The Adaptive Nature of Memory

By Qiong Zhang for Class 85-426, CMU Spring 2018

What are the problems we face in memory and their probabilistic nature

- [Last class] Given blurred memory trace, how to reconstruct original memory as accurately as possible? e.g. Can you draw the apple logo?

We can infer about it based on some of our experiences.

- [This class] How does our brain decide whether to retain or forget a piece of memory? e.g. Why you remember better what happened yesterday than a day last week?

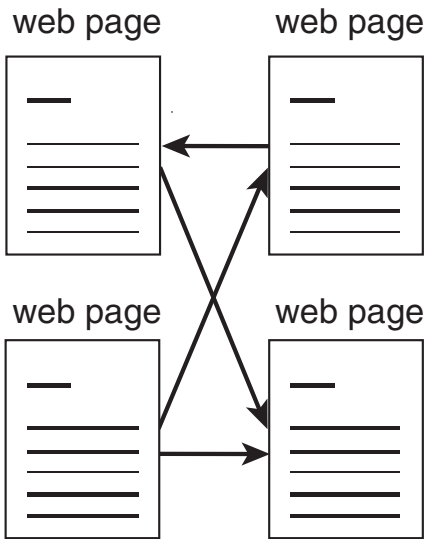
It depends on the probability that we'll need it again.

Human Memory is similar to library collections

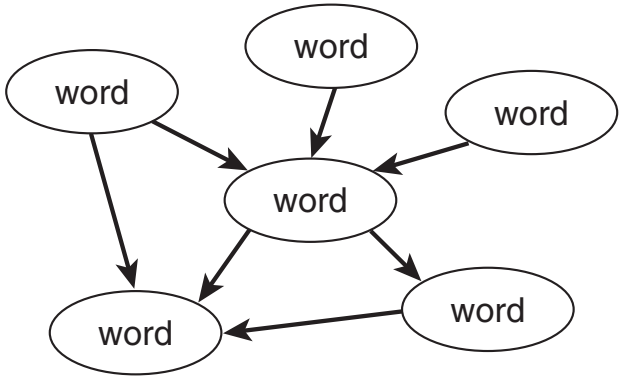
- How to store information/books
- When to discard information/books when it is no longer needed
- How to effectively retrieve information/books when there is demand

Information retrieval algorithm can be applied to model human memory

(a) World wide web



(b) Semantic network



Task: find a page given a query word

Task: find a word given an initial letter

- PageRank is developed by Google research to efficiently do this task
- We are more likely to need memories that are more central in such networks.

(Griffiths, Steyvers, & Firl, 2007)

PageRank models human responses better than conventional predictors

Table 15.1. Most frequent responses in the fluency task for the letter 'd' and the rankings given by PageRank, In-degree and KF frequency.

Human responses		**PageRank**		**In-degree**		**KF Frequency**	
DOG	(19)	DOG	(19)	DOG	(19)	DO	(2)
DAD	(16)	DARK	(3)	DEATH	(1)	DOWN	(4)
DOOR	(5)	DRINK	(1)	DRINK	(1)	DAY	(2)
DOWN	(4)	DOWN	(4)	DIRTY	(0)	DEVELOPMENT	(0)
DARK	(3)	DEATH	(1)	DARK	(3)	DONE	(1)
DUMB	(3)	DOOR	(5)	DOWN	(4)	DIFFERENT	(0)
DAY	(2)	DAY	(2)	DIRT	(0)	DOOR	(5)
DEVIL	(2)	DIRTY	(0)	DEAD	(0)	DEATH	(1)
DINOSAUR	(2)	DIRTY	(0)	DANCE	(0)	DEPARTMENT	(0)
DO	(2)	DEAD	(0)	DANGER	(1)	DARK	(3)

Note: The numbers between parentheses are frequencies in human responses. All responses are restricted to the words in the word association norms by Nelson *et al*. (1998).

Rational analysis: a rational explanation tells why the mind does what it does (Anderson et al., 1989, 1990)

- Instead of proposing directly potential mechanisms humans use to solve a problem ("What")
- Human cognition/memory <u>optimally</u> solves the problems that it faces ("Why")

Mar's three levels:

- Computational level
- Algorithm level
- Implementation level

Rational analysis: a rational explanation tells why the mind does what it does (Anderson et al., 1989, 1990)

- Instead of proposing directly potential mechanisms humans use to solve a problem
- Human cognition/memory <u>optimally</u> solves the problems that it faces

Mar's three levels:

- Computational level
- Algorithm level
- Implementation level

Rational analysis: a rational explanation tells why the mind does what it does (Anderson et al., 1989, 1990)

- Instead of proposing directly potential mechanisms humans use to solve a problem
- Then, human cognition/memory <u>optimally</u> solves the problems that it faces

Under constraints S,
stop retrieving a memory if :

PG<C under constraints

P: probability of memory relevance
G: gain if goal is achieved
C: cost in retrieving the memory

Under constraints S,
stop looking for a book if :

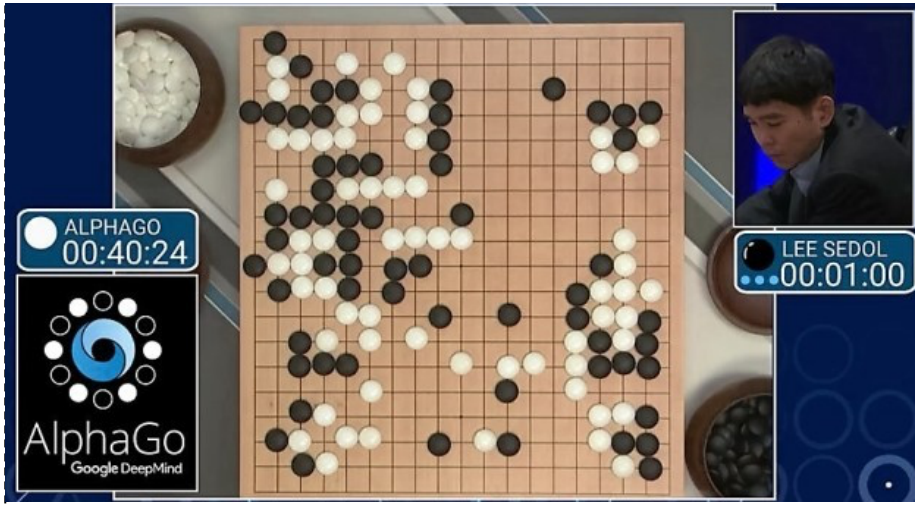
PG<C

P: probability of book relevance
G: score in the exam next day
C: time spent in looking for it

Under constraints S,
stop retrieving a memory if :

$$PG<C$$

P: probability of memory relevance
G: gain if goal is achieved
C: cost in retrieving the memory



Why constraints S is important

*Why AlphaGo does not behave like human players even with similar gain and cost specified in the game?

Under constraints S,
stop retrieving a memory if :

$$PG<C$$

P: probability of memory relevance
G: gain if goal is achieved
C: cost in retrieving the memory

When C is changed in PG<C

A memory task the strategy of which strongly depends on changes in the cost

- Fluency task:

Can you list as many animals as you can?

G: performance measure (i.e. total items)

C: time spent in recalling each item

P: every eligible answer is relevant

- Fluency task:

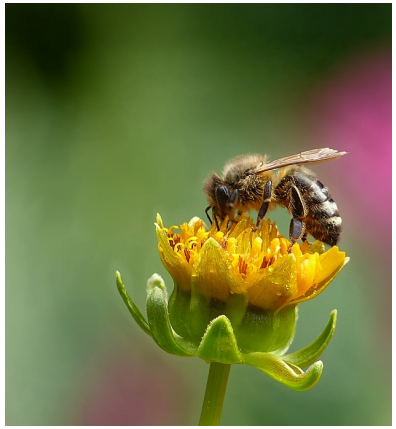
Can you list as many animals as you can?

Tendency to give “patches” of answers

Cat, dog, hamster, giraffe, lion, zebra, rhino, ostrich, whale, shark, guppy, goldfish …

The question is when to switch the patch and if human memory does it in an optimal way?

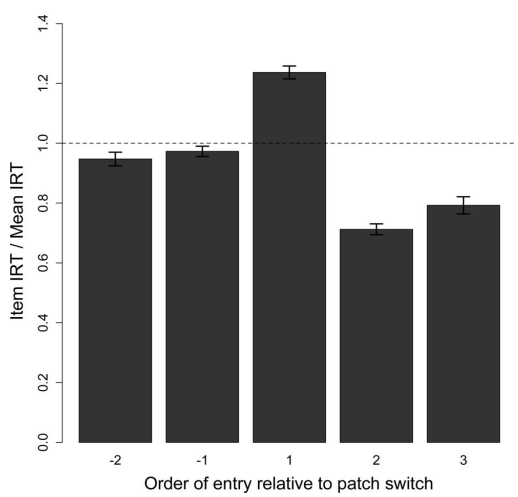
Patch switching in human memory search is similar to the problem of animal foraging



http://vestalgrove.blogspot.com/2017/08/somme-woods-east-progress-report.html

Patching switching pattern in memory is similar to optimal foraging theory

- Over evolution, animals learn to forage by switching at a optimal point to maximize overall average reward
- In optimal foraging, <u>animals switch patches at the point when the current cost is larger the average cost</u>



(Hills, Jones, & Todds, 2012)

- Similarly, in human memory search, <u>patch-switching takes place when the current retrieval time is longer than the average retrieval time</u>

PG<C

Under constraints S,
stop retrieving a memory if :

$$PG<C$$

P: probability of memory relevance
G: gain if goal is achieved
C: cost in retrieving the memory

When P is changed in PG<C

Retrieve previously un-retrievable information after a shift in memory relevance

Stop retrieving a memory if :

PG<C

P: probability of memory relevance

G: gain if goal is achieved

C: cost in retrieving the memory

Materials. The experimental passage was a narrative about what two boys did at one of the boys' homes while they were skipping school. It contained a number of points of interest to a burglar or a real estate prospect. The story was 373 words long and contained 72 idea units which previously had been rated for their relative importance to a burglar and to a prospective homebuyer.

	Information cluster			
	Burglar		Homebuyer	
First/second perspective	First recall	Second recall	First recall	Second recall
Homebuyer/burglar	.51	.61	.59	.48
Burglar/homebuyer	.68	.36	.40	.50

(Anderson & Pichert, 1978)

Rational analysis – need probability

H is the hypothesis that a piece of memory is need
E is evidence
1st term is the history factor: frequency and recency
2nd term is the context factor: how strong the cue j in E is associated with H

$$\frac{P(H|E)}{P(\overline{H})|E)} = \frac{P(H)}{P(\overline{H})} \prod_{j \in E} \frac{P(j|H)}{P(j|\overline{H})}$$

- It holds for both human memory and environmental statistics
- Rational analysis: human memory respond to the demands in the environment

What is a cue j

- Study phase:

Apple-Basket; Flower-Glass; Teacher-Classroom

- Test phase:

Apple _; Classroom _; Glass _.

A few terms to clarify

$$Odds = \frac{\Pr ob}{1 - \Pr ob}$$

$$\Pr ob = \frac{Odds}{1 + Odds}$$

Prob ranges from 0-1, Odds from 0 to infinity

The effect of frequency/practice

The effect of recency/decay/retention

The History Factor 1: recency

$$\frac{P(H|E)}{P(\bar{H})|E)} = \frac{P(H)}{P(\bar{H})} \prod_{j \in E} \frac{P(j|H)}{P(j|\bar{H})}$$

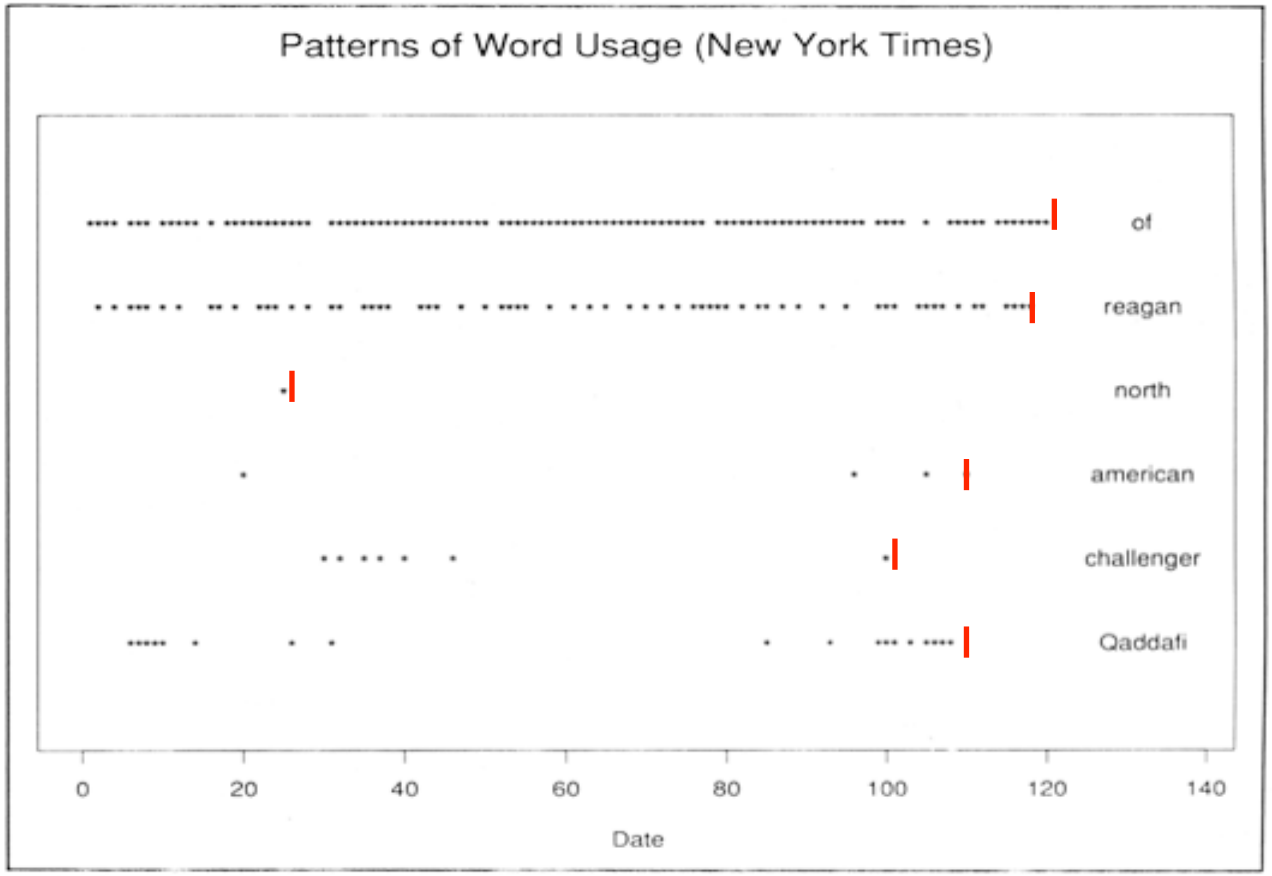
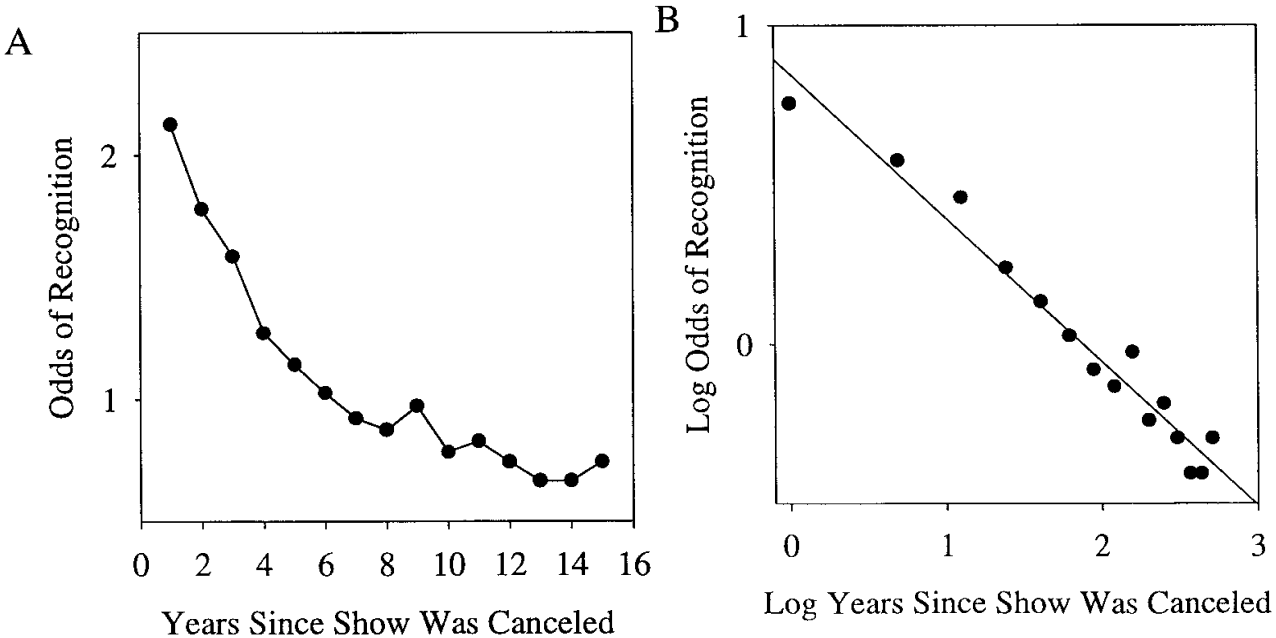


Fig. 5. Patterns of usage of various words in the *New York Times* data base over a 100-day period.

The History Factor 1: recency

$$\frac{P(H|E)}{P(\overline{H})|E)} = \frac{P(H)}{P(\overline{H})} \prod_{j \in E} \frac{P(j|H)}{P(j|\overline{H})}$$

- Recency influences need probability. e.g. your word doc lists most recent files on top.
- Experimentally in human memory, it is a power function.



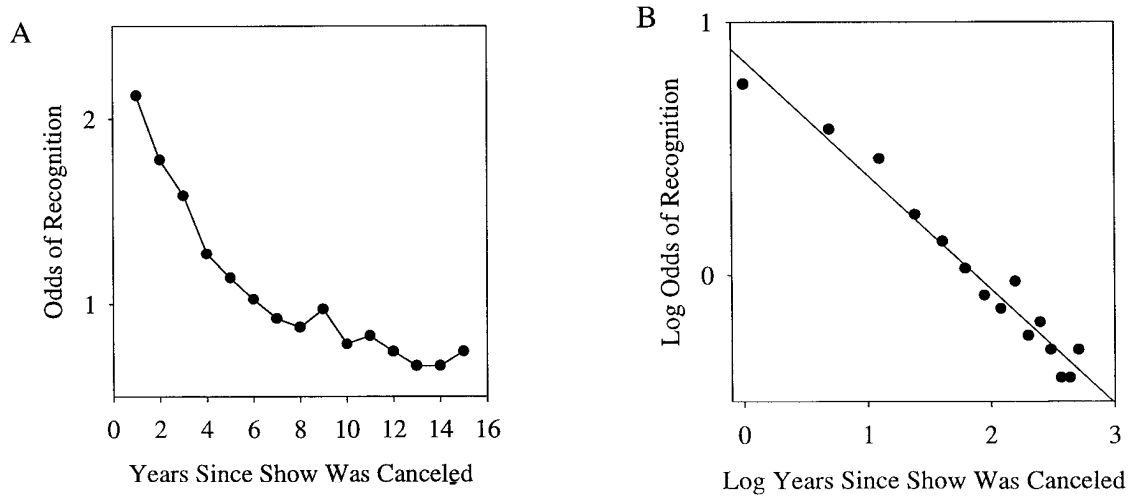
Power function:
$f(x) = a\,x^b$

Take log transform:
$\log(f(x)) = \log(ax^b)$
$= \log(a) + b*\log(x)$

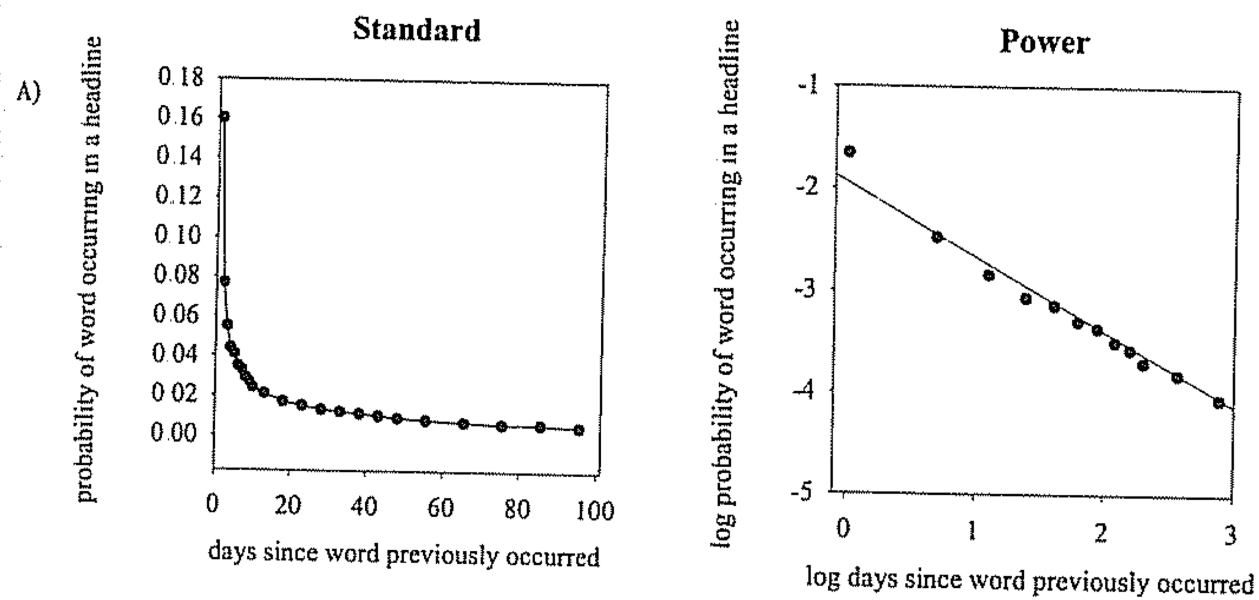
The History Factor 1: recency

- Similar patterns are observed in the environment.

Human Memory

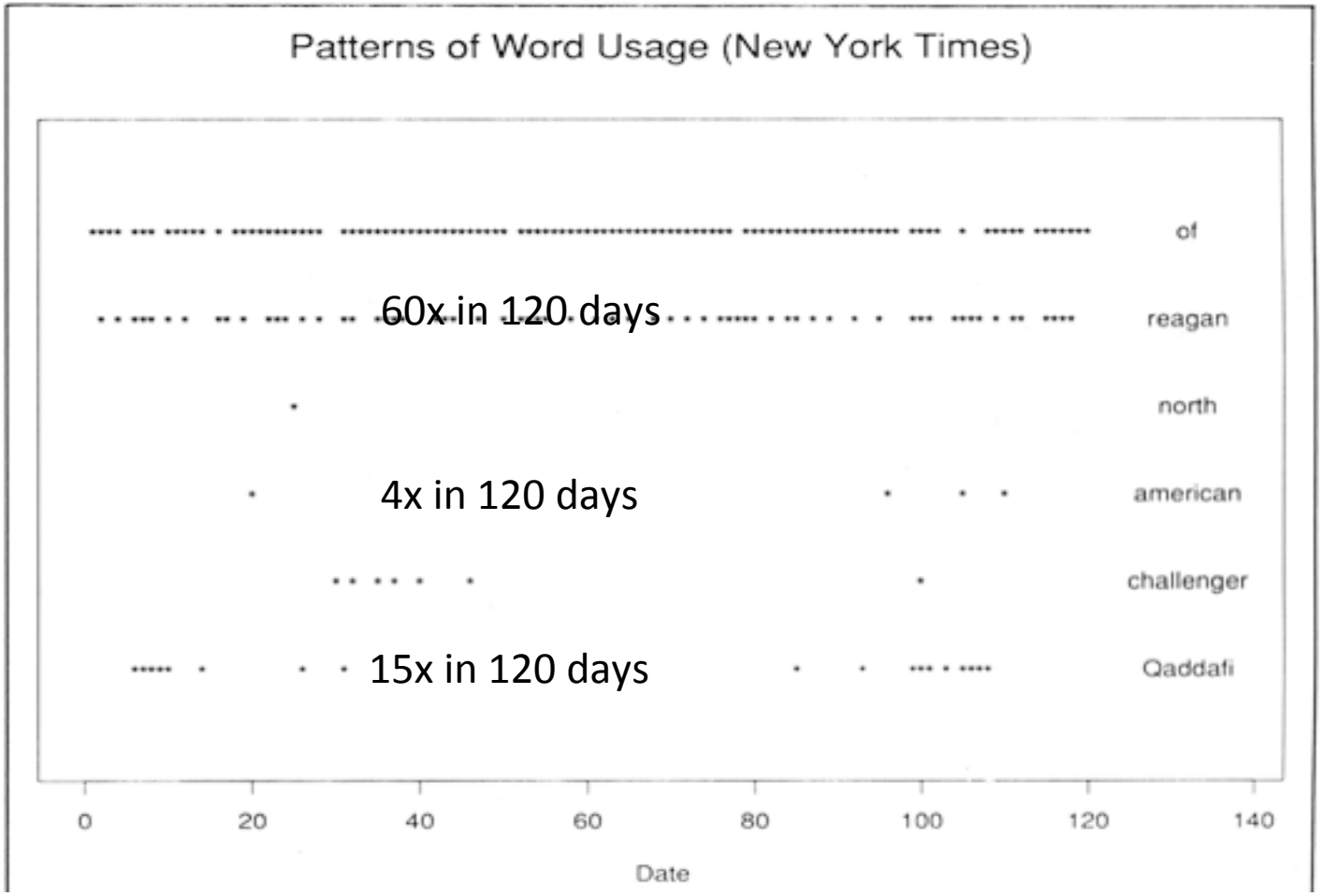


Environment



The History Factor 2: frequency

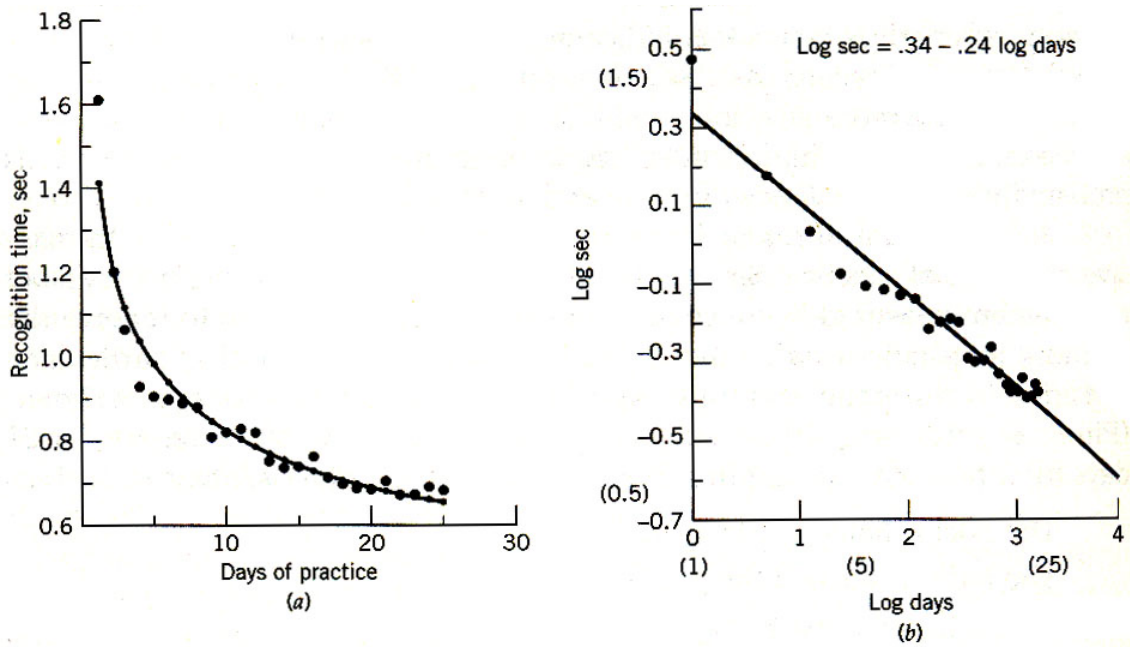
$$\frac{P(H|E)}{P(\overline{H})|E)} = \frac{P(H)}{P(\overline{H})} \prod_{j \in E} \frac{P(j|H)}{P(j|\overline{H})}$$



The History Factor 2: frequency

$$\frac{P(H|E)}{P(\overline{H})|E)} = \frac{P(H)}{P(\overline{H})} \prod_{j \in E} \frac{P(j|H)}{P(j|\overline{H})}$$

- Frequency influences need probability. e.g. your browser suggests you the most visited websites.
- Experimentally in human memory (the practice effect), it is also a power function.
- Note that the y-axis is latency instead of need probability

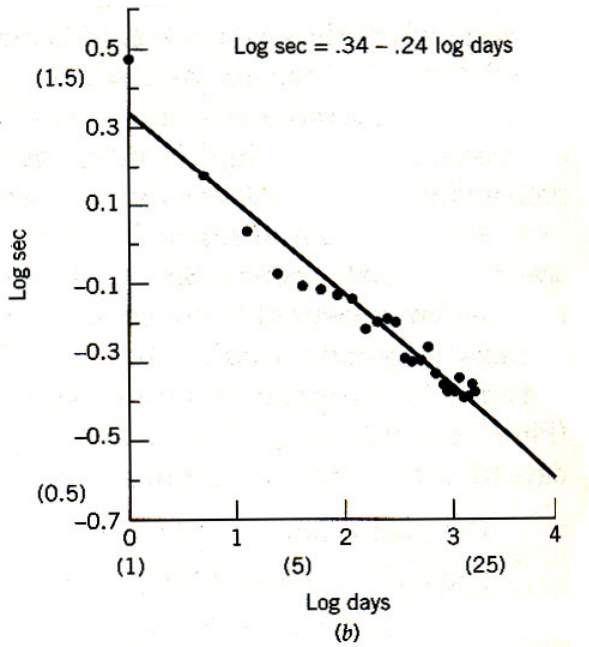


Pirolli, P. L., & Anderson, J. R. (1985). The role of practice in fact retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 136–153.

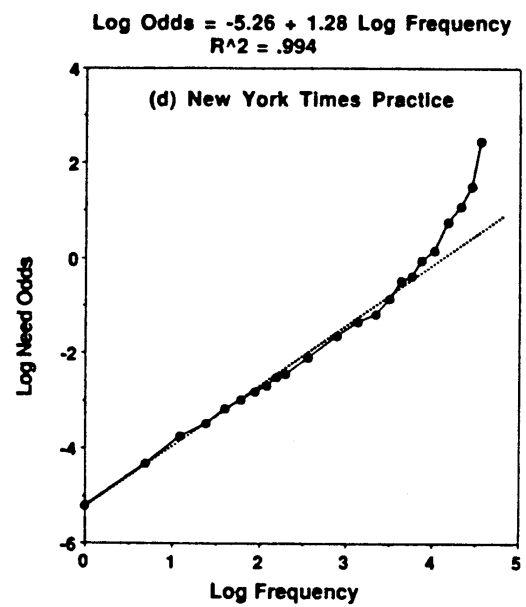
The History Factor 2: frequency

- Similar patterns of frequency are observed in the environment.

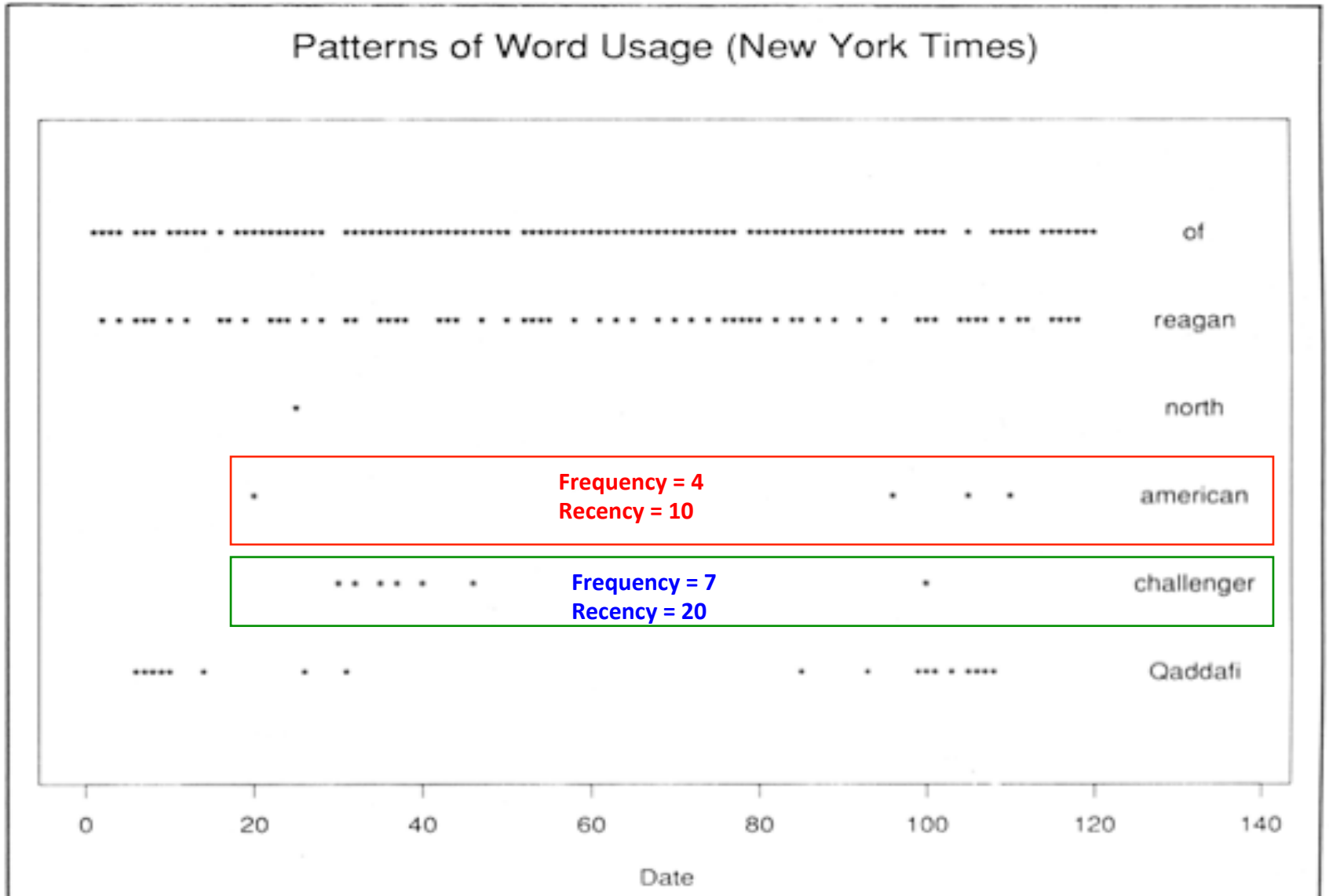
Human Memory



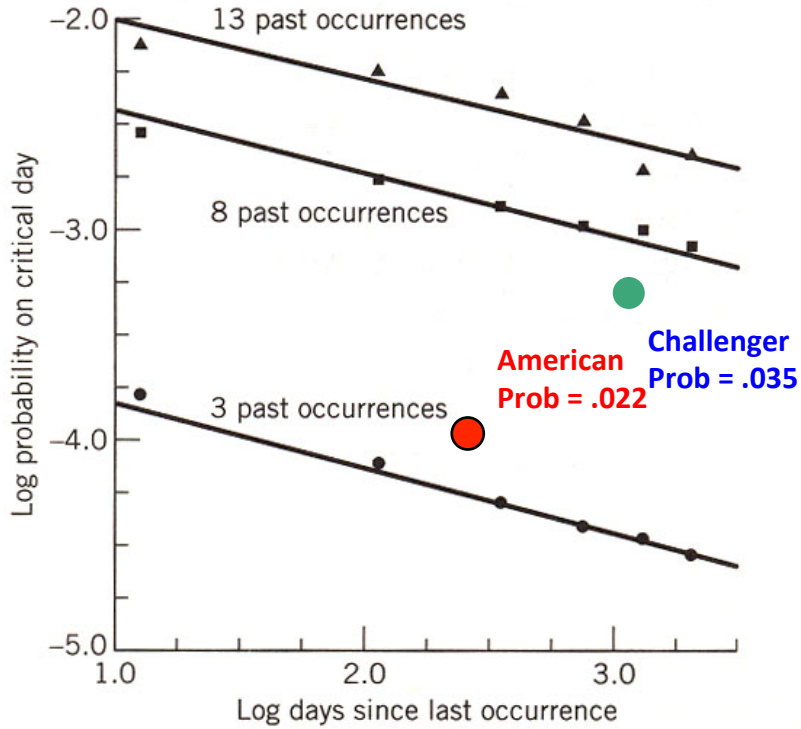
Environment



Combine recency and frequency



Recency and frequency add up in the history factor



Ln Odds = A + b*Ln(Frequency)-c*Ln(Recency)

A ~ -4.8 b ~ 1.2 c ~ 0.3

The Context Factor

$$\frac{P(H|E)}{P(\bar{H})|E)} = \frac{P(H)}{P(\bar{H})} \prod_{j \in E} \frac{P(j|H)}{P(j|\bar{H})}$$

Table 34.1 Associative rates in the *New York Times* database.

	$p(\text{AIDS}) = .018$	
Associates	$p(\text{AIDS}/\text{associate})$	$\frac{p(\text{AIDS}/\text{associate})}{p(\text{AIDS})}$
virus	.75	41.0
spread	.54	29.4
patients	.40	21.8
health	.27	14.6

What is a cue j

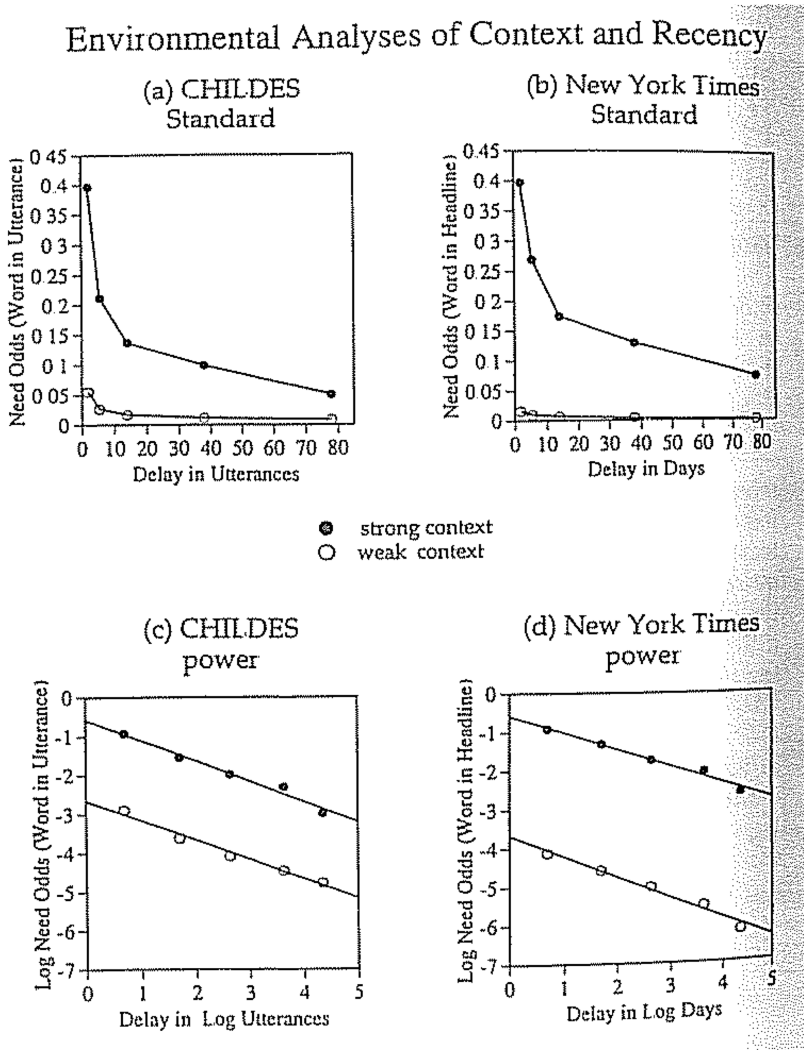
- Study phase:

Apple-Basket; Flower-Glass; Teacher-Classroom

- Test phase:

Apple _; Classroom _; Glass _.

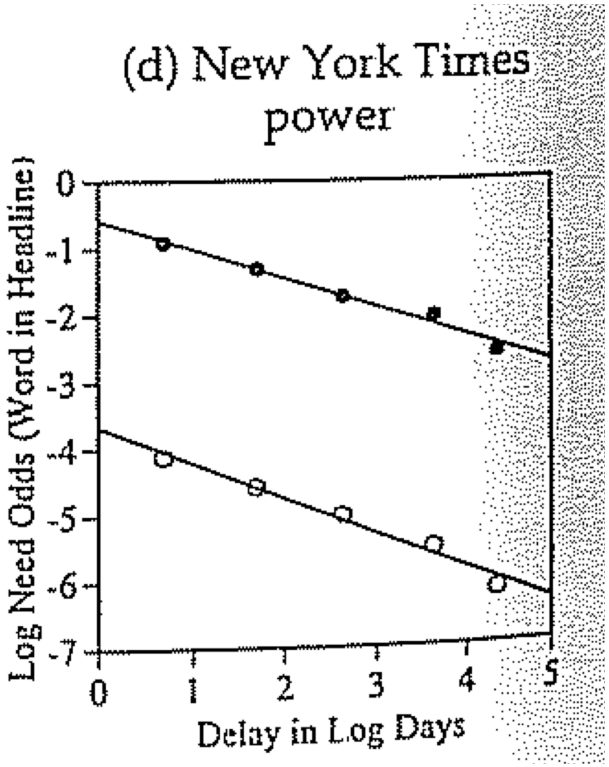
Context factor adds up with history factor



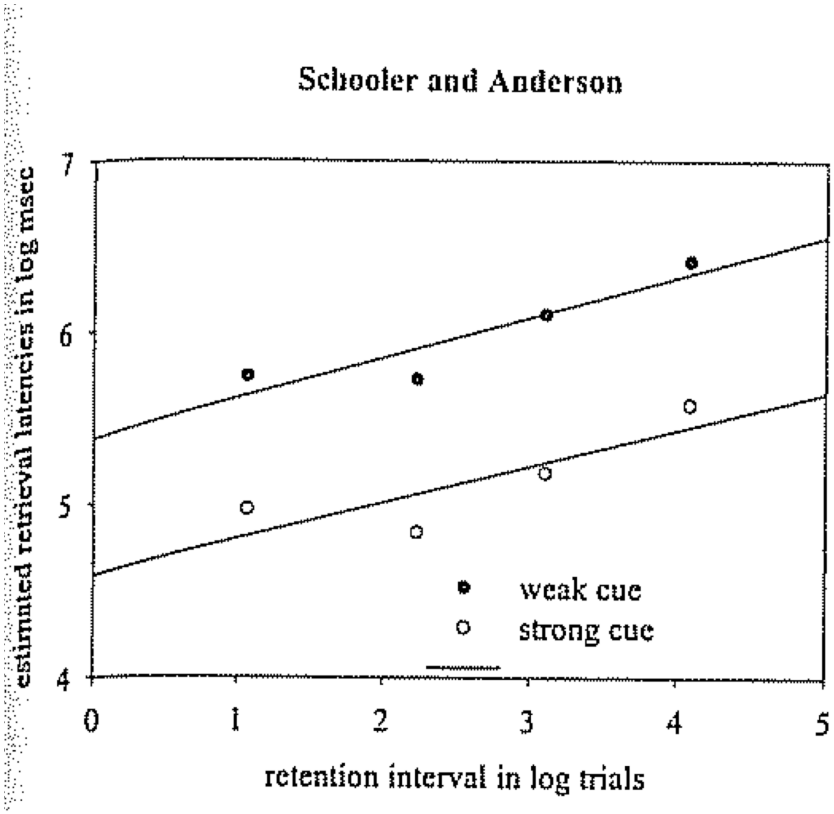
$$\log\left(\frac{P(H|E)}{P(\overline{H}|E)}\right) = \log\left(\frac{P(H)}{P(\overline{H})}\right) + \sum_{j \in E} \log\left(\frac{P(j|H)}{P(j|\overline{H})}\right) \quad (4)$$

Similar results in human memory predicted by the rational analysis

Environment



Human Memory

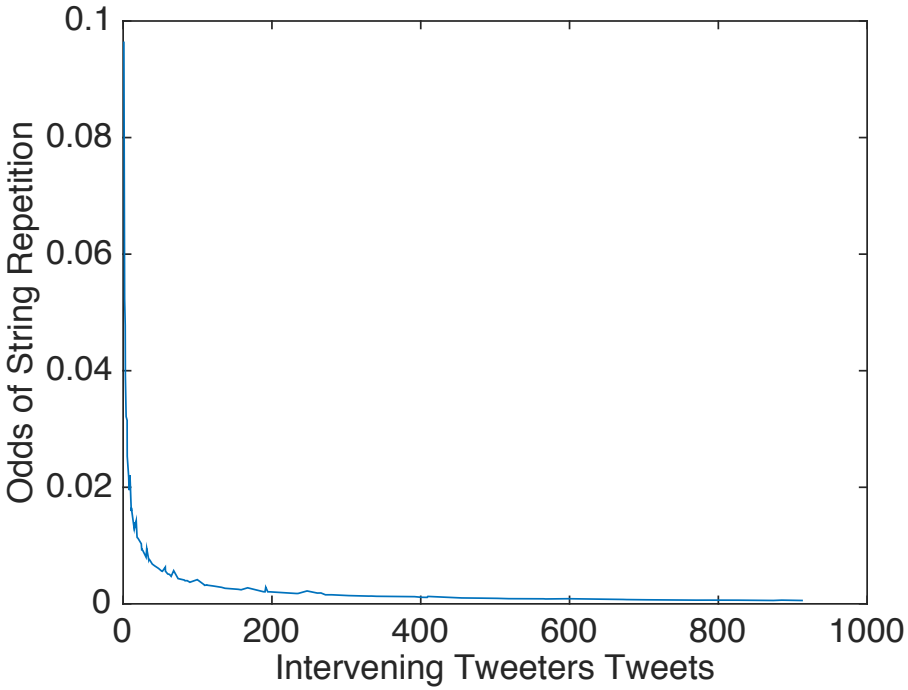


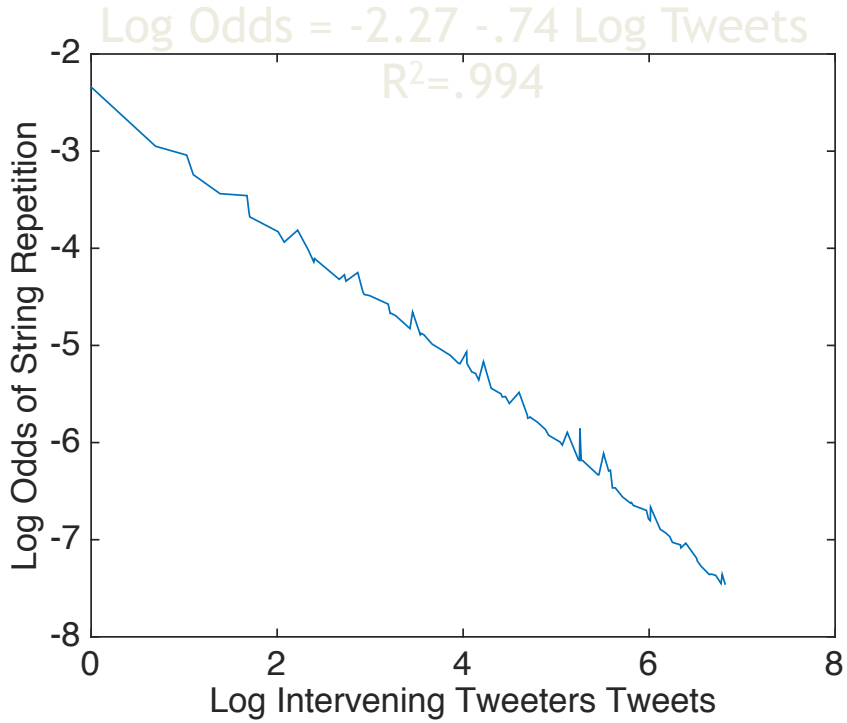
Does the statistical pattern in the environment give rise to pattern in memory?

OR

Does the pattern in human memory create the statistics in the environment?

Recency effect in Tweets from Top 500 Tweeters
July 11, 2007 – Jan 7, 2014





- Tweets are generated by human
- But human memory is influenced by environment when deciding on a tweet