A Neural Network Model of the Past Tense

- The past tense in English – some interesting observations
- A neural network model of the past tense

Slide credits: Frank Keller, Frank Mollica, Chris Lucas, Mirella Lapata

Review

- Human language involves two different kinds of “mental tissue”:
  - a finite lexicon of words/word-parts, stored in and retrieved from memory;
  - a finite grammar of rules (productive, combinatorial, recursive).

- These two mechanisms produce an infinite set of sentences.

How do we figure out if speakers really store words and apply rules in their heads?

Various ways to get evidence:

- look at large amounts of text or speech (corpus data)
- study how humans process language in real time (eye-tracking, brain imaging)
- look at what sort of errors speakers make
- track language acquisition in children

The Past Tense in English

- Can use this as a case study to look at how rules might be represented and learned
- Past tense rule: just add -ed to the end of present tense form.
  - Set of regular verbs is open-ended (probably tens of thousands in the mental lexicon of an educated adult).
  - New regular verbs enter English every year.

jog	→	jogged
walk	→	walked
play	→	played
kiss	→	kissed

spam	→	spammed
google	→	googled
blog	→	blogged
nuke	→	nuked

regular verbs

But... we also have irregular verbs

- Linguists use an asterisk (*) to mark things which are not part of the language, or, at least, which make native speakers uncomfortable, or are meaningless.
- Some past tense forms don't just add -ed to the end of the present tense form.
- Irregular past tense inflection is chaotic and idiosyncratic.
- Irregular verbs are a closed list, of 150–180 members.
- There have been no recent additions (not since sneak–snuck arrived during the 19th century).

buy	→	*buyed	→ bought
hold	→	*holded	→ held
steal	→	*stealed	→ stole
go	→	*goed	→ went

Some irregular verbs seem to have their own rules

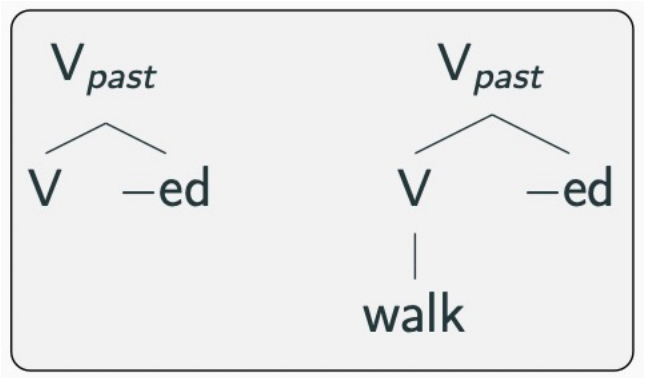
- sing → sang
- ring → rang

- cling → *clang → clung

- think → *thank → thought

A Simple Theory of Regular and Irregular Verbs

- Regular past tense forms are created by a rule.
- Irregular past tense forms are stored and retrieved as words.



sound: *hold*

meaning:

part of speech: V

sound: *held*

meaning:

part of speech: V_{past}

Two mechanisms for past tense formation…

- a productive rule for regular past tense forms;
- irregular past tense forms stored as words;

- PROBLEM WITH OUR THEORY: Why don't they get in each other's way?
- held vs. *holded, stole vs. *stealed

- Simple Answer:
- If a past tense verb form is stored in memory as a word, the rule is blocked. If no past tense form is stored, then the rule may be applied (e.g., snarfed, moshed, ricked).

Evidence from language acquisition?

- Errors preschool children make in their spontaneous speech!
  - It was neat – you should have sawn it!
  - Doggie bat me [bit].
  - The cheerios got aten by the Marky.
  - I know how to do that. I truck myself [tricked].
  - This is the best place I ever sot [sat].

- Most children make at least some errors of this kind.
- Such errors persist well into their school-age years.
- Children have never heard adults using past tense forms like swang or shuck. Must be constructing these forms creatively, by analogy with verbs they already know.

Not just past tense errors! Children also…

- overuse the plural suffix -s (mans, foots, tooths, mouses)
- overuse the third person sing suffix -s (haves, do's, be's)
- overuse the comparative -er and superlative suffixes -est (specialer,
- powerfullest, gooder)
- overuse the ordinal suffix -th on numerals (oneth, twoth)

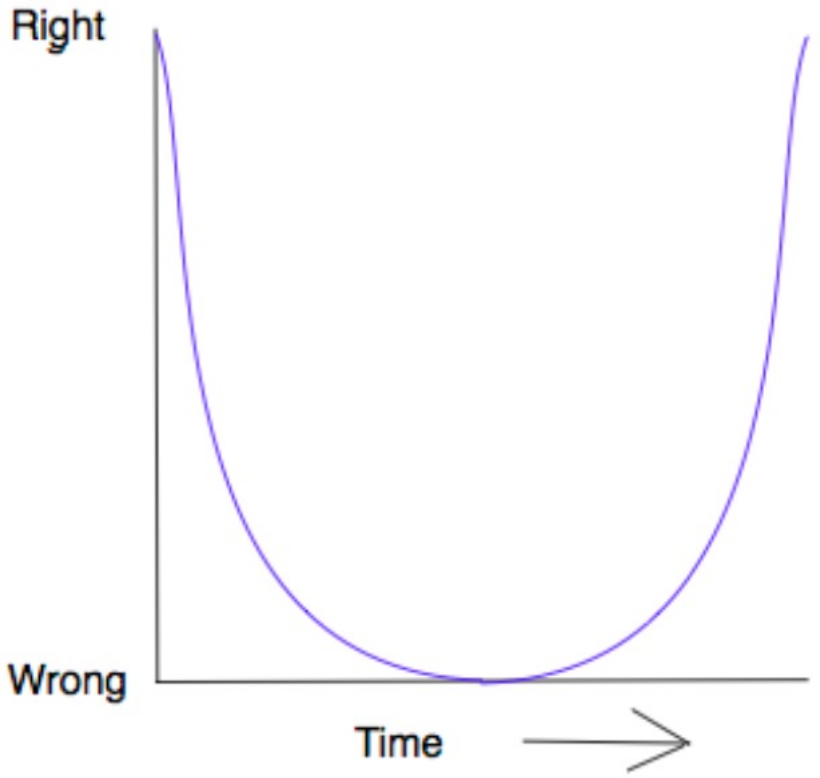
- Children find regularity in the oddest places.
  - Parent: No booze in the house! Child: What's a "boo"?
  - Child: "It did! It snew!" [After being told it was going to snow.]

Children’s verb use over time: 3 stages

- Stage 1 - children produce both regular and irregular past tense forms with very few errors.
- Stage 2 - after a certain amount of time, the error rate appears to increase significantly;
  - children add regular past tense suffix -ed to irregular verb stems even with verbs whose past tense forms they had previously mastered.
- Stage 3 the error rate slowly decreases, as the child gets older, until almost no errors are made.

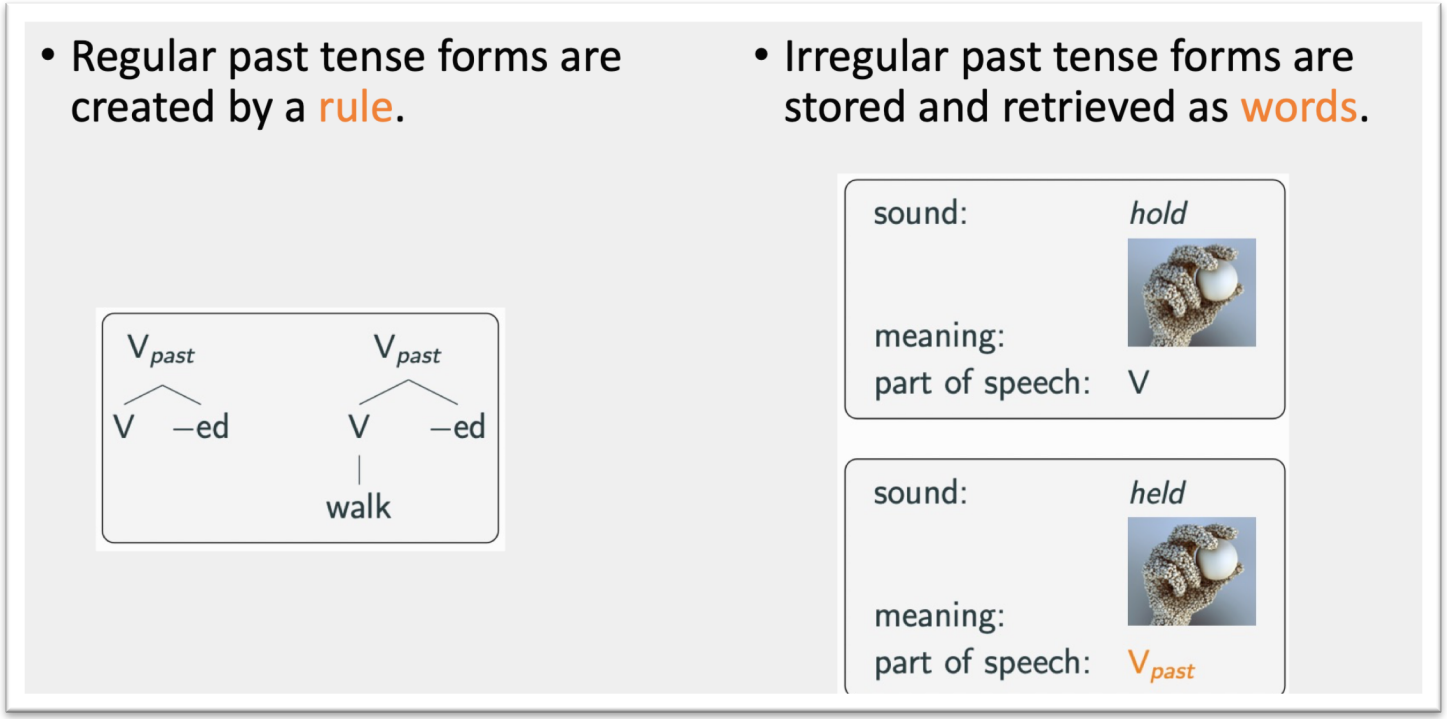
U-shaped pattern of learning

- Child uses spoke, then speaked, and later again spoke.



A puzzle...

- If this is our theory...



- How are rules represented in the brain?
- If the rule is just sitting there, why can't we easily articulate it?
  - Many mental rules seem to be like this!
  - What is this?
  - How did you decide?
- Rules like this are implicit
- **How do we represent something implicit?**

On Learning the Past Tenses of English Verbs

D. E. RUMELHART and J. L. McCLELLAND

to the paper!