

# **Quasi-regular domain: English past tense**

**core problem**

- **the ability to apply qualitatively different operations to different inputs**

**test case**

- **selection of different inflectional operations to English verbs**

**Simplification to selection task  
common in empirical data**

# Past tense phenomena described by the dual route

- Irregulars are affected by similarity

speakers occasionally extend irregular patterns to verbs that are highly similar to irregular families

A continuous effect of similarity has been measured experimentally: subjects frequently (44%) convert *spling* to *splung* (based on *string sling*, et cetera), less often (24%) convert *shrink* to *shrunk*, and rarely (7%) convert *sid* to *sug*

Steven Pinker

“Rules of Language” *Science* 253

The experimental data clearly point to a morphological class consisting of verbs bearing a family resemblance to a prototype. On the one hand, the data show that the verbs that can form a past tense by using the vowel /ʌ/ are not a random set, but that the likelihood of a verb forming a past tense this way correlates with its phonological similarity to a prototype. On the other hand, we have also demonstrated that this class of verbs does not have discrete boundaries, or phonological features whose presence is absolutely required, but rather consists of both central and marginal members.

Bybee and Moder

“Morphological Classes as Natural Categories” *Language* 59 251

# Past tense phenomena described by the dual route

- **Regulars are not affected by similarity**

Work by Prasada and Pinker(1993) indicates that generalization by similarity does not occur with regulars:

the primary goal of the experiments reported here is to test the prediction that humans can apply the regular process as a default operation to stems with any sound ... by seeing how people inflect novel monosyllabic verb roots that do not resemble English regular verbs that people have had experience with.

page 7

For the pseudo-regular classes, suffixed forms were readily supplied in all cases: the subjects gave an average of 9.5, 9.2 and 8.9 suffixed responses out of a possible 10 for the prototypical, intermediate and distant items. The slight decrease in the number of suffixed responses produced as a function of similarity was not significant [ $F_1(2,46) = 1.96, NS$  ;  $F_2(2,27) = 2.51, NS$ ]. The subjects' ratings for these responses produced as a function of similarity did not monotonically decline from prototypical to intermediate to distant: 5.5, 5.2, 5.3, respectively, a difference that was only marginally significant.

page 26

The results [also] show that the similarity of a novel form to familiar regular forms does not significantly affect the likelihood of subjects' producing a regular suffixed form.

page 28

# Question to be answered

Support for the dual route comes in the form of contrasting correlations between

prototypicality of phonological form and willingness to generalize to regular operation  $\neq$  prototypicality of phonological form and willingness to generalize to irregular operation

Do such contrasting correlations constitute unique support for a dual-route model?



# Results

- Training

At iteration 694 of training, there was no step of size greater than  $10^{-50}$  which would reduce error any further.

At this point the network gets 420/422 of the training set right.

The two it gets wrong are:

“drag→drug” (listed as class 6)

output 9 = 0.00128438, output 6 = -0.0771544

and

“light→lighted” (listed as class 9)

output 4 = -0.0173876, output 9 = -0.0206404

# Results

- **Analysing correlations**

For each class  $c$

for each testing pattern  $p$

we determined the Euclidean distance from

$p$  to the nearest neighbor of  $p$  that is of class  $c$

**In irregular classes, these distances correlated negatively with activation of the class output unit — and were often significant.**

**When all irregular classes were taken together, they showed a significant inverse correlation with output unit activation.**

**In the regular class (9) the correlation with output unit activation was insignificant.**

**The difference between the correlations in the irregular classes and the (lack of) correlation in the regular class was significant.**

# Results

- Correlation of distance to appropriate-class nearest neighbor with activation of corresponding output unit

n=60	Z	P <
class1 r = 0.0837	0.6341	0.526
class2 r = -0.2495	-1.9246	0.0543
class3 r = -0.4050	-3.2437	0.0012
class4 r = -0.1106	-0.8392	0.4014
class5 r = -0.1508	-1.1476	0.2511
class6 r = -0.5209	-4.3609	1.2954e-05
class7 r = -0.0729	-0.5516	0.5812
class8 r = -0.3588	-2.8353	0.0046
class9 r = -0.1596	-1.2155	0.2242
classes 1-8 collapsed, n=480		
r = -0.2505	-5.5912	2.2544e-08
Difference of correlations between collapsed 1-8 and 9		
	-31.2229	5.2062e-214

# Conclusion

- Formalizing similarity to irregular families as “distance to nearest neighbor of class  $X$ ,” and
- interpreting the model’s output unit activations as willingness to generalize to class  $X$ ,

we have exhibited, in a single route model, the kinds of contrasting correlations between phonological similarity and willingness to generalize that have been used by proponents of the dual route model as unique support for that model.

---

Current neural net training methods can override similarity biases in data like the English past tense.

---