# Psych 156A/ Ling 150: Psychology of Language Learning

Lecture 12
Words & Rules

## Announcements

Quiz 4: Avg 12.4 out of 16

HW4 due today

HW5 assigned today, due next Thursday (5/15/08)

Quiz 5 on Tuesday (5/13/08)

## Words & Rules

Computational Problem: Identifying word *affixes* that signal meaning.
= Identify the rules for altering word forms in order to signal meaning.

Example: What do you have to change about the verb to signal the past tense in English? (There are both regular and irregular patterns.)

| | | |
|---|---|---|
| blink~blinked (+ed) | confide~confided (+ed) | drink~drank ("ih" --> "ey") |
| rub~rubbed (+ed) | hide~hid ("aye" --> "ih") | think~thought ("ink" --> "ought") |

## Rules for Words: Summary from Last Time

When learning how to form the past tense, children behave as if they are extracting a regular past tense rule.

When children over-apply the regular past tense rule, overregularization errors appear. This often leads to a U-shaped learning trajectory on their performance with the past tense forms of verbs.

There is evidence for children extracting irregular rules as well, based on their performance with the past tense forms of verbs.

## Psychological Reality of Rules: Debate

Yang (2002), Chomsky & Halle (1968): Rules, No Words

There are rules for everything, both regular and irregular patterns.

Pinker & Ullman (2002): Words and Rules

There is a rule for the regular pattern (+ed), but irregular verbs are stored in an associative memory. There is no abstraction of irregular patterns like *drink~drank* and *sink~sank*. To use an irregular past tense form, a speaker simply retrieves the appropriate irregular form from memory.

Rumelhart & McClelland (1986), McClelland & Patterson (2002): Words, No Rules

There is associative memory for everything. The mind never explicitly uses a rule to transform a verb into its past tense.

## Support for Rules, No Words

Analysis of children's performance in the CHILDES database (Yang 2002)

Irregular past tense verb forms benefit if the child encounters many other verbs that use the same rule. The frequency of the rule influences the child's performance.

Example rule & performance:

"No Change" rule: hurt~hurt, cut~cut

children's success: ~80%

many other verbs in this class: *hit, quit, split, slit, spit, bid, rid, forbid, spread, wed, let, set, upset, wet, shut, put, burst, cast, cost, thrust*

## Support for Rules, No Words

Analysis of children's performance in the CHILDES database (Yang 2002)

Irregular past tense verb forms benefit if the child encounters many other verbs that use the same rule. The frequency of the rule influences the child's performance.

*Logic of argument:*

(1) children benefit from irregular rule's use

(2) this would not happen if children's minds don't have an irregular rule

Therefore, children's minds must have an irregular rule. So, irregular verbs are not just memorized individually. They have irregular rules the same way regular verbs use the regular rule.


## Words and Rules

Idea (Pinker (1995), Pinker & Ullman (2002))

Producing a past tense form is a process:

- Intended form: VERB + past tense
- Root form of VERB: VERB
- If irregular VERB, past tense:
  - IRREGULAR PAST (retrieve from memory)
- If regular VERB, past tense:
  - VERB + ed (apply regular rule)

## Words and Rules

Idea (Pinker (1995), Pinker & Ullman (2002))

Producing a past tense form is a process:

- Intended form: VERB + past tense
- Root form of VERB: VERB
- If irregular VERB, past tense:
  - IRREGULAR PAST (retrieve from memory)
- If regular VERB, past tense:
  - VERB + ed (apply regular rule)

Lexicon:
Looking up a word in memory

## Words and Rules

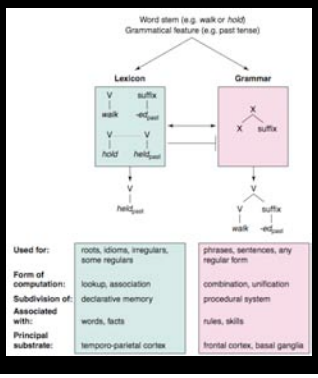
Idea (Pinker (1995), Pinker & Ullman (2002))

Producing a past tense form is a process:

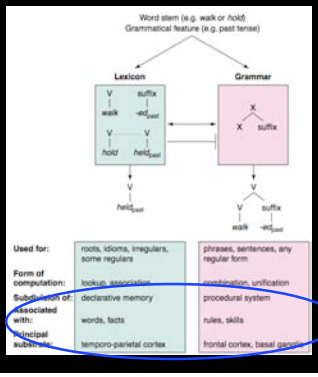
- Intended form: VERB + past tense
- Root form of VERB: VERB
- If irregular VERB, past tense:
  - IRREGULAR PAST (retrieve from memory)
- If regular VERB, past tense:
  - VERB + ed (apply regular rule)

Grammar: Apply a rule

## Words and Rules: Lexicon vs. Grammar




## Words and Rules: Lexicon vs. Grammar



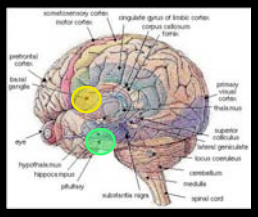

Neurological division of words and rules

## Words and Rules: Neurological Basis

Declarative/Procedural Hypothesis (Pinker & Ullman 2002):

lexical/irregular, hippocampus & medial lobe structures = declarative

grammatical/regular, basal ganglia & frontal cortex = procedural




## Declarative/Procedural Hypothesis Predictions

1) Separable memory

Irregulars - psychological, linguistic, neuropsychological traces of lexical memory

Regulars - psychological, linguistic, neuropsychological traces of grammatical processing

2) "Elsewhere" rule for +ed

When memory fails for irregulars, use +ed rule for past tense.

## Neurological Evidence: Declarative/Procedural Hypothesis

Studies on patients with brain lesions

agrammatism

anomia

Agrammatism: problems with grammar of language (rules)

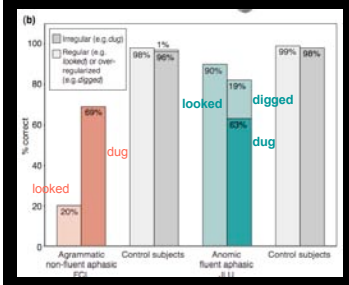
Prediction: These patients do worse on regular +ed rule than irregulars.

Anomia: problems with remembering words (lexical access)

Prediction: These patients do worse on irregulars than +ed rule.

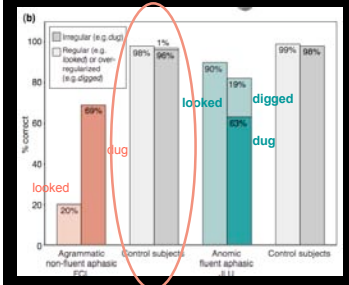
## Neurological Evidence: Declarative/Procedural Hypothesis

Pinker & Ullman (2002)




## Neurological Evidence: Declarative/Procedural Hypothesis

Pinker & Ullman (2002)

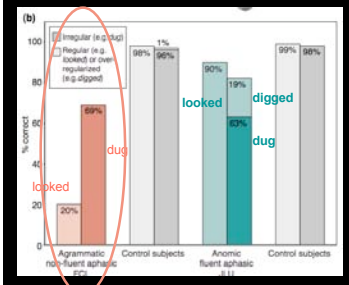



Control subjects:

At ceiling performance (near 100%) for producing the correct past tense for both irregular verbs (dig~dug) and regular verbs (look~looked).

## Neurological Evidence: Declarative/Procedural Hypothesis

Pinker & Ullman (2002)



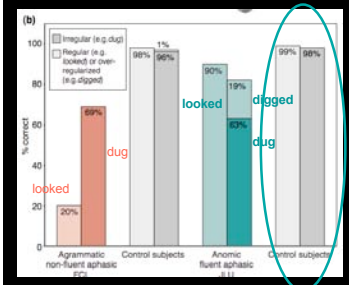

Agrammatic subject:

Poor performance comparatively, but much worse on producing the correct past tense form for regular verbs and no overregularizations for irregular verbs.

Worse at rules

## Neurological Evidence: Declarative/Procedural Hypothesis

Pinker & Ullman (2002)

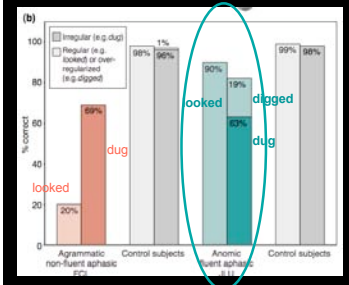



Control subjects:

At ceiling (near 100% performance) for producing both regular and irregular past tense forms.

## Neurological Evidence: Declarative/Procedural Hypothesis

Pinker & Ullman (2002)




Anomic subject:

Not so bad comparatively (over 80% production), but better at regular verbs (look~looked) than irregular verbs (dig~dug). Also, produced many overregularzations (dig~digged).

Good at rules, not so good at irregulars.

## Neurological Evidence: Declarative/Procedural Hypothesis

There seems to be a double dissociation between performance on regular verbs and performance on irregular verbs. We can find patients who are good at regulars, but poor at irregulars. We can also find patients who are good at irregulars, but poor at regulars.

This lends support to the idea that the past tense of regular and irregular verbs may be generated differently. Regular verbs may be making use of more rule-like brain structures and irregular verbs may be making use of more associative-memory-like structures.

## More Neurological Evidence: Declarative/Procedural Hypothesis

More results: Patients with Alzheimer's Disease, Parkinson's Disease, Huntington's Disease

1) Alzheimer's: impaired lexical knowledge (can't remember words) & impaired irregular verbs

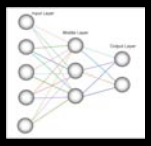
2) Parkinson's: impaired grammatical knowledge (can't use rules of language) & impaired regular verbs

3) Huntington's: unsuppressed basal ganglion (~grammatical brain structure) & overuse of -ed rule (*dugged, walkeded*)

## Words, No Rules

Rumelhart & McClelland (1986), McClelland & Patterson (2002)

Pattern associators learn via a gradual adjustment of simple processing units. They represent the mind's ability to retrieve the correct past tense form without ever using a rule. Also, they can easily capture the regularity in the irregular past tense forms (drink~drank, sink~sank, shrink~shrank).



## Words, No Rules: What About the Neurological Evidence?

Because neural networks can be mapped to brains, networks can have "lesions" in them the same way that brains do, by selectively removing a section of a functional network.

## Words, No Rules: What About the Neurological Evidence?

Because neural networks can be mapped to brains, networks can have "lesions" in them the same way that brains do, by selectively removing a section of a functional network.

However, it is hard to get the double dissociation pattern observed in human patients. No matter where a neural network is lesioned, the network's performance on irregulars (dig~dug) suffers more than its performance on regulars (look~looked). (It always behaves like an anomic patient, not like an agrammatic patient.)

Point: There must be something additional besides this kind of associative memory in human brains.

## Words, No Rules: Issues with Novel Forms

Human Behavior (both adult and child): the ability to generate an appropriate past tense ending for a novel word (like "wug")

wug~wugged (regular past tense rule)

Neural network behavior: Unless the network has specifically built in a section that applies the past tense rule, it will not generate appropriate past tense forms for words it has never encountered before.

Example: Network is trained on English verbs, but never has seen "mail". When forced to generated a past tense form, it produces "membled" (something humans would never do).

## Summary: Words And/Or Rules

To understand how acquisition works, we need to know what the final knowledge is.

For the English past tense, there seems to be evidence that at least some verbs (if not all verbs) use rules in order to form their past tense. Evidence comes from both child performance data and adult neurological data.

Therefore, children need to learn the appropriate rules.

Question:
When do children figure out that they need a rule for certain groups of verbs?

## Words To Rules?

Idea: The point of using rules for past tense forms would be that it's easier in some sense -- as opposed to simply storing each verb and its associated past tense individually.

| | | | | |
|---|---|---|---|---|
| look | looked | | look | |
| kiss | kissed | | kiss | |
| lurch | lurched | vs. | lurch | +ed |
| laugh | laughed | | laugh | |
| dance | danced | | dance | |

harder

easier

## Words To Rules?

Idea: The point of using rules for past tense forms would be that it's easier in some sense -- as opposed to simply storing each verb and its associated past tense individually.

If a particular transformation (rule) occurs a lot (like +ed), it's said to be productive. Productive rules make sense to store because they're used for a lot of different verbs.

Question: What determines if a rule is productive? That is, how does a child decide that a rule is used enough to be worth storing?

## Productive Rules


Yang (2005): Productivity of a rule depends on some kind of cost-benefit analysis for how many words follow the rule and how many words don't.

Specifically, the child keeps track of how many exceptions there are for a particular rule. If there are too many exceptions, it's easier to just not have a rule.

Rule: *ing --> *ang

Verbs that follow the rule: ring~rang, sing~sang, …

Verbs that don't follow the rule: sting~sung, bring~brought, …

## Productive Rules


Important: a rule can be productive while still having exceptions. The big question is simply how many exceptions is too much?

+ed Rule: [any verb] --> [any verb]+ed

Verbs that follow the rule: look~looked, kiss~kissed, laugh~laughed,…
How many?

Verbs that don't follow the rule: sting~sung, bring~brought, drink~drank, ring~rang, keep~kept…
How many?

## Productive Rules

Yang (2005): What matters is how long it takes to access the right past tense form.

There are two options when some verbs follow a rule and some items don't.
(1) Store all the exceptions to the rule, and then the rule. If the verb needed isn't among the exceptions, apply the rule.
(2) Just store all the verbs and their past tense forms individually. (Treat all the verbs as exceptions.)

Tolerance Principle: If it takes longer (on average) to find the right past tense form when both the exceptions and the rule are stored (option 1), just store all the verbs separately (option 2).

## Productive Rules

Yang (2005): What matters is how long it takes to access the right past tense form.

Note: Exceptions to rule are in order of frequency

Rule: *ing-*ang

If word = *sting* then *stung* (freq 100)
Else if word = *swing* then *swung* (freq 80)
Else if word = *ding* then *dinged* (freq 10)
Else if word = *cling* then *clung* (freq 8)
Else Apply *ing --> *ang

## Productive Rules

Yang (2005): What matters is how long it takes to access the right past tense form.

Note: Exceptions to rule are in order of frequency

Rule: *ing-*ang

If word = *sting* then *stung* (freq 100)
Else if word = *swing* then *swung* (freq 80)
Else if word = *ding* then *dinged* (freq 10)
Else if word = *cling* then *clung* (freq 8)
Else Apply *ing --> *ang

swing?
--> swung

Time units: 2

## Productive Rules

Yang (2005): What matters is how long it takes to access the right past tense form.

Note: Exceptions to rule are in order of frequency

Rule: *ing-*ang

If word = *sting* then *stung* (freq 100)
Else if word = *swing* then *swung* (freq 80)
Else if word = *ding* then *dinged* (freq 10)
Else if word = *cling* then *clung* (freq 8)
Else Apply *ing --> *ang

ring?
--> rang

Time units: 5 + rule application

## Tolerance Principle Prediction

Default +ed rule can only be productive if it applies to the vast majority of verbs it could apply to.

Regular rule: [any verb] --> [any verb] + ed

There are 150 irregular verbs, which are exceptions to the regular rule. They fit the [any verb] context, however.

By the Tolerance Principle (which is mathematically precise and has a formula), there need to be at least 1000 regular verbs for it to be faster to have a rule. This seems to be true (we have a lot of regular verbs).

## Tolerance Principle in Child Learning

1) Child identifies possible rule.
   Ex: (*ing --> *ang)

2) Child (unconsciously) checks current vocabulary with Tolerance Principle to see if it's better to store a rule + exceptions, or just all exceptions (individual verbs and their associated past tense forms).

3) Child repeats with each new word type encountered. (Productivity of rules can change, based on how many exceptions the child is aware of at any given time.)


## Productivity Predictions

Depending on the verbs they have encountered, children may believe certain rules are productive while other rules are not.

Prediction for English regular +ed rule:
Children who have this rule should know many more regular verbs than irregular verbs. (This seems to be true.)

## Summary: Storing Rules vs. Words

It makes sense from a processing standpoint for children to store rules if these rules are used a lot - that is, if they are productive rules. Otherwise, it will be easier to simply store individual words and their associated past tense forms.

One way children might decide if a rule should be stored is based on how many verbs follow the rule vs. how many verbs do not follow the rule.

Evidence from children's input and behavior suggests that children who know the regular past tense rule in English (and who may overregularize) do know more verbs that follow the regular rule (regular verbs) than verbs that do not follow the rule (irregular verbs).

## Questions?