• Today: Item-based knowledge
• Wed: (language internal) generalizations and constraints on generalizations
• Mon July 20: Constructions
• Wed July 22: Accounting for cross-linguistic tendencies

Prefabs:

• The following phrases are clearly stored in our “mental lexicon”:
  – Kick the bucket
  – Hang up the phone

The grass is always greener
Persian: Monghe hamsayeh ghaz ast
“the neighbor’s hen is a goose!”

He struts like a peacock.
Persian: Mesle inke az kune fil ohade
“as if he fell of the ass of an elephant”

Do you believe me?
Persian: Harfe u m baveer mikanld?
“do you believe his words?”

How are you?
Persian: Hale shoma shetone?
“How is your health?”

Memory for details
Tens of thousands of words, idioms and compositional “prefabs” are learned (Pawley and Syder 1983; Jackendoff 2002; Dabrowska 2004)

Language acquisition (e.g., Akhtar and Tomasello 1997; Baker 1979; Bates and MacWhinney 1987; Bowerman 1982; Brauner 1976; Grogen et al. 1989; Ingram and Thompson 1996; Lieven et al. 1997; Tomasello 2000, 2003; Wannacott, Newport and Tanenhaus 2008)


Recall and recognition memory for verbatim language is well above chance (Gurevich, Johnson and Goldberg, to appear).

Detailed visual patterns retained, even if they are not attended to or remembered explicitly (DeSchepper and Treisman 1996)

idioms and “prefabs”

You’ve got to be kidding!  Double whammy
wear out <one’s> welcome  Eat, drink and be merry
What’s up?  Excuse <poss> French
What for?  Face the music
shoot the breeze  sooner or later
Are you all right?  What did you say?
Tell me what happened.  Can I come in?
I’m sorry to hear that.  Need any help?
It just goes to show  I see what you mean.
Blithering idiot

The grass is always greener
Persian: Monghe hamsayeh ghaz ast
“the neighbor’s hen is a goose!”

He struts like a peacock.
Persian: Mesle inke az kune fil ohade
“as if he fell of the ass of an elephant”

Do you believe me?
Persian: Harfe u m baveer mikanld?
“do you believe his words?”

How are you?
Persian: Hale shoma shetone?
“How is your health?”

English: black eye
German: ein blaues Auge
“blue eye”

Slept like a dog/log.
German: wie ein Murmelthier schlafen
“slept like a woodchuck”

Someone believes himself to be God’s gift to the world
French: se croire sorti de la cuisse de Jupiter
“believed himself to have sprung from Jupiter’s thigh”
Longer prefab sentences:

How are you going to do that?
Once you’ve done that the rest is easy.
I see what you mean.
I’ll believe it when I see it.
It just goes to show, you can’t be too careful.
You’re not allowed to do that.
It’s a free country isn’t it?
Tell me what happened.
I’m sorry to hear that.

Prefab sentence “stems”

• NP be-TENSE sorry to keep-TENSE you waiting.

• Why DO-tense n’t NP, pick on someone pro,gen own size!

Speakers are faster and more accurate at repeating utterances that they hear with high frequency (when lexical frequency and length are controlled for). (2 & 3 year olds: Bannard and Matthews, to appear, Psych. Science; adults: Bod 1998).

<table>
<thead>
<tr>
<th>high frequency</th>
<th>low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>out of the way</td>
<td>go to the top</td>
</tr>
<tr>
<td>we haven’t got any</td>
<td>we haven’t got enough</td>
</tr>
<tr>
<td>a lot of noise</td>
<td>a lot of juice</td>
</tr>
</tbody>
</table>


– 200 words account for 80% of the input of four children (in analysis of each of several children’s 14K corpora). Mintz, Newport, Bever (2002)

The creative capacity to produce novel language exists, but we often don’t take full advantage

• It’s one twenty five.
vs Conventions in other languages?
• It’s 1:30 minus 5.
• It’s five minutes to 1:30.
• It’s 25 past 1.
Knowledge of input statistics

In language acquisition (e.g., Akhtar and Tomasello 1997; Baker 1979; Bates and MacWhinney 1987; Bowerman 1982; Braine 1976; Gropen et al. 1989; Ingram and Thompson 1996; Lieven et al. 1997; Tomasello 2000, 2003; Wannacott, Newport and Tanenhaus 2008)

In adult language processing (Ford, Bresnan and Kaplan 1982; Jurafsky forthcoming; MacDonald, Pearlmutter and Seidenberg 1993; Garnsey et al. 1997; Tracewski et al. 1993; Pierrehumbert 2000; Losiewicz 1992; Baayan et al. 1997; Bod 1998; Bybee 2000; Gahl and Garnsey 2004; Booij 2002)

• Children are conservative in their language-learning; make most over-regularizations relatively late.

Tomasello and Brooks, Cognitive Linguistics (1998)

Adult Model Always Intransitive:
It fudded. The dough won’t fud. The dough’s fudding in the machine.

Transitive Biasing Question:
What are you doing (to the dough)?
[encouraging: I’m fudding it]

Results
Only 4 out of 32 two-year-old children (12.5%) produced a transitive SVO utterance

Brooks and Tomasello Developmental Psychology (1999)

Adult Model Always Passive:
It’s being floosed by the horsie. It’s being floosed.

Active Biasing Question:
What’s the horsie doing (to it)?
[encouraging: He’s floosing it]

Results
12 out of 48 three-year-old children (25%) produced a transitive SVO utterance

Percentage of children that produce transitive utterances using novel verbs in different studies. (Tomasello, Cognitive, 2009)

Akhtar (1999)

• Elmo the car daxed. (SOV)

• Elmo the car pushed. (SOV: known verb)
Percentage of utterances in which children “corrected” weird word order to SVO with familiar and unfamiliar verbs (Akhtar 1999; Abbott-Smith et al. 2001)

Knowledge of statistical biases

She /confirmed/ the date of our visit.  (DO complement)
She /confirmed/ the date was soon.  (Clausal complement)

DO-biased verbs: confirm, emphasize, hear, understand, write
S-biased verbs: argue, believe, claim, suspect, suspect

Sentences with matching bias are processed faster and with greater accuracy (Clifton et al. 1984; Ferreira and Henderson 1990; Trueswell et al. 1993; MacDonald et al. 1994; Garnsey et al. 1997)

Phonological reductions (/t,d/ deletions) are more likely in matching bias contexts than in non-matching contexts (Gold and Gahl 2004, Language).

“Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist.

...In addition, patterns are stored even if they are fully predictable as long as they occur with sufficient frequency” (Goldberg 2006: 5)

Just how detailed is our memory for language?

How could we know that any item had “sufficient frequency” if some memory trace of it were not stored to enable the frequency of it to be recorded?

Common wisdom holds that people don’t remember the exact form of utterances, only the semantic “gist.”

• “the original form of the sentence is stored only for the short time necessary for comprehension to occur” (Sachs 1967)

• “One of the most robust findings in psycholinguistics is that people cannot reliably recall sentence structures” (Loebell and Bock 2003)

• “Research on memory for verbal materials has demonstrated that sentences are quickly transformed into an underlying abstract meaning and that the original surface structure is lost” (Holtgraves, 2008:361).

Verbatim recognition of sentences is known to exist under certain conditions...

• If people are told they will be asked to recognize the formal properties of sentences (Johnson-Laird and Stevenson 1970; Reyna and Kierman 1994)

• In “highly interactive” contexts (Kearns, MacWhinney, Mayhew 1977; Murphy & Shapiro 1994)

• If few sentences are given and recognition is tested immediately (Reyna and Kierman 1994)
But do people retain explicit verbatim memory in more naturalist contexts?

Verbatim Memory studies

- Undergrads hear one of two versions of a 300 word story. (between subjects).
- They are not warned that their memory will be tested.

STUDY #1: RECOGNITION MEMORY

Results

72% correct (“yes” to matching and “no” to non-matching)
Chance rate: 50%

<table>
<thead>
<tr>
<th>Probability</th>
<th>Matching</th>
<th>Non-Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>“yes”</td>
<td>.86 (hits)</td>
<td>.41 (false alarms)</td>
</tr>
<tr>
<td>“no”</td>
<td>.14 (misses)</td>
<td>.59 (correct rejections)</td>
</tr>
</tbody>
</table>

\[d'= 1.42 : t(23)=14.08, p< .01\]

STUDY #2: RECOGNITION MEMORY, controlling for lexical effects

- Story version 1: I am strong enough to fight four bad guys at once!
- Story version 2: Fighting four guys at once is easy for someone as strong as me.
STUDY #2: RECOGNITION MEMORY, controlling for lexical effects

Results

59% correct ("yes" to matching and "no" to non-matching)
Chance rate: 50%

<table>
<thead>
<tr>
<th></th>
<th>Matching</th>
<th>Non-Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of &quot;yes&quot;</td>
<td>.81 (hits)</td>
<td>.63 (false alarms)</td>
</tr>
<tr>
<td>Probability of &quot;no&quot;</td>
<td>.19 (misses)</td>
<td>.37 (correct rejections)</td>
</tr>
</tbody>
</table>

$d^\prime = .70$
$t(23) = 5.25, p<0.01$

(Gurevich, Johnson and Goldberg, to appear, Language and Cognition)

• What of those older studies?
• It turns out that whenever #’s were given, they hinted at the existence of verbatim memory (Sachs 1967; Jarvella 1973)
• Their aim was to compare verbatim with gist memory.

STUDY #3: Immediate RECALL

Stories contained: 46 clauses

Subjects listened to the story and then, without warning, were asked to "retell the story to me." Pictures were shown as a cue.

– Subjects produced on average 29 clauses.

(Gurevich, Johnson and Goldberg, to appear, Language and Cognition)

Study #3: Recall Results

Stories contained: 46 clauses

Subjects produced on average 29 clauses.

Significant main effect for match/coincidental-match

$F(1,16) = 22.14, p < 0.01.$

Study #4: Recall; Immediate and 2 Day Delay

Study #5 Implicit recall @ week delay

Open possibility that the subjects were remembering their own productions (not the original description)

More naturalistic context of language use

Longer delay
(1 week as opposed to 2 days)
**Exposure Phase**

Cover: “Interested in how well people communicate without seeing each other”

Confederate always “randomly” chosen to be the describer

Two comparable “scripts”

**Assessment Phase**

~ 1 week later

2nd Confederate

Dialogue recorded

---

**Summary of verbatim recall results**

People spontaneously able to recall significant amount verbatim, even in fairly naturalistic context in which:
- They are not warned they will need to remember sentences
- They hear a relatively long story (300 words)
- The context is non-“interactive”
- Even after a week-long delay.

(Gurevich, Johnson and Goldberg, to appear, Language and Cognition)

---

**Negative priming of novel, unattended figures**

DeSchepper and Treisman 1996

Task: Does the purple shape match the blue shape?

Prime trial:

Test trial:
Negative priming of novel, unattended figures
DeSchepper and Treisman 1996

Task: Does the purple shape match the blue shape?

Prime trial:

Slowdown in response when previously ignored shape becomes the subsequent target shape.

Test trial:

Memory for details

tens of thousands of words, idioms and compositional “prefabs” are learned (Pawley and Syder 1983; Jackendoff 2002; Dobrovolsk 2004)


Recall and recognition memory for verbatim language is well above chance (Garevich, Johnson and Goldberg, to appear).

Detailed visual patterns retained, even if they are not attended to nor remembered explicitly (DeSchepper and Treisman 1996)

Generalizations necessarily made

Otherwise languages could be a collection of item-specific factoids:

Pat saw Chris.
Pat Chris kissed.
Hate Pat Chris.

Usage-based model:

We retain impressive amount of item-specific knowledge.

We also categorize (generalize) the input we hear into patterns based on form and function….

Stay tuned…