Modeling Domain-Narrowing Analogical Change

Constantine Lignos

Computer and Information Science
Institute for Research in Cognitive Science
University of Pennsylvania

11/7/2012

Introduction: A case for modeling

Why model language change?

- Change happens

  Unfortunately, or luckily, no language is tyrannically consistent. All grammars leak. (Sapir, 1921)

- Modeling forces us to build a mechanistic explanation of *how*
- Advances in child data availability and acquisition modeling allow for explicit models

Analogical change

- Analogical change: learning process causes a generalization to spread from one set of forms to others
- Generalization's domain narrows (Kiparsky, 1993):
  1. Applies to phrase
  2. Applies to word
  3. Applies to stem

Postnasal plosive deletion

- *b/g* → ø / N_]$_{\sigma}$ (N = homorganic nasal) (Borowsky, 1993)
- Ex.: sing: *sɪŋ *sɪŋg,* sing-er: *sɪŋəɹ *sɪŋgəɹ,* finger: *fɪŋgəɹ *fɪŋəɹ*
  - Tricky cases in comparatives that retain /*g*/: younger, longer, stronger

Still allowed in novel forms

Modeling goals

- Build the *simplest* model that can help us understand the conditions required for change
- Model change as multi-generational acquisition: what generalization would each successive generation learn?
- Ask: what conditions enabled a specific analogical change to proceed?

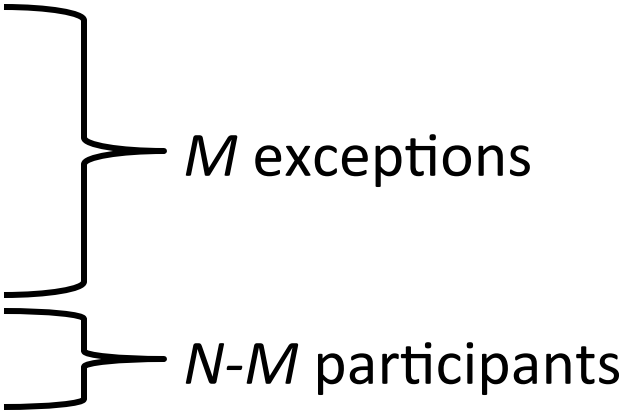
The productivity criterion

Productivity and generalization (Yang, 2005)

- What makes a worthwhile generalization? Real-time processing criterion
- A productive rule can tolerate a number of exceptions *M* only if treating them as exceptions leads to an lower expected real-time processing time than simply listing all items sorted by frequency
- Organization requires irregulars be listed before regulars
  - Controlling for frequency, idioms (Cutler & Swinney 1979) and irregulars (Clashen et al. 2004, Penke & Krause 2005) faster than regulars
  - Example of elsewhere condition

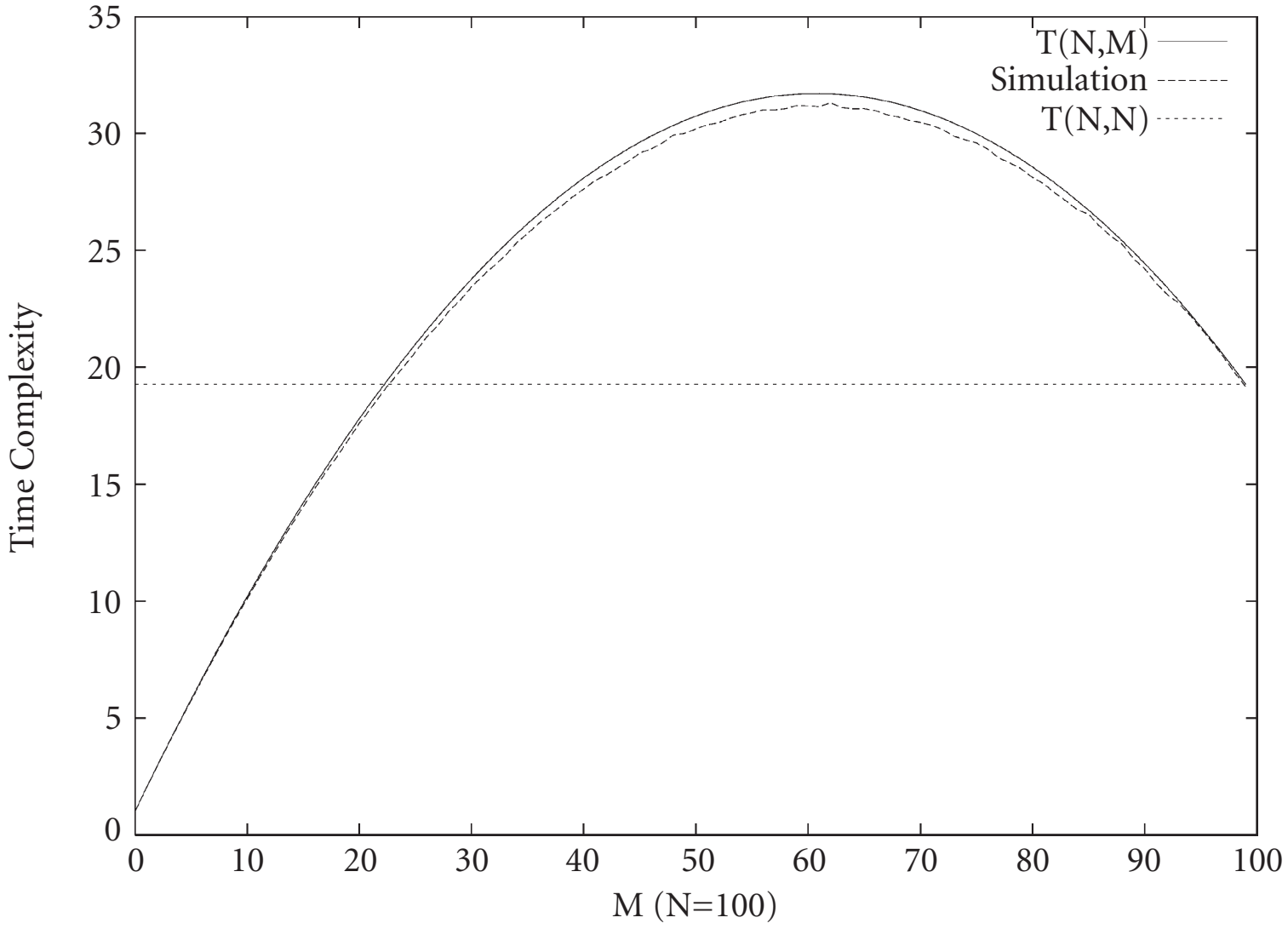
Exception lookup

IF form == x THEN x'
ELSE IF form == y THEN y'
ELSE IF form == z THEN z'
…
else DEFAULT



- Mathematically:
  - *N*- # of items that meet structural description of rule
  - *M*- # of items that meet structural description of rule but are exceptions
  - Criterion: $M < N / ln(N)$ assuming a Zipfian world (see Yang 2005 for proof)

Cost of storing exceptions (Yang, 2005)

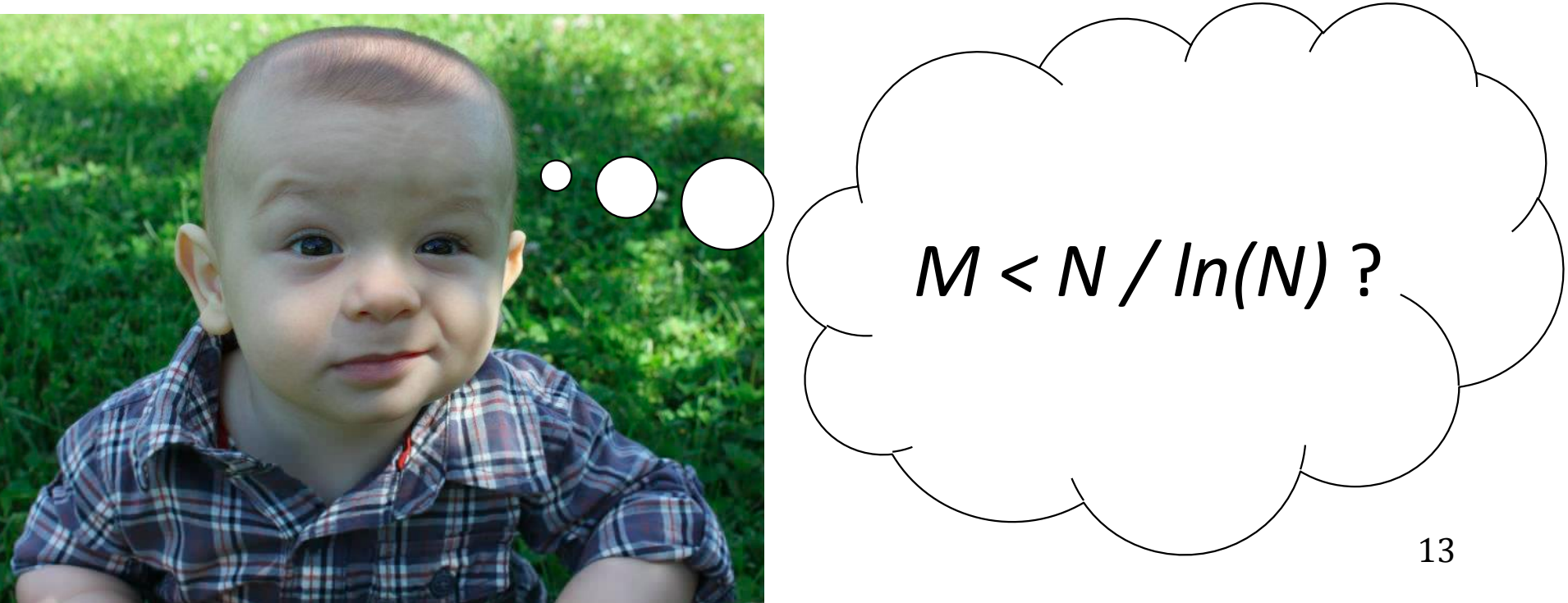


Application to English past tense

- Approximately 150 irregular verbs in English
  - Number of exceptions M = 150, thus N > 1000 for there to be a productive system
  - Number of participants (N - M) must be greater than 850
- Corpus analysis using COBUILD shows 1,252 regular past tense forms with frequency > 1 per million

Wide applications (Yang *et al.*, 2012)

- Non-productive systems predict paradigmatic gaps
  - Participles for English irregulars (*forgo*-**forwent*/**forgoed*)
  - Verbal gaps in Spanish and Russian
- Predict overregularization (U-shaped learning) during acquisition (Marcus et al., 1992)



The phenomenon

Stratal-cyclic models

- We see *overapplication* in cases like *sɪŋ.ɪŋ (singing)* and *lɔŋ.ɪʃ (longish)*
- On surface, */g/* not in coda → evidence of earlier application (Bermúdez-Otero, 2011):

  [sɪŋ~~g~~][ɪŋ~~g~~]
  [sɪŋ][ɪŋ]
  sɪŋ.ɪŋ

- Under stratal-cyclic models (e.g., Lexical Phonology, Stratal OT):
  - Morphological and phonological operations are interleaved
  - Phonological processes can apply at several levels/strata:
    - Stem, word, phrase

Deletion at each level

- Phrase level: applies where the /g/ cannot resyllabify onto following content, remaining in the coda
  pro.lon~~g~~]$_{PL}$
  pro.lon~~g~~ that]$_{PL}$
  pro.lon.g it]$_{PL}$
- Word level: additionally applies where the /g/ is in the coda in a complete word
  pro.lon~~g~~]$_{WL}$
  pro.lon~~g~~]$_{WL}$ it
  pro.lon.g]$_{SL}$ *-er*]$_{WL}$
  sin.g]$_{SL}$ *-in~~g~~*]$_{WL}$
- Stem level: applies where the /g/ is in the coda before level 2 affixes:
  sin~~g~~]$_{SL}$]$_{WL}$
  sin~~g~~]$_{SL}$ *-ing*]$_{WL}$

The change (Bermúdez-Otero, 2011)

	Stage 0	1	2	3
*elo**ng**ate*	ŋg	ŋg	ŋg	ŋg
*prolo**ng**-er*	ŋg	ŋg	ŋg	ŋ
*prolo**ng** it*	ŋg	ŋg	ŋ	ŋ
*prolo**ng*** ‖	ŋg	ŋ	ŋ	ŋ

Step 1 Step 2

Modeling results

Simulation

- Assembled all US English CHILDES data, transcribed using CMUDict
  - 2.8 million tokens in total
- Goal of simulation is to evaluate whether productivity can explain progression of change
- Hypotheses:
  - Step 1: level of ambiguity between **phrase/word** levels…
  - Step 2: level of ambiguity between **word/stem** levels…
  - …are high enough that reanalysis will occur

Step 1

- Innovation at the word level:

	phrase-level /*g*/-deletion (conservative)	word-level /*g*/-deletion (innovative)
$[_{PL}[_{WL}$ sin*g*-er]]	*g*	*g*
$[_{PL}[_{WL}$ sin*g*] $[_{WL}$ aloud]]	*g*	∅
$[_{PL}[_{WL}$ sin*g*]]	∅	∅

- If this change is to proceed, number of exceptions to a word-level deletion rule must not exceed tolerance
- Source of apparent exceptions is resyllabification preventing deletion

First problem: counting types and tokens

- Productivity is traditionally computed over *types*, unique words in the input
- However, in this case we see variation in *tokens*; each occurrence of *sing* can be different
- Some baseline strategies from dealing with this:
  - Conservative: a word type is an exception if it ever doesn't participate
  - Aggressive: a word type is a participant if it ever participates
  - Cautious: only count types that are completely consistent, e.g. always participate or never participate

Resyllabification poses a problem

- Unrestricted phrase level resyllabification prevents a productive generalization:

	Participants	**Exceptions**	**Tolerance**
Conservative	378	**821**	169
Aggressive	1002	**197**	169
Cautious	378	**197**	90

- But do we have evidence of restrictions?

Returning to the original account

The naſal ſound, like the other liquids, though by nature depreſſive, is no leſs capable, we already know, of direct combination. Not indeed that it ever precedes, except in *length* and *ſtrength*, either in the ſame or in another ſyllable, any direct conſonant but its own guttural, however painted, in *ſank*, *ſanction*, *anchor*, *banquet*, *ſphinx*, &c.

As the *g* is always underſtood before the direct guttural, ſo is it before the depreſſive, when this has to articulate either a vowel or a liquid; which it does not only if the vowel or liquid follow in the ſame word, but even, upon ſolemn occaſions, if either feebly commence the word following in immediate connexion and dependance. Thus *ſinking* and *ſinging*, *ancle* and *angle*, *anker* and *anger*, *Tancred* and *angry*, &c. and ſo *ſing aloud*, *ſpring eternal*, *ſtrong and mighty*, &c. as if *ſingking* and *ſingguing*, *ſinggue aloud*, &c. But in different words it muſt indeed be a very ſtrong, though not an impoſſible articulation, which expreſſes a final *g* before an initial *l* or *r*: as in *young Leander*, *long repoſe*.

A closer look at Elphinston's formal register

- "Upon solemn occasions […] if either feebly commence the word following" (formal register):
  - sin[*g*] aloud, prolon[*g*] it, stron[*g*] and mighty, sprin[*g*] eternal
  - Given as equivalent to word/stem level cases
- "But in different words it must indeed be a very strong, though not an impossible articulation, which expresses a final *g* before an initial *l* or *r*"
  - youn[*g*] Leander, lon[*g*] repose
- Analysis:
  - Always require "feeble" (unstressed) following syllable
  - Potential restriction to creating onset, not maximizing

Restricted phrase-level resyllabification

- With both stress and no-maximization restrictions:

	Participants	**Exceptions**	**Tolerance**
Conservative	378	671	150
Aggressive	1002	**147**	163
Cautious	378	147	83

- Summary: restrictions on phrase-level resyllabification were essential to change proceeding

Step 2

- Innovation at the stem level:

	word-level /*g*/-deletion (conservative)	stem-level /*g*/-deletion (innovative)
$[_{PL}\,[_{WL}\,[_{SL}$ sin*g*]-er]]	*g*	∅
$[_{PL}\,[_{WL}\,[_{SL}$ sin*g*]] $[_{WL}$ aloud]]	∅	∅
$[_{PL}\,[_{WL}\,[_{SL}$ sin*g*]]]	∅	∅

- If this change is to proceed, number of exceptions to a stem-level deletion rule must not exceed tolerance
- Source of apparent exceptions is suffixed stems

Step 2

- Easy transition, no matter how you count:

	Participants	Exceptions	Tolerance
Conservative	1074	**77**	163
Aggressive	1083	**68**	163
Cautious	1074	**68**	162

- *M* far below tolerance predicts that word level application without stem level will be unstable and rapidly change
  - Consistent with no account of a stable period
- Lingering question: what stops the learner from reanalyzing the underlying form entirely?

Predictions

- In languages with more aggressive phrase level resyllabification, processes will have difficulty moving from phrase to word level
- In languages with fewer bare stems surfacing, processes will not progress to the stem level at all
  - Dutch final coda devoicing (Booij,1997), Spanish nominals (Bermúdez-Otero, in press)
- Further test cases needed: phrase level and word level rules that stay where they are

From simulation to formal prediction

- With simulation results in hand, we can work toward a more analytical approach
  - Ideally, predicting outcome based on predicted stem/word overlap computed from lexical resources
- Simulation of changes that:
  - Remain at phrase level
  - Continue from phrase to word level, but stop before stem
- Investigate learner's strategy regarding conflicting information for given word types: is it just frequency?
  - Only anecdotal analysis so far

Conclusions

- Gives first mechanistic account of how such a change can proceed
- Predicts that languages with different levels of domain ambiguity and different syllabification restrictions will allow different changes
- For this change to have happened, the learner must have relatively eager to reanalyze

Acknowledgments

- Many thanks to:
  - Gene Buckley
  - Ricardo Bermúdez-Otero
  - Charles Yang