Nature vs. nurture

Infants have domain-specific mechanisms

- •Math
- •Physics
- •Psychology
- •Animacy
- •Causality

the ONION

a.v. club | store | subscriptions | books | personals

Mobile/PDA | Print Edition | Text Version

NEWSLETTERS | CONTESTS

VOLUME 31 ISSUE 19 AMERICA'S FINEST NEWS SOURCE 21 MAY 1997

Home: Archives: Science

SCIENCE

Study Reveals: Babies Are Stupid

LOS ANGELES—A surprising new study released Monday by UCLA's Institute For Child Development revealed that human babies, long thought by psychologists to be highly inquisitive and adaptable, are actually extraordinarily stupid.

The study, an 18-month battery of intelligence tests administered to over 3,500 babies, concluded categorically that babies are "so stupid, it's not even funny."

Above: Despite their relatively large cranial capacities, babies such as this one are so ... able to ... squeak toys from

... not grasp even this

... drowned.

Some of the babies tested were actually so stup... toys. Others, unable to use such primitive instrum... skills, simply starved to death, despite being surro... Gerber-brand baby-food products.

... of Micronaut space ... insufficient motor ... giving

Infants have domain-general mechanisms

- •Habituation
- •Associative learning
- •Conditioning
- •Imitation

"Over 90 percent of them, when poked, failed to make even rudimentary attempts to defend themselves."

Rethinking innateness (Elman et al., 1996)

What does it mean to say a behavior (or knowledge) is **innate**?

- Brain is highly structured at birth and continues to develop/mature through adolescence
  - But not enough room in the genome to code complex knowledge (about 99% shared between humans and chimps/bonobos)
- Experience is rich and massive and begins before birth
- All complex behavior emerges from the **interaction** of structure and experience

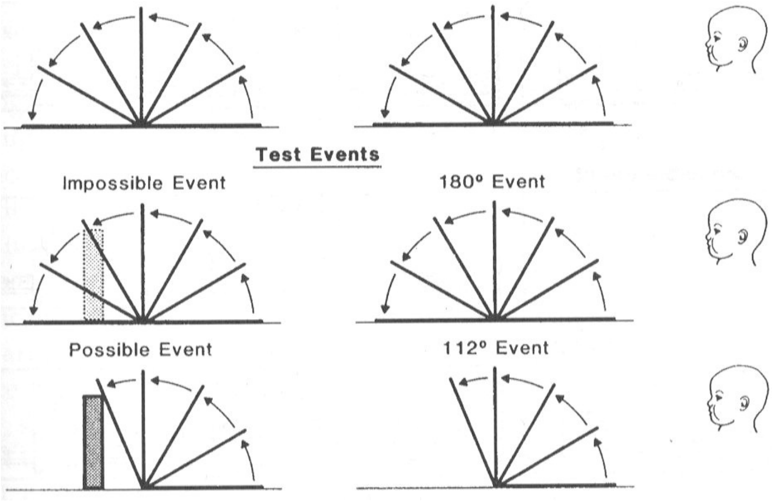
"That which is inevitable need not be innate."
— Jean Piaget

Object permanence: A-not-B error

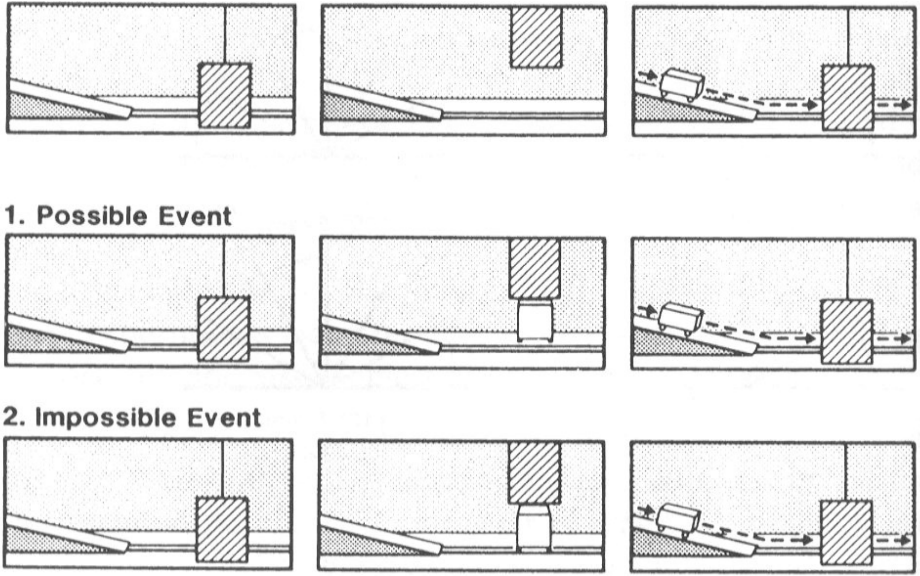
Object permanence: A-not-B error

Object permanence

- When a toy is moved and hidden, young children will continue to search in the old location, even though much younger infants "know" that hidden objects continue to exist



Object permanence (cont.)



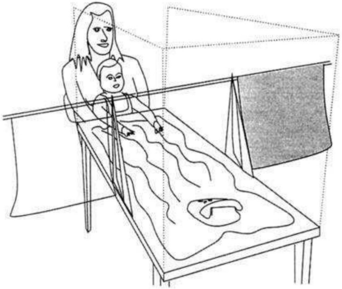
Munakata et al. (1997)

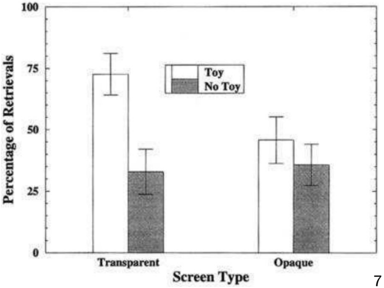
Puzzle: Why do infants show sensitivity to occluded objects and show an ability to retrieve objects but fail to retrieve occluded objects?

- Perhaps infants “know” the occuluded object exists but lack the means-ends ability to retrieve it.

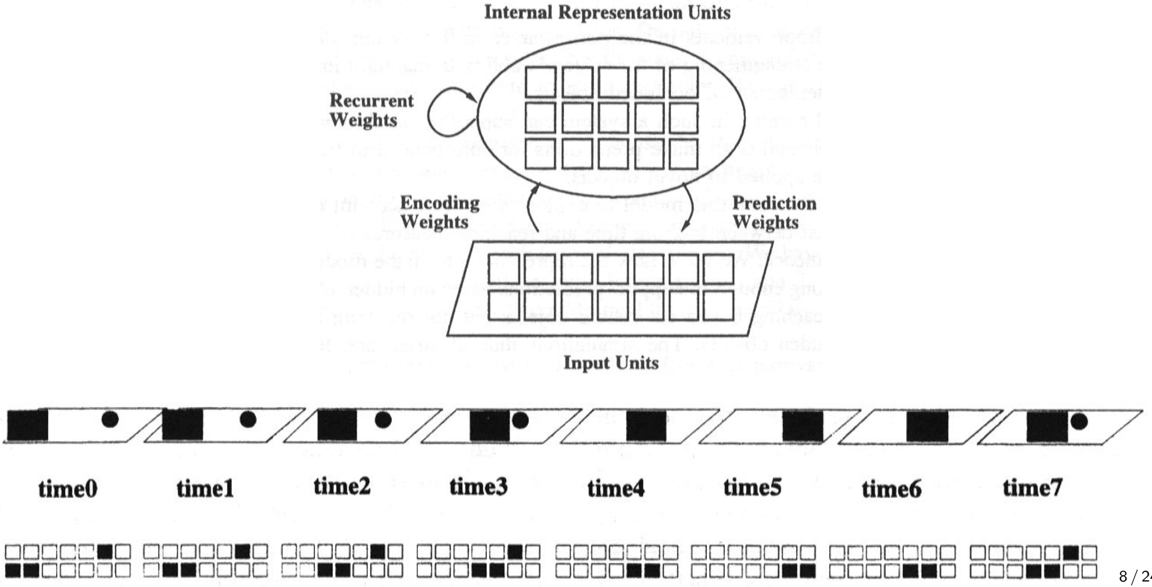
Munakata et al. trained 7 mo infants to retrieve distant objects lying on a towel, separated by a transparent or opaque barrier

- Infants had no difficulty retrieving object viewed through transparent barrier
- Lack of “object permanence” when reaching (vs. looking) not due to differences in means-ends demands

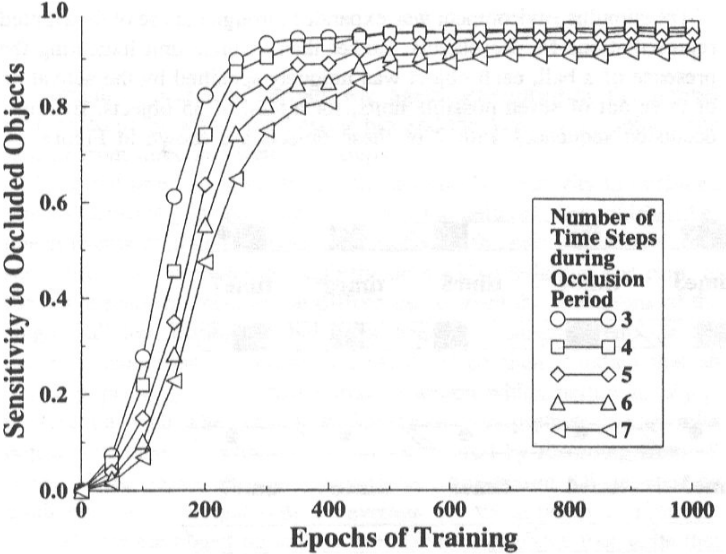




Model (Munakata et al., 1997)

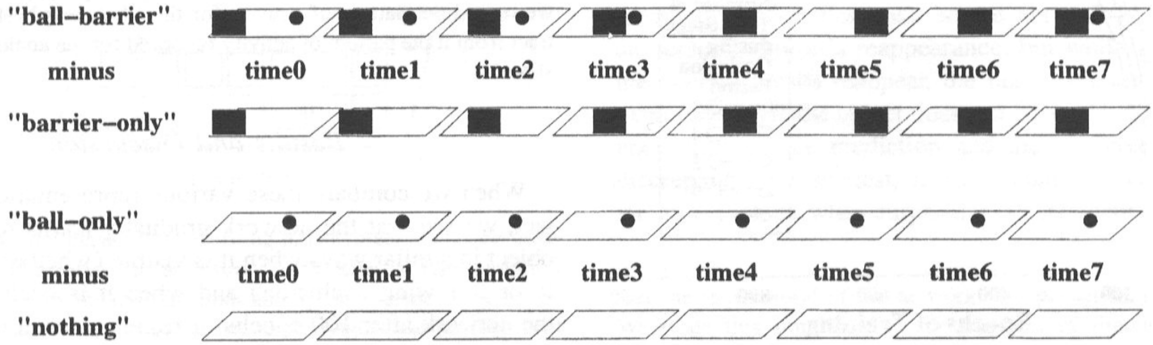


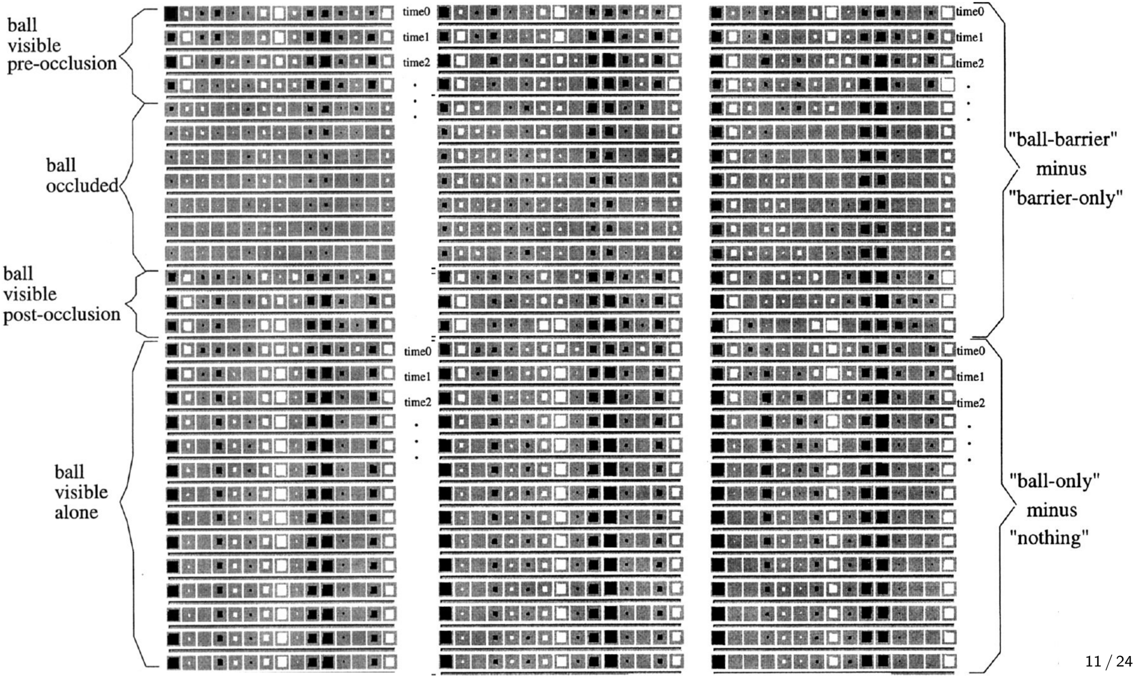
Training



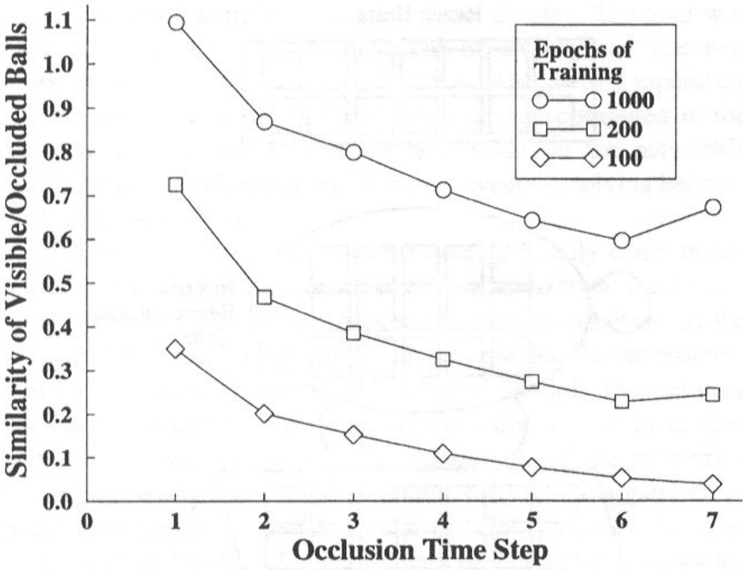
Sensitivity to Occluded Objects
1.0
0.8
0.6
0.4
0.2
0.0
0
200
400
600
800
1000
Epochs of Training
Number of Time Steps during Occlusion Period
3
4
5
6
7

Analysis of hidden representations

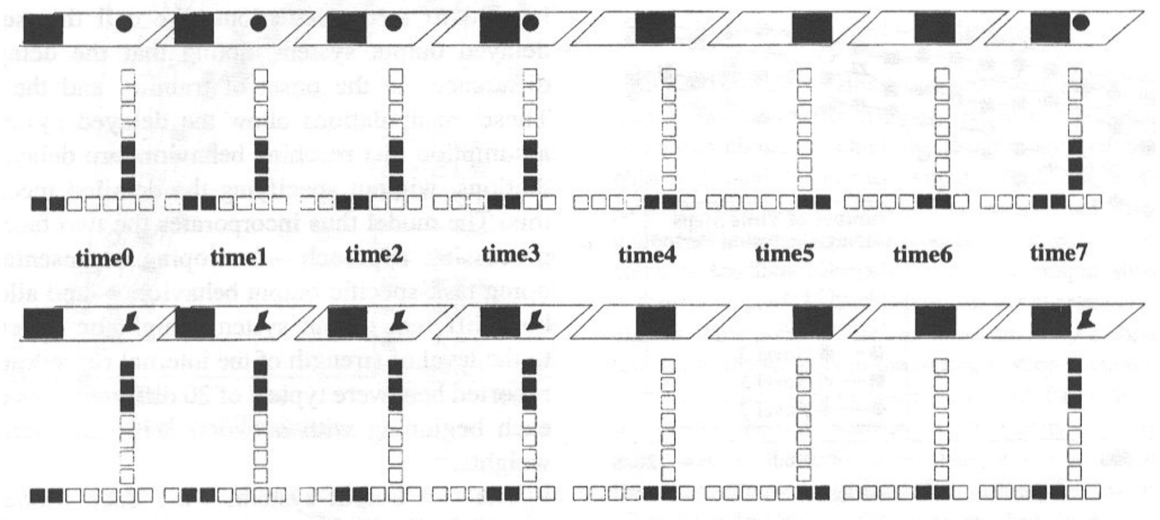



ball visible pre-occlusion
ball occluded
ball visible post-occlusion
ball visible alone
time0
time1
time2
"ball-barrier" minus "barrier-only"
"ball-only" minus "nothing"

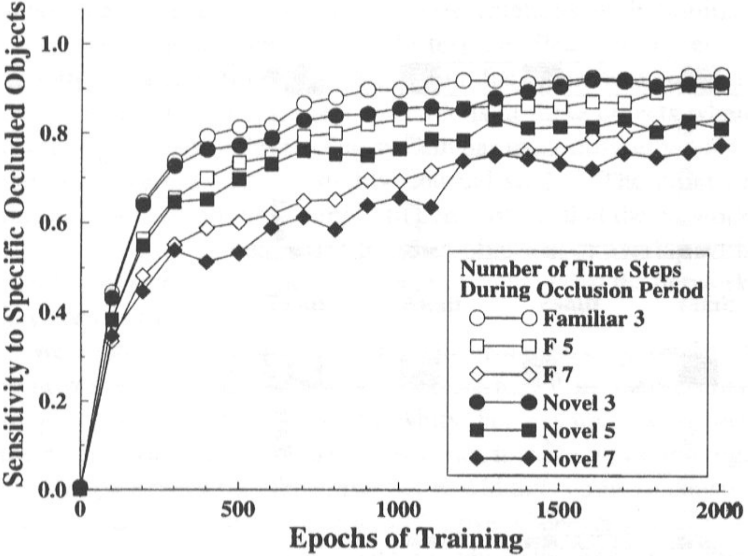
Persistence of object information over occlusion steps



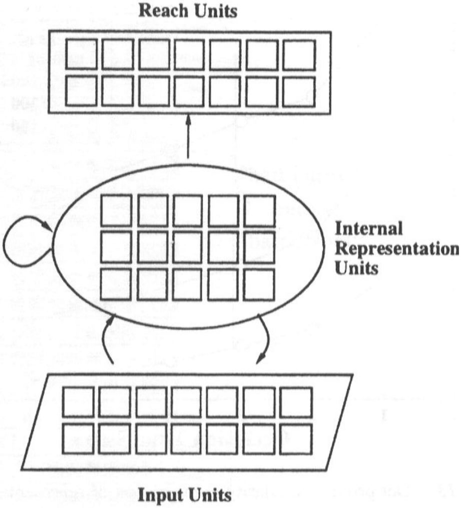
Memory for object "features"

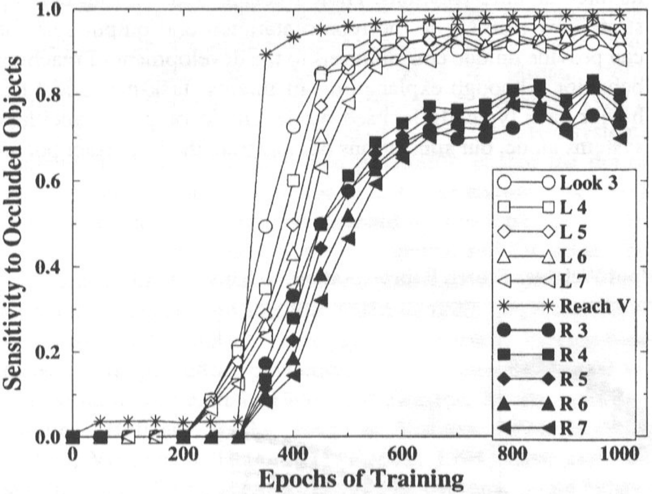


Memory for familiar vs. unfamiliar objects

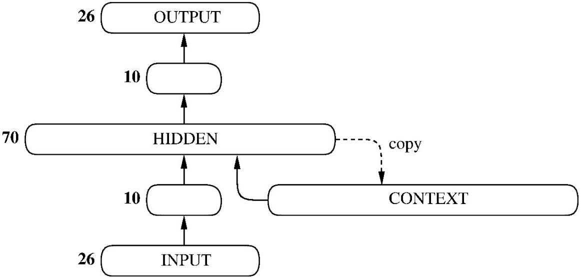


Dissociation of looking vs. reaching





A “critical period” in language learning? (Elman, 1991)



S → NP VI . | NP VT NP .
NP → N | N RC
RC → who VI | who VT NP | who NP VT
N → boy | girl | cat | dog | Mary | John | boys | girls | cats | dogs
VI → barks | sings | walks | bites | eats | bark | sing | walk | bite | eat
VT → chases | feeds | walks | bites | eats | chase | feed | walk | bite | eat

- Simple recurrent network trained to predict next word in English-like sentences
  - Context-free grammar, number agreement, variable verb argument structure, multiple levels of embedding
  - 75% of sentences had at least one relative clause; average length of 6 words.
  - e.g., **Girls who cat who lives chases walk dog who feeds girl who cats walk .**
- After 20 sweeps through 4 sets of 10,000 sentences, mean absolute error for new set of 10,000 sentences was **0.177** (cf. initial: 12.45; uniform: 1.92)

The importance of "starting small" (Elman, 1991)

- Training was successful **only** when *"starting small"*
  - Trained on only simple sentences before gradually introducing embedded sentences
  - Trained on full language but with initially limited memory that gradually improved

- Consistent with Newport's (1990, *Cog. Sci.*) *"less is more"* hypothesis
  - Child language acquisition is helped rather than hindered by maturational limits on cognitive resources

Alternative Hypothesis: Need to start small was exaggerated by lack of important *soft* constraints inherent in natural language

- SRN's learn long-distant dependencies better when intervening material is partially correlated with distant information (Cleeremans et al., 1989, *Neural Comp.*)
- Soft **semantic constraints**—distributional biases on noun-verb co-occurrences across clauses—provide such correlations

Simulation 1: Semantic constraints (Rohde & Plaut, 1999)

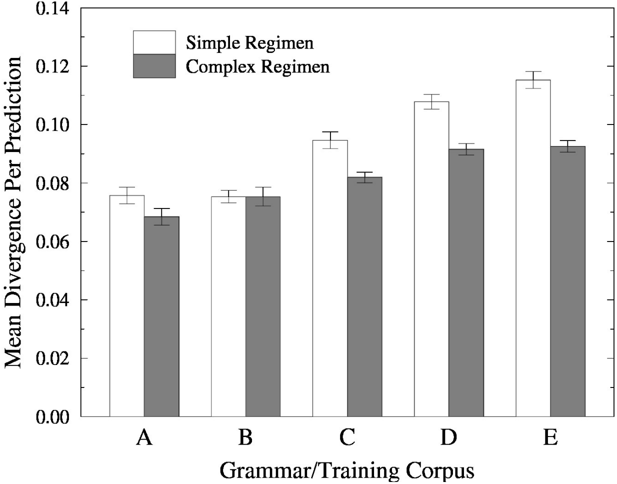
- Replication of Elman (1993) simulation with addition of constraints on verb arguments
- Parametric variation of reliability of semantic constraints across clauses
  (**A** = none, ..., **E** = 100% reliable)
- Minor improvements in technical aspects of simulation (e.g., error function, initialization)

Verb	Intransitive Subjects	Transitive Subjects	Objects if Transitive
chase	–	any	any
feed	–	human	animal
bite	animal	animal	any
walk	any	human	only dog
eat	any	animal	human
bark	only dog	–	–
sing	human or cat	–	–

Compared two **training regimens**:

- **Complex**: Trained on full language throughout
  - 25 epochs through 10,000 sentences (75% complex)
- **Simple**: Trained incrementally
  - 5 epochs on simple sentences; 5 on 25% complex; 5 on 50% complex; 10 on 75% complex

Results: Prediction error

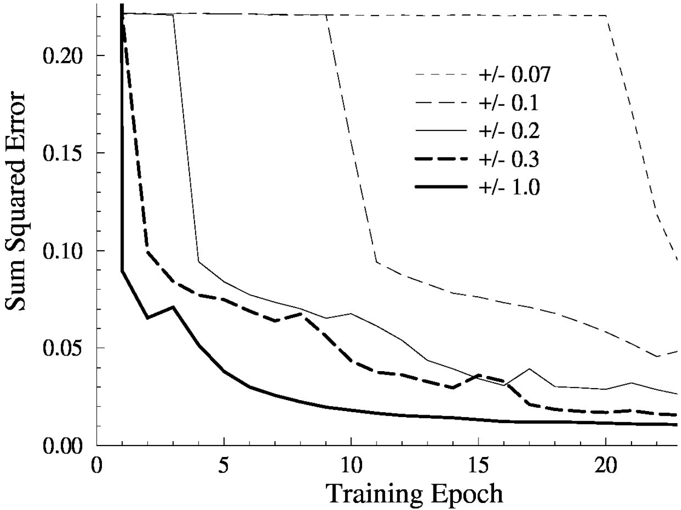


- **Disadvantage** for “starting small” that increases with reliability of semantic constraints

Relation to Elman's (1993) results

Exact replication, varying **magnitudes of initial random weights**

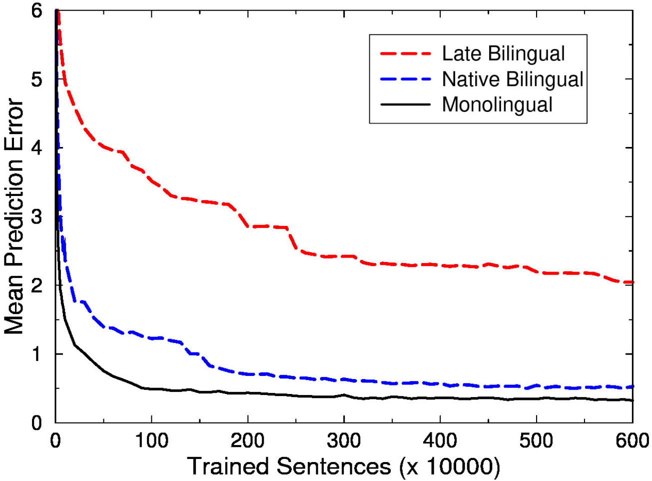
- Simulation 1 used $\pm\mathbf{1.0}$; Elman used $\pm\mathbf{0.001}$
- Very small initial weights prevent effective accumulation of error derivatives



Simulation 2: Native vs. late bilingual acquisition

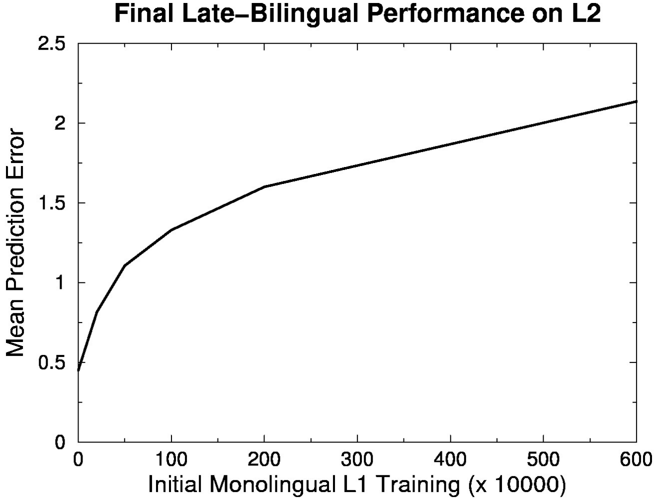
- **Languages**
  - *English*: Analogous to language from Simulation 1
  - *German*: German vocabulary ("hund" vs. "dog"), gender marking, case-marking in masculine, verb-final relative clauses
  - Phoneme-based input and output representations
- **Training Conditions**
  - *Monolingual*: Trained on either English or German
    - 6 million sentence presentations sampled from corpus of 50,000 sentences
  - *Native Bilingual*: Trained on both English and German (50/50)
    - 6 million sentence presentations sampled from two corpora of 50,000 sentences each
    - Language selected randomly every 50 sentences
  - *Late Bilingual*: Monolingual training followed by bilingual training
- **Testing**
  - Late Bilingual tested on L2 (new sample of 5,000 sentences)
  - All results counterbalanced for English vs. German

Results: Acquisition



- Initial monolingual training impedes subsequent bilingual acquisition
- Native bilingual acquisition is only slightly worse than monolingual acquisition

Results: Early-bilingual acquisition



- Even relatively brief exposure to monolingual L1 impacts subsequent L2 acquisition

Language learning: Conclusions

- Introducing soft semantic constraints aids learning of pseudo-natural languages by simple recurrent networks
  - No need to manipulate training environment or cognitive resources
  - Networks *inherently* learn local dependences before longer distance ones
- Critical-period effects may reflect *entrenchment* of representations that have learned to perform other tasks (including other languages)
  - No need to introduce additional maturational assumptions (e.g., "less is more")