sensorimotor operations.
- We store concrete episodes in working memory as prepared sensorimotor sequences.

Semantic representations as sequences

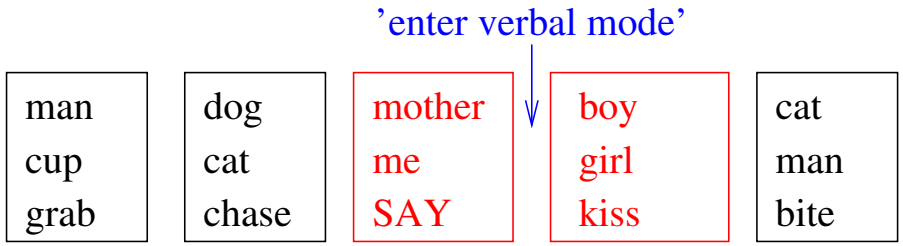
Recall: I've already argued that neural representations of concrete episodes in the world are structured as *sequences*.

- Experiencing an episode in the world requires a canonically structured sequence of sensorimotor operations.
- We store concrete episodes in working memory as prepared sensorimotor sequences.

Extending this idea, perhaps a communicative action is represented as a sequence of *two whole episode representations*.

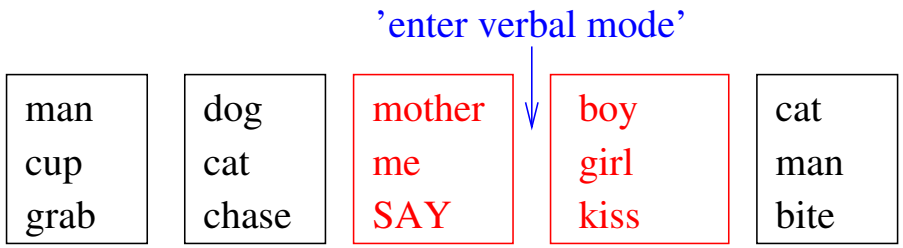
A model of communicative action representations

This scheme offers some interesting solutions to the problems involved in representing mental states in neural networks.



A model of communicative action representations

This scheme offers some interesting solutions to the problems involved in representing mental states in neural networks.



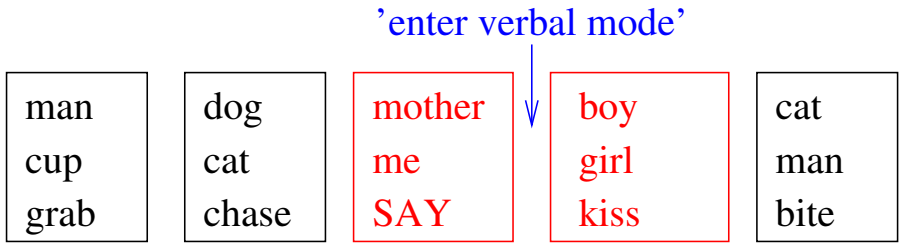
Problem 1: it's hard to represent *nested* propositions in a neural network. ('Mum says [the boy kissed the girl]')

Solution: a mental state ascription has the form of a *sequence*.

- The embedding proposition occupies the same medium as the embedded proposition. But at a later time.

A model of communicative action representations

This scheme offers some interesting solutions to the problems involved in representing mental states in neural networks.



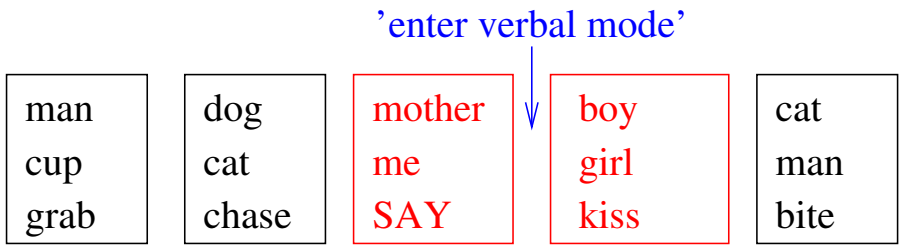
Problem 2: it's hard to represent the relation between a propositional attitude and its content.

Solution: the transition to verbal mode explains this.

- It explains why the content needn't describe the actual world.
- It explains how the talk action is *about* its content.
- It explains why mental state ascriptions are intensional.

A model of communicative action representations

This scheme offers some interesting solutions to the problems involved in representing mental states in neural networks.



Note that these two solutions are related!

- It's only *because* talk actions are represented as sequences that we can change mode halfway through.

Summary

How infants learn words (empirical data):

- Preliminaries: word-sized phonological representations and concepts
- A small passive vocabulary
- Fast expansion of passive and active vocabulary size, perhaps correlated with development of social/pragmatic skills

Computational models of early word learning:

- Cross-situational learning is probably involved
- Joint attention is probably involved
- Perhaps vocabulary development bootstraps development of the concept of a communicative action