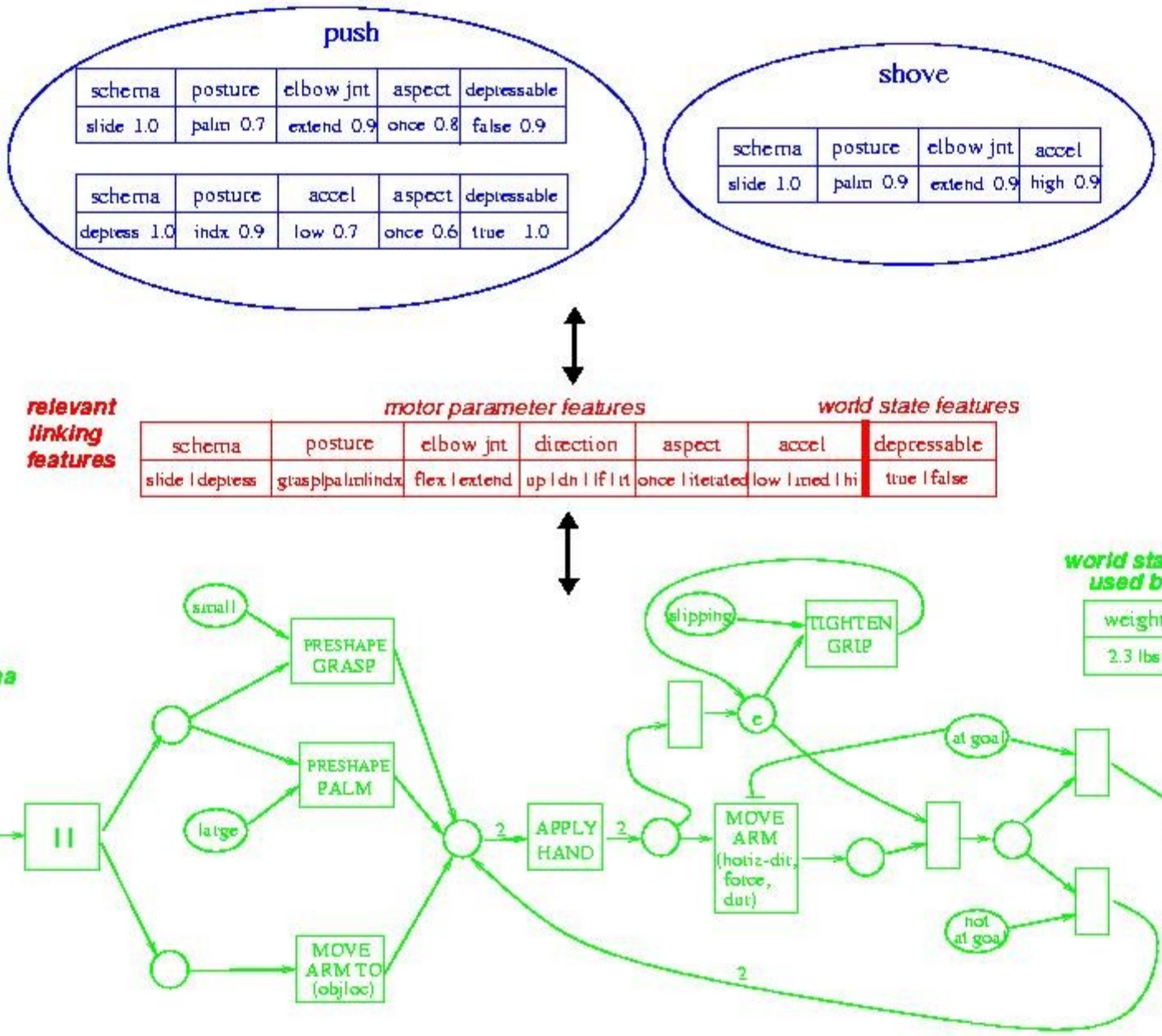
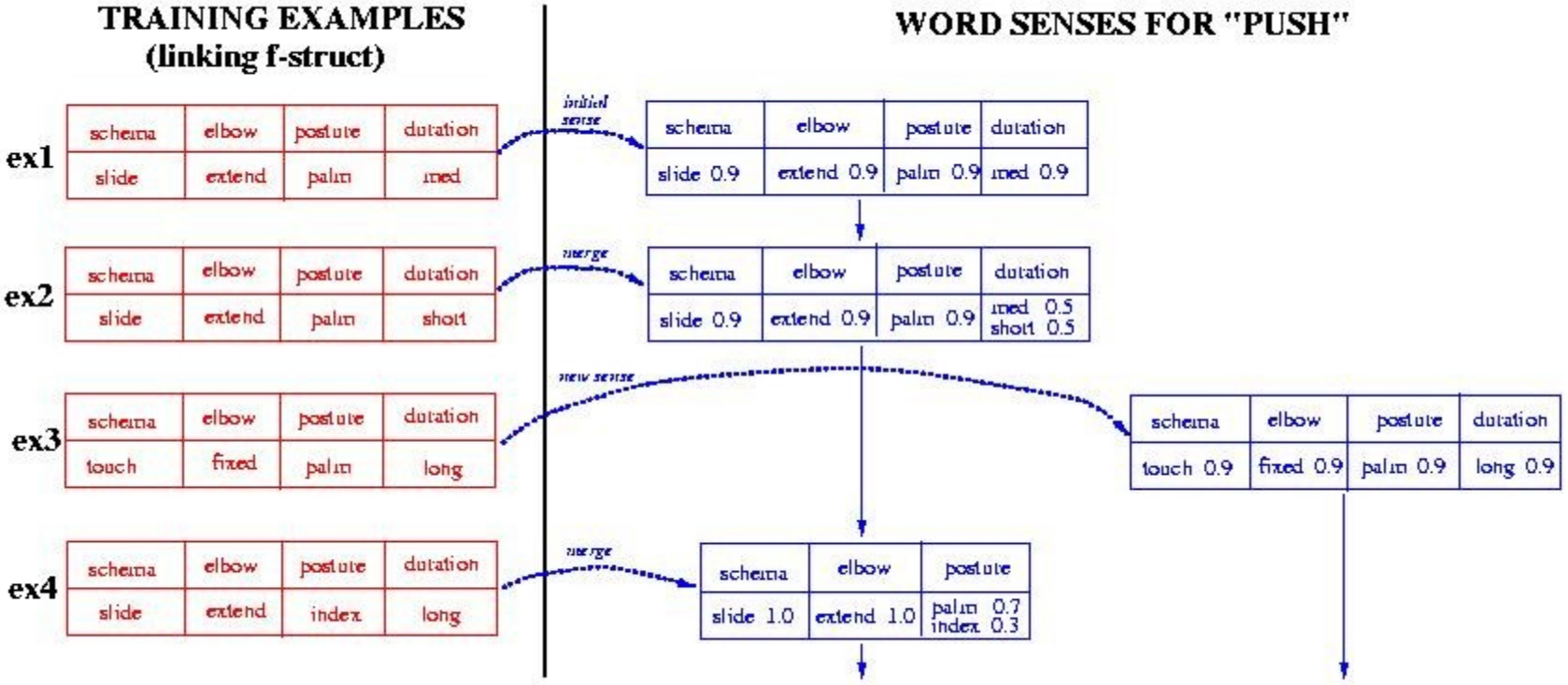
System Overview



Learning Two Senses of PUSH



Model merging based on Bayesian MDL

Training Results
David Bailey

English

- 165 Training Examples, 18 verbs
- Learns optimal number of word senses (21)
- 32 Test examples : 78% recognition, 81% action
- All mistakes were close *lift* ~ *yank*, etc.
- Learned some particle CXN,e.g., *pull up*

Farsi

- With identical settings, learned senses not in English

COMMAND-OBEYING PERFORMANCE

- If the command is *push*:

  and the initial world state is:

  {size=small elongated=false depressible=true contact=false}

  then these linking features are set:

  {schema=depress posture=index accel=med dur=short}

  but if the initial world state is:

  {size=large elongated=false depressible=false contact=false}

  then these linking features are set:

  {schema=slide elbow=extend accel=med dir=away aspect=once}

- Overall command-obeying performance (judged by subsequent labelling of the action) is 81%.

LEARNED SENSES OF *PUSH*

push74 (12 ex) {
size {SMALL=0.785 large=0.214}
elongated {true=0.071 FALSE=0.928}
depressible {TRUE=0.928 false=0.071}
contact {true=0.071 FALSE=0.928}
schema {slide=0.004 lift=0.004 rotate=0.004
DEPRESS=0.983 touch=0.004}
posture {grasp=0.055 wrap=0.055 pinch=0.056 palm=0.055
platform=0.055 INDEX=0.722}
elbow {flex=0.333 extend=0.333 fixed=0.333}
force {low=0.466 med=0.2 high=0.333}
accel {zero=0.062 low=0.312 MED=0.5 high=0.125}
dir {away=0.166 toward=0.166 up=0.166 down=0.166
left=0.166 right=0.166}
aspect {once=0.5 iterated=0.5}
dur {SHORT=0.6 med=0.266 long=0.133}
}

push79 (8 ex) {
size {small=0.4 large=0.6}
elongated {true=0.3 FALSE=0.7}
depressible {true=0.2 FALSE=0.8}
contact {TRUE=0.7 false=0.3}
schema {slide=0.369 lift=0.006 rotate=0.006
depress=0.127 touch=0.490}
posture {grasp=0.142 wrap=0.071 pinch=0.071 PALM=0.5
platform=0.071 index=0.142}
elbow {flex=0.1 extend=0.5 fixed=0.4}
force {low=0.181 MED=0.545 high=0.272}
accel {zero=0.166 low=0.333 med=0.333 high=0.166}
dir {AWAY=0.384 toward=0.076 up=0.076 down=0.076
left=0.230 right=0.153}
aspect {ONCE=0.9 iterated=0.1}
dur {short=0.363 med=0.090 long=0.545}
}

push78 (21 ex) {
size {small=0.565 large=0.434}
elongated {TRUE=0.608 false=0.391}
depressible {true=0.043 FALSE=0.956}
contact {true=0.086 FALSE=0.913}
schema {SLIDE=0.990 lift=0.002 rotate=0.002
depress=0.002 touch=0.002}
posture {grasp=0.481 wrap=0.037 pinch=0.037 palm=0.370
platform=0.037 index=0.037}
elbow {flex=0.041 EXTEND=0.916 fixed=0.041}
force {low=0.333 med=0.416 high=0.25}
accel {zero=0.04 low=0.32 MED=0.48 high=0.16}
dir {AWAY=0.518 toward=0.037 up=0.037 down=0.037
left=0.222 right=0.148}
aspect {ONCE=0.739 iterated=0.260}
dur {short=0.333 med=0.333 long=0.333}
}

Bayes Formula

$$\Pr(A, B) = \Pr(A)\Pr(B \mid A) = \Pr(B)\Pr(A \mid B) \quad \text{- by definition}$$

Dividing both sides by Pr (A) we get:

$$\Pr(B \mid A) = \Pr(B)\Pr(A \mid B) / \Pr(A)$$

Suppose that B represents admission to Berkeley with values y, n and that A represents Academic standing with values v(very high), h (high) and m(mediocre). We can use Bayes formula to compute the probability that someone with v academics will be admitted:

$$\Pr(B = y \mid A = v)$$

Suppose 10,000 people apply and 1,000 are admitted. Suppose we also know that half the people who are admitted have very high grades, but only 20% of all applicants are v. Then the calculation becomes:

$$\Pr(B = y \mid A = v) = (.1)(.5) / (.2) = .25$$

So a student with v academics has a .25 chance of being admitted.

Learning by Model Merging

The learning problem is to find the model that has best overall fit to the training set. That is, to find word definitions that will yield the highest conditional probability for all the training data.

This isn't quite right, because we also need to account for the complexity of the model itself. A model that had a separate word sense for each example would have the best fit, but would not generalize at all. This can be fixed technically by computing a prior Pr for each model related to its size and optimizing the posterior Bayes probability of the sample data, given a class of possible models.

Model merging is a heuristic for trying to learn Bayes optimal models. We start with each example being a separate model element, here a word sense. As more data comes in, model elements (word senses) are merged. This will yield a somewhat worse fit to the data, but reduce the size of the estimated model.

We saw an informal version of Model Merging in the slide showing the incremental learning of two senses of "push". Computational Assignment 4 asks for a program that uses Model Merging to learn a simple form of grammar, again from sample inputs. This will come up when we discuss theories of natural language learning.

Bayesian Model Selection – Word Learning

Knowing a verb can be modeled as finding the best label for an experience - here acting with one hand on a table. The experience is characterized by parameterized X-schemas and frames. After training, the system should (like Regier's) pick the best word to describe each experience. This can be formalized as:

$$\underset{v}{\text{argmax}} \; Pr(v \mid l, m) = \prod_{l} Pr(v \mid m) * Pr(l \mid v, m)$$

where m = model, v = verb, l = linking feature values and
Pr(v | m) is the prior Pr (frequency) of word v in the model
Pr(l | v, m) is the likelihood of feature values l, given word v

For example, "feel" might be learned as having prior probability Pr(v=feel | m) = .1 and the feature likelihoods Pr(l | v, m)

- Schema: touch(.9)
- Direction: any(.2)
- Elbow: extend(.8), flex (.2)
- Force: low(.9)
- Object Size: medium(.5) large(.5)

An experience with the feature values l as follows

- Schema: touch
- Direction: left
- Elbow: extend
- Force: low
- Object Size: medium

would yield

$Pr(v= feel \mid l, m) = .1^5*(.9*.2*.8*.9*.5) = .00000648$

Which would almost certainly be the best.

In fact, we add logarithms instead of multiplying.

Event Structure and Metaphor

Jerome Feldman
feldman@icsi.berkeley.edu
CS182/CogSci110/Ling109
Spring 2007

General and Domain Knowledge

- Conceptual Knowledge and Inference
  - Embodied
  - Language and Domain Independent
  - Powerful General Inferences
  - Ubiquitous in Language
- Domain Specific Frames and Ontologies
  - FrameNet, OWL ontologies
- Metaphor links domain specific to general
  - E.g., France slipped into recession.

Conceptual Metaphor Provides Embodied Reasoning For Abstract Concepts

Virtually all abstract concepts (if not *all*) have conventional metaphorical conceptualizations — normal everyday ways of using concrete concepts to reason systematically about abstract concepts.

Most abstract reasoning makes use of embodied reasoning via *metaphorical mappings from concrete to abstract domains*

What Are Conceptual Metaphors?

In NTL, conceptual metaphors are structured connectionist “maps” — circuits linking concrete source domains to abstract target domains.

In the fit of NTL to Neuroscience, such metaphorical maps would be neural circuits in the brain linking sensory-motor regions to other regions.

We claim therefore that, in such cases, the sensory-motor system is **directly engaged** in abstract reasoning.

Metaphorical Grasping

There is a conceptual metaphor, *Understanding Is Grasping,* according to which one can *grasp ideas.*

One can begin to grasp an idea, but not quite get a hold of it.

If you fail to grasp an idea, it can go right by you — or over your head!

If you grasp it, you can turn it over in your mind.

You can't hold onto an idea before having grasped it.

In short, reasoning patterns about physical grasping can be mapped by conceptual metaphor onto abstract reasoning patterns.

We use metaphors everyday

- The council *attacked every weak point* of his proposal.
- I don't know how to *put my thoughts into words*.
- I've been feeling quite *depressed* of late.
- "*Washington* remains stuck in talks with Russia and France over the failure to secure a second U.N. resolution"
- My summer plans are still *up in the air*.
- I *see* what you mean.
- Something *smells fishy*, but I can't quite *put my finger on it*.

What is the basis for metaphors?

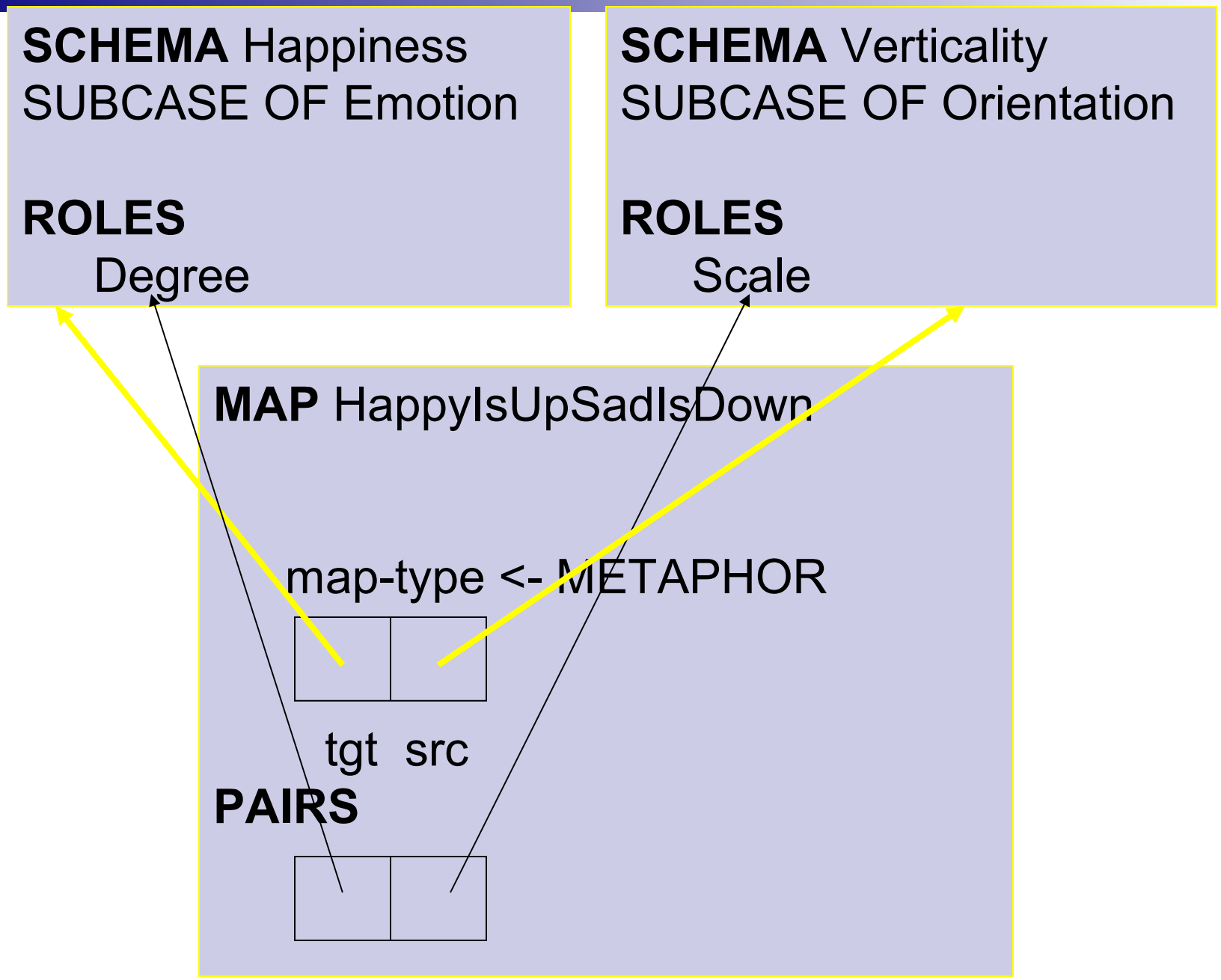
- metaphor is understanding one thing in terms of another
- specifically, we reason about abstract concepts through our sensory-motor experience.
- that means we have:
  - correlation
  - inference

Metaphors, defined

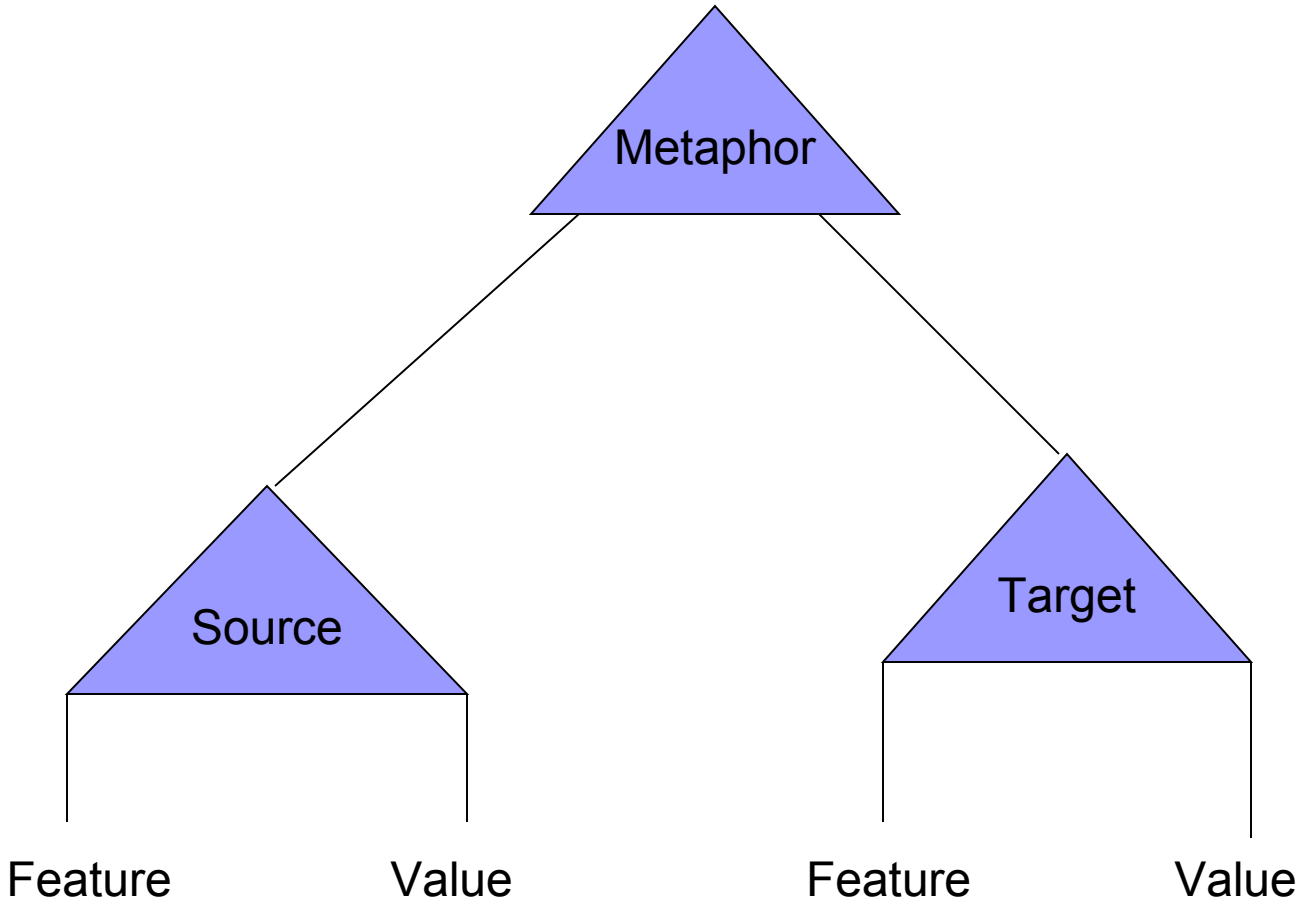
- Formally, metaphors are *mappings* from a *source domain* to a *target domain*
- both the source and target domains are structured by schemas and frames
- Take a simple example:

  I've been feeling quite *depressed* of late.

  (Happy is Up; Sad is Down)


SCHEMA Happiness
SUBCASE OF Emotion
ROLES
Degree
SCHEMA Verticality
SUBCASE OF Orientation
ROLES
Scale
MAP HappyIsUpSadIsDown
map-type <- METAPHOR
tgt src
PAIRS

Metaphors with Triangle Nodes



How are these metaphors developed?

- Conflation Hypothesis:
  Children hypothesize an early meaning for a source domain word that *conflates* meanings in both the literal and metaphorical senses
  - experiencing warmth and affection when being held as a child
  - observing a higher water level when there's more water in a cup

A few primary metaphors

- The conflation hypothesis works for metaphors that have an experiential basis, i.e. primary metaphors
  - Affection Is Warmth
  - Important is Big
  - Categories are Containers
  - Knowing is Seeing
  - Time Is Motion

Affection is Warmth

- Subjective Judgment: Affection
- Sensory-Motor Domain: Temperature
- Example: They greeted me *warmly*.
- Primary Experience: Feeling warm while being held affectionately.

Important is Big

- Subjective Judgment: Importance
- Sensory-Motor Domain: Size
- Example: Tomorrow is a *big* day.
- Primary experience: As a child, important things in your environment are often big, e.g., parents, but also large things that exert a force on you

Categories are Containers

- Subjective Judgment: Perception of Kinds
- Sensory-Motor Domain: Space
- Example: Are tomatoes *in* the fruit or vegetable category?
- Primary Experience: Things that go together tend to be in the same bounded region

Knowing is Seeing

- Subjective Judgment: Knowledge
- Sensory-Motor Domain: Vision
- Example: I *see* what you mean.
- Primary Experience: Getting information through vision

Chris Johnson's Thesis

- Predicts 3 stages of acquisition:
  - source domain word within the source domain
  - constructions that have double-meaning
  - constructions that are specific to the target domain
- e.g.

  “Can you see what’s in here?” (stage 2)

  “I see what you mean” (stage 3)

Time is Motion

- Subjective Judgment: The passage of time
- Sensory-Motor Domain: Motion
- Example: Time *flies*.
- Primary Experience: Experiencing the passage of time as one moves or observes motion

Dual Metaphors for Time

1. Time is stationary and we move thru it
   - The finals are *just around the corner*.
   - Don't look *back* on what you have done.
2. Time is a moving object
   - My spring break *went by* so quickly.
   - *Come* what may.

Another Time Metaphor

- Time AS A Resource
  - This method will **save** time.
  - Time is **money.**
  - She's **wasting** her time.
  - How **long** do we have to wait?
  - **Use** your time well.
  - He is making up for **lost** time

Complex Metaphors

- Complex metaphors combine a number of primary metaphors
- Examples:
  - Event Structure Metaphor
  - Metaphors of Ideas

Event Structure Metaphor

- Here were some sentences we saw last week:
  - Day by day, we are *moving closer* to victory.
  - US forces ready to *resume final push* into Baghdad.
  - US Economy *on the verge of falling back* into recession after *moving forward* on an *anemic* recovery.

Event Structure Metaphor

- States are Locations
- Changes are Movements
- Causes are Forces
- Causation is Forced Movement
- Actions are Self-propelled Movements
- Purposes are Destinations
- Means are Paths
- Difficulties are Impediments to Motion
- External Events are Large, Moving Objects
- Long-term, Purposeful Activities are Journeys

ESM (1)

- States are Locations
  - I'm *in* oblivion.
  - They're so *in* love.
- Changes are Movements
  - He finally *got out* of his depression.
  - She *went* crazy.

ESM (2)

- Causes are Forces
- Causation is Forced Movement
  - That incident *pushed* him over the edge.
  - I was *dragged* into this project.
  - The economy was *brought* to a halt.
- Actions are Self-Propelled Movements
  - I *went ahead* with the most obvious choice.

Entailments

- Aids to Action are Aids to Motion
  - It's all *downhill* from here.
- Manner of Action is Manner of Motion
  - We're *skipping* right along.
- Careful Action is Careful Motion
  - He is *treading on thin ice*.
- Speed of Action is Speed of Motion
  - Things have *slowed to a crawl*.

ESM (3)

- Purposes are Destinations
  - There's a *long way to go*.
  - We're *going in circles*.
- Purposeful Action is Self-Propelled Motion To a Destination
  - Don't give up just yet. We're *getting there*.
  - Work *towards* a better future.

Entailments

- Making Progress Is Forward Movement
- Amount of Progress is Distance Moved
- Undoing Progress is Backward Movement
- Expected Progress is a Travel Schedule
- Starting a Purposeful Action is Starting out on a Path
- Achieving a Purpose Is Reaching The End of the Path
- Lack of Purpose is Lack of Direction
- Lack of Progress is Lack of Movement

ESM (4)

- Means are Paths
  - However you want to *go about it* is fine with me.
  - Do it *this way*
- Difficulties are Impediments to Motion
  - He's trying to *get around* the regulations.
  - It's been a *rough* ride.

ESM (5)

- External Events are Large Moving Objects
- Special Case 1: Things
  - How're *things* going?
  - *Things* took a turn for the worse.
- Special Case 2: Fluids
  - You gotta *go with the flow*.
  - I'm just trying to *keep my head above water*.

ESM (6)

- External Events are Large, Moving Objects
- Special Case 3: Horses
  - *Keep a grip* on the situation.
  - Don't let things *get out of hand*.

ESM (7)

- Long-term, Purposeful Activities are Journeys
  - Intermediate purposes are intermediate destinations
  - Ultimate purpose is the ultimate destination
  - Progress is movement towards a destination
  - Achieving the purpose is reaching the ultimate destination

Metaphors of Ideas/Mental Entities

- Ideas are Living/Moving Entities
  - *The ideas swam furiously in the inspector's head*.
- Ideas are shared spaces/locations
  - *Shared beliefs, common ground*
- Ideas are physical/manipulable entities
  - *I turned the idea over in my head.*
  - *I pushed aside those memories.*
  - *I finally grasped the idea*.

Metaphors of Ideas

- Ideas are Possessions
  - *He inherited his ideas*
  - *I have a new idea*
- Ideas are Resources
  - *He **ran out of** ideas.*
  - *Let's **pool** our ideas.*
  - *We've **used up** all our ideas.*
  - *That's a **useless** idea*
- Ideas are external entities
  - *The idea took hold of me*
  - *I shied away from those memories*

How about other abstract concepts

- Love
- Friendship
- Justice

Metaphors for Love

love-as-patient metaphor?
love-as-physical-force metaphor?
love-as-bond metaphor?
love-as-captive-animal metaphor?
love-as-commodity metaphor?
love-as-fire metaphor?
love-as-fluid-in-container metaphor?
love-as-hidden-object metaphor?
love-as-insanity metaphor?
love-as-journey metaphor?
love-as-magic metaphor?
love-as-natural-force metaphor?
love-as-nutrient metaphor?
love-as-opponent metaphor?
love-as-rapture metaphor?
love-as-unity metaphor?
love-as-war metaphor?

Examples

- They have a **strong, healthy** marriage. (patient)
- They **gravitated** to each other immediately. (force)
- She found love in all the wrong places. (hidden object)
- He **poured out** his affections on her. (fluid in container)
- She **couldn't hold in** her love for him any longer (fluid in container)
- She was **overcome** by love. (external force)
- Love **took complete control over** him. (external force)
- She **pursued** him **relentlessly**. (war)
- He **made an ally** of her mother (war)

Metaphor and On-line Processing

- Is Metaphor purely a linguistic map or does it play a role in on-line processing?
- Some initial data
  - People don't take more time to process metaphoric senses of a word compared to literal senses.
- But is the metaphor actually accessed in processing?

Dual Metaphors for Time

1. Time is stationary and we move thru it
   - The finals are *just around the corner*.
   - Don't look *back* on what you have done.
2. Time is a moving object
   - My spring break *went by* so quickly.
   - *Come* what may.

Experiment on Time Metaphors

Metaphoric structuring: understanding time through spatial metaphor.

Lera Boroditsky

Cognition (75) 2000, 1-28

Ego Moving versus Time Moving

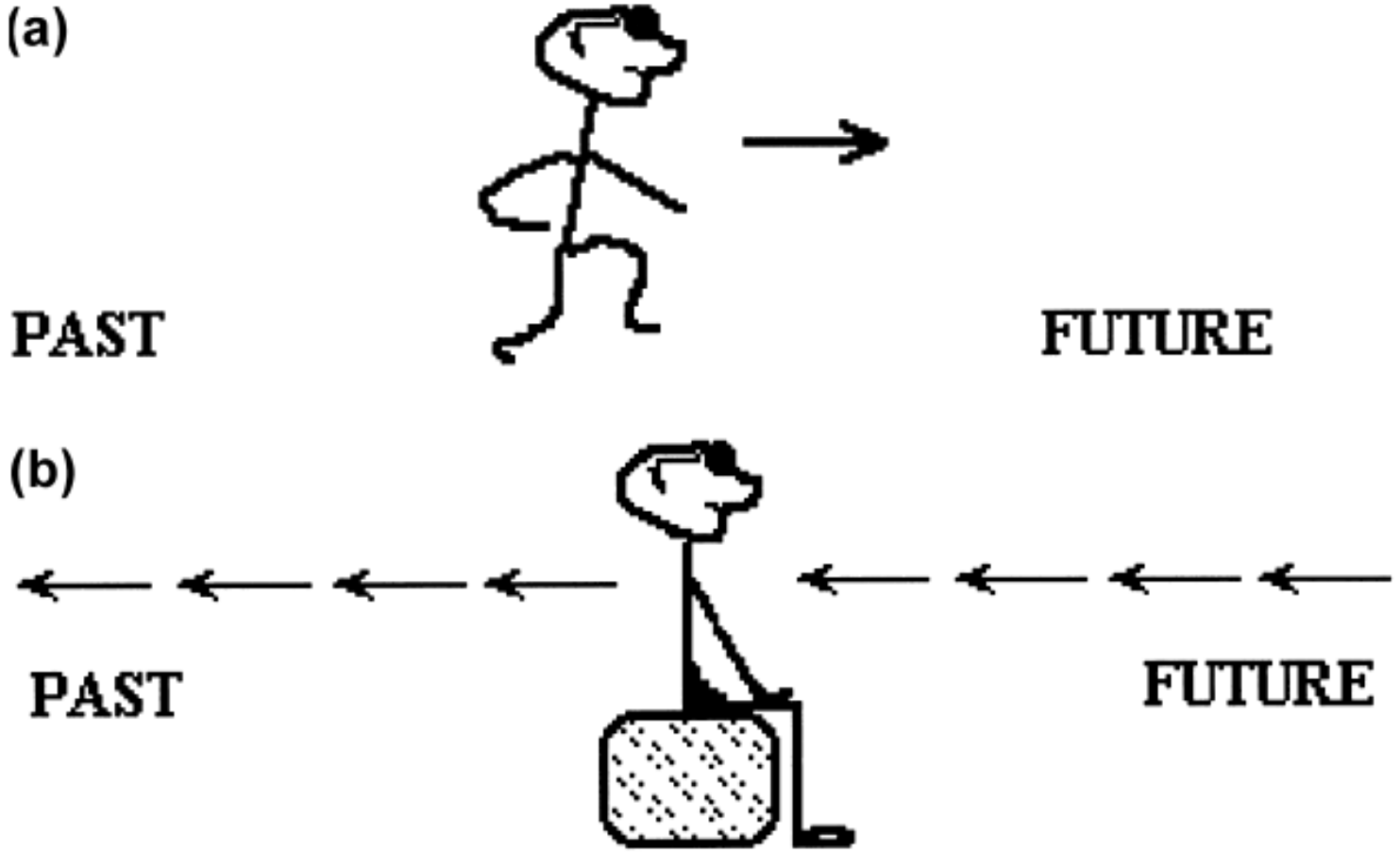
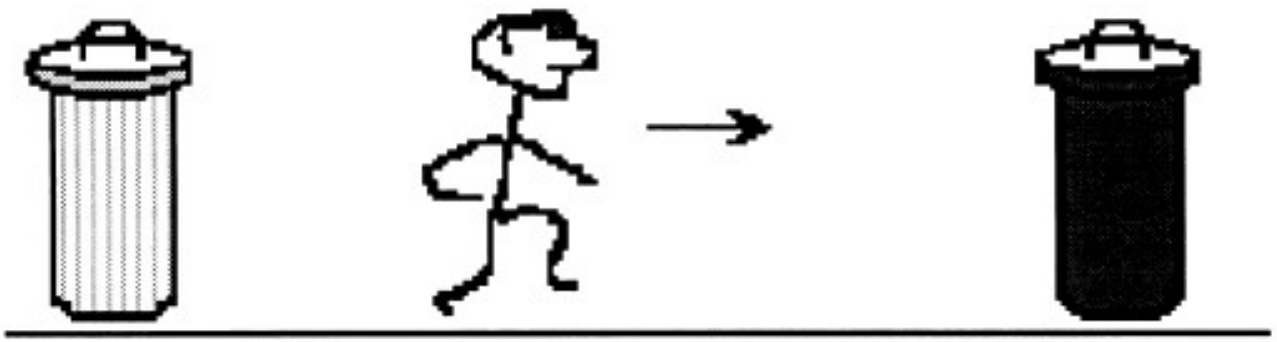


Fig. 1. (a) Schematic of the ego-moving schema used to organize events in time. (b) Schematic of the time-moving schema used to organize events in time.

Ego Moving and Object Moving Spatial Primes

(a)



The dark can is in front of me.

(b)

The light widget is in front of the dark widget.

Do people use Spatial Metaphors for time?

- First, participants answered several priming questions about spatial relations of objects in pictures.
  - These pictures used either the ego-moving or the object moving spatial schemas.
- Then, participants interpreted an ambiguous temporal statement such as `**Next Wednesday's meeting has been moved forward two days'.**
- If the above statement is interpreted using the ego-moving schema, then forward is in the direction of motion of the observer, and the meeting should now fall on a Friday.
- In the time-moving interpretation, however, forward is in the direction of motion of time, and the meeting should now be on a Monday.

Hypothesis

- If space and time do share some relational structure, then participants primed in the ego-moving spatial perspective should thus think that the meeting will be on Friday.
- Participants primed in the object-moving perspective should prefer the time-moving interpretation and think that the meeting will be on Monday.
- However, if the domains of space and time do not share any relational structure, then spatial primes should have no effect on the way participants think about time.

Results

PRIME	Meeting is Monday	Meeting is Friday
Ego Moving	26.7%	73.3%
Object Moving	69.2%	30.8%

Results: Discussion

- So far, only the effect of spatial thinking on thinking about time has been examined. But what if the experiment was reversed?
- Would making people think about time in a particular way affect how they think about space?
- If the mapping is asymmetric (as proposed by the Metaphoric View), then solving a problem about time should necessarily access and prime the appropriate way of thinking about space.

Experiment 2

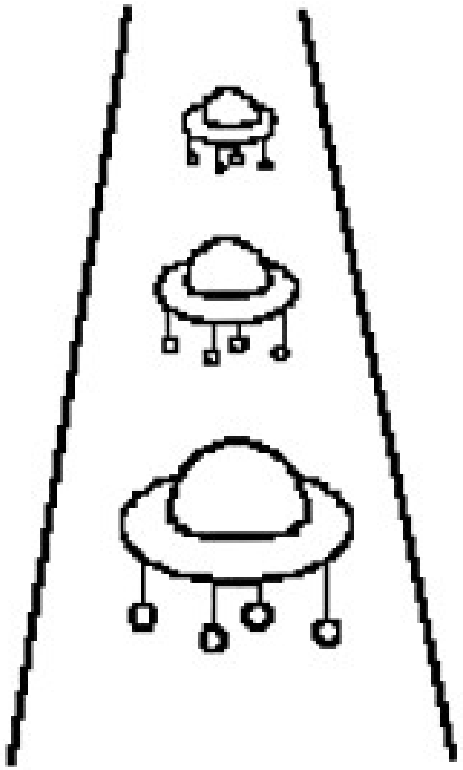
- In Experiment 2 participants answered ambiguous questions about spatial and temporal scenarios.
- Each target question followed several prime questions that used either the ego-moving schema or the object/time-moving schema.
- For some of the participants, spatial primes preceded target questions about time.
- For others, temporal primes preceded target questions about space.
- There were also two control groups for whom spatial primes preceded spatial targets, and temporal primes preceded temporal targets.

Temporal Primes and Target

- `On Thursday, Saturday is before us'), and half employed the time-moving schema (e.g. `Thursday comes before Saturday').

- Next Wednesday’s meeting has been moved forward two days. Which day is it now?

Ambiguous Spatial Target



Which one of these widgets is ahead?

(please circle one)

Fig. 4. Ambiguous spatial target used in Experiment 2.

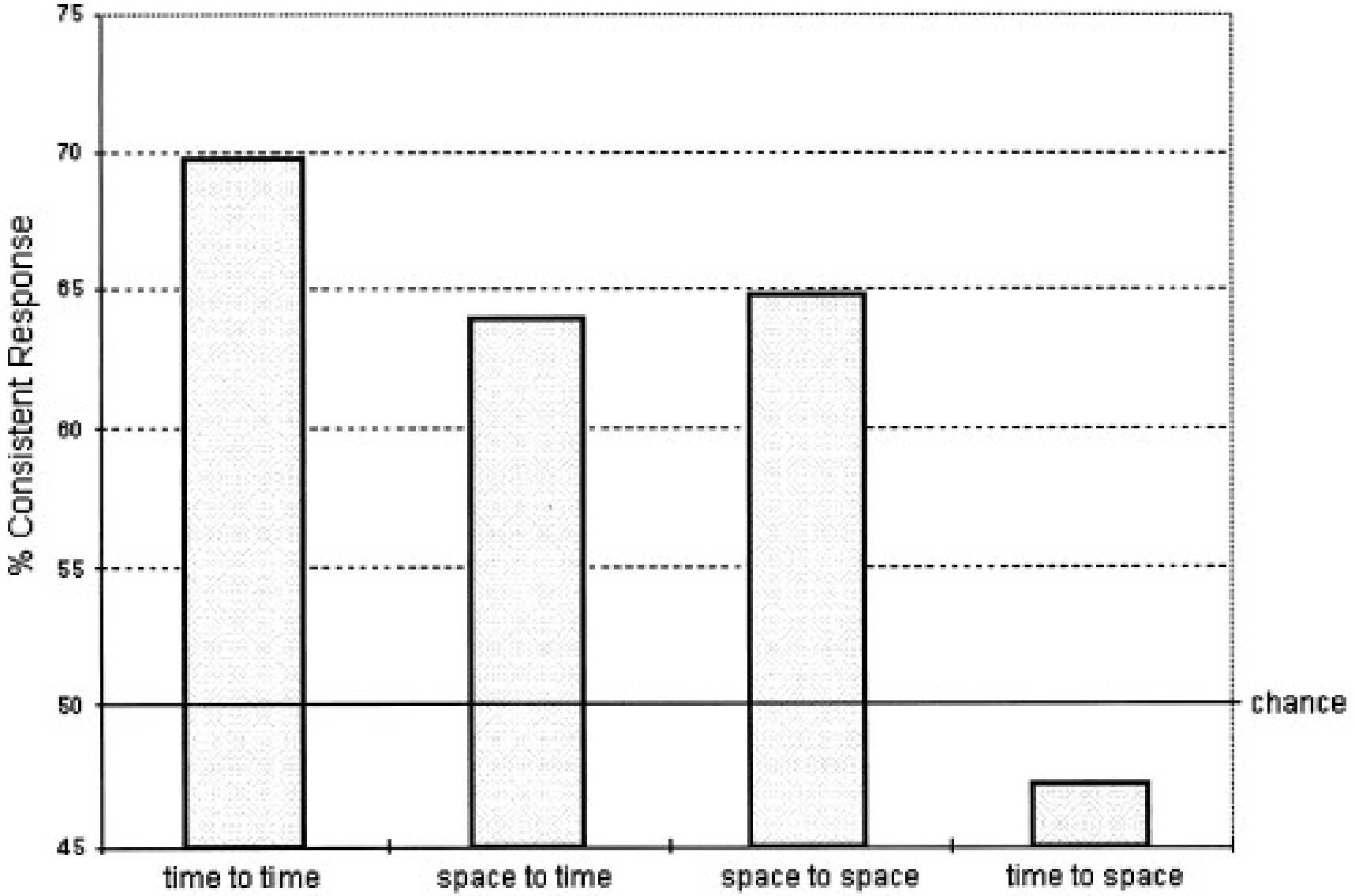


Fig. 5. Results of Experiment 2 are shown in terms of % Consistent Response (plotted on the ordinate) as a function of Transfer Type (plotted on the abscissa). There was a significant effect of consistency in all but the time-to-space condition. Chance is at 50%.

Are there language specific effects of metaphor

- Do different metaphor systems about time lead to different conceptualizations?

Time expressions in English

- In English, we predominantly use front/back terms to talk about time.
- We can talk about the good times *ahead* of us or the hardships *behind* us.
- We can move meetings *forward,* push deadlines *back*.
- On the whole, the terms used to order events are the same as those used to describe asymmetric horizontal spatial relations
  - (e.g., ‘‘he took three steps *forward*’’ or ‘‘the dumpster is *behind* the store’’).

Mandarin time expressions

- In Mandarin, front/back spatial metaphors for time are also common(Scott, 1989).
- Mandarin speakers use the spatial morphemes *qia´n* (“front”) and *ho`u* (“back”) to talk about time.
- Mandarin speakers also systematically use vertical metaphors to talk about time (Scott, 1989). The spatial morphemes *sha`ng* (“up”) and *xia`* (“down”) are frequently used to talk about the order of events, weeks, months, semesters, and more.
- Earlier events are said to be *sha`ng* or “up,” and later events are said to be *xia`* or “down.”

Question

- So, do the differences between the English and Mandarin ways of *talking* about time lead to differences in how their speakers *think* about time?
- This question can be expanded into
  - Does using spatial language to talk about time have implications for on-line processing?

Lera Boroditsky's experiment

- Mandarin and English speakers were asked to answer a spatial priming question followed by a target question about time.
- The spatial primes were either about horizontal spatial relations between two objects or about vertical relations.
- After solving a set of two primes, participants answered a TRUE/FALSE target question about time.
  - Is March *earlier* than April

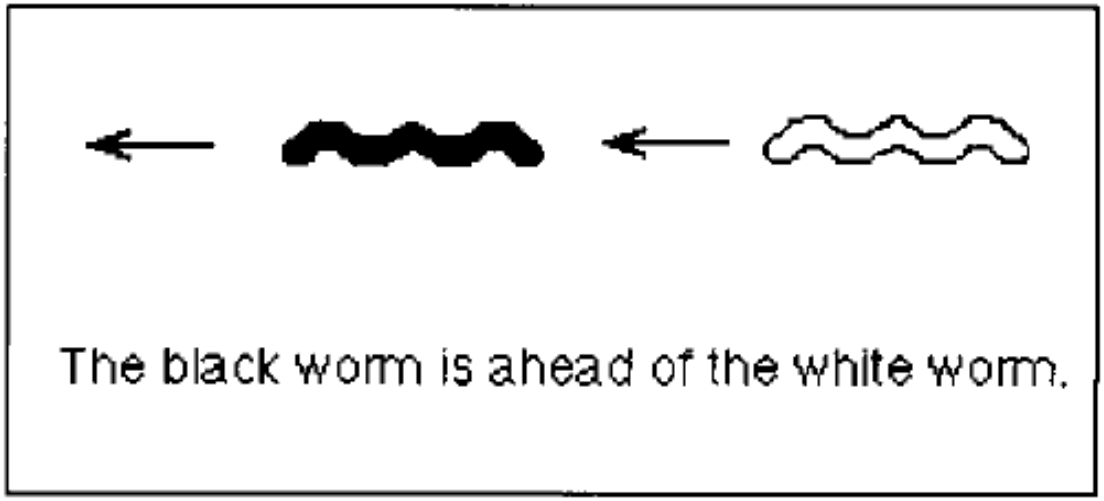


FIG. 3a. Example of a horizontal spatial prime used in Experiments 1 and 3.

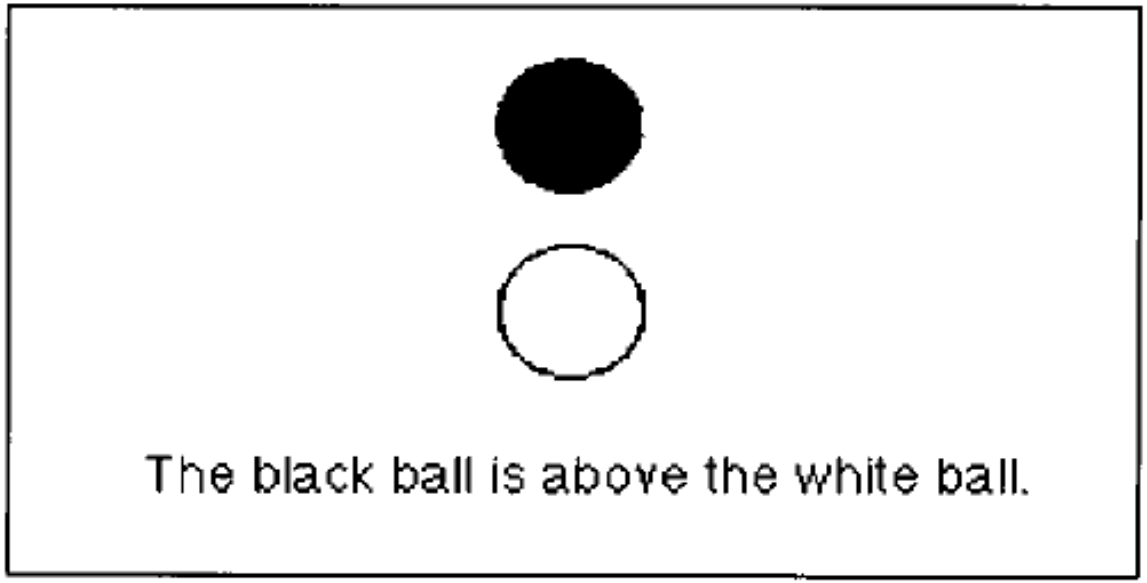


FIG. 3b. Example of a vertical spatial prime used in Experiments 1 and 3.

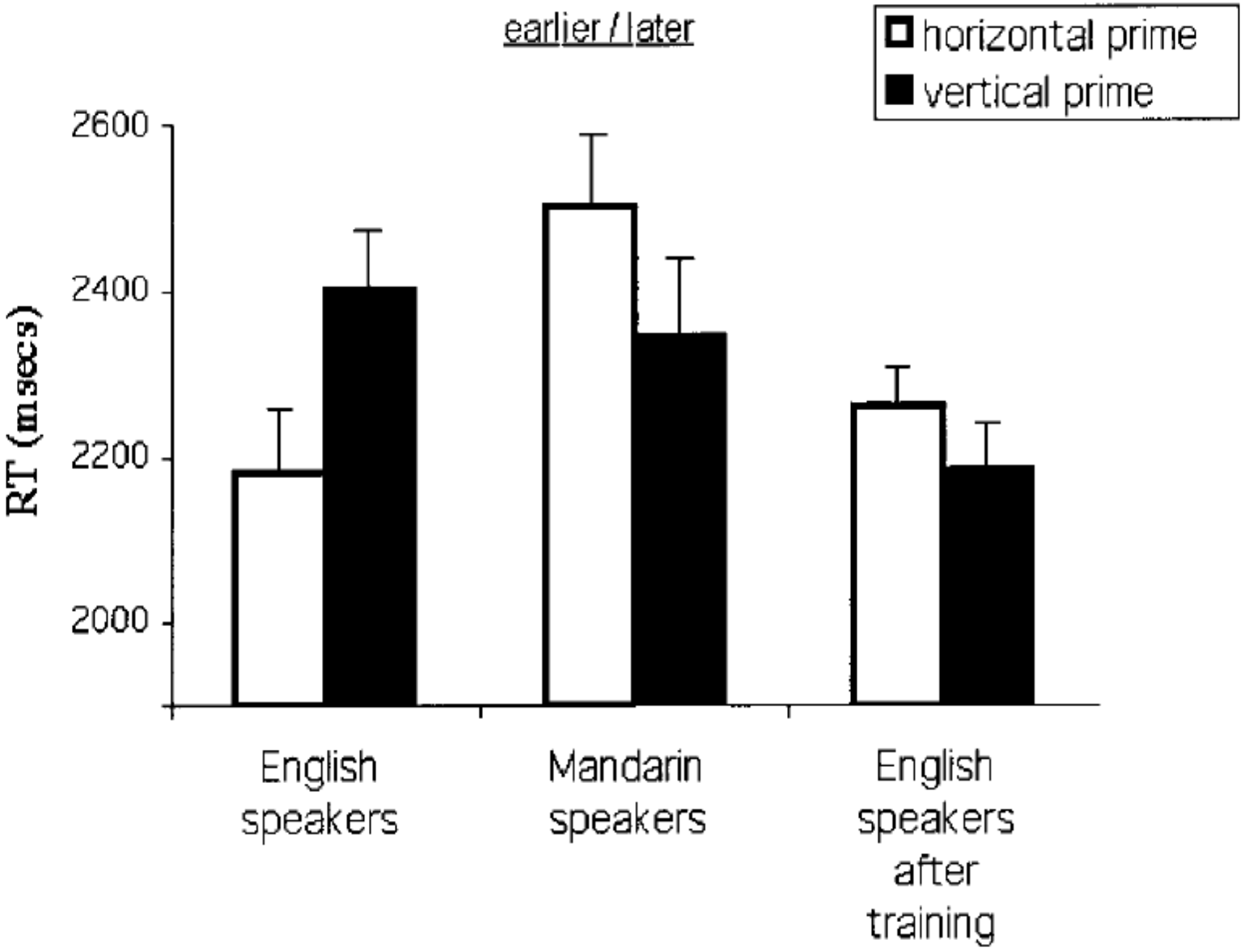


FIG. 4b. Experiments 1 and 3: Response times to purely temporal *earlier/later* questions about time following either a horizontal or a vertical prime are plotted for English speakers, Mandarin speakers, and English speakers who had been trained to talk about time vertically.

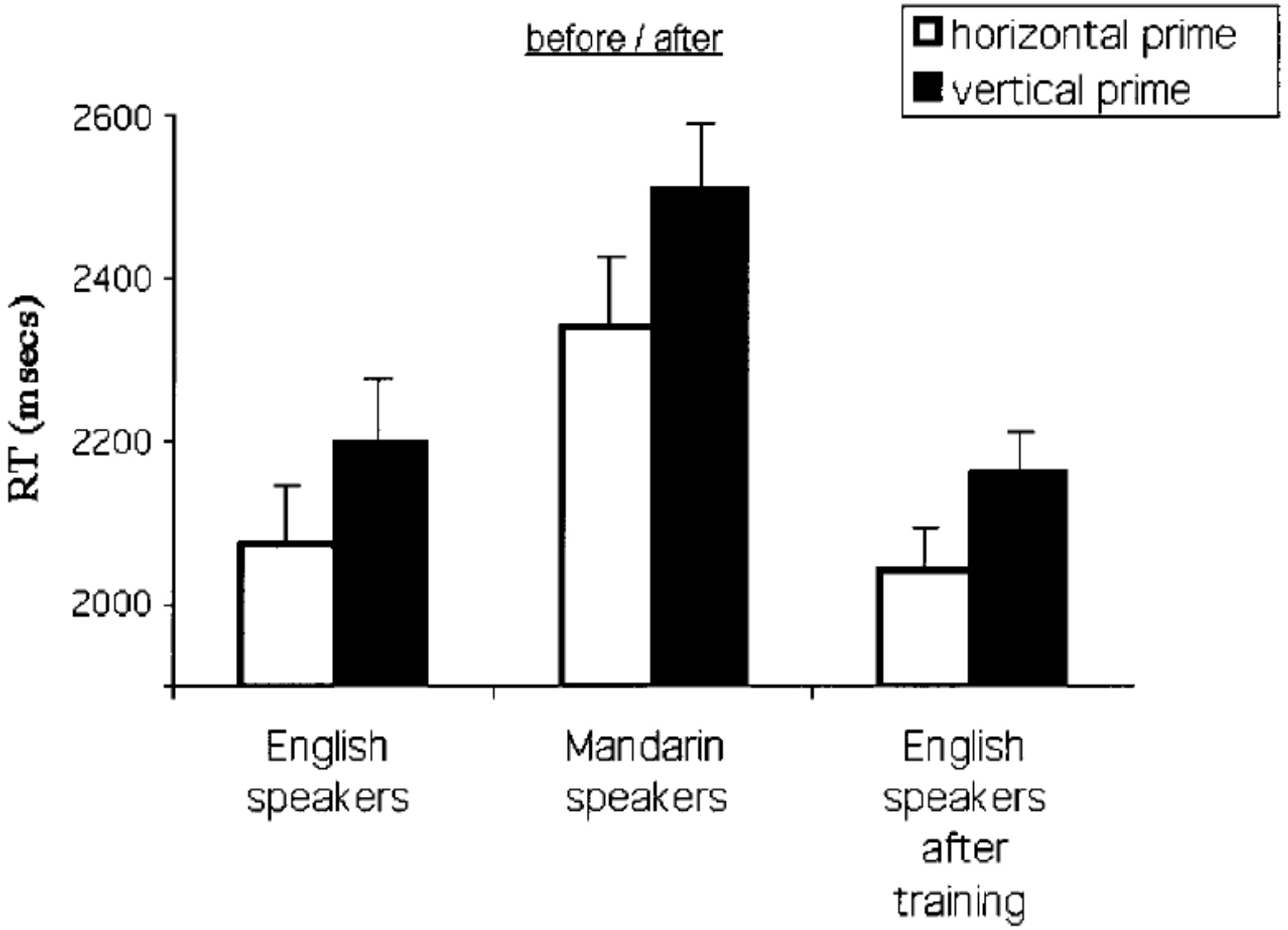


FIG. 4a. Experiments 1 and 3: Response times to spatiotemporal *before/after* questions about time following either a horizontal or a vertical prime are plotted for English speakers, Mandarin speakers, and English speakers who had been trained to talk about time vertically.

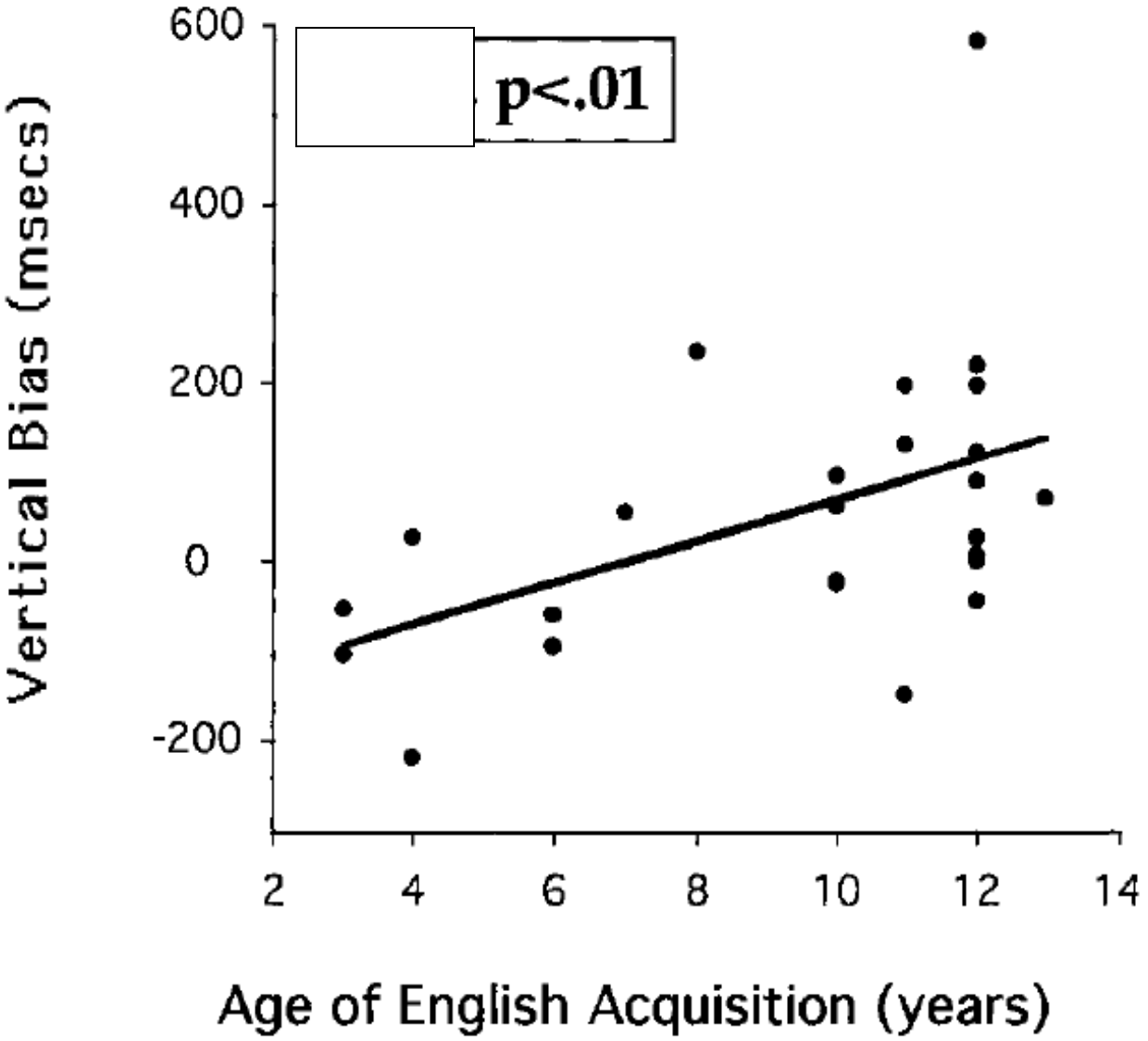


FIG. 6. Experiment 2: Results from 25 Mandarin speakers. Vertical Bias in milliseconds is plotted as a function of Age of Acquisition of English in years. Vertical Bias equals the difference in reaction time between targets following horizontal primes and targets following vertical primes.

Results discussion

- English speakers were faster to verify that ‘‘March comes *earlier* than April’’ after horizontal primes than after vertical primes. This habit of thinking about time horizontally was predicted by the preponderance of horizontal spatial metaphors used to talk about time in English.
- The reverse was true for Mandarin speakers. Mandarin speakers were faster to verify that ‘‘March comes *earlier* than April’’ after vertical primes than after horizontal primes. This habit of thinking about time vertically was predicted by the preponderance of vertical time metaphors in the Mandarin.