Acquiring the Latin Past Participles

Synchronic and Diachronic Implications

Jordan Kodner
University of Pennsylvania

LSA, Jan 3, 2020
New Orleans

Outline

- **The Classical Latin Past Participles**
- Acquiring Morphological Generalizations
- Language Acquisition in the Past
- Predictability of the Past Participles
- Synchronic and Diachronic Implications

Classical Latin Principal Parts and Conjugations

- Traditionally classified into 4.5 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’

The Principal Parts and Conjugations

- **Stems are not reliably derivable from one another**

1st PP	2nd PP	3rd PP	4th PP
amō	*amāre*	*amāvī*	*amātus*
sonō	*sonāre*	*sonuī*	*sonitus*
moneō	*monēre*	*monuī*	*monitus*
maneō	*manēre*	*mānsī*	*mānsus*
teneō	*tenēre*	*tenuī*	*tentus*
audiō	*audīre*	*audīvī*	*auditus*
pellō	*pellere*	*pepulī*	*pulsus*
capiō	*capere*	*cēpī*	*captus*
ferō	*ferre*	*tulī*	*lātus*

The Principal Parts and Conjugations

- Stems are not reliably derivable from one another

Verbs with similar stems in one column may not have similar stems in the others

1st PP	2nd PP	3rd PP	4th PP
amō	*amāre*	*amāvī*	*amātus*
sonō	*sonāre*	*sonuī*	*sonitus*
moneō	*monēre*	*monuī*	*monitus*
maneō	*manēre*	*mānsī*	*mānsus*
teneō	*tenēre*	*tenuī*	*tentus*
audiō	*audīre*	*audīvī*	*auditus*
pellō	*pellere*	*pepulī*	*pulsus*
capiō	*capere*	*cēpī*	*captus*
ferō	*ferre*	*tulī*	*lātus*

“Regularity” of the Conjugations

- Many past participles are not predictably derivable from the present stem
- Traditionally noted that 1st is overwhelmingly regular, 2nd and 4th are mostly regular, 3rd is not[1]

Conjugation[1]	# Verbs	# “Regular”	% “Regular”	Form
1st	360	345	96%	-ātus
2nd	120	90	75%	-itus/-tus
3rd	170	60	35%	-itus
4th	60	40	67%	-ītus

[1] eg Aronoff 1994, [2] Table from Laurent 2003 expanded from Aronoff 1994

“Regularity” of the Conjugations

- Many past participles are not predictably derivable from the present stem
- Traditionally noted that 1st is overwhelmingly regular, 2nd and 4th are mostly regular, 3rd is not[1]

What counts as regular?

Conjugation[1]	# Verbs	# “Regular”	% “Regular”	Form
1st	360	345	96%	-ātus
2nd	120	90	75%	-itus/-tus
3rd	170	60	35%	-itus
4th	60	40	67%	-ītus

[1] eg Aronoff 1994, [2] Table from Laurent 2003 expanded from Aronoff 1994

The Elsewhere Condition

Listing vs Derivations

- A common trade-off in theoretical morphology
- “Regular” patterns are derived, “irregulars” are listed exceptions

The Elsewhere Condition

Listing vs Derivations

- A common trade-off in theoretical morphology
- “Regular” patterns are derived, “irregulars” are listed exceptions

Applied to the Classical Latin PPtcs,

- Which pptcs really are productively derived?
- Is the pptc derived from the present, perfect, or neither?
- What other than the theme vowel cues speakers?

Leveraging Child Language Acquisition

- Determination of productive patterns is a central question in acquisition
- Exemplified by the English “Past Tense Debate”[1]
  - How are patterns and exceptions learned?
  - How are developmental trajectories explained?

[1] Rumelhart & McClelland 1986, Pinker & Prince 1988, Pinker 1994, Albright & Hayes 2006, Yang 2005, *and many more*

Leveraging Child Language Acquisition

- Determination of productive patterns is a central question in acquisition
- Exemplified by the English "Past Tense Debate"[1]
  - How are patterns and exceptions learned?
  - How are developmental trajectories explained?

Virtually everyone agrees:

it isn't just token frequency (and derived measures)![2]

[1] Rumelhart & McClelland 1986, Pinker & Prince 1988, Pinker 1994, Albright & Hayes 2006, Yang 2005, *and many more*
[2] Aronoff 1976, MacWhinney 1978, Bybee 1985, Baayen 1993, Elman 1998, Pierrehumbert 2003, Yang 2016

Leveraging Child Language Acquisition

- Determination of productive patterns is a central question in acquisition
- Exemplified by the English "Past Tense Debate"[1]
  - How are patterns and exceptions learned?
  - How are developmental trajectories explained?

Virtually everyone agrees:

it isn't just token frequency (and derived measures)![2]

→ Quantitative corpus analysis alone won't cut it

→ Should work through the implications of some concrete learning mechanism

[1] Rumelhart & McClelland 1986, Pinker & Prince 1988, Pinker 1994, Albright & Hayes 2006, Yang 2005, *and many more*
[2] Aronoff 1976, MacWhinney 1978, Bybee 1985, Baayen 1993, Elman 1998, Pierrehumbert 2003, Yang 2016

Outline

- The Classical Latin Past Participles
- **Acquiring Morphological Generalizations**
- Language Acquisition in the Past
- Predictability of the Past Participles
- Synchronic and Diachronic Implications

The Tolerance Principle[1]

- A concrete model for the acquisition of linguistic generalization
- Developed in the context of the Past Tense Debate

Example Applications

- Is *+ed* the default past for English verbs?
- Is vowel mutation as in *sing~sang* productive among similar verbs?

[1] Yang 2016

The Tolerance Principle

- An evaluation metric[1] over linguistic hypotheses
- Is derived from
  - an Elsewhere Condition for 'rules' and 'exceptions'[2]
  - frequency-rank correlated lexical access[3]
  - Generally Zipfian input distributions
- Received psychological backing from artificial language learning experiments[4]

[1] Chomsky 1955, 1965, Chomsky & Halle 1968, [2] Anderson 1969, *inter alia,* [3] Murray & Forster 2004, [4] Schuler et al 2017

The Tolerance Principle

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

The Tolerance Principle

Given a hypothesized generalization *R* operating over a class *C*, 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

$$e < \theta$$

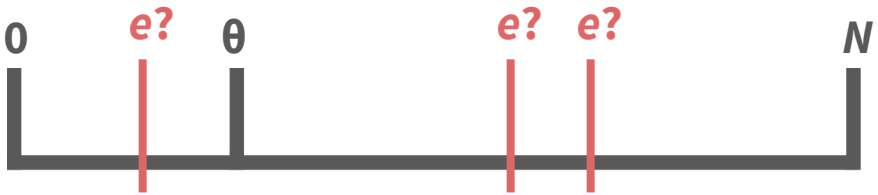
$$\theta = N / \ln N$$

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*

Visualization of the Tolerance Principle

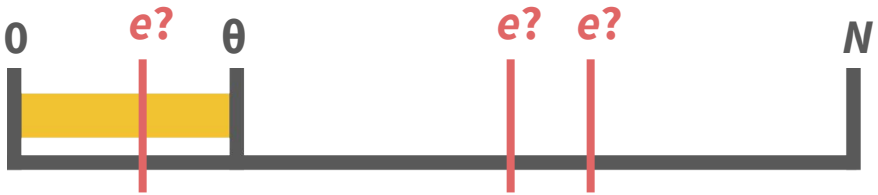
N = types it should apply to
e = types that are exceptions
θ = tolerance threshold



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

Visualization of the Tolerance Principle

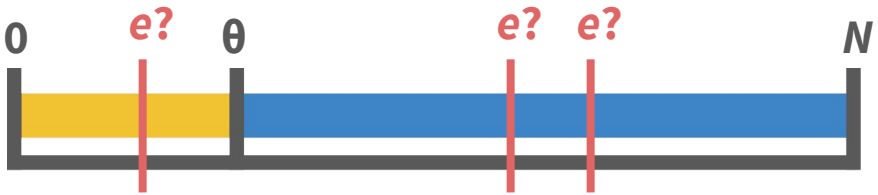
N = types it should apply to
e = types that are exceptions
θ = tolerance threshold



If *e* is below θ,
acquire generalization

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 generalization
Otherwise, do not generalize

Visualization of the Tolerance Principle

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

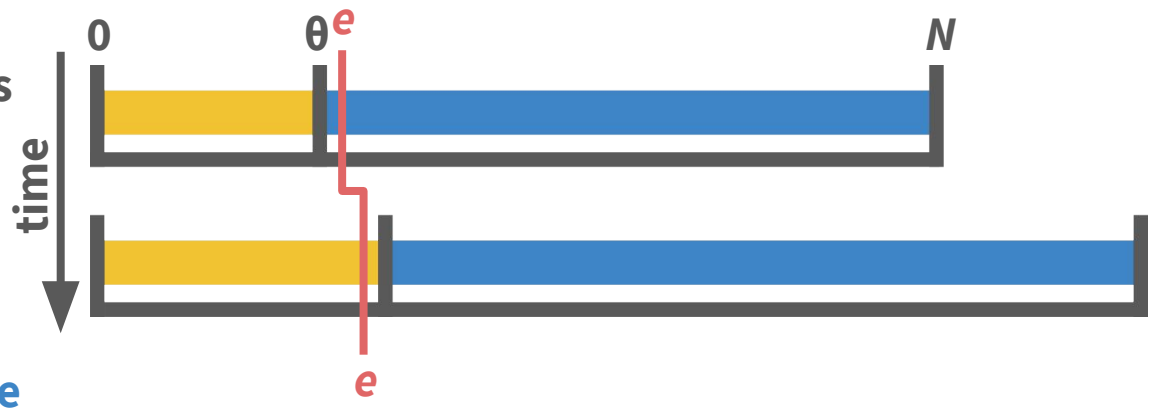
If *e* is below θ,
acquire generalization
Otherwise, do not generalize

- *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 generalization
Otherwise, do not generalize

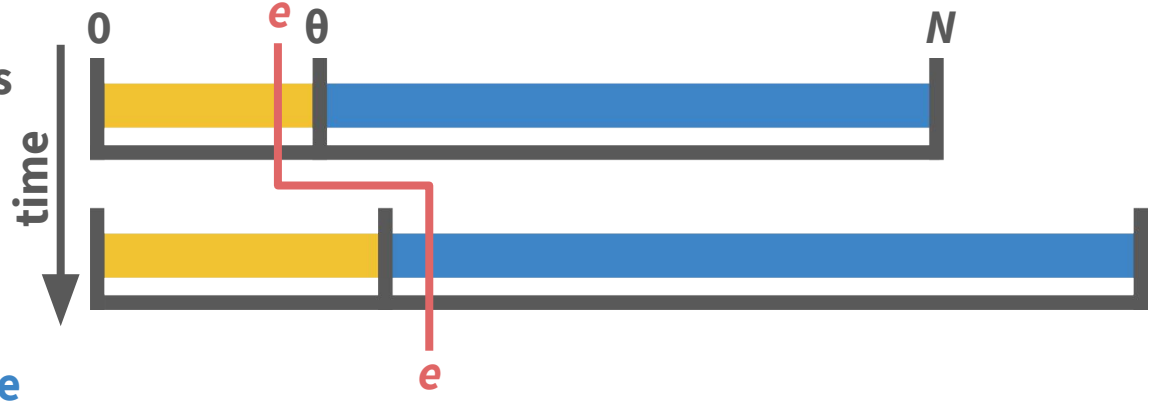


- *N* grows over an individual's development, θ grows more slowly
- If θ grows faster than *e*, a generalization may fall into productivity

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 generalization
Otherwise, do not generalize



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

Child Lexical Knowledge

- Learners' vocabularies grow over the course of development
- There is significant individual variation, but consistent trends
- **Only on the order of 10^2** for English and German learners by around age 3
- Children have the foundations for language-specific grammars by this point

A roughly 1 per million frequency cutoff applied to the larger CHILDES corpora yields lexicons like these[1]

Language	Estimated |Vocab|
English 2;10-3;0[2]	525-1,116
German 2;6[3]	$\mu = 429, \sigma > 100$

[1] Nagy & Anderson 1984, [2] Hart & Risley 2003, [3] Szagun et al 2006

Outline

- The Classical Latin Past Participles
- Acquiring Morphological Generalizations
- **Language Acquisition in the Past**
- Predictability of the Past Participles
- Synchronic and Diachronic Implications

Four Features of Native Language Acquisition

1. **All children receive unique input yet exhibit gross developmental uniformity**[1]
2. The type frequency of a pattern is crucial for acquisition of generalizations, as opposed to token frequency or attestation of initial items[2]
3. Token frequencies correlate with relative order of acquisition[3]
4. Early learner vocabularies are small[4]

[1] Labov 1972, [2] Aronoff 1976, MacWhinney 1978, Bybee 1985, Baayen 1993, Elman 1998, Pierrehumbert 2003, Yang 2016, [3] Goodman 2008, [4] Hart & Risley 1995, 2003, Szagun et al. 2006

Four Features of Native Language Acquisition

1. All children receive unique input yet exhibit gross developmental uniformity[1]
2. The type frequency of a pattern is crucial for acquisition of generalizations, as opposed to token frequency or attestation of initial items[2]
3. Token frequencies correlate with relative order of acquisition[3]
4. Early learner vocabularies are small[4]

[1] Labov 1972, [2] Aronoff 1976, MacWhinney 1978, Bybee 1985, Baayen 1993, Elman 1998, Pierrehumbert 2003, Yang 2016, [3] Goodman 2008, [4] Hart & Risley 1995, 2003, Szagun et al. 2006

Four Features of Native Language Acquisition

1. All children receive unique input yet exhibit gross developmental uniformity[1]
2. The type frequency of a pattern is crucial for acquisition of generalizations, as opposed to token frequency or attestation of initial items[2]
3. Token frequencies correlate with relative order of acquisition[3]
4. Early learner vocabularies are small[4]

[1] Labov 1972, [2] Aronoff 1976, MacWhinney 1978, Bybee 1985, Baayen 1993, Elman 1998, Pierrehumbert 2003, Yang 2016, [3] Goodman 2008, [4] Hart & Risley 1995, 2003, Szagun et al. 2006

Four Features of Native Language Acquisition

1. All children receive unique input yet exhibit gross developmental uniformity[1]
2. The type frequency of a pattern is crucial for acquisition of generalizations, as opposed to token frequency or attestation of initial items[2]
3. Token frequencies correlate with relative order of acquisition[3]
4. Early learner vocabularies are small[4]

[1] Labov 1972, [2] Aronoff 1976, MacWhinney 1978, Bybee 1985, Baayen 1993, Elman 1998, Pierrehumbert 2003, Yang 2016, [3] Goodman 2008, [4] Hart & Risley 1995, 2003, Szagun et al. 2006

Four Features of Native Language Acquisition

1. All children receive unique input yet exhibit gross developmental uniformity[1]
2. The type frequency of a pattern is crucial for acquisition of generalizations, as opposed to token frequency or attestation of initial items[2]
3. Token frequencies correlate with relative order of acquisition[3]
4. Early learner vocabularies are small[4]

As a result,

- Applying a frequency cutoff to lemmas in CDS approximates a “typical” child
- Insight taken by type frequency-based models of acquisition[5]

[1] Labov 1972, [2] Aronoff 1976, MacWhinney 1978, Bybee 1985, Baayen 1993, Elman 1998, Pierrehumbert 2003, Yang 2016, [3] Goodman 2008, [4] Hart & Risley 1995, 2003, Szagun et al. 2006, [5] Nagy & Anderson 1984, Yang 2016

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

Can non-CDS be substituted for CDS to study the relevant problem?

[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

Can non-CDS be substituted for CDS to study the relevant problem?

Yes, for the purposes of lexical acquisition[2]

[1] Labov 1972 as applied to linguistics, Walkden 2019, [2] Kodner 2019

Data Set

Perseus Corpus

- Scraped all Old and Classical Latin texts from website HTML
  - 3rd BC - AD 2nd inclusive
  - ~3.5mil tokens
- More than available by download - undocumented “feature” :-\

Largest plain text OL/CL corpus?

Data Set

Perseus Corpus

- Scraped all Old and Classical Latin texts from website HTML
  - 3rd BC - AD 2nd inclusive
  - ~3.5mil tokens
- More than available by download

Post-Processing

- POS-tagged and lemmatized with modified CLTK library
  - 1,292 unique verb lemmas when derivational prefixes removed
- Scraped Latin Wiktionary verbs to match lemmas to principal parts

Data Set

Perseus Corpus

- Scraped all Old and Classical Latin texts from website HTML
  - 3rd BC - AD 2nd inclusive
  - ~3.5mil tokens
- More than available by download

Post-Processing

- POS-tagged and lemmatized with modified CLTK library
  - 1,292 unique verb lemmas when derivational prefixes removed
- Scraped Latin Wiktionary verbs to match lemmas to principal parts
- Manually compared ~100 principal parts to Oxford Latin Dictionary

Latin Wiktionary is surprisingly accurate!

Outline

- The Classical Latin Past Participles
- Acquiring Morphological Generalizations
- Language Acquisition in the Past
- **Predictability of the Past Participles**
- Synchronic and Diachronic Implications

Applying the Tolerance Principle

Over several possible generalizations

- Theme vowels → pptc forms
- Other present generalizations → pptc forms
- Perfect generalizations → pptc forms
- Present + perfect → pptc form

Applying the Tolerance Principle

Over several possible generalizations

- Theme vowels → pptc forms
- Other present generalizations → pptc forms
- Perfect generalizations → pptc forms
- Present + perfect → pptc form

Theory independent interpretation

- Generalizations over surface phonotactics “rightmost vowel is /a:/”
- Or generalizations over morphemes “ThV is *-ā-*”

Applying the Tolerance Principle

Over several possible generalizations

- Theme vowels → pptc forms
- Other present generalizations → pptc forms
- Perfect generalizations → pptc forms
- Present + perfect → pptc form

Theory independent interpretation

- Generalizations over surface phonotactics "rightmost vowel is /a:/"
- Or generalizations over morphemes "ThV is *-ā-*"

Modeling early and late learners

- Multiple frequency cutoffs
- Verbal vocab sizes $n = 100, 500, 1000$

Example Calculation

Is *-ātus* the productive pptc derivation for verbs with ThV *ā*?

Example Calculation

Is *-ātus* the productive pptc derivation for verbs with ThV *ā*?

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

- ***N*=221 *ā* verbs**
- ***e*=13 *ā* verbs with non *-ātus* pptcs**

Example Calculation

Is *-ātus* the productive pptc derivation for verbs with ThV *ā*?

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

- *N*=221 *ā* verbs
- *e*=13 *ā* verbs with non *-ātus* pptcs

Exceptions are tolerable if

$$e < \theta$$

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

Example Calculation

Is *-ātus* the productive pptc derivation for verbs with ThV *ā*?

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

- *N*=221 *ā* verbs
- *e*=13 *ā* verbs 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

Productive Present → PPtc by Theme Vowel

Theme Vowel	PPtc	Example	At *n*=100?	At 500?	At 1,000?
ā (1st)	*-ātus*	*vocāre ~ vocātus*	YES	YES	YES
ē (2nd)	*-ĭtus*	*habēre ~ habitus*	no	no	no
ē (2nd)	*-tus*	*docēre ~ doctus*	no	no	no
e (3rd non-*iō*)	*-ĭtus*	*reddere ~ redditus*	no	no	no
e (3rd non-*iō*)	*-tus*	*scribere ~ scriptus*	no	no	no
i (3rd *-iō*)	*-tus*	*capiō ~ captus*	YES	YES	YES
e or *i* (all 3rd)	*-ĭtus*	*" ~ "*	no	no	no
e or *i* (all 3rd)	*-tus*	*" ~ "*	no	no	no
ī (4th)	*-ītus*	*audīre ~ audītus*	YES	marginal*	no
ī (4th)	*-tus*	*venīre ~ ventus*	YES	no	no

Individual Development →

* within 1 of threshold

Productive Present → PPtc more Narrowly

Present	PPtc	Example	At *n*=100?	At 500?	At 1,000?
-[*a, o*]*veō*	-[*au, ō*]*tus*	*faveō ~ fautus*	-	YES	YES
-[Velar]*eō*	*-tus*	*doceō ~ doctus*	-	no	no
-[not Velar]*eō*	*-itus*	*debeō ~ debitus*	marginal*	no	no
-[not Velar]*eō*	*-tus*	*teneō ~ tentus*	no	no	no
-vere	*-ūtus*	*solvere ~ solūtus*	YES	marginal*	marginal*
-[*ll, rr*]*ere*	-[*l,r*]*sus*	*currō ~ cursus*	-	marginal*	no
other 3rd	*-ĭtus*	*reddere ~ redditus*	no	no	no
other 3rd	*-tus*	*scribere ~ scriptus*	no	no	no

Individual Development →

* within 1 of threshold

Productive Perfect → PPtc

Perfect	PPtc	Example	At *n*=100?	At 500?	At 1,000?
-āv-	*-ātus*	*amāvī ~ amātus*	YES	YES	YES
-īv-	*-ītus*	*dormīvī ~ dormītus*	YES	YES	YES
-ēv-	*-ētus*	*flēvī ~ flētus*	YES	YES	marginal*
-u-	*-ĭtus*	*valuī ~ valitus*	no	no	no
-u-	*-tus*	*tenuī ~ tentus*	no	no	no
-[Velar]*u-*	*-tus*	*līquī ~ līctus*	-	no	no
-[not Velar]*u-*	*-ĭtus*	*dēbuī ~ dēbitus*	no	no	no
-[not Velar]*u-*	*-tus*	*peruī ~ pertus*	no	no	no
-s-	*-tus*	*scripsī ~ scriptus*	no	no	no
-Cs-	*-tus*	*iūnxī ~ iūnctus*	YES	YES	YES
bare or stem change	*-ĭtus*	*lēgī ~ lēctus*	no	no	no

Individual Development →

* within 1 of threshold

Productive Perfect + Present → PPtc

Perfect	PPtc	Example	At *n*=100?	At 500?	At 1,000?
-vere + -u-	*-ūtus*	*volvere ~ voluī ~ volūtus*	YES	YES	YES

Individual Development →

- Only makes a difference for once class, but it is ***-utu***
- Only an option when a learner happens to know both stems

Summary

If derivations are only possible from the present,

- **Productive pptc derivation for 1st (*-ātus*), 3rd-*iō* (*-tus*)**
- **Marginal for *faveō*-type (*-autus*/*-ōtus*) and *solvō*-type (*-ūtus*)**

Summary

If derivations are only possible from the present,

- Productive pptc derivation for 1st (*-ātus*), 3rd-*iō* (*-tus*)
- Marginal for *faveō*-type (*-autus*/*-ōtus*) and *solvō*-type (*-ūtus*)
- No productive pptc derivation for 2nd, 3rd-*ō*, 4th
- No broadly productive *-ĭtus* or *-tus*

Summary

If derivations are only possible from the present,

- Productive pptc derivation for 1st (*-ātus*), 3rd-*iō* (*-tus*)
- Marginal for *faveō*-type (*-autus*/*-ōtus*) and *solvō*-type (*-ūtus*)
- No productive pptc derivation for 2nd, 3rd-*ō*, 4th
- No broadly productive *-ĭtus* or *-tus*

If derivations is possible from the perfect,

- The above + productive deriv for *-īvī* (most of 4th; *-ītus*), *-ēvī* (*-ētus*), *-Csī* (*-tus*)
- Solidly productive *-ūtus* for *solvō*-types

Summary

If derivations are only possible from the present,

- Productive pptc derivation for 1st (*-ātus*), 3rd-*iō* (*-tus*)
- Marginal for *faveō*-type (*-autus*/*-ōtus*) and *solvō*-type (*-ūtus*)
- No productive pptc derivation for 2nd, 3rd-*ō*, 4th
- No broadly productive *-ĭtus* or *-tus*

If derivations is possible from the perfect,

- The above + productive deriv for *-īvī* (most of 4th; *-ītus*), *-ēvī* (*-ētus*), *-Csī* (*-tus*)
- Solidly productive *-ūtus* for *solvō*-types
- No broadly productive pptc derivation for *-uī*-perfect verbs
- Still no broadly productive *-ĭtus* or *-tus*

Outline

- The Classical Latin Past Participles
- Acquiring Morphological Generalizations
- Language Acquisition in the Past
- Predictability of the Past Participles
- **Synchronic and Diachronic Implications**

The System from Latin to Proto-Romance

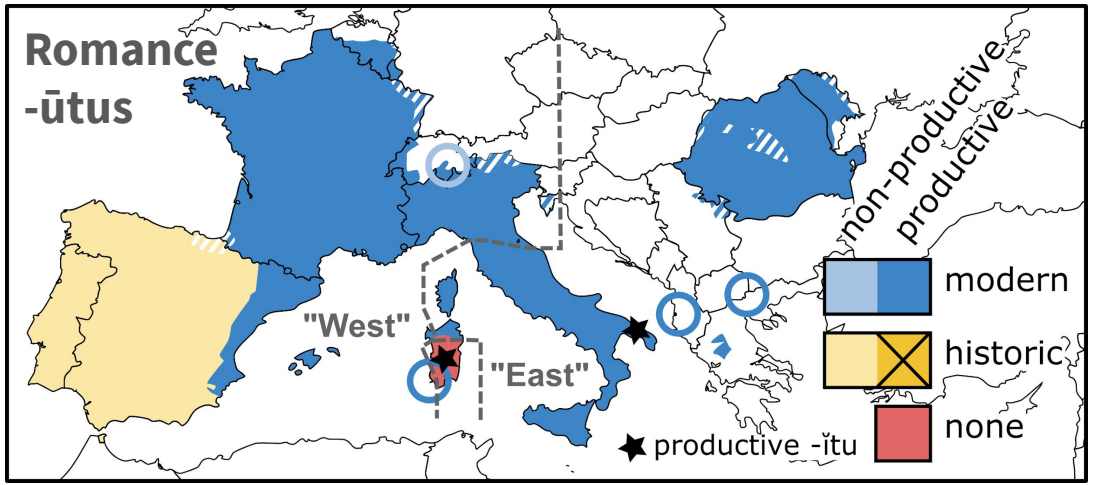
Varied across the Latin-speaking world, but in general…

- Novel verbs tended to have regular pptcs[1]
- "Regular" ***-ɑtu***, ****-itu***, ****-utu*** < ***-ātus***, ***-ītus*** (not ***-ĭtus***), ***-ūtus*** expanded at the expense of ***-itus***, ***-tus***, and others[2]
- The rise of ****-utu*** is mysterious given that it is rare in CL
- Perfects (→ preterites) were often regularized, often in ****-ui*** < ***-uī***[3]

[1] Laurent 2003, [2] ibid., [3] ibid.

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

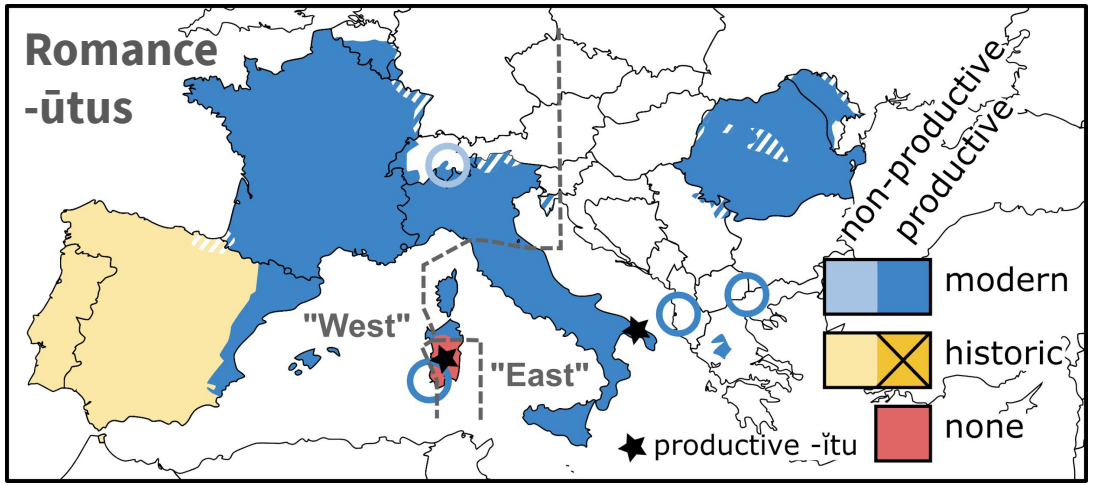
- Reflexives of ***-ūtus*** constitute the default for at least some class in most Romance languages
  - They are present but apparently non-productive in Surselvan (Rhaeto-Romance; Switzerland)
- Reflexes are attested in Old Spanish and Portuguese but have been lost
  - Their only reflexes are in adjectives eg, ***agudo***, ***menudo***



[1] data compiled from Laurent 2003

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

- Reflexives of ***-ūtus*** constitute the default for at least some class in most Romance languages
  - They are present but apparently non-productive in Surselvan (Rhaeto-Romance; Switzerland)
- Reflexes are attested in Old Spanish and Portuguese but have been lost
  - Their only reflexes are in adjectives eg, *agudo*, *menudo*
- ***-ĭtus*** *remains productive in Apulian and Sardinian*
  - /i/ merged with /i:/ in Sardinian, causing *-ĭtus* to fall together with *-ītus*



[1] data compiled from Laurent 2003

Diachronic Implications

Developments in Late Latin

- Three productive LL pptcs: **-atu* < *-ātus*, **-itu* < *-ītus*, **-utu* < *-ūtus*
- *-ĭtus* and *-tus* were unproductive in CL and reduced to irregulars
- *-ūtus* was productive for a small class
- But the only productive option for *-uī* perfects!
- It spread first among *-uī* perfects
- No competition, "a big fish in a small pond"

Implications

Listing and Rules

- An externally motivated model guides theoretical analysis
- Predicts much more listing than a linguist relying on intuitions might

Implications

Listing and Rules

- An externally motivated model guides theoretical analysis
- Predicts much more listing than a linguist relying on intuitions might

The relationship between stems

- If pptcs are derived from perfects
  - More can be derived by rule
  - Accounts for diachronic leveling of the perfect and pptc
- To do so, either perfect stems exist as representational objects
  or multiple step root → perfect “stem” → pptc derivations are required

End.

With support from

Thank you

- Charles Yang
- Mitch Marcus
- Don Ringe
- Rolf Noyer
- Rebecca Starr Lurie
- Mitcho Erlewine

How are past particples derived?

- Are regular pptcs influenced by the present or perfect, or all memorized?
- Diachronic evidence for both

  present → pptc: nasal infix spread

  perfect → pptc: perfect analogies

[1] Aronoff 1994

The Nasal Infix

- Inherited from PIE, inserted into present stems
- Some continue to work like this in Latin[1]
- But some have analogized to the perfect and pptc

Type	Present	Perfect	PPtc
Inherited	*fundō*	*fūdī*	*fūsus*
Pres, Perf	*fingō*	*fīnxī*	*fictus*[2]
All	*iungō*	*iunxī*	*iūnctus*
Pres, PPtc	*pungō*	*pupugī*	*pūnctus*

[1] Poultney 1937, [2] but Italian *finto*

The Nasal Infix

- Inherited from PIE, inserted into present stems
- Some continue to work like this in Latin[1]
- But some have analogized to the perfect and pptc
- Only evidence for present → pptc derivation if absent in the perfect
  - At most two examples of this…
  - Otherwise, can present → perfect → pptc

Type	Present	Perfect	PPtc
Inherited	*fundō*	*fūdī*	*fūsus*
Pres, Perf	*fingō*	*fīnxī*	*fictus*[2]
All	*iungō*	*iunxī*	*iūnctus*
Pres, PPtc	*pungō*	*pupugī*	*pūnctus*
	tundō	*tutudī*	*tū(n)sus*

[1] Poultney 1937, [2] but Italian *finto*

Perfect Analogies

- Some pptcs have clearly been reworked on the basis of the perfect[1]

	cernō	*crēvī*	*crētus* (expected *certus* retained as adj)
	sternō	*strāvī*	*strātus*
?	*sonāre*	*sonuī*	*sonitus*

- Continues into Late Latin: eg **-utu* pptcs typically correspond to **-ui* perfects

[1] Table from Laurent 2003, p. 22

The System from Proto-Romance to Romance

Spanish, for example, shows the most regularization[1]

- **Regularization continued**
  - ***-ado***, ***-ido***, and ***-udo*** existed in Old Spanish
  - Only ***-ado***, ***-ido*** remain productive
- **A handful of irregular pptcs remain, many relegated to adjectival meaning**
  - ***hecho***, ***puesto***, ***suelto***, ***visto***, ***vuelto***, etc, not all inherited
  - ***teñir*~*teñido*** 'dyed' but adj ***tinto*** 'dyed red' < ***tinctus***, etc
  - OS had more eg ***querer*~*quisto***, ***prender*~*preso*** < ***prehensus***

[1] Laurent 2003 ch. 4.7

Past Participle Gaps and Meanings

- Past participles are typically passive
- But not all verbs have past participles[1]
  - Sometimes due to semantics (eg, statives have no pptcs)
  - Sometimes they’re more properly paradigmatic gaps

    eg *bibō*, but *pōtus not* **bibitus*, *feriō*, but *percussus* not **ferītus*

- Some pptcs are active rather than passive[2]
  - Expected for deponents
  - But applies to some non-deponents as well

    eg *locūtus* (deponent) ‘having spoken,’ *iūrātus* ‘having sworn’

[1,2] Laurent 2003, [2] Embick 2000

Cross-Language Lexical Comparisons

- Compared lexical composition of modern CDS and historical corpora
- Calculated number of verb types across corpora with similar meanings

For corpus-derived lexicons A and B
where A and B are unordered sets,

$$similarity = |A \cap B| / \min(|A|, |B|)$$

Cross-Language Corpora

- **English CDS** - verb lemmas in CHILDES Brown (and Brent for comparison)
- **Spanish CDS** - verb lemmas in combined CHILDES FernAguado, Hess, OreaPine, Remedi, Romero, SerraSole
- **Classical Latin** - verb lemmas in all Perseus online 3rd BC - 2nd AD (inclusive)

Corpus	Freq Cutoff	Lexicon size (*n*)
English CDS Brown	< 17	260
English CDS Brent	< 17	257
Spanish CDS	< 11	263
Latin	< 666	260

[1] Credit to Don Ringe for extracting them

Cross-Language Comparisons

- Baselines: English-English (within-language) English-Spanish (cross-language)
- English-English unsurprisingly has the highest overlap
- Latin comparisons fall in between English-Spanish and English-English

Latin Perseus contains the same kind of high frequency verbs that CDS does

Comparison	% Overlap
English - EN Brent	81.71%
English - Spanish	73.07%
English - Latin	75.77%
Spanish - Latin	78.62%

Paradigm Saturation

- **Paradigm Saturation**[1] **- the proportion of a verb's possible inflected forms which are actually attested in a corpus**
- **A measure of data sparsity**
- **Mean saturations tend to be low**
- **Obeys Zipfian distribution**

[1] Chan 2008

Paradigm Saturation Data

- All POS-tagged, lemmatized, morpho feature annotated
- **CDS** - English (Brown), Spanish
- and German (CDS Leo[1])
- **Modern** - UD[2] English, Finnish, German, Spanish, Turkish
- **Historical** - UD Gothic, Latin
- Order 10^5 verb tokens

Corpus	Lang	# V Tokens	# V Types	Ratio
CDS	English	94,768	916	103.46
CDS	Spanish	96,686	879	110.00
CDS	German	81,351	641	126.91
Modern	English	53,796	3,225	16.67
Modern	Spanish	85,861	5,019	17.11
Modern	German	21,835	2,826	7.73
Modern	Finnish	63,891	3,476	18.38
Modern	Turkish	12,064	968	12.46
Historic	Gothic	12,749	1,172	10.88
Historic	Latin	99,066	2,2833	34.97

[1] Behrens 2006, [2] Nivre et al 2018

Paradigm Saturation Data

- All POS-tagged, lemmatized, morpho feature annotated
- **CDS** - English (Brown), Spanish
- and German (CDS Leo[1])
- **Modern** - UD[2] English, Finnish, German, Spanish, Turkish
- **Historical** - UD Gothic, Latin
- Order 10^5 verb tokens
- **CDS token/type ratios are on the order of 10x higher**

Corpus	Lang	# V Tokens	# V Types	Ratio
CDS	English	94,768	916	103.46
CDS	Spanish	96,686	879	110.00
CDS	German	81,351	641	126.91
Modern	English	53,796	3,225	16.67
Modern	Spanish	85,861	5,019	17.11
Modern	German	21,835	2,826	7.73
Modern	Finnish	63,891	3,476	18.38
Modern	Turkish	12,064	968	12.46
Historic	Gothic	12,749	1,172	10.88
Historic	Latin	99,066	2,2833	34.97

[1] Behrens 2006, [2] Nivre et al 2018

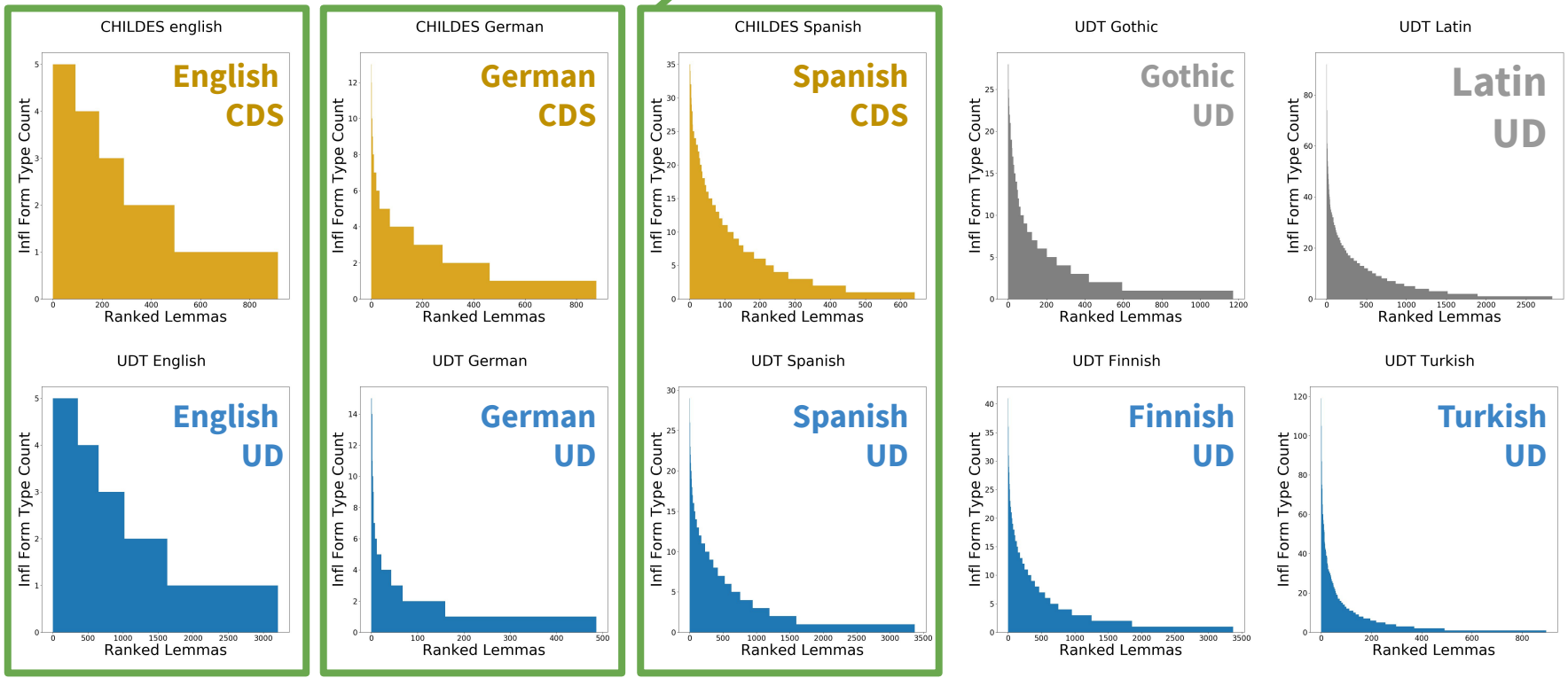
Paradigm Saturations

- **CDS saturations only slightly higher than modern equivs**
- **Despite difference in token/type ratios**
- **Historical corpora similar to modern ones**
- **Saturation appears related to paradigm size if anything**

Corpus	Lang	|Paradigm|	Max Sat.	Mean Sat.	Med Sat.
CDS	English	5	100%	43.23%	40.00%
CDS	Spanish	29	44.83%	7.59%	6.90%
CDS	German	67	52.24%	8.31%	4.48%
Modern	English	5	100%	42.80%	40.00%
Modern	Spanish	67	43.28%	4.91%	1.49%
Modern	German	29	51.72%	5.83%	3.45%
Modern	Finnish	150	27.33%	2.46%	1.33%
Modern	Turkish	120	99.17%	4.83%	1.67%
Historic	Gothic	52	53.85%	6.31%	3.85%
Historic	Latin	113	81.42%	5.90%	2.65%

Zipfian Distributions

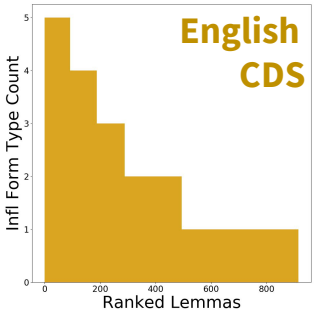
CDS and UD distributions correspond by language



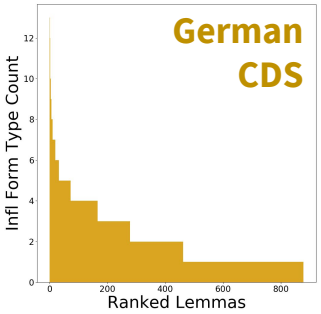
Zipfian Distributions

Historical distributions look like modern ones

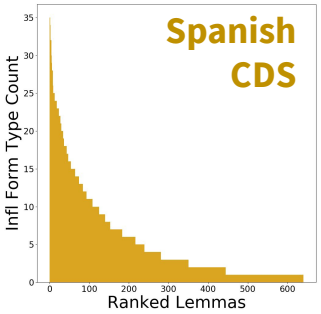
CHILDES english



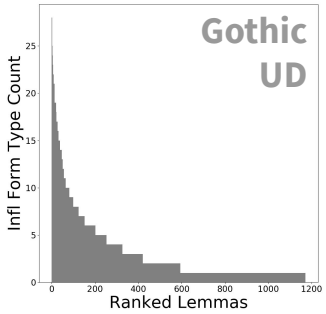
CHILDES German



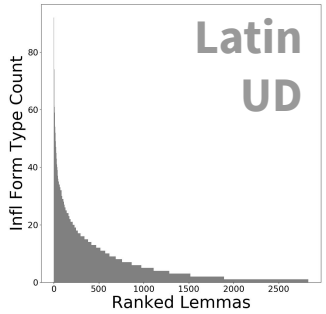
CHILDES Spanish



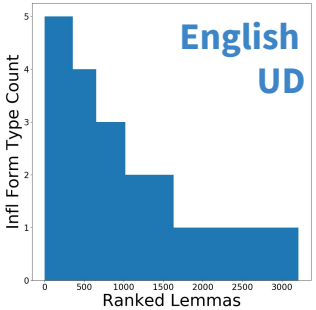
UDT Gothic



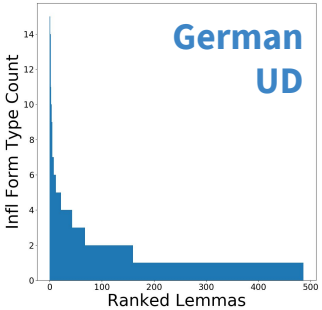
UDT Latin



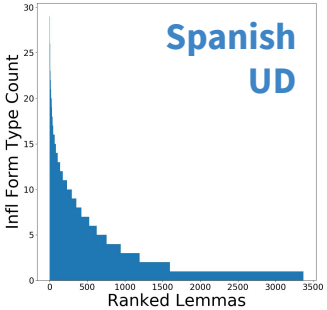
UDT English



UDT German

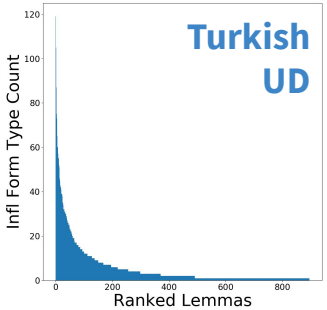


UDT Spanish



UDT Finnish

UDT Turkish



Language Change by Language Acquisition

- Child language acquisition is one of the primary drivers of language change[1]
- Not a new idea (Schleicher 1861, Paul 1880, etc)
- Children are both innovators and propagators of change

[1] Schleicher 1861, 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, *inter alia*

Language Change by Language Acquisition

- Child language acquisition is one of the primary drivers of language change[1]
- Not a new idea (Schleicher 1861, Paul 1880, etc)
- Children are both innovators and propagators of change
- Minor learning "errors" over successive generations → major population-level change

[1] Schleicher 1861, 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, *inter alia*

The Paradox of Language Change

- Term coined by Niyogi & Berwick 1997
- As I see it, a central problem in the study of language change

If children are so good at language acquisition, why are they so bad at it?

Transmission is not strictly linear and generational

- Children mature in communities and receive input from multiple speakers
- Young children learn sociolinguistic variables[1]
- Children attend to input from older children[2] who are not linguistically mature
- Not inconsistent with the adolescent peak[3] of many continuous changes

[1] Labov 1989, Anderson 1990, [2] Manly 1930, Weinreich, Labov & Herzog 1968 p 145, Roberts and Labov 1995, Labov 2001, [3] Eckert 1989, Labov 2001,

Some learning targets are unclear or absent

- One cannot acquire language from input alone due to **Poverty of the Stimulus**
- UG is proposed to render learning possible in the face of the PoS[1]
- But many language specific patterns must still be acquired from the input[2]

Input is both richer and poorer than typically acknowledged

- Evidenced by the successes and failures of modern NLP[3]
- Zipfian and other long-tailed distributions for all manner of linguistic features
  - Most lexical items appear only once even in massive corpora
  - Zipfian distributions mean sparsity is consistently worse than our intuitions about sparsity

[1] Chomsky 1959, 1980, [2] eg Baker's Paradox (Baker 1979), [3] eg the successes of distributional semantics vs the failures of coreference

Abject Poverty

Occasionally the PoS is so great that UG cannot ensure that all learners converge on the same grammar

- Forms in even moderately complex paradigms may never appear in the input[1]
- Paradigmatic gaps occur when learners fail to learn a generalization for unattested input[2]
- Some syntactic ‘parameters’ cannot be set consistently[3]

[1] Chan 2008, Lignos & Yang 2017, [2] Yang 2016, [3] Han et al 2007

Moving Targets

Variation is a normal and unavoidable part of acquisition

- Even in “monolingual” environments[1]
- Children learn from multiple adults and each other

Change is formally inevitable[2]

- Given categorical representations[3] and “trivial” variation
- The population composition must change over time

[1] contra Meillet, Meissel 2011, [2] Niyogi & Berwick 1997, [3] Singleton & Newport 2004, Schuler et al 2017, Sneller et al 2018

What causes innovation?

“Errors” presuppose a target. Innovations need not be due to “errors”

What causes innovation?

“Errors” presuppose a target. 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

What causes innovation?

“Errors” presuppose a target. 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 even trivial variation

The Sibling Effect

- Why might children not overcome their overgeneralizations?

Imagine big sister Alice and little brother Bob

- Alice is currently producing innovative *$\bar{e}$ pasts in Class IV
  - Plausible given how Class IV *$\bar{e}$ is tenable late
  - Bob may hear these forms
- Bob is receiving both adult conservative IV pasts and Alice's
- How does this effect Bob?

The Sibling Effect

Can Bob identify Alice's innovation?

- Bob is likely not hear adult-produced tokens for any given low frequency Class IV verb until much later
- Since Alice is mostly consistent with adults, he cannot tell if she is innovating

Will Bob adopt Alice's innovation?

- Even young children orient toward peers
- Bob may prefer Alice's forms over his parents
- He may later learn adult forms as sociolinguistic variant doublets