

Housekeeping

Semantic Memory

April 9, 2009

- Should have completed Lexical Decision already.
- Next experiment is Statistical Learning, due Thursday.
- Read Chapter 10, Ch 8.
- Last exam last in final exam period.

More Housekeeping

- Give back 1st Exp report.
- 2nd Exp report due (on paper in class) last day of class (4/28).
- Late Exp Reports up to final exam.
- Reports about experiments conducted prior to the first experiment report will be marked down.

Important stuff from last time

- Semantic priming

Overview of Today's Material

- Semantic memory: Tasks and concepts
 1. Is priming semantic?
 2. Computational models of semantic learning.

Semantic Priming

Lexical decision task

Why lexical decision?

Basic Results (table 10.3)

What does it mean that RT is faster if preceded by a related word?

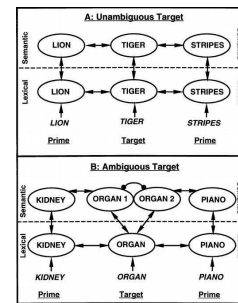
Is semantic priming strategic or automatic?

Fig 10.6

Is it really *semantic* ?

- Experimental controls
- Balota & Paul, 1996.

Balota & Paul, 1996



Balota & Paul, 1996

Table 1
Sample Word Stimuli

Condition	Word stimuli		
	Prime 1	Prime 2	Target
Category label targets (Experiment 1)			
RR	COPPER	BRONZE	METAL
UR	ORDER	BRONZE	METAL
RU	COPPER	WOOL	METAL
UU	ORDER	WOOL	METAL
Ambiguous targets (All experiments)			
RR	KIDNEY	PIANO	ORGAN
UR	WAGON	PIANO	ORGAN
RU	KIDNEY	SODA	ORGAN
UU	WAGON	SODA	ORGAN
Unambiguous targets (Experiments 2-6)			
RR	LION	STRIPES	TIGER
UR	FUEL	STRIPES	TIGER
RU	LION	SHUTTER	TIGER
UU	FUEL	SHUTTER	TIGER

Note. The one-related-prime and two-related-prime conditions were created by replacing related primes with an unrelated word in the first, second, or both prime positions. R = related; U = unrelated.

Balota & Paul, 1996

Table 8
Mean Predicted and Observed Summation Priming Effects and the Difference Between These Two Measures as a Function of Experiments 1-5 Together and Experiment 6 by Itself

Experiment and target type	n	Summation priming effects		Difference
		(RU + UR) = predicted	(RR) = observed	
Experiment 1	40	(5 + 17) = 22	34	12 ^{**}
Ambiguous		(13 + 24) = 37	34	-3 ^{**}
Unambiguous				
Experiment 2	48	(9 + 19) = 28	32	4 ^{**}
Ambiguous		(11 + 27) = 38	44	6 ^{**}
Unambiguous				
Experiment 3	64	(6 + 11) = 17	15	-2 ^{**}
Ambiguous		(8 + 11) = 19	13	-6 ^{**}
Unambiguous				
Experiment 4	48	(3 + 7) = 10	11	1 ^{**}
Ambiguous		(8 + 10) = 18	14	-4 ^{**}
Unambiguous				
Experiment 5	48	(17 + 17) = 34	36	2 ^{**}
Ambiguous		(26 + 18) = 44	46	2 ^{**}
Unambiguous				
Weighted Average	248	(8 + 14) = 22	25	3 ^{**}
Ambiguous		(13 + 17) = 30	29	-1 ^{**}
Unambiguous				
Experiment 6	32	(184 + 163) = 347	200	-147 ^{**}
Ambiguous		(141 + 122) = 263	262	-1 ^{**}
Unambiguous				

^{**}p > .25; ^{**}p < .01.

How do we learn what words mean?

How do we learn what words mean?

- High school graduates can identify the meaning of between 40k and 100k words
- This works out to about 9-15 words per day from age 2 onwards.
- Explicit instruction (vocab tests) accounts for maybe a couple 100 words a year.
- Can't be speech—the vocabulary of spoken English is much smaller

- Must be from reading—school age students (read for about an hour or two a day) may encounter as many as 50 new words a day—and manage to identify a meaning for about 10-15 of them!

How is that possible?

Dictionary?

Context?

Computational models attempt to build semantic meaning from context

Basic idea

- Have a computer take in a mass quantity of text
- Do some more or less fancy math
- At the end, have vectors (or something like that) for each word that are related to each other in the way words are related to each other.

HAL: Hyperspace analog of language (Lund & Burgess, 1996)

- Run a sliding window over a corpus of text.
- Build a co-occurrence matrix that describes how often two words appear together in the window.
- Treat the rows of the co-occurrence matrix as vectors.

HAL results

Vectors are spaced kind of the way meanings of words are. For instance, words from a category have vectors that are closer together than words that aren't in a category together.

Some computational models of semantic memory

- Hyperspace Analog of Language (HAL)
- Latent semantic analysis (LSA)
- Topics model (no acronym)

Co-occurrence in HAL

“The horse raced past the barn fell”

	barn	fell	horse	past	raced	the
PERIOD	4	5	0	2	1	3
barn	0	0	2	4	3	6
fell	5	0	1	3	2	4
horse	0	0	0	0	0	5
past	0	0	4	0	5	3
raced	0	0	5	0	0	4
the	0	0	3	5	4	2

Why does this work

HAL retains information about the contexts in which a word appears. Two words are similar if they occur together or if they occur with similar words.

LSA: Latent semantic analysis (Landauer & Dumais, 1997)

- Build a word by contexts matrix: you have now as many dimensions as contexts (tens of thousands)
- Do some fancy math to “throw away” some dimensions.
- Works best with 300 dimensions.

An aside

LSA is actively being developed into a number of commercial applications.

Some topics

FIELD	SCIENCE	BALL	JOB
MAGNETIC	STUDY	GAME	WORK
WIRE	KNOWLEDGE	TEAM	CAREER
NEEDLE	WORK	FOOTBALL	EXPERIENCE
CURRENT	RESEARCH	BASEBALL	EMPLOYMENT
COIL	CHEMISTRY	PLAYERS	OPPORTUNITIES
POLES	TECHNOLOGY	FIELD	WORKING
IRON	MANY	BASKETBALL	TRAINING
COMPASS	MATHEMATICS	COACH	SKILLS
LINES	BIOLOGY	HIT	POSITIONS
CORE	FIELD	TENNIS	FIELD

Conclusions from the computational models

- Word order comes in, word “meaning” comes out.
- “Meaning” in some sense is just placing a word in relation to all the others.
- We can learn *a lot* just from context (think of all the things these models *can't* do)

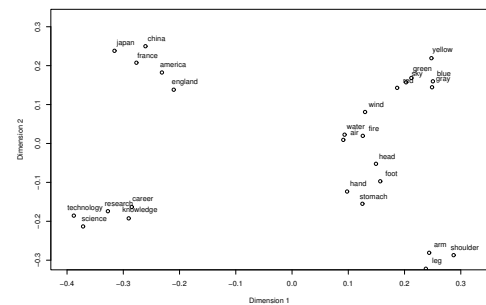
LSA: Let's play around a little

<http://lsa.colorado.edu/>

The topics model

- The dimensions in HAL are just words.
- The dimensions in LSA are not interpretable.
- Instead of a spatial model, constructs a set of topics.

A learning model of semantic memory



Assignment

- Read Chapter 10 if you haven't already.
- Read Chapter 8.
- Statistical learning next experiment.