Analogical Change as Rule Learning Gone Wrong:

The Lengthened *ē-grade in Proto-Germanic Strong Verbs

Penn

Jordan Kodner - U. of Pennsylvania

BGLR - April 07, 2018 - Berkeley, CA

Outline

- PGmc. Strong Verbs
  - The Paradigm and the *ē-grade
  - Previous Accounts
- Studying Proto-Germanic Children?
- Acquiring Paradigms
- Analogical Change: Accounting for *ē
  - Class V -> Class IV
  - but not -> Class III
- Time Permitting...
  - Accounting for *ē within Class V
  - More details about the acquisition algorithm
  - Future work

PGmc. Strong Verbs

PGmc Strong Verbs

- Overall typical for a Germanic language
- 4 stems: present, past 3sg, past, past participles
- Seven classes (I-VII)
- Classes I-VI phonologically determined
- Mostly traceable back to PIE ablaut
- A few hundred securely reconstructable

Strong Verb Paradigm

	Root	Present	Past 3sg	Past	PParticiple	Trans
I	*-iC-	*bītanǫ	*bait	*bitun	*bitanaz	‘bite’
II	*-euC-	*teuhanǫ	*tauh	*tugan	*tuganaz	‘pull’
III	*-eCC-	*helpanǫ	*halp	*hulpun	*hulpanaz	‘help’
IV	*-eR-	*beranǫ	*bar	*bērun	*buranaz	‘carry’
V	*-eT-	*gebanaz	*gab	*gēbun	*gebanaz	‘give’
VI	*-aC-	*faranǫ	*fōr	*fōrun	*faranaz	‘travel’

C = Consonant; R = Sonorant; T = Obstruent

The Lengthened *ē-Grade

- Not derived from PIE by regular sound change

	Present	Past 3sg	Past	PParticiple
I	e-grade	*o*-grade	zero-grade	zero-grade
II	e-grade	*o*-grade	zero-grade	zero-grade
III	e-grade	*o*-grade	zero-grade	zero-grade
IV	e-grade	*o*-grade	*ē*-grade	zero-grade
V	e-grade	*o*-grade	*ē*-grade	*e*-grade

Previous Accounts[1]

- Rectifying stems after reduplication was lost (eg **g^{h}eg^{h}b-* → **gb-*) (Streitberg 1896, Schumacher 2005)
- Some kind of old aorist (Sverdrup 1927, Prokosch 1939, Cowgill 1957)
- Compensatory lengthening (Hirt 1931)
- Length analogy with Class VI ō-grade (eg Kuryłowicz 1968, Meid 1971, Bammesberger 1986)
- Brugmann 1913's second perfect formation (Matzel 1970, Meid 1971)
- Analogical spread from **etaną* 'eat' (Kortlandt 1992, Schumacher 1998, 2005, Mottausch 2000, Ringe 2006, Mailhammer 2007)
- From the nominal system (Bammesberger 1994, 1996)

[1] (Mailhammer, 2007)

Analogical Change

- Most of these accounts are **analogical change**
- We can reason about (and dismiss some of) them based on what we know about analogy
  - Humbolt's Universal, Kuryłowicz's Laws...
  - The notion that analogy is connected to productivity

Can we develop a concrete mechanism for analogy that lets us test out the assumptions of individual accounts more directly?

Identifying a Mechanism

- If analogy is something children do, let's look at children for insights
- The challenge is reasoning about children in a (pre)historic context

Studying Proto-Germanic Children?

Reconstructed vs Child Lexicons

Makes sense to apply acquisition findings only if reconstructed lexicons can stand in for child lexicons

Must show that the known PGmc lexicon “falls within the space of” child lexicons

Size

- Typical 3-year-olds know a couple thousand lemmas[1]
- There are a a couple thousand securely reconstructable PGmc roots

[1] (Hart & Risley, 2003)

Size

- Typical 3-year-olds know a couple thousand lemmas[1]
- There are a couple thousand securely reconstructable PGmc roots

[1] (Hart & Risley, 2003)

Contents

Are the reconstructed verbs “the same” verbs that children would know?

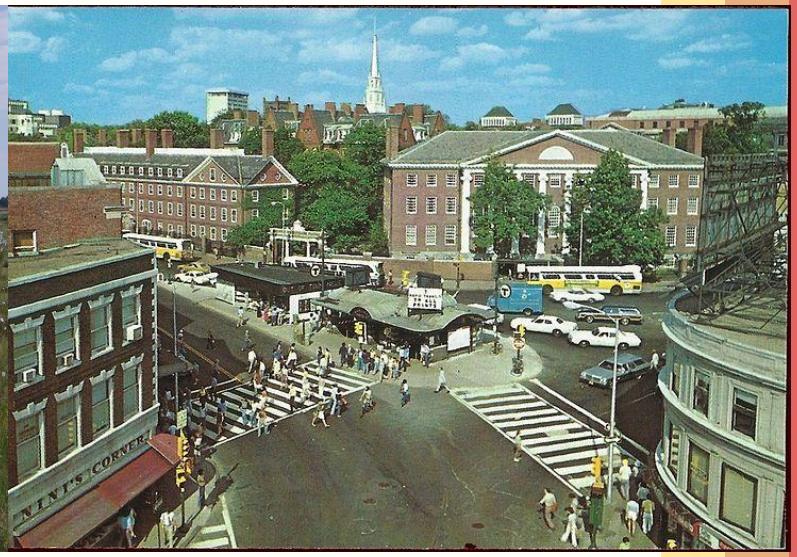
- Extracted 258 securely reconstructed PGmc strong verbs[1]
- Extracted all 358 verbs appearing ≥10 times in the Brown subset of English CHILDES
- Calculated the number of PGmc verbs with English translations by class

[1] (Ringe from Seebold 1979)

English CDS → PGmc Results

	#PGmc	#EN→PGmc	%
I	41	30	73.2
II	40	29	72.5
III	51	45	88.2
IV	16	13	81.3
V	28	21	75.0
VI	29	23	79.3
VII	53	41	77.4
Total	258	202	**78.3**

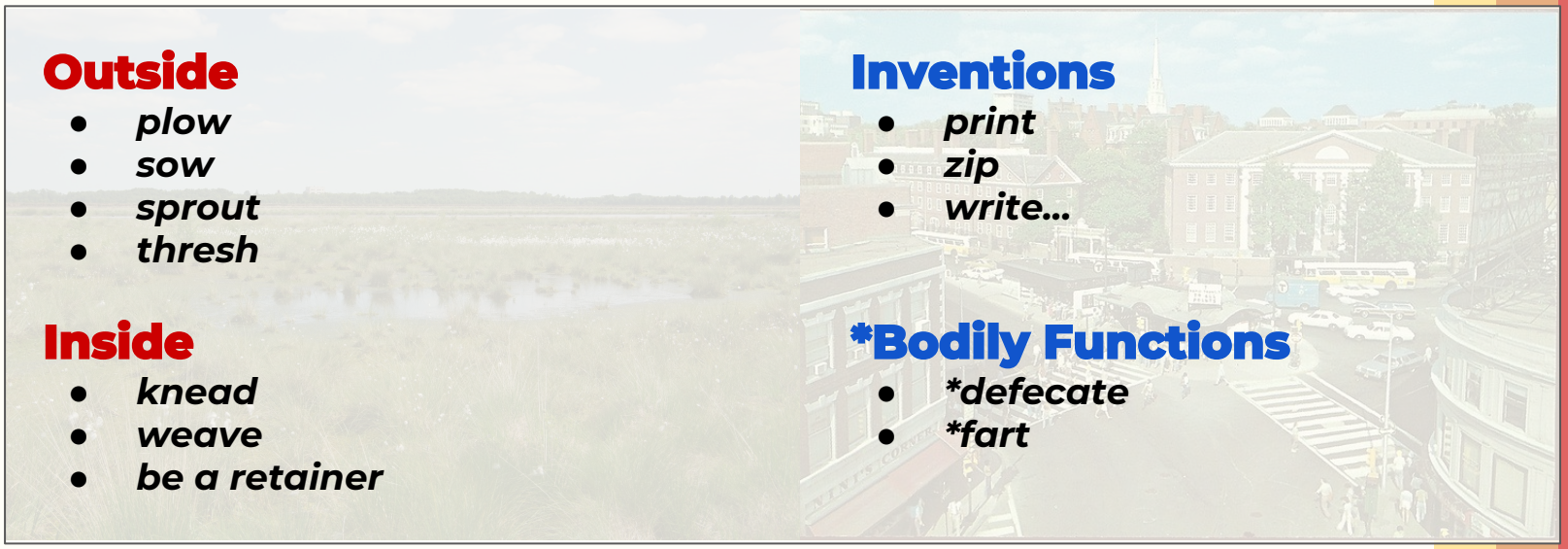
Explanations for Missing Verbs



*Germanic Urheimat, 1st Millenium BC

Cambridge, MA, c. 1970

Explanations for Missing Verbs



Outside
- *plow*
- *sow*
- *sprout*
- *thresh*

Inside
- *knead*
- *weave*
- *be a retainer*

Inventions
- *print*
- *zip*
- *write…*

***Bodily Functions**
- **defecate*
- **fart*

***Germanic Urheimat, 1st Millenium BC**

Cambridge, MA, c. 1970

All Results

	#PGmc	#EN→PGmc	%	#ES→PGmc	%	#EN→ES	%
I	41	30	73.2	30	73.2		
II	40	29	72.5	33	82.5		
III	51	45	88.2	35	68.6		
IV	16	13	81.3	12	75.0		
V	28	21	75.0	21	75.0		
VI	29	23	79.3	21	72.4		
VII	53	41	77.4	34	64.2		
Total	258	202	**78.3**	186	**72.1**	234	**77.8**

All Results

	#PGmc	#EN→PGmc	%	#ES→PGmc	%	#EN→ES	%
I	41	30	73.2	30	73.2		
II	40	29	72.5	33	82.5		
III	51	45	88.2	35	68.6		
IV	16	13	8[illegible]	[illegible]	75.0		
V	28	21	75	21	75.0		
VI	29	23	79.3	21	72.4		
VII	53	41	77.4	34	64.2		
Total	258	202	**78.3**	186	**72.1**	234	**77.8**

Acquiring Paradigms

Learning Rules vs Exceptions

- Given some pairs that appear to follow a pattern, and some that violate the pattern
- Is it better to learn
  - one general rule that has exceptions?
  - multiple more specific rules with fewer exceptions?

How Many English Past Rules?

- ***+ed*** **is obvious**

How Many English Past Rules?

- *+ed* is obvious
- What about *-iN(C) → -aN(C)*?
  - *sing~sang*, *swim~swam*, *drink~drank*, etc.
  - but not *wing~winged*, *sting~stung*, *bring~brought*, etc.

How Many English Past Rules?

- *+ed* is obvious
- What about *-iN(C) → -aN(C)*?
  - *sing~sang*, *swim~swam*, *drink~drank*, etc.
  - but not *wing~winged*, *sting~stung*, *bring~brought*, etc.
- (Ignoring other small classes), two options:
  - **One rule**:
    - *+ed* with *-iN(C)→-aN(C)* as exceptions
  - **Two rules**:
    - *+ed* with no exceptions
    - *-iN(C)→-aN(C)* with exceptions

The Tolerance Principle[1]

- Model of productivity learning
- Based on economy of lexical access - **Is it more efficient to assume some pattern is productive?**

[1] (Yang, 2016)

The Tolerance Principle[1]

- Model of productivity learning
- Based on economy of lexical access - **Is it more efficient to assume some pattern is productive?**
- Many applications
  - Modern English strong verbs
  - English diatones
  - German noun plurals
  - Russian and Polish genitives
  - English and Mandarin numeracy
  - etc.

[1] (Yang, 2016)

The Tolerance Principle[1]

- Model of productivity learning
- Based on economy of lexical access - **Is it more efficient to assume some pattern is productive?**
- Many applications
  - Modern English strong verbs
  - English diatones
  - German noun plurals
  - Russian and Polish genitives
  - English and Mandarin numeracy
  - etc.
- Calculated over ***type frequencies*** (counts in a lexicon), not ***token frequencies*** (counts in a corpus)

Representation

- Lexical items have rules governing derivations -or- are memorized as word-derivation pairs

Representation

- Lexical items have rules governing derivations -or- are memorized as word-derivation pairs
- **Rules = productivity**
- **Memorization = non-productivity**

Representation

- Lexical items have rules governing derivations
  -or- are memorized as word-derivation pairs
- **Rules = productivity**
- **Memorization = non-productivity**
- So learning a rule is tantamount to hypothesizing productivity
- Which option is better for a given case?

How Many English Past Rules?

- One-Rule is tantamount to deciding that *-iN(C) → -aN(C)* is **non-productive**
  - ie, the child should assume *+ed* for new *-iN(C)* words
- Two-Rules is equivalent to deciding that *-iN(C) → -aN(C)* is **productive**
  - ie one should assume *-iN(C)→-aN(C)* for new *-iN(C)* words

Formalism

- ***N*** = # of lemmas in class under consideration

Formalism

- ***N*** = # of lemmas in class under consideration
- ***e*** = # of exceptions in that class

Formalism

- N = # of lemmas in class under consideration
- e = # of exceptions in that class
- Learn a rule if e is tolerable:

$$e < N / \ln N$$

- Otherwise, try a narrower generalization

Formalism

- N = # of lemmas in class under consideration
- e = # of exceptions in that class
- Learn a rule if e is tolerable:

$$e < N / \ln N$$

- Otherwise, try a narrower generalization
- If that fails too, memorize everything

N over time

- *N* is the number of lemmas learned **so far**
- So as the child learns more, tolerable *e* changes
- So children can **temporarily** propose productivity then grow out of it

N over time

- *N* is the number of lemmas learned **so far**
- So as the child learns more, tolerable *e* changes
- So children can **temporarily** propose productivity then grow out of it
- **Quantitatively explains observed overgeneralization errors in child speech**

Analogy as Overproductivity

- **Analogy**
  = overproductivity
  = learning rules with overly wide generalizations
- This happens routinely when a child forms hypotheses on too little data
- But they almost always grow out of it
- **“Almost always” → analogical change**

Accounting for *ē with Acquisition

The Eat Analogy

- ***etanǫ, *ēt, *ētun, *etanaz*** **‘*eat*’ is the only Class V verb with *ē by regular sound change**

  PIE *h_1e-$h_1$ód- > *ēt- > PGmc *ēt-
  PIE *h_1e-h_1d-ˊ > PGmc *ēt-

Steps to the Eat Analogy

- *ē spread from ‘eat’ to all Class V verbs
- *ē spread from Class V to Class IV

Point 2 is well accepted (eg Matzel 1970, Bammesberger 1986, Mottausch 2000, Ringe 2006)

Steps to the Eat Analogy

- ***ē* spread from ‘eat’ to all Class V verbs**
- ***ē* spread from Class V to Class IV**
- ***ē* didn’t spread to Class III**
- ****u* didn’t spread to Class V from IV**
- **Class V past participles didn’t spread**

Steps to the Eat Analogy

- *ē spread from ‘eat’ to all Class V verbs
- *ē spread from Class V to Class IV
- *ē didn’t spread to Class III
- **u didn’t spread to Class V from IV
- Class V past participles didn’t spread

Generalization from V to IV+V

The child has three options:

- Propose Class IV defined by *-*eR*- and Class V by *-e*T*- with few exceptions
- Propose Class IV+V defined by *-eC- with V as the rule and IV as exceptions
- Propose Class IV+V defined by *-eC- with IV as the rule and V as exceptions

Generalization from V to IV+V

- **Mature learners should reject IV+V.**
- **Class IV verb pasts and past participles create too many exceptions.**

Generalization from V to IV+V

- **Mature learners should reject IV+V.**
- **Class IV verb pasts and past participles create too many exceptions.**

- ***N*** **= |IV+V| = 44**

Generalization from V to IV+V

- **Mature learners should reject IV+V.**
- **Class IV verb pasts and past participles create too many exceptions.**

- N = |IV+V| = 44
- e = |IV| = 18

Generalization from V to IV+V

- **Mature learners should reject IV+V.**
- **Class IV verb pasts and past participles create too many exceptions.**

- $N = |\text{IV+V}| = 44$
- $e = |\text{IV}| = 18$
- $N / \ln N = 11.6$

Generalization from V to IV+V

- **Mature learners should reject IV+V.**
- **Class IV verb pasts and past participles create too many exceptions.**

- N = |IV+V| = 44
- e = |IV| = 18
- $N / \ln N$ = 11.6

18 > 11.6.
IV+V FAILS!

Generalization from V to IV+V

- But imagine a younger child.
- If that child knows, say, 9 Class V verbs and 5 Class IV verbs so far,

Generalization from V to IV+V

- **But imagine a younger child.**
- **If that child knows, say, 9 Class V verbs and 5 Class IV verbs so far,**

- $N = 5+9 = 14$
- $e = 5$
- $N / \ln N = 5.3$

Generalization from V to IV+V

- **But imagine a younger child.**
- **If that child knows, say, 9 Class V verbs and 5 Class IV verbs so far,**

- $N = 5+9 = 14$
- $e = 5$
- $N / \ln N = 5.3$

5 < 5.3.
IV+V SUCCEEDS!

Generalization from V to IV+V

- ~26% of learner states are IV+V with V rule
  - Avenue for analogy of V forms into IV
- ~02% of learner states are IV+V with IV rule
  - It was much more likely for *ē to spread to from V to IV than for **u to spread from IV to V

Generalization from V to IV+V

Attested Evidence

- Daughters disagree about some IV and V past participles
  - ON *drepinn* , OE *drepen* vs Beowulf 2981 *dropen*
  - WGmc (OHG) treden, cnedan vs ON *troða*, OSw *knodha* *Goth trudan*
  - Large numbers of V→IV in OHG

Summary

- A large minority of learners would try to inflect Class IV verbs with Class V forms, at least for a while
- The other direction was rare
- This provides an avenue for analogical levelling

No Generalization to III+IV+V

- IV+V is defined by *-eC- and III is defined by *-eCC-
- Would a III+IV+V defined by *-eC(C)- work?

No Generalization to III+IV+V

- **IV+V is defined by *-eC- and III is defined by *-eCC-**
- **Would a III+IV+V defined by *-eC(C)- work?**

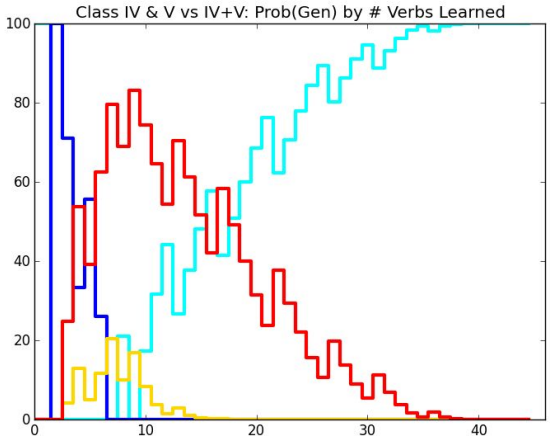
Not for mature learners. Not even close...

- N = |III+IV+V| = 96
- e = |III| = 52
- N / ln N = 21.0

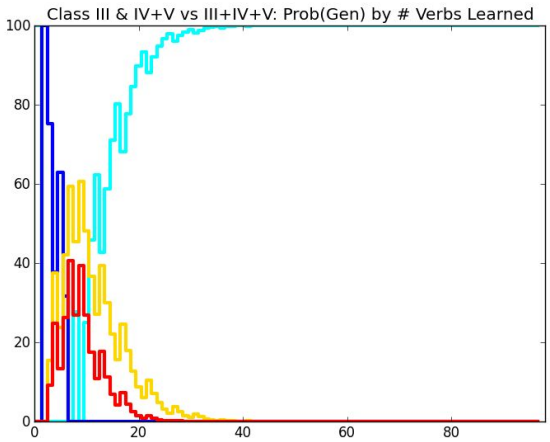
52 >> 21.0.

Generalization to III+IV+V

- ~3% of learner states are III+IV+V with IV+V rule
- Very unlikely that *ē would spread from IV and V to III
- Contrast with 26% of states spreading *ē from V to IV

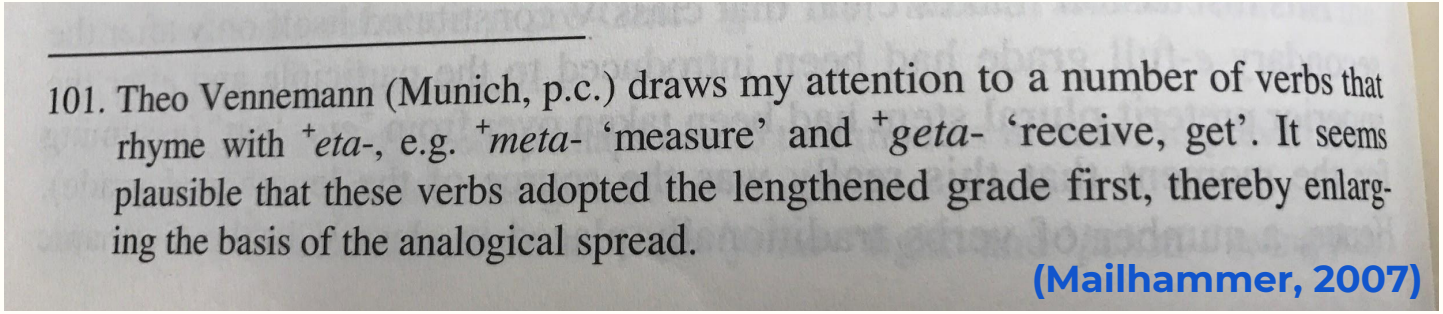


Red: %learners generalizing V -> IV by vocab size



Red: %learners generalizing IV+V -> III by vocab size

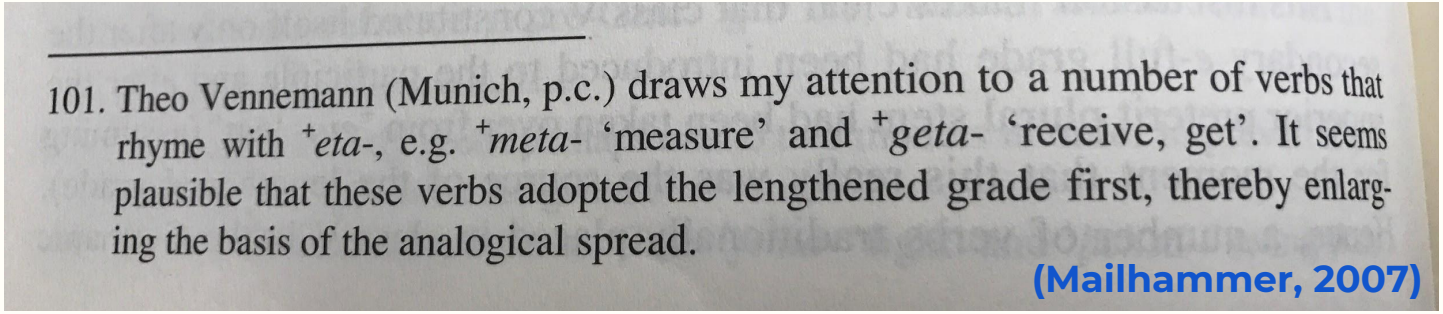
From *etanq to Class V



101. Theo Vennemann (Munich, p.c.) draws my attention to a number of verbs that rhyme with $^{+}$*eta*-, e.g. $^{+}$*meta*- 'measure' and $^{+}$*geta*- 'receive, get'. It seems plausible that these verbs adopted the lengthened grade first, thereby enlarging the basis of the analogical spread.

(Mailhammer, 2007)

From *etanq to Class V



101. Theo Vennemann (Munich, p.c.) draws my attention to a number of verbs that rhyme with +*eta*-, e.g. +*meta*- 'measure' and +*geta*- 'receive, get'. It seems plausible that these verbs adopted the lengthened grade first, thereby enlarging the basis of the analogical spread.

(Mailhammer, 2007)

With 4 verbs (et-*, **met-*, **get-*, **fet-*) we can use the Tolerance Principle.**

Could *ē claw its way up from 4 verbs to all Class V verbs?

Subgeneralizations in V

Generalizations between **et-* and **eT-*

Generalization	*N*	*N* / ln *N*	*e* = *N*-4
*-e[-voi -cont -son]-	7	3.59	3
*-e[-voi -son]-	19	6.45	15
*-e[-voi COR]-	11	4.58	7
*-e[-cont -son]-	12	4.83	8
*-e[-son COR]-	12	4.83	8

Generalization from *-*et*- to V

- It could have spread from the 4 verbs to Class V verbs with voiceless stops:
  - **lekaną, *rekaną, *wrekaną*
- And from there to broader generalizations until it reached **-eT-*

Conclusions

- Well-reconstructed lexicons can be investigated like child lexicons
- This gives us insights into the mechanisms for analogical change
- Applied to PGmc strong verbs, this method supports the plausibility of the Eat Analogy

End

Acknowledgements:

- Don Ringe
- NDSEG Fellowship

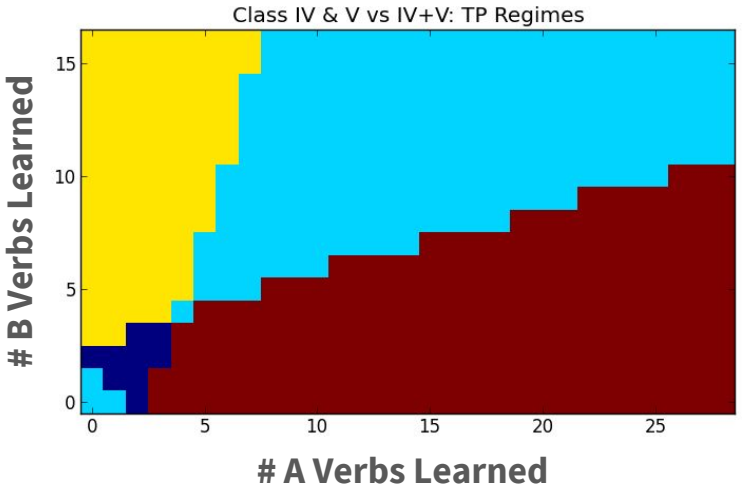
Contact:

- jkodner@sas.upenn.edu
- ling.upenn.edu/~jkodner

Likelihood of Overgeneralizations

Given two classes A and B of sizes *K* and *N-K* and a plausible generalization between them, there are 4 possible outcomes

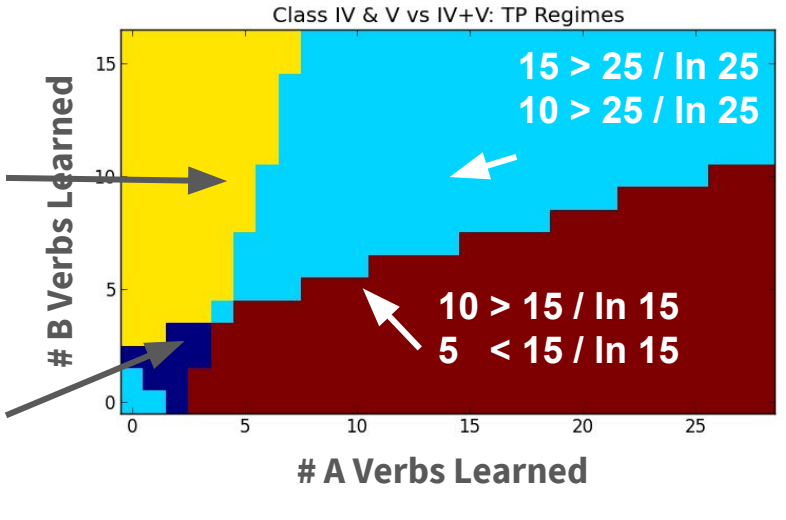
- **Separate rules for A and B**
- **Rule A for A+B**
- **Rule B for A+B**
- **Rule A or B for A+B**



Likelihood of Overgeneralizations

Given two classes A and B of sizes *K* and *N-K* and a plausible generalization between them, there are 4 possible outcomes

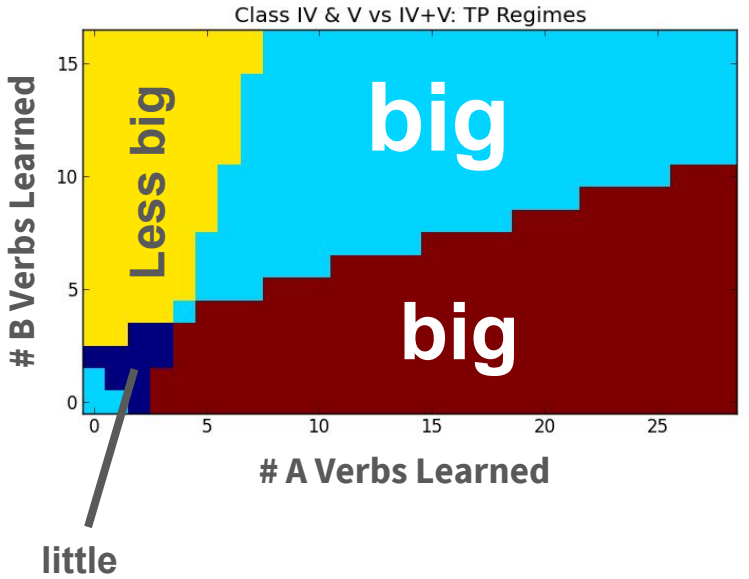
- **Separate rules for A and B**
- **Rule A for A+B**
- **Rule B for A+B**
- **Rule A or B for A+B**



Likelihood of Overgeneralizations

Given two classes A and B of sizes *K* and *N-K* and a plausible generalization between them, there are 4 possible outcomes

- **Separate rules for A and B**
- **Rule A for A+B**
- **Rule B for A+B**
- **Rule A or B for A+B**

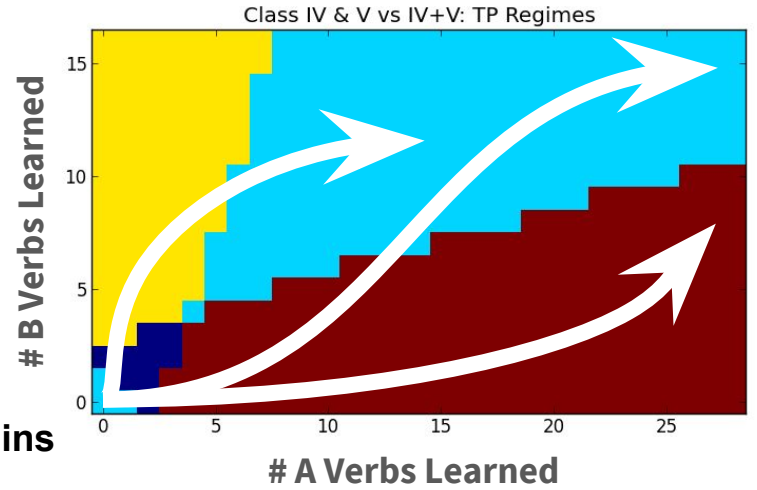


Likelihood of Overgeneralizations

Children progress along paths through this space

- Separate rules for A and B
- Rule A for A+B
- Rule B for A+B
- Rule A or B for A+B

Mature learner
at N = |A+B|



Fresh child begins
at N=0

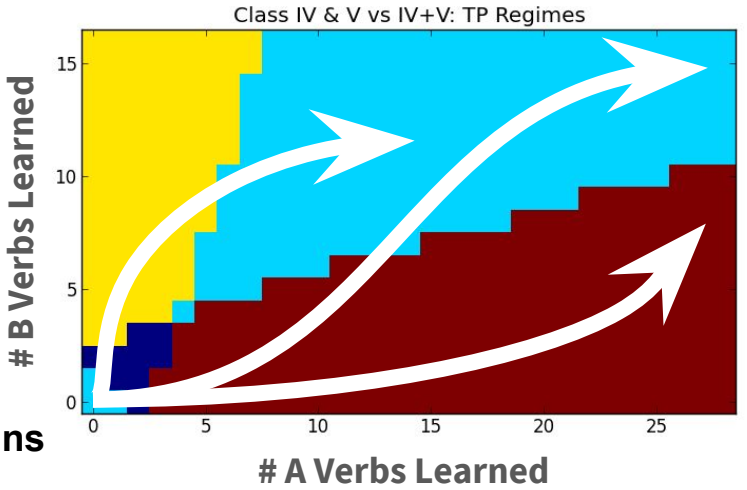
Likelihood of Overgeneralizations

Children progress along paths through this space

but not all paths are equally likely!

- Separate rules for A and B
- Rule A for A+B
- Rule B for A+B
- Rule A or B for A+B

Mature learner
at N = |A+B|



Fresh child begins
at N=0

Likelihood of Overgeneralizations

Likelihood of landing in each state modeled as a hypergeometric distribution ie drawing marbles without replacement[1]

[1]Unweighted marbles approximated when both classes have similar frequency distributions?

Likelihood of Overgeneralizations

Likelihood of landing in each state modeled as a hypergeometric distribution ie drawing marbles without replacement[1]

- $N = |A \cup B|$
- $K = |A|$
- $n = |\subseteq A \cup B$ learned so far$|$
- $k = |\subseteq A$ learned so far$|$
- $n\text{-}k = |\subseteq B$ learned so far$|$

[1]Unweighted marbles approximated when both classes have similar frequency distributions?

Likelihood of Overgeneralizations

Likelihood of landing in each state modeled as a hypergeometric distribution ie drawing marbles without replacement[1]

- $N = |A \cup B|$
- $K = |A|$
- $n = |\subseteq A \cup B \text{ learned so far}|$
- $k = |\subseteq A \text{ learned so far}|$
- $n\text{-}k = |\subseteq B \text{ learned so far}|$

$$P(X = k) = f(k; N, K, n)$$

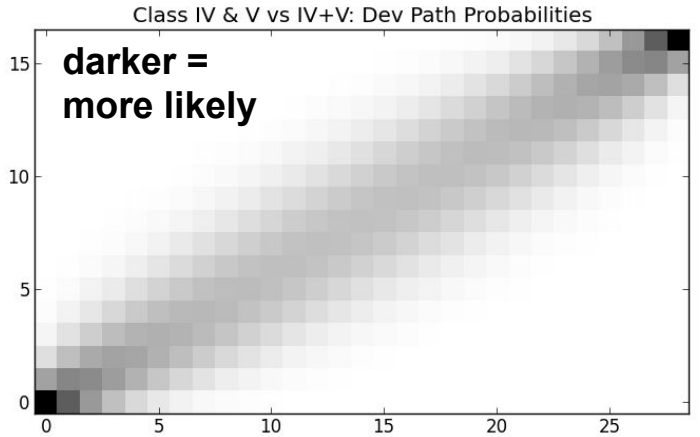
$$= \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

[1]Unweighted marbles approximated when both classes have similar frequency distributions?

Likelihood of Overgeneralizations

Likelihood of landing in each state modeled as a hypergeometric distribution ie drawing marbles without replacement[1]

- $N = |A \cup B|$
- $K = |A|$
- $n = |\subseteq A \cup B \text{ learned so far}|$
- $k = |\subseteq A \text{ learned so far}|$
- $n\text{-}k = |\subseteq B \text{ learned so far}|$

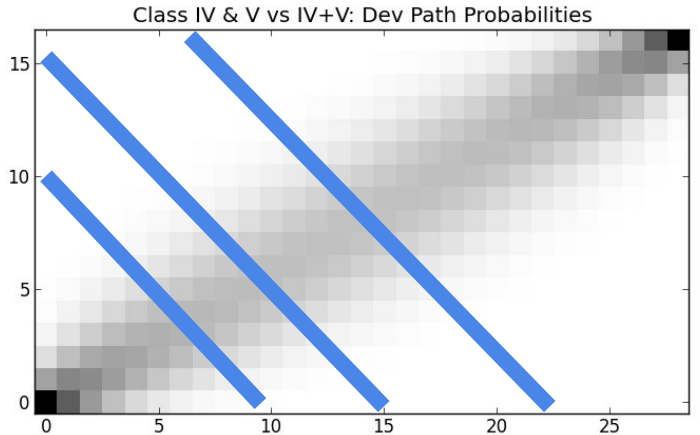


[1]If one class tends to be much more common than the other, this “line” will bow up or down

Likelihood of Overgeneralizations

Likelihood of landing in each state modeled as a hypergeometric distribution ie drawing marbles without replacement[1]

- $N = |A \cup B|$
- $K = |A|$
- $n = |\subseteq A \cup B \text{ learned so far}|$
- $k = |\subseteq A \text{ learned so far}|$
- $n\text{-}k = |\subseteq B \text{ learned so far}|$



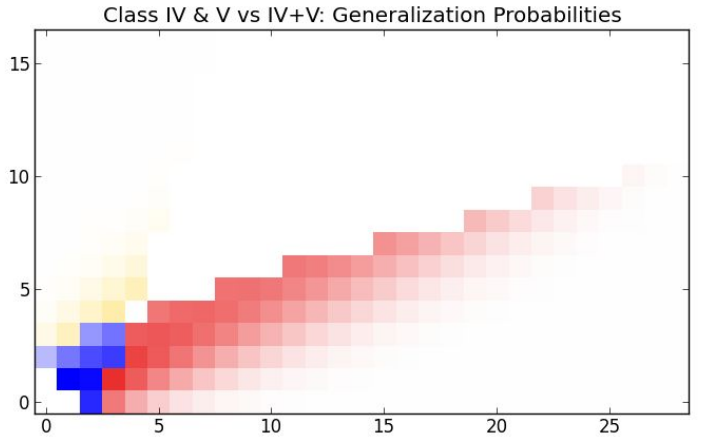
diagonals (constant *n*) each sum to 1

[1]If one class tends to be much more common than the other, this "line" will bow up or down

Likelihood of Overgeneralizations

Composing the previous two plots visualizes likelihood of generalizing

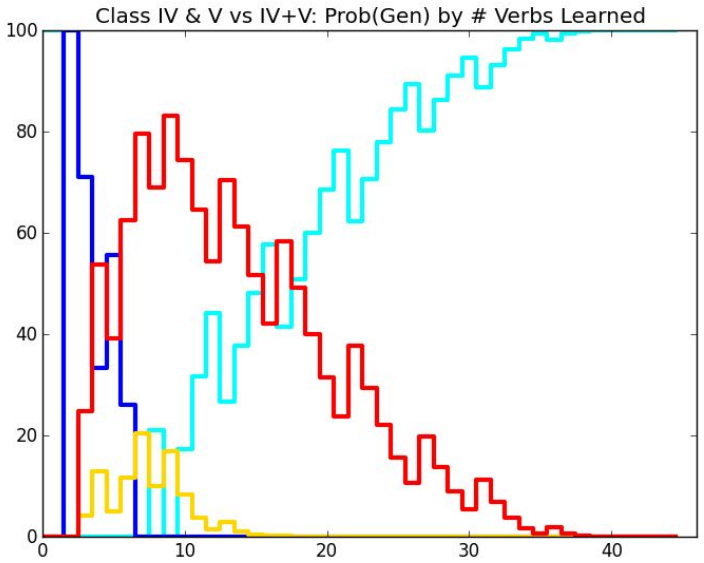
- **Rule V for IV+V**
- **Rule IV for IV+V**
- **Rule V or IV for IV+V**



Likelihood of Overgeneralizations

Plotting likelihood by *n* of each state

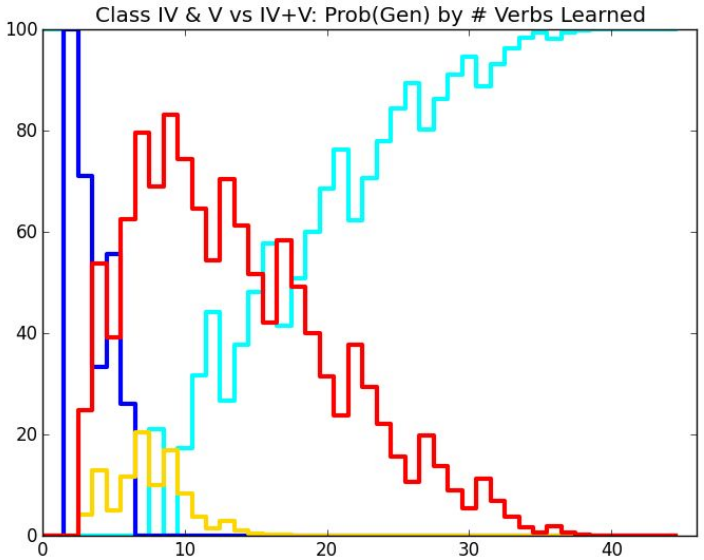
- Separate rules for V and IV
- Rule V for IV+V
- Rule IV for IV+V
- Rule V or IV for IV+V



Likelihood of Overgeneralizations

Area under the curves ≈ proportion of time spent in state[1] ≈ proportion of learners in state[2]

- **64.3% (wins by the end)**
- **27.2% (dominant early, trails)**
- **2.2% (present early only)**
- **6.4% (dominant early only)**

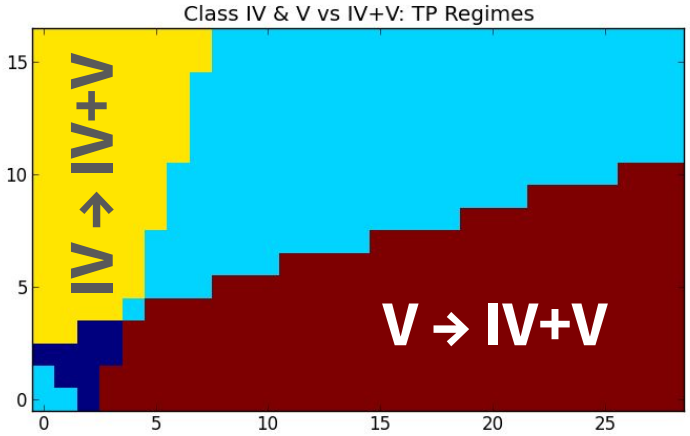


[1]Related to learning rate
[2]Related to population structure

Comparing V → IV+V and IV+V → III+IV+V

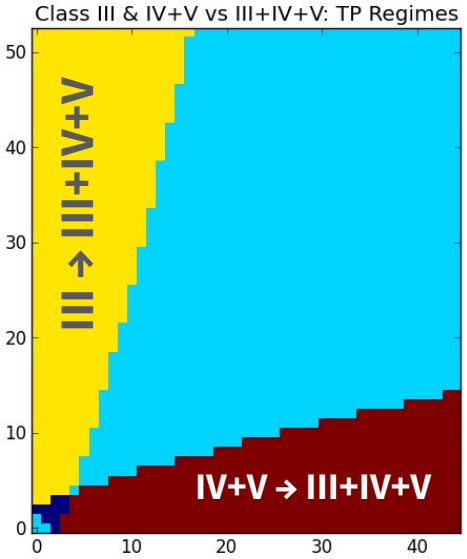
V → IV+V

|IV| = 16, |V| = 28



IV+V → III+IV+V

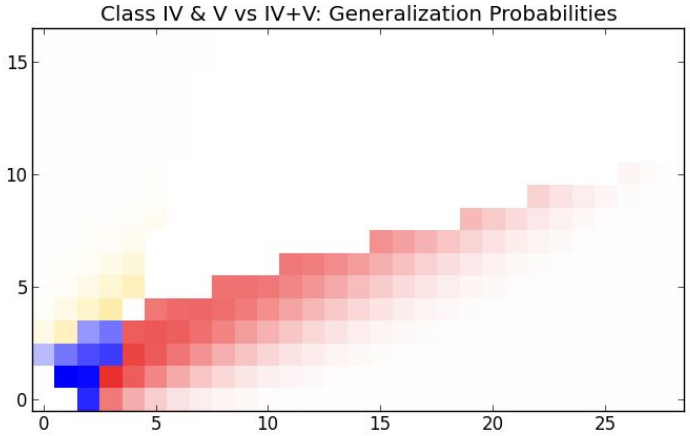
|III| = 52, |IV+V| = 44



Comparing V → IV+V and IV+V → III+IV+V

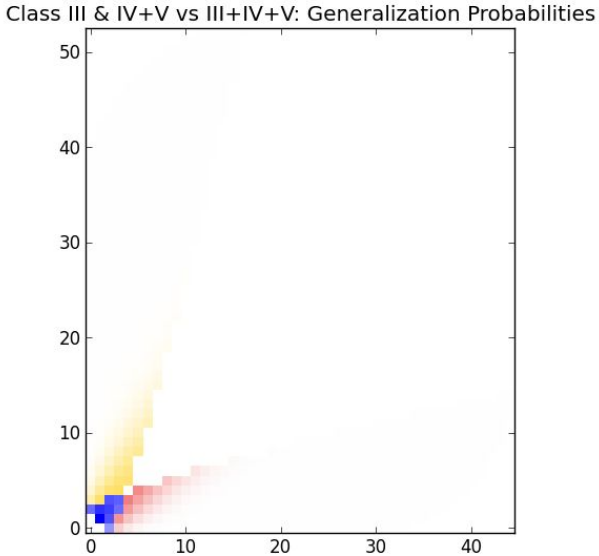
V → IV+V

|IV| = 16, |V| = 28



IV+V → III+IV+V

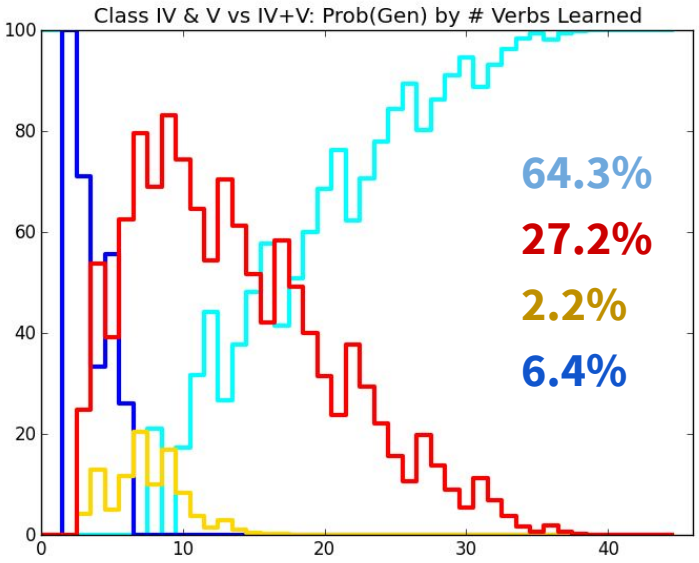
|III| = 52, |IV+V| = 44



Comparing V → IV+V and IV+V → III+IV+V

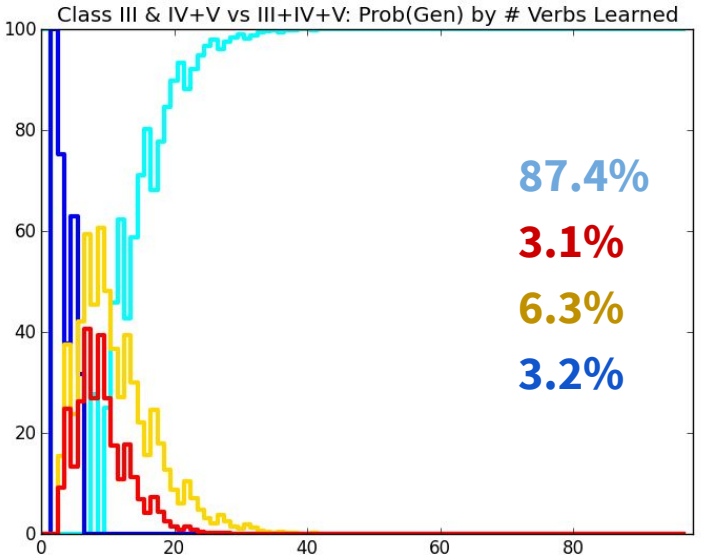
V → IV+V

|IV| = 16, |V| = 28



IV+V → III+IV+V

|III| = 52, |IV+V| = 44



Comparing V → IV+V and IV+V → III+IV+V

- Overgeneralizations provides the avenue for analogy
- Some overgeneralizations are more likely than others
- Given the Proto-Germanic lexicon,

- V→IV+V is much more likely than IV→IV+V (27.2 vs 2.2%)
  - Why the analogy was from V to IV rather than vice-versa
- IV+V→III+IV+V and III→III+IV+V were also unlikely (3.1,6.3%)
  - Why further generalization did not happen

The Paradox of Language Change[1]

[1]Niyogi & Berwick 1995

The Paradox of Language Change[1]

If children are so good at acquiring language, why are they so bad at it?[2]

[1]Niyogi & Berwick 1995
[2]A paraphrase of Niyogi & Berwick 1995

The Paradox of Language Change[1]

If children are so good at acquiring language, why are they so bad at it?[2]

We've shown which overgeneralizations are more likely to occur, but we haven't explained why they persisted in adult speakers

[1]Niyogi & Berwick 1995
[2]A paraphrase of Niyogi & Berwick 1995

Learner Errors

Blame the Child

- The learner does not act correctly on its input
- “a buggy algorithm”
- Hard-coded ε parameter (cf Griffiths, Kirby, etc)

Learner Errors

Blame the Child

- The learner does not act correctly on its input
- “a buggy algorithm”
- Hard-coded ε parameter (cf Griffiths, Kirby, etc)

Blame the Environment

- The learner acts correctly but is dealt a bad input sample
- “garbage in, garbage out”
- Change in the face of trivial variation (cf Niyogi & Berwick)

The Sibling Effect

- Children rarely receive input from a single source grammar
- Trivial variation is ever-present in the input

Imagine two incompetent peers Alice & Bob

- Alice is currently overgeneralizing and Bob is listening
- Bob receives “correct” adult tokens and Alice’s tokens
- What does Bob do?

The Sibling Effect

Is Bob Skeptical?

- Can Bob recognize Alice's incompetence?
- If so, can Bob ignore her?

The answers to these predict different behaviors

Can Bob Recognize Alice's Incompetence?

- Only if Bob has heard an adult-produced token

Can Bob Recognize Alice's Incompetence?

- **Only if Bob has heard an adult-produced token**
  - **Alice is only somewhat untrustworthy**
  - **Without an adult reference, when can he assume that Alice is wrong?**

Can Bob Recognize Alice's Incompetence?

- Only if Bob has heard an adult-produced token
  - Alice is only somewhat untrustworthy
  - Without an adult reference, when can he assume that Alice is wrong?
- Likelihood of ***not*** hearing an adult token varies by domain
  - Not hearing (many tokens of) a phone?

Can Bob Recognize Alice's Incompetence?

- **Only if Bob has heard an adult-produced token**
  - **Alice is only somewhat untrustworthy**
  - **Without an adult reference, when can he assume that Alice is wrong?**
- **Likelihood of *not* hearing an adult token varies by domain**
  - **Not hearing (many tokens of) a phone?** **~impossible**

Can Bob Recognize Alice's Incompetence?

- **Only if Bob has heard an adult-produced token**
  - **Alice is only somewhat untrustworthy**
  - **Without an adult reference, when can he assume that Alice is wrong?**
- **Likelihood of *not* hearing an adult token varies by domain**
  - **Not hearing (many tokens of) a phone?** **~impossible**
  - **Not hearing some corner of verbal inflection?**

Can Bob Recognize Alice's Incompetence?

- **Only if Bob has heard an adult-produced token**
  - Alice is only somewhat untrustworthy
  - Without an adult reference, when can he assume that Alice is wrong?
- **Likelihood of *not* hearing an adult token varies by domain**
  - Not hearing (many tokens of) a phone? ~impossible
  - Not hearing some corner of verbal inflection? sure

Can Bob Recognize Alice's Incompetence?

- **Only if Bob has heard an adult-produced token**
  - Alice is only somewhat untrustworthy
  - Without an adult reference, when can he assume that Alice is wrong?
- **Likelihood of *not* hearing an adult token varies by domain**
  - Not hearing (many tokens of) a phone? ~impossible
  - Not hearing some corner of verbal inflection? sure
  - Not hearing some obtuse syntactic construction?

Can Bob Recognize Alice's Incompetence?

- **Only if Bob has heard an adult-produced token**
  - Alice is only somewhat untrustworthy
  - Without an adult reference, when can he assume that Alice is wrong?
- **Likelihood of *not* hearing an adult token varies by domain**
  - Not hearing (many tokens of) a phone? ~impossible
  - Not hearing some corner of verbal inflection? sure
  - Not hearing some obtuse syntactic construction? yes.

Can Bob Ignore Alice?

- I don’t know…

Can Bob Ignore Alice?

- I don’t know…
  - How bad do Alice’s mistakes have to be?
  - Does relative age matter? Are 3yo’s cool to 2yo’s?

Can Bob Ignore Alice?

- I don’t know…
  - How bad do Alice’s mistakes have to be?
  - Does relative age matter? Are 3yo’s cool to 2yo’s?
- Likely dependent on on the domain again
  - Morphological doublets

Germanic Inflectional Doublets

A persistent feature of the family

- Post-PGmc IV/V confusions
- Weak Verbs in Old/Middle English
- Modern English
    - dived/dove, sneaked/snuck, brought/brang, saw/seen…

The Sibling Effect Effect

- If Bob accepts Alice’s overgeneralized tokens of IV+V,

Short-term

Long-term

The Sibling Effect Effect

- If Bob accepts Alice's overgeneralized tokens of IV+V,

Short-term

- Do these decrease the number of exceptions?
- If anything, these work in favor of IV+V

Long-term

The Sibling Effect Effect

- If Bob accepts Alice's overgeneralized tokens of IV+V,

Short-term

- Do these decrease the number of exceptions?
- If anything, these work in favor of IV+V

Long-term

- Even if Bob matures into IV and V, will adult Bob occasionally produce IV verbs with V's *ē?
- If so, next generation will receive **competent** IV *ē

Explicanda

Positives

- **Why did **ē* spread from *eat* to V?**
- **Why did **ē* spread from V to IV?**

Negatives

- **Why was the spread not IV to V?**
- **Why did **ē* not spread from IV+V to III?**
- **Why did **u* not spread from III to IV or V?**
- **Why did the past 3sg and pparticiple stem vowels not spread?**

Explicanda

Positives

- **Why did **ē* spread from *eat* to V?**
- Why did **ē* spread from V to IV?

Negatives

- Why was the spread not IV to V?
- Why did **ē* not spread from IV+V to III?
- Why did **u* not spread from III to IV or V?
- **Why did the past 3sg and pparticiple stem vowels not spread?**

Why did only Class V’s past stem form spread?

- The other stems ***could*** spread and have (cf WGmc)

Why did only Class V's past stem form spread?

- The other stems *could* spread and have (cf WGmc)
  - So this isn't really a spread/didn't spread dichotomy
  - It's more of a happened/kinda almost happened dichotomy

Why did only Class V's past stem form spread?

- The other stems *could* spread and have (cf WGmc)
  - So this isn't really a spread/didn't spread dichotomy
  - It's more of a happened/kinda almost happened dichotomy

My idea is not much better than the classic stories

Why did only Class V's past stem form spread?

- The other stems *could* spread and have (cf WGmc)
  - So this isn't really a spread/didn't spread dichotomy
  - It's more of a happened/kinda almost happened dichotomy

My idea is not much better than the classic stories

- Are the past 3sg and pparticiple stems more or less frequent than past?
  - Influences how early forms are heard/learned
  - Could affect the TP and the Sibling Effect