Over-Regularization Unites Disparate Paradoxical Diachronic Patterns in Phonology, Morphology, and Syntax

Jordan Kodner
Stony Brook University

AMC Symposium
Edinburgh, December 5, 2022

Language Change by Language Acquisition

- First language acquisition is one of the primary drivers of language change[1]
- Taken to play a role in both innovation and propagation

The general idea

- Minor “errors” in acquisition accrue over successive generations
- This eventually yields population-level change, which may be dramatic

Prediction

- Trends in child innovations should mirror historical developments
- At least in the domains that are driven by acquisition

[1] Paul 1880, Sweet 1899, Halle 1962, Kiparsky 1965, Andersen 1973, Baron 1977, Lightfoot 1979 *et seq*, Labov 1989, Niyogi 1996 *et seq*, Kroch 2005, Yang 2002 *et seq*, van Gelderen 2011, Cournane 2017, Kodner 2020, *inter multa alia*

Prima Facie Support

Leveling in the English Past Tense

- A classically observed parallel between acquisition and change
- Mostly “regularization” towards weak *-ed* pasts

  ME *help~halp* → ModE *help~helped*

  ME *werke~wroghte* → ModE *work~worked*
- Child production errors are overwhelmingly overregularizations like these

  English estimates:[1] 4-10% overreg. ≫ 0.2% “over-irregularization”

  Spanish estimates:[2] 5% overreg. ≫ 0.1% “over-irregularization”

Suggests acquisition ‘errors’ as a source diachronic innovation

[1]Pinker & Prince 1994, Xu & Pinker 199, Maslen et al 2004, [2]Clahsen et al 1992, Mayol 2007

Prima Facie Support - But it's complicated

Leveling in the English Past Tense

- A classically observed parallel between acquisition and change
- Mostly "regularization" towards weak *-ed* pasts

  ME *help~halp* → ModE *help~helped*

  ME *werke~wroghte* → ModE *work~worked*
- Child production errors are overwhelmingly overregularizations like these

  English estimates:[1] 4-10% overreg. ≫ 0.2% "over-irregularization"

  Spanish estimates:[2] 5% overreg. ≫ 0.1% "over-irregularization"
- But then why do "irregularization" changes happen?

  dig~diggede → *dig~dug* *stician~sticode* → *stick~stuck*

[1]Pinker & Prince 1994, Xu & Pinker 199, Maslen et al 2004, [2]Clahsen et al 1992, Mayol 2007

Prima Facie Support - But it's complicated

Leveling in the English Past Tense

- A classically observed parallel between acquisition and change
- Mostly "regularization" towards weak *-ed* pasts

ME *help~halp* → ModE *help~helped*

ME *werke~wroghte* → ModE *work~worked*

And don't children grow out of these innovations?

- Why/when/how would these innovations gain traction in a population?

[1]Pinker & Prince 1994, Xu & Pinker 199, Maslen et al 2004, [2]Clahsen et al 1992, Mayol 2007

Actuation and the Paradox of Language Change[2]

If children are so good at acquiring language,
how are they so bad at it?

[1] term coined by Niyogi & Berwick 1997

Actuation and the Paradox of Language Change[1]

If children are so good at acquiring language,
how are they so bad at it?

Helps to have a precise definition of actuation[2]…

Actuation = Innovation + uptake into the speech community

(The hand-off from an individual-level process to a population-level one)

[1] term coined by Niyogi & Berwick 1997, [2] definition paraphrased from Labov, Yager & Steiner 1972

Actuation and the Paradox of Language Change[1]

If children are so good at acquiring language,
how are they so bad at it?

Helps to have a precise definition of actuation[2]…

Actuation = Innovation + uptake into the speech community

(The hand-off from an individual-level process to a population-level one)

…and precise models of the relevant aspects of acquisition

I focus on the Tolerance Principle[3] because its recent track-record

[1] term coined by Niyogi & Berwick 1997, [2] definition paraphrased from Labov, Yager & Steiner 1972, [3] Yang 2005, 2016

The Tolerance Principle (Yang 2005, 2016)

- A concrete model for the acquisition of linguistic generalization
- A cognitively-motivated evaluation metric over linguistic hypotheses
- Separates the algorithmic aspects of acquisition from the representations over which generalizations are formed

Has been applied to a wide range of generalization-learning tasks

- Inflection in Arabic, Cree, English, Frisian, German, Icelandic, Polish, Spanish…
  (Yang 2005, 2016, Belth et al 2021, Björnsdóttir 2021, Munshi 2021, Merkuur 2021, Henke 2022,…)
- Dutch, English, and Latin derivational morphology (Yang 2016, van Tuijl and Coopmans 2021, Kodner 2022)
- Argument structure constraints in English, Icelandic, and Korean
  (Yang 2016, Irani 2019, Lee & Kodner 2019, Nowenstein et al 2020, Pearl & Sprouse 2021)
- 'Root infinitive' phenomenon (or lack thereof) in English, French, Hebrew and Spanish (Payne 2022)
- Phonological 'rules' in English (Sneller et al 2018, Richter 2021, Dresher and Lahiri 2022)
- Variation in Scottish *amn't* (Thoms, Adger, Heycock, Jamieson & Smith)

and many more…

The Tolerance Principle (Yang 2005, 2016)

- A concrete model for the acquisition of linguistic generalization
- A cognitively-motivated evaluation metric over linguistic hypotheses
- Separates the algorithmic aspects of acquisition from the representations over which generalizations are formed

Has been applied to a wide range of generalization-learning tasks

And has gained backing from a range of psycholinguistic experiments

(Schuler et al 2017, Koulaguina and Shi 2019, Emond & Shi 2021)

The Tolerance Principle (Yang 2005, 2016)

Given a hypothesized generalization operating over some class, quantitatively define the number of exceptions below which the generalization is tenable

N = number of types that should obey the generalization

e = number of types that do not obey the generalization

θ = max # of exceptions that can be tolerated

Exceptions are tolerable if

$$e < \theta$$

$$\theta = N / \ln N$$

N and *e* Vary over Individual Development

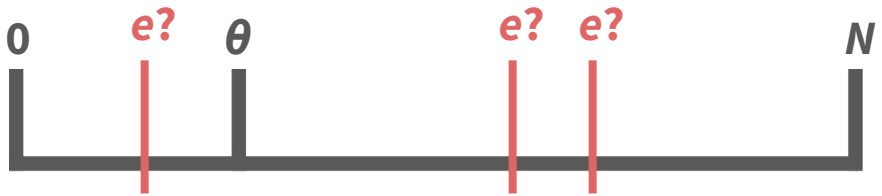
- *N* and *e* are properties of each individual
- *N* is the number of class members a child has learned so far
- *N* and *e* grow as the learner's vocabulary grows
- Can learn generalizations over small *N* not possible over large *N*

Concrete example: "Form past tense by suffixing *-ed*"

- Say *N* is the number of verbs a child knows so far
- And *e* is the number of verbs with irregular pasts known so far

Visualization of the Tolerance Principle

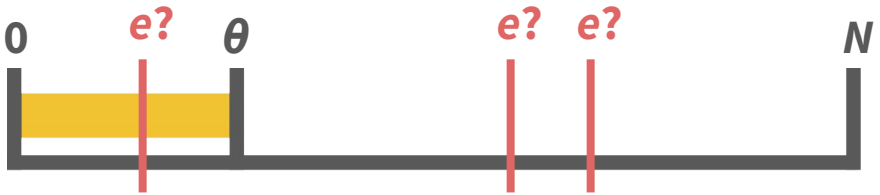
N = types it should apply to
e = types that are exceptions
θ = tolerance threshold



e falls in [0,*N*] and may be less than or greater than *θ*

Visualization of the Tolerance Principle

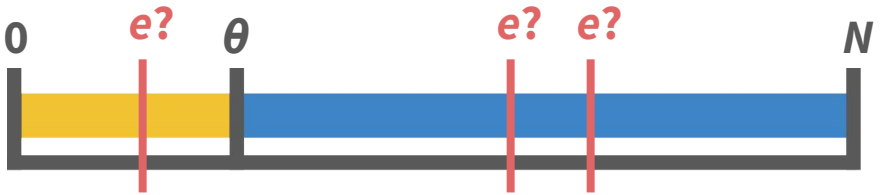
N = types it should apply to
e = types that are exceptions
θ = tolerance threshold



If *e* is below *θ*,
acquire pattern as rule

Visualization of the Tolerance Principle

N = types it should apply to
e = types that are exceptions
θ = tolerance threshold



If *e* is below *θ*,
acquire pattern as rule
Otherwise, do not form rule

Visualization of the Tolerance Principle

N = types it should apply to
e = types that are exceptions
θ = tolerance threshold

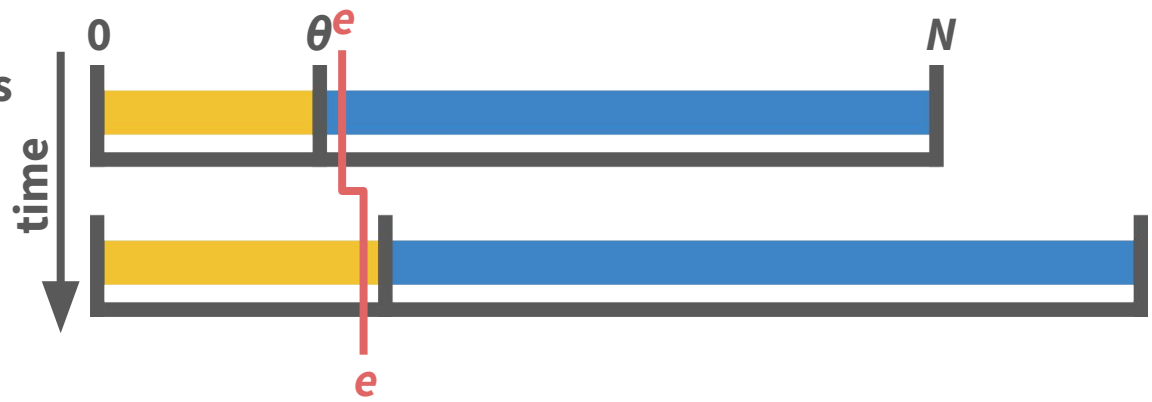
If *e* is below *θ*,
acquire pattern as rule
Otherwise, do not form rule

- *N* grows over an individual's development, *θ* grows more slowly

Visualization of the Tolerance Principle

N = types it should apply to
e = types that are exceptions
θ = tolerance threshold

If *e* is below *θ*,
acquire pattern as rule
Otherwise, do not form rule

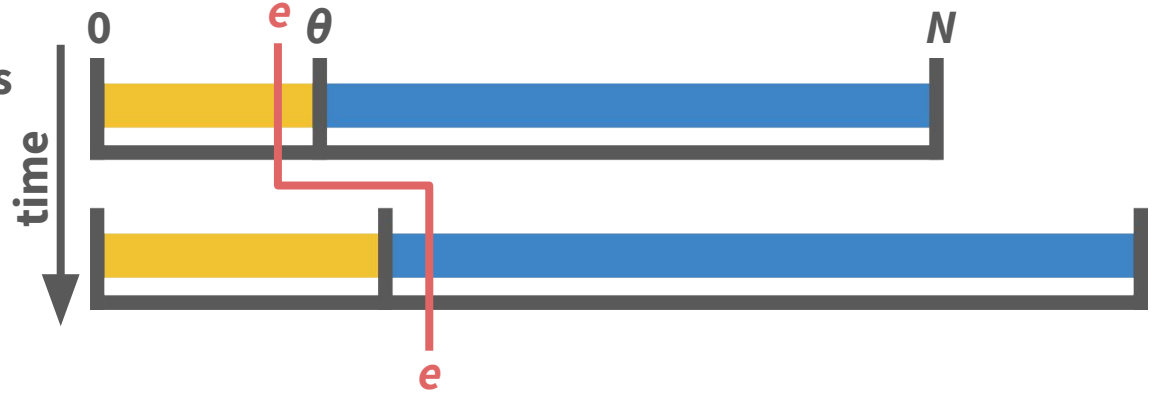


- *N* grows over an individual's development, *θ* grows more slowly
- If *θ* grows faster than *e*, a pattern may fall into productivity

Visualization of the Tolerance Principle

N = types it should apply to
e = types that are exceptions
θ = tolerance threshold

If *e* is below *θ*,
acquire pattern as rule
Otherwise, do not form rule



- *N* grows over an individual's development, *θ* grows more slowly
- If *θ* grows faster than *e*, a pattern may fall into productivity
- If *e* grows faster than *θ*, a pattern may fall out of productivity

The Tolerance Principle (Yang 2005, 2016)

- A concrete model for the acquisition of linguistic generalization
- A cognitively-motivated evaluation metric over linguistic hypotheses
- Separates the algorithmic aspects of acquisition from the representations over which generalizations are formed

A concrete quantitative model makes concrete predictions

- May or may not be intuitive or surface-obvious predictions
- Offers explanations for a range of child-driven changes as part of actuation

Successful Applications to Language Change

Syntax-Semantics / Morphosyntax

- Rise (and partial retreat) of the *to*-Dative in Middle English (Kodner, 2020)
- "Dative Sickness" in Modern Icelandic (Nowenstein et al., 2020)
- Subject-experiencer psych verbs in Middle English (Trips & Rainsford, 2022)

Morphology / Morphophonology

- Analogical extension of past participles in Late Latin and Romance (Kodner, 2022)
- "Irregularized" past tense forms in Early Modern English (Ringe & Yang, 2022)
- "Elsewhere reversal" in Iranian Armenian perfectives (w/ Hossep Dolatian)

Phonology

- Emergence of "transparent" /aɪ/-Raising (Kodner & Richter, 2020)
- Shift towards a nasal /æ/-tensing system in Philadelphia (Sneller, Fruehwald & Yang, 2018)
- "Rule reversal" in Middle High German (Richter, 2021)
- Secondary split in Menominee vowels (Richter, 2021)

Successful Applications to Language Change

Syntax-Semantics / Morphosyntax

- Rise (and partial retreat) of the *to*-Dative in Middle English (Kodner, 2020)
- "Dative Sickness" in Modern Icelandic (Nowenstein et al., 2020)
- Subject-experiencer psych verbs in Middle English (Trips & Rainsford, 2022)

Morphology / Morphophonology

- Analogical extension of past participles in Late Latin and Romance (Kodner, 2022)
- "Irregularized" past tense forms in Early Modern English (Ringe & Yang, 2022)
- "Elsewhere reversal" in Iranian Armenian perfectives (w/ Hossep Dolatian)

Phonology

- Emergence of "transparent" /aɪ/-Raising (Kodner & Richter, 2020)
- Shift towards a nasal /æ/-tensing system in Philadelphia (Sneller, Fruehwald & Yang, 2018)
- "Rule reversal" in Middle High German (Richter, 2021)
- Secondary split in Menominee vowels (Richter, 2021)

Will touch on four today

Successful Applications to Language Change

Syntax-Semantics / Morphosyntax

- Rise (and partial retreat) of the *to*-Dative in Middle English (Kodner, 2020)
- "Dative Sickness" in Modern Icelandic (Nowenstein et al., 2020)
- Subject-experiencer psych verbs in Middle English (Trips & Rainsford, 2022)

Morphology / Morphophonology

- Analogical extension of past participles in Late Latin and Romance (Kodner, 2022)
- "Irregularized" past tense forms in Early Modern English (Ringe & Yang, 2022)
- "Elsewhere reversal" in Iranian Armenian perfectives (w/ Hossep Dolatian)

Phonology

- Emergence of "transparent" /aɪ/-Raising (Kodner & Richter, 2020)
- Shift towards a nasal /æ/-tensing system in Philadelphia (Sneller, Fruehwald & Yang, 2018)
- "Rule reversal" in Middle High German (Richter, 2021)
- Secondary split in Menominee vowels (Richter, 2021)

Poster tomorrow!

Acquisition in the Past

- Children in the past must have acquired language in the same way that modern children do - this is straightforward uniformitarianism[1]
- We can reason about acquisition in the past in the same way we do now

[1] Labov 1972 as applied to linguistics, Walkden 2019

Acquisition in the Past

- Children in the past must have acquired language in the same way that modern children do - this is straightforward uniformitarianism[1]
- We can reason about acquisition in the past in the same way we do now

But where can we get data about acquisition in the past?

- We can't run experiments on subjects who are no longer alive
  With appropriate caution, we can project experimental results back to the past
- We can't do corpus or modeling work on ancient child-directed speech
  With appropriate caution, we can use historical corpora for certain data
  (Kodner, 2019)

[1] Labov 1972 as applied to linguistics, Walkden 2019

Using Historical Data

Though CDS-derived and non-CDS derived lexicons differ in terms of exact lexical makeup and other superficial corpus stats…

- **They are quantitatively indistinguishable over several linguistic dimensions**
  Including the ones relevant here today
- **When using the same processing steps are applied to extract lexicon estimates**

With appropriate processing, historical and modern adult-derived corpora may be reasonably used to approximate child linguistic experience (Kodner, 2019, *Glossa*)

Prima Facie Support - But it's complicated

Leveling in the English Past Tense

- Classically observed parallel between acquisition and change
- Mostly "regularization" towards weak *-ed* pasts

  ME *help~halp* → ModE *help~helped*

  ME *werke~wroghte* → ModE *work~worked*

- Child production errors are overwhelmingly overregularizations like these

  English estimates:[1] 4-10% overreg. ≫ 0.2% "over-irregularization"

  Spanish estimates:[2] 5% overreg. ≫ 0.1% "over-irregularization"

- But then why do "irregularization" changes happen?

  dig~diggede → *dig~dug* *stician~sticode* → *stick~stuck*

[1]Pinker & Prince 1994, Xu & Pinker 199, Maslen et al 2004, [2]Clahsen et al 1992, Mayol 2007

Prima Facie Support - But it's complicated

Leveling in the English Past Tense

- Classically observed parallel between acquisition and change
- Mostly "regularization" towards weak *-ed* pasts
  - ME *help~halp* → ModE *help~helped*
  - ME *werke~wroghte* → ModE *work~worked*
- Child production errors are overwhelmingly overregularization
  - English estimates:[1] 4-10% overreg. ≫ 0.2% "over-irregularization"
  - Spanish estimates:[2] 5% overreg. ≫ 0.1% "over-irregularization"
- **But then why do "irregularization" changes happen?**
  - *dig~diggede* → *dig~dug*
  - *stician~sticode* → *stick~stuck*

Ringe & Yang 2022
apparent irregular extensions in English past tense

[1]Pinker & Prince 1994, Xu & Pinker 199, Maslen et al 2004, [2]Clahsen et al 1992, Mayol 2007

Productivity in the Past

Regularity should be evaluated when the change occurred

- These changes didn't happen yesterday. They happened in Early Mod English
- Oxford English Dictionary - records first written attestations
- PPCEME[1] - Used to identify high frequency verbs of the time

[1] Kroch, Santorini, & Delfs 2004

Productivity in the Past

Regularity should be evaluated when the change occurred

- These changes didn't happen yesterday. They happened in Early Mod English
- Oxford English Dictionary - records first written attestations
- PPCEME[1] - Used to identify high frequency verbs of the time

Dig~Digged → Dig~Dug

Phonologically related forms at the time

bring~brought	*stick~stuck*	*sing~sang/sung*
dig~digged	*sting~stung*	*spring~sprang/sprung*
pick~picked	*fling~flung*	*ring~rung*
nick~nicked	*strike~struck*	sling~*slung*
	wring~wrung	

Rule: "/ɪ/ → /ʊ/"

$N = 14$

$e = 4 < \theta = 4.3$

Over-regularization was possible

Productivity in the Past

Regularity should be evaluated when the change occurred

- These changes didn't happen yesterday. They happened in Early Mod English
- Oxford English Dictionary - records first written attestations
- PPCEME[1] - Used to identify high frequency verbs of the time

Why not *click~*cluck*, *flick~*fluck*?

No words innovated "/ɪ/ → /ʊ/" pasts if they entered the lexicon after past *sang*, *sprang*, *rang* thoroughly replaced past *sung*, *sprung*, *rung*

→ with the lost of these three *-ung* past forms, the generalization was no longer productive, so these forms would not be generated by learners.

[1] Kroch, Santorini, & Delfs 2004

The rise of < *-ūtus* Past Participles in Late Latin and Romance

Classical Latin Principal Parts and Conjugations

- Traditionally classified into 4½ conjugations distinguished by 4 principal parts
- Conjugations correspond to theme vowels, principal parts to stems

Principal parts

1. present active indicative 1sg
2. present active infinitive
3. perfect active indicative 1sg
4. past participle (or supine)

Conj.	ThV	1st PP	2nd PP	3rd PP	4th PP	Meaning
		present stem		perfect	pptc	
1st	*ā*	*amō*	*amāre*	*amāvī*	*amātus*	‘love’
2nd	*ē*	*moneō*	*monēre*	*monuī*	*monitus*	‘warn’
3rd	*e*	*legō*	*lēgere*	*lēgī*	*lēctus*	‘choose’
3rd *-iō*	*i*	*capiō*	*capere*	*cēpī*	*captus*	‘take’
4th	*ī*	*audiō*	*audīre*	*audīvī*	*audītus*	‘hear’

Complex Forms of the Past Participle

Verbs with similar forms for one stem may not have similar forms for the others

We can find patterns!
But which patterns "matter?"

Present		Perfect	PPtc	Meaning
amō	*amāre*	*amāvī*	*amātus*	'love'
sonō	*sonāre*	*sonuī*	*sonitus*	'sound'
moneō	*monēre*	*monuī*	*monitus*	'warn'
maneō	*manēre*	*mānsī*	*mānsus*	'stay'
teneō	*tenēre*	*tenuī*	*tentus*	'hold'
audiō	*audīre*	*audīvī*	*audītus*	'hear'
pellō	*pellere*	*pepulī*	*pulsus*	'push'
capiō	*capere*	*cēpī*	*captus*	'take'
ferō	*ferre*	*tulī*	*lātus*	'carry'
5 forms		7 forms	7 forms	

Conjugations and PPtcs by Type Count

Out of the most frequent in the OL/CL part of the Perseus collection,

- 1st conjugation is largest and most homogeneous
- 3rd conjugation is second largest and most heterogeneous
- *-itus* and *-tus* are the most common pptcs outside the 1st conjugation

Conjugation	# Verbs	Top freq		% Top	Next most		% Top two
1st	541	*-ātus*	528	97.6%	*-itus*	6	98.7%
2nd	65	*-itus*	25	38.5%	*-tus*	17	64.6%
3rd	215	*-tus*	80	37.2%	*-itus*	19	46.6%
4th	55	*-ītus*	34	61.8%	*-tus*	13	87.3%

A Diachronic Mystery

Developments in Late Latin[1]

- Three productive LL pptcs: **-atu* < *-ātus*, **-itu* < *-ītus* (not *-ĭtus*), **-utu* < *-ūtus*
- *-ĭtus* and *-tus* lose out to **-itu* and **-utu*

-ūtus ousted statistically predominant competitors

- In CL, *-ūtus* applied to only about a dozen verbs (eg *solvō ~ solūtus*)
- *-itus* and *-tus* were the most common for 2nd/3rd conjugation verbs!
- *-ūtus* spread first among *-uī* perfects (common in 2nd/3rd conjugations)
- *-tus* is the source of inherited modern irregulars (eg, *scritto*, *escrito*, < *scrip-tus*)

Why should they have lost out to this upstart **-utu*?

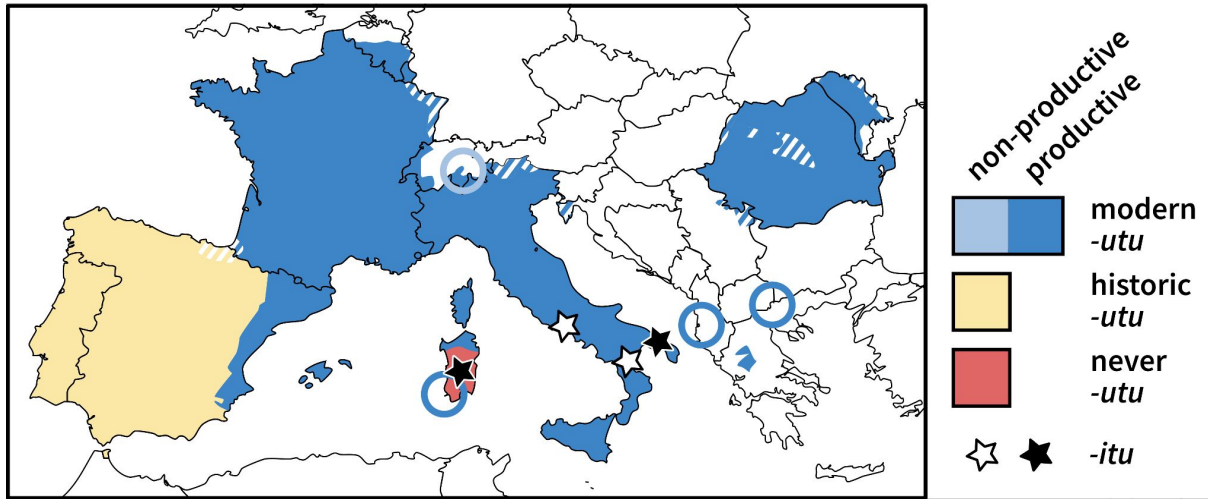
[1] Laurent 2003 §2.14

Reflexes of *-ūtus* and *-ĭtus* in Attested Romance[1]

- Reflexives of *-ūtus* constitute the default (apparently productive) for at least some class in most Romance languages
- Reflexes are attested in Old Spanish and Old Portuguese but have been lost
- *-ĭtus* remains productive in ★Apulian and ★Sardinian

  /i/ merged with /i:/ in ★Sardinian, causing *-ĭtus* to fall together with *-ītus*

Past Participle Reflexes of **-utu* and **-itu* in Romance



[1] data compiled from Laurent 2003

Example Calculation

Is *-ātus* the productive pptc derivation for verbs with ThV *ā* at *n*=500?

A typical child who knows *n*=500 verbs knows

- *N*=221 ThV *ā* verbs
- *e*=13 with non-*-ātus* pptcs
- *θ*=40.94 tolerance threshold

Exceptions are tolerable if

$$13 < 40.9$$

$$\theta = N / \ln N$$

-ātus is productive for *ā* verbs at *n*=500

Summary results for Past Participles

All Productive Patterns

- 1st → *-ātus*
- 3rd-*iō* → *-tus*
- *-sī* perfect → *-tus*
- *-īvī* perfect → *-ītus*
- *-ēvī* perfect → *-ētus*
- *faveō*-type → *-autus*/*-ōtus*
- *solvō*-type → *-ūtus*

Selected Unproductive Patterns

- 2nd → *-tus, -itus…*
- 3rd → *-tus, -itus…*
- 4th → *-tus, -itus, -ītus…*
- *-uī* perfect → *-tus, -itus…*
- bare perfect → *-tus, -itus…*

Summary results for Past Participles

All Productive Patterns

- 1st → *-ātus*
- 3rd-*iō* → *-tus*
- *-sī* perfect → *-tus*
- *-īvī* perfect → *-ītus*
- *-ēvī* perfect → *-ētus*
- *faveō*-type → *-autus*/*-ōtus*
- *solvō*-type → *-ūtus*

Selected Unproductive Patterns

- 2nd → *-tus, -itus…*
- 3rd → *-tus, -itus…*
- 4th → *-tus, -itus, -ītus…*
- *-uī* perfect → *-tus, -itus…*
- bare perfect → *-tus, -itus…*

Large gaps here

- *-tus* is not productive for any large class.
  -itus is not for any class.
- 2nd and 3rd conjugation and *-uī* perfect verbs are mostly uncovered by productive patterns

Summary results for Past Participles

All Productive Patterns

- 1st → *-ātus*
- 3rd-*iō* → *-tus*
- *-sī* perfect → *-tus*
- *-īvī* perfect → *-ītus*
- *-ēvī* perfect → *-ētus*
- *faveō*-type → *-autus*/*-ōtus*
- *solvō*-type → *-ūtus*

These are the only *-uī* perfect verbs with productively derived past participles

Selected Unproductive Patterns

- 2nd → *-tus, -itus…*
- 3rd → *-tus, -itus…*
- 4th → *-tus, -itus, -ītus…*
- *-uī* perfect → *-tus, -itus…*
- bare perfect → *-tus, -itus…*

Large gaps here

- *-tus* is not productive for any large class. *-itus* is not for any class.
- 2nd and 3rd conjugation and *-uī* perfect verbs are mostly uncovered by productive patterns

Pathway for the Rise of Romance *-*utu*

- Latin *-ūtus* was productive “within its sphere” while its statistically predominant competitors like *-itus* and *-tus* were not!

An Account

- Productively derived forms tend expand at the expense of the unproductive
- Unproductive patterns are at the mercy of attestation and memorization
  → Always at risk at being pushed out by productive patterns
- *-ūtus* was productive for a small class, the only option for *-uī* perfects!
- It spread first among *-uī* perfects, pushing out unproductive *-ĭtus* and *-tus*

No competition, “a big fish in a small pond”

Secondary Split in 20th Century Menominee

from Richter (2021)

Ō-Raising in Menominee

Menominee *Ō*-RAISING: *ō* > *ū* / _[+hi vowel or glide later in the word]

Unraised	Raised
kōn ‘snow’	*kūnyak* ‘lumps of snow’
watōp ‘alder’	*watūpyak* ‘alders’
āteqnōhkew ‘he tells a sacred story’	*āteqnūhkuwew* ‘he tells them a sacred story’
pōset ‘when he embarks’	*pūsetwaq* ‘when they embark’

Ō-Raising in Menominee

Menominee *Ō*-RAISING: *ō* > *ū* / _[+hi vowel or glide later in the word]

Unraised	Raised
kōn ‘snow’	*kūnyak* ‘lumps of snow’
watōp ‘alder’	*watūpyak* ‘alders’
āteqnōhkew ‘he tells a sacred story’	*āteqnūhkuwew* ‘he tells them a sacred story’
pōset ‘when he embarks’	*pūsetwaq* ‘when they embark’

Not all *ō*/*ū* alternate - roots with high vowels, onomatopoeia, borrowings

sūniyan ‘coin’ ~ *sūniyak* ‘coins’

ōhōpīwēkat ‘there is a whooping sound’

čapūq ‘sploosh!’

Cōh ‘Joe’ vs *Cūh* ‘Jew’

Vowel Shortening and Short Vowel Neutralization

Menominee has a series of vowel shortening rules

- **Famously complex.[1] Some shortening is typologically unusual**
  Under some conditions, long vowels are shortened in open syllables[2]
- **Underlying *ō* and raised *ū* may be surface as *o* and *u***

A phonetic change:

Short vowel neutralization spread during the 20th Century[3]

[1] Bloomfield 1962, [2] Bloomfield 1962, Buckley 2000, Cudworth 2019, Milligan 2005, Miner 1975, Hayes 1995, [3] Cudworth 2019, Milligan 2005

Learning Whether *ū* is Phonemic

A learner's hypothesis H0: "If surface *ū* then underlying *ū*"

- **If upheld, *ū* is phonemic, if not, it is derived allophonically**
- **Observed alternations provide evidence against H0**
- **Failure to reject H0 → a split between *ō* and *ū***

Loss of Short Vowel Contrasts Triggers the Split

A learner's hypothesis H0: "If surface *ū* then underlying *ū*"

- **Failure to reject H0 → a split between *ō* and *ū***
- **When the distinction between short *u* and *o* turned unreliable → accepted H0**
- **Data from the Menominee Dictionary[1]**

[1] McCauley & Menominee Tribe of WI 2012, compiled from consultations with Menominee elders and Bloomfield 1975

Loss of Short Vowel Contrasts Triggers the Split

A learner's hypothesis H0: "If surface *ū* then underlying *ū*"

- Failure to reject H0 → a split between *ō* and *ū*
- When the distinction between short *u* and *o* turned unreliable → accepted H0
- Data from the Menominee Dictionary[1]

"Pre-Neutralized" State

ū~*ō* and ~*o* count against H0

N = 231, *e* = 88 → *e* > *θ* = 42.4.

→ *ū* is not underlying

"Neutralized" State

Only *ū*~*ō* counts against H0

short *u* and *o* are not distinguished

N = 231, *e* = 31 → *e* < *θ* = 42.4.

→ *ū* is underlying → the split.

[1] McCauley & Menominee Tribe of WI 2012, compiled from consultations with Menominee elders and Bloomfield 1975

The Rise of the *to*-Dative in Middle English

Rise of the *to*-Dative in Middle English[1]

The Change

Old English

- No (or heavily restricted) *to*-dative[1]
- DO-IO and IO-DO double object[1]
- Overt DAT-ACC case marking[2]

Modern English

- *to*-dative competes with IO-DO
- Semantic, phonological, and apparently arbitrary restrictions[3]
- No overt DAT-ACC distinction

[1] Visser 1963, Mitchell 1985, Allen 1995, [2] with exceptions. Allen 1995 p. 29, [3] verb sensitive approach: Levin 2008, Rappaport Hovav & Levin 2008, other approaches: Pinker 1989, Goldberg 1995. Exceptions to exceptions: Storm 1977, Pinker 1989, Gropen et al 1989, Levin 1993

Rise of the *to*-Dative in Middle English[1]

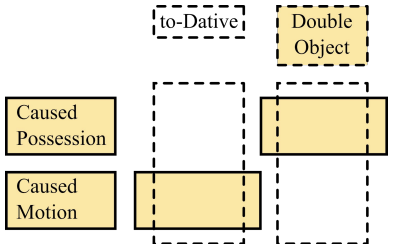
The Problem

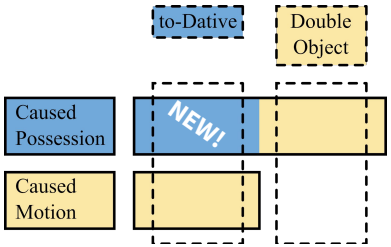
- **Change was rapid across the lexicon**
  → semantic expansion was too rapid to be attested in writing[1]
- **Poor temporal/geographical correlation with loss of case marking and DO-IO**
  Problems extend to equivalent constructions in North Germanic
  → a morphology/syntax trade-off is hard to justify
- **The *to*-dative actually exceeded the modern distribution before retreating[2]**
  → Can the advance and retreat be accounted for with the same mechanism?

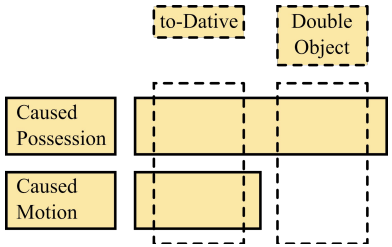
[1] Elter 2018, Kodner 2020, [2] Visser 1963

Timeline of the English *to*-Dative

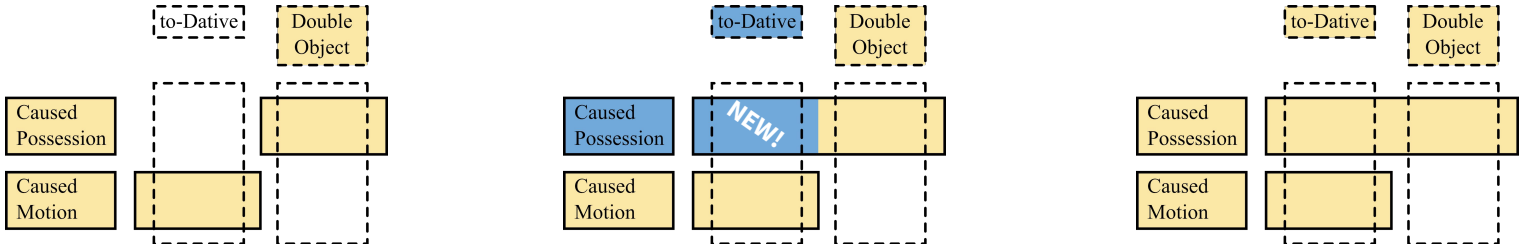
Visualization: "Broad-range" semantic classes
(cf Rappaport Hovav & Levin)

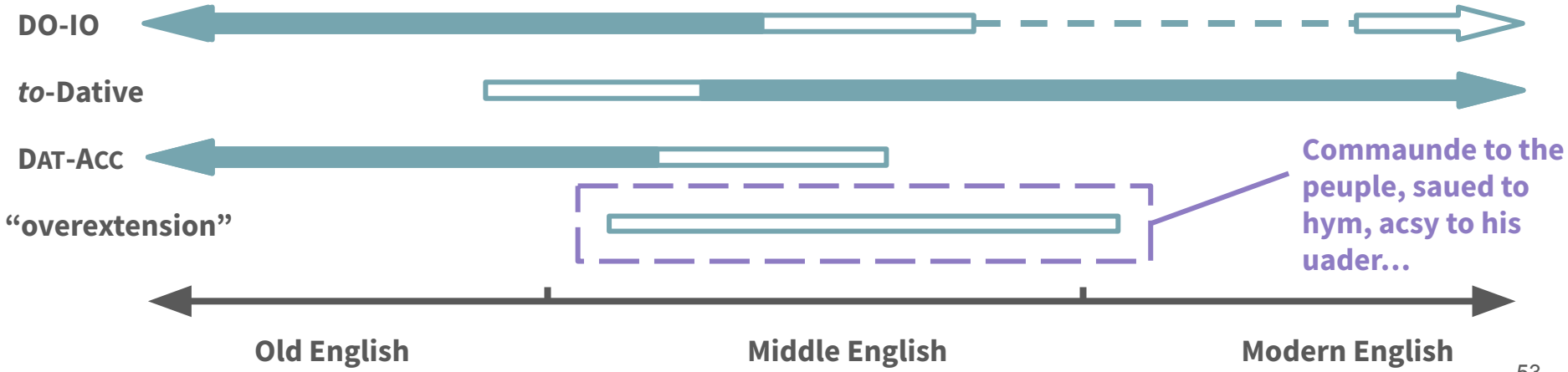






Timeline of the English *to*-Dative





Acquiring the Modern Dative Alternation[1]

Consider narrow generalizations: one for each narrow-range class[2]

- Each class has its own *N*, *m*, θ according to that child's experience

- These numbers are estimated from text corpora for a "typical" child
- A frequency cutoff gives a child-like lexicon size and composition

[1] Yang 2016, [1] Gropen et al 1989, Levin 1993

Summarizing the Rise

Doub Obj + to-Dat	Generalize
CLASS 1	YES
CLASS 2	YES

to-Dative Only	Generalize
CLASS 3	YES
CLASS 4	YES

Doub Object Only	Generalize
CLASS 5	no

- Application to verbs attested in Middle English (PPCME2)[1] yields the "overextended" distribution attested in Middle English.
- Adding verbs first attested during the 15th century effects the retreat to the modern distribution.

As a concrete mechanism for learner actuation

Explanatory and Empirical Advantages

- **Quantitative predictions about the relationship between the input, language acquisition, and the actuation of change**
  → Falsifiable with further empirical investigation
- **Serves as an explanatory model for actuation events**
  → A way to asymptotically approach the actuation point
- **Suggests where acquisition/corpus/socio research can look next**

Not all Change is Driven by Acquisition

To a 1st-degree approximation, children are responsible for discrete rather than continuous changes

Discrete Changes

- Categorical properties of the grammar
- New or lost structures or constructions
- May be fixed over individuals' lifetimes[1]
- The realm of child language acquisition

Continuous Changes

- The stereotypical subjects of variationist sociolinguistics
- Positions in the vowel space, usage frequencies, optionality
- Variable over lifetimes
- Not only child language acquisition

[1] Andersson 1995, Sankoff & Blondeau 2007, Nycz 2013

Discrete and Continuous Changes

Two Sides of One Coin

- Once a discrete innovation enters the population, it becomes variation[1]
- Variationism concerns [continuous=] distribution of discrete choices[2]
- So do competing grammars in historical syntax and morphology[3]

Actuation = Innovation + uptake into the speech community

(The hand-off from an individual-level process to a population-level one)

[1] Kroch 2005, [2] Sankoff 1988, [3] Kroch 1994

Discrete and Continuous Changes

Two Sides of One Coin

- Once a discrete innovation enters the population, it becomes variation[1]
- Variationism concerns [continuous=] distribution of discrete choices[2]
- So do competing grammars in historical syntax and morphology[3]

Actuation = Innovation + uptake into the speech community

(The hand-off from an individual-level process to a population-level one)

[1] Kroch 2005, [2] Sankoff 1988, [3] Kroch 1994

Transmission is not Strictly Linear and Generational

- Children mature in communities and receive input from multiple speakers
- Community input is formally necessary for attested dynamics of change[1]
- Young children learn sociolinguistic variables[2]
- Children attend to input from older children[3] who are not linguistically mature
- Multiple competing targets may be present in the input

Everybody receives input from multiple grammars

"Monolingual" "Multilingual"

Multi-idiolect multi-dialectal traditional multilingual

[1] Niyogi & Berwick 2009, [2] Labov 1989, Anderson 1990, [3] Manly 1930, Weinreich, Labov & Herzog 1968 p 145, Roberts and Labov 1995, Labov 2001 p449, Nardy, Chevrot & Barbu 2014

Learner Innovation ≠ Learner Error

Innovations need not be due to “errors”

Errors - “Blame the Child”

- The learner does not act correctly on its input “a buggy algorithm”
- errors presuppose appropriate evidence and an available target

Non-errors - “Blame the Environment”

- The learner acts correctly but is dealt a bad input sample
- Even for a good algorithm, “garbage in, garbage out”
- Change in the face of severely underspecified input or even trivial variation

Conceptualizing the Hand-Off

Solution to the Paradox of Language Change

- **Children are good at acquisition, but it's still hard!**
- **Learning targets are obscured by**
  - **Ambiguous surface constructions**
  - **Variation of all kinds in the input**
  - **Severe skew and sparsity in the input**

→ **So even a "perfect" learner can initiate change - "blame the environment"**

A thought experiment: "Sibling-Induced Change"

“Sibling-Induced Change”

Imagine two young children, Alice is slightly older than Bob

- Alice is currently producing innovative forms
- Bob is receiving both conservative adult input and Alice’s

How does this effect Bob?

“Sibling-Induced Change”

Can Bob identify Alice’s innovation?

- Alice is mostly consistent with adults
- Bob may rarely if ever hear a conservative token corresponding Alice’s
- If Bob never hears a conservative token, he cannot know if Alice is innovating

“Sibling-Induced Change”

Can Bob identify Alice’s innovation?

- Alice is mostly consistent with adults
- Bob may rarely if ever hear a conservative token corresponding Alice’s
- If Bob never hears a conservative token, he cannot know if Alice is innovating

Will Bob adopt Alice’s innovation?

- In cases of severe sparsity, yes. What choice does he have?
- In other cases,

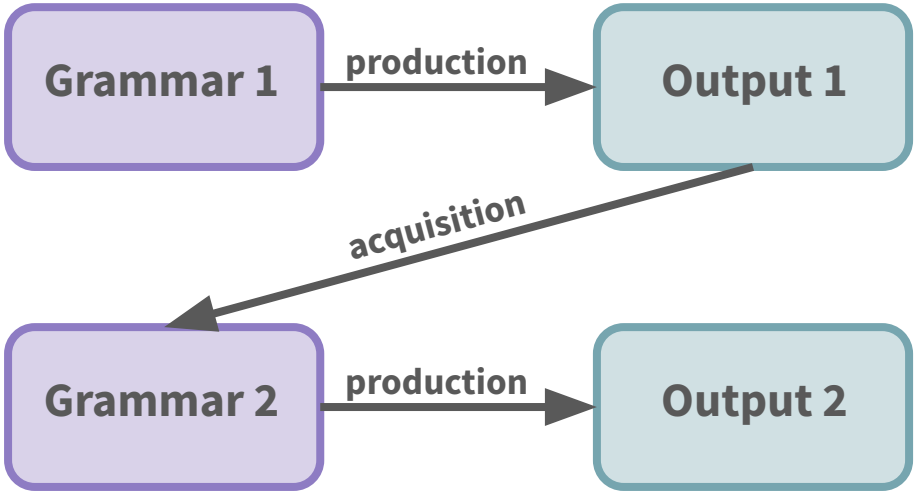
  Even young children orient toward peers

  Bob may prefer Alice’s forms over his parents

  He could learn both! (Competing grammars and sociolinguistic variation)

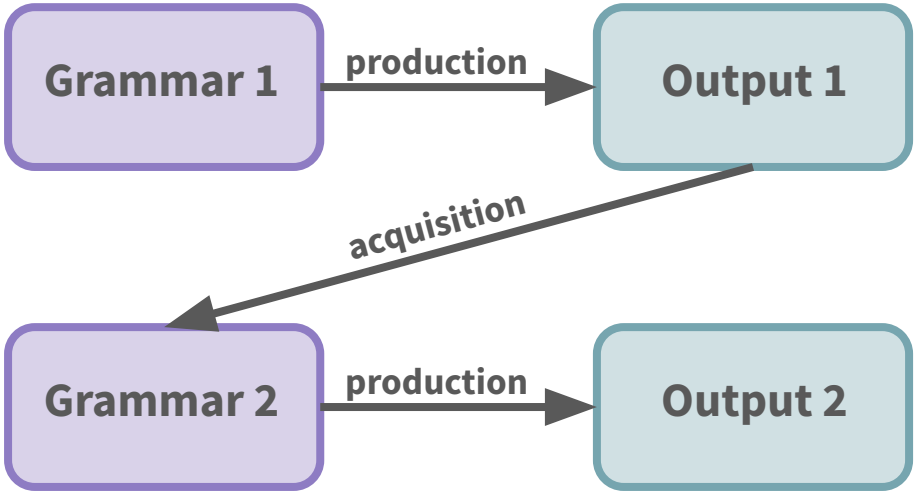
Z-Model of Language Acquisition and Change[1]

- A cycle of error-prone abductive and inductive learning
- Amenable to many interpretations



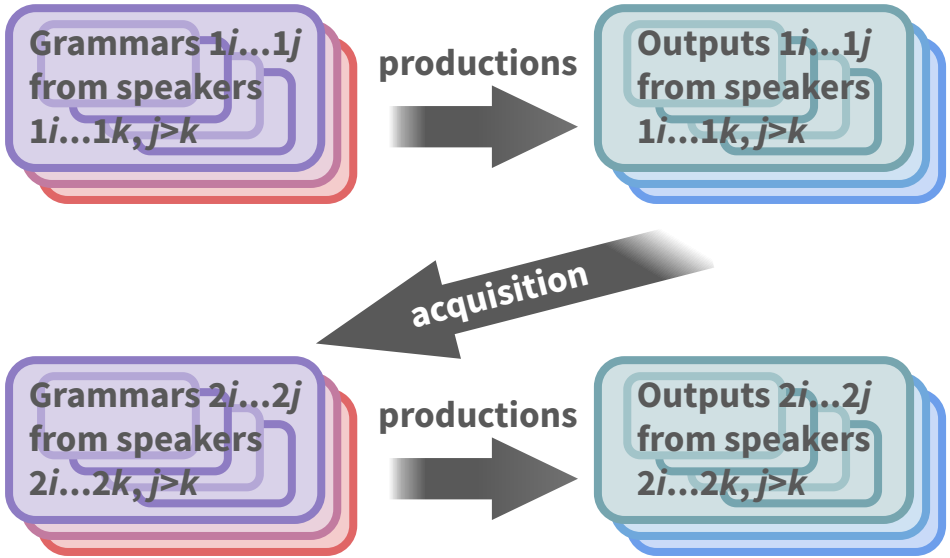
[1] Andersen 1973

Insufficiency of the Z-Model



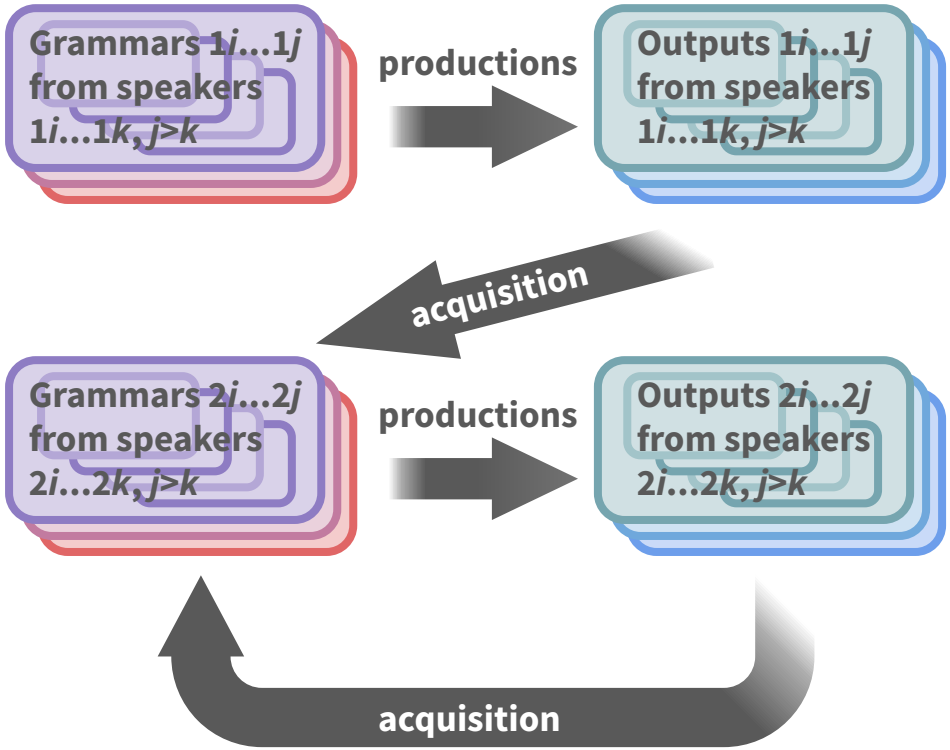
Insufficiency of the Z-Model

- **Individual production**
  - Variation across social settings
  - Variation over lifetimes
- **Community Embedding**
  - Variation across people
  - Everyone receives many inputs



Insufficiency of the Z-Model

- **Individual production**
  - Variation across social settings
  - Variation over lifetimes
- **Community Embedding**
  - Variation across people
  - Everyone receives many inputs
- **Gradual Maturation**
  - Transmission isn't just generational
  - Acquisition takes time
  - Immature learners influence others
  - "Sibling-Induced Change"



Additional Predictions

Relationship between learning trajectories and change

- **Innovations need to occur/be sustained late enough to be transmitted to peers**
  → Errors that occur early should not be transmitted, even if frequent
- **Late childhood innovations correspond to common trajectories of change**
  morphological overregularization, changes in modal semantics,
  certain innovations in argument structure, certain phonological rules…
- **Early childhood innovations do not correspond to common changes**
  Consonant harmony, dramatic phonotactic simplification + reduplication,
  instant total loss of inflection…

Proofs-of-Concept

1. As a baseline for paradigm trade-offs

- **It is sufficient on its own to reproduce**
  - Correlations between token frequency and irregularity
  - Correlations between paradigm size and irregularity
- **A much richer model than iterated learning**
  - Includes a population ← change is population-level!
  - Does not privilege generational transmission

2. Modeling semantic change in Chinese Classifier Systems (Kali & Kodner 2022)


Addressing the Symposium Title Question:
Some types of phonological and syntactic change
share underlying mechanisms of actuation
The End.
Thank you!