Getting the Right Stuff Wrong:
Modeling the Acquisition of Inflectional Morphology

Sarah Brogden Payne
Stony Brook University
sarah.payne@stonybrook.edu

CUNY Computational Linguistics Speaker Series
March 1, 2023

Introduction: Morphological Acquisition

- Children **learn inflectional morphology**
  - From highly ***sparse, skewed*** input
  - On <1000 word types
  - Despite ***exceptions***
  - With complex systems of ***allomorphy***

Challenging problem!

Introduction: Morphological Acquisition

- Children make **systematic errors** cross-linguistically
  - ***Overregularization:*** e.g. "feel-feeled"
  - ***Omissions of Marking:*** e.g. "Papa have it"

  } **Almost all errors**

- Why **these** errors and not others?
- What do the errors tell us about:
  - Acquisition?
  - The resulting grammar?

  } **Models of morphological acquisition should address these questions**

Introduction: Mechanistic Accounts

- To provide **mechanistic accounts of error patterns,** models should:
  - Make the ***same errors as children***
  - On ***developmentally-plausible training data***
  - Be ***interpretable***
- **Neural models** struggle with this
  - ***Unnatural error patterns:*** over-irregularizations common
  - ***Data-hungry:*** large and/or saturated data
  - ***Not interpretable***

Introduction: Proposal

Abduction of Tolerable Productivity (ATP)

- Accounts for **overregularization errors**
- Accounts for **developmental regression**
- Models the **mapping of inflectional classes to form**

- **Recursive search** with **Tolerance-Sufficiency Principle**
- Trained on **developmentally-plausible** data
- Give **interpretable accounts** of errors

Sufficient Contrast Learner (SCL)

- Accounts for **omission errors**
- Accounts for **cross-linguistic** differences in acquisition
- Models the acquisition of **inflectional classes**

Outline

- **What makes a plausible model?**
  - Nature of the input
  - Developmental Findings
- **Previous work**
  - The Past Tense Debate
  - The Past Tense Debate: Reprise
- **Proposal**
  - ***ATP:*** mapping features to form
  - ***SCL:*** learning inflectional classes
- **Future work**

What makes a plausible model?

The Nature of the Input

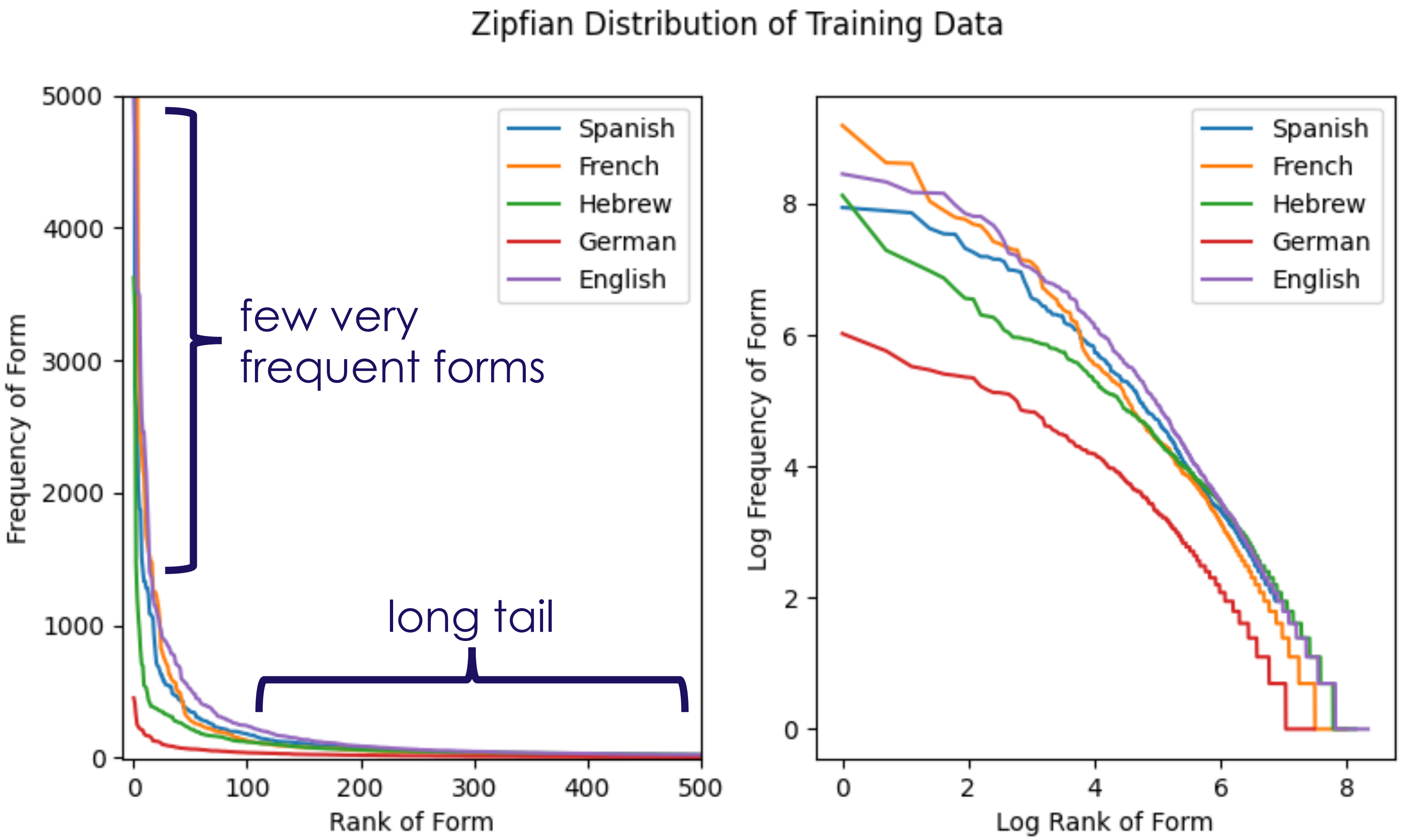
Input Sparsity: Zipf's Law

- **Zipf's law:** word ***rank*** inversely proportional to ***frequency***

$$f(r) \propto \frac{1}{r}$$

- **Consequences:**
  - A few forms occur very frequently
  - Most occur very rarely (long tail)

Input Sparsity: Zipf's Law



(data from Payne et al 2021, Belth et al 2021, Payne 2022, and Payne 2023)

Input Sparsity: Paradigm Saturation

- Long-tailed distributions in morphology: **Paradigm Saturation**
  - How many possible inflected forms does a lemma actually occur in?

$$saturation = \frac{\#\, seen}{\#\, possible}$$

	Present	Preterite	Imperfect	Conditional	Future
1SG	*amo*	*amé*	*amaba*	*amaría*	*amaré*
2SG	*amas*	*amaste*	*amabas*	*amarías*	*amarás*
3SG	*ama*	*amó*	*amaba*	*amaría*	*amará*
1PL	*amamos*	*amamos*	*amábamos*	*amaríamos*	*amaremos*
2PL	*amáis*	*amasteis*	*amabais*	*amaríais*	*amaréis*
3PL	*aman*	*amaron*	*amaban*	*amarían*	*amarán*

(Chan 2008, Lignos & Yang 2016)

Input Sparsity: Paradigm Saturation

- Long-tailed distributions in morphology: **Paradigm Saturation**
  - How many possible inflected forms does a lemma actually occur in?

$$saturation = \frac{\#\, seen}{\#\, possible}$$

$$= \frac{\mathbf{7}}{\#\, possible}$$

	Present	Preterite	Imperfect	Conditional	Future
1SG	*amo*		*amaba*		*amaré*
2SG		*amaste*			
3SG	*ama*		*amaba*		
1PL	*amamos*				
2PL					
3PL					

(Chan 2008, Lignos & Yang 2016)

Input Sparsity: Paradigm Saturation

- Long-tailed distributions in morphology: **Paradigm Saturation**
  - How many possible inflected forms does a lemma actually occur in?

$$saturation = \frac{\#\ seen}{\#\ possible}$$

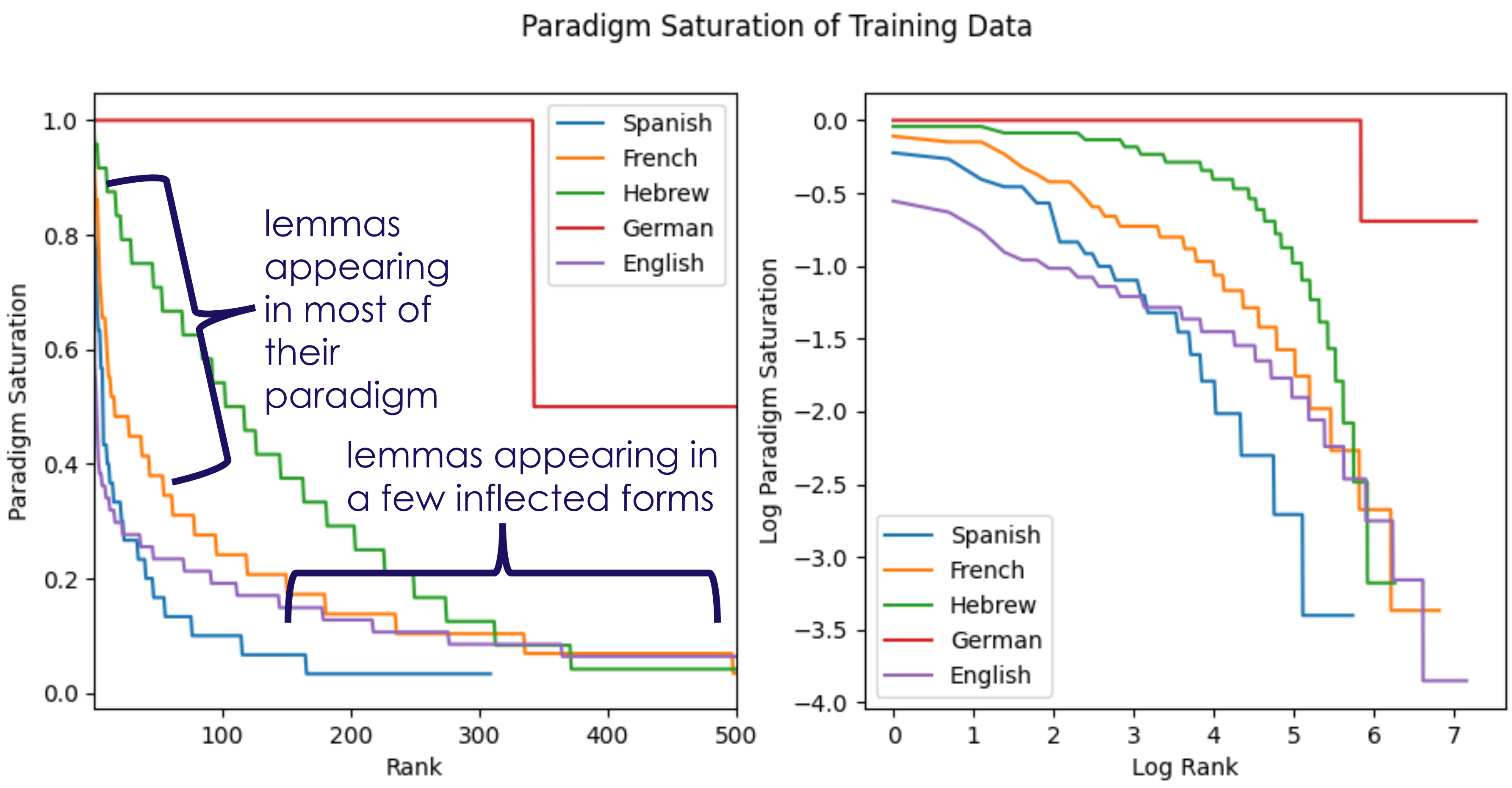
$$= \frac{\mathbf{7}}{\#\ possible}$$

$$= \frac{\mathbf{7}}{\mathbf{26}} \approx 27\%$$

	Present	Preterite	Imperfect	Conditional	Future
1SG	***amo***	*trabajé*	***amaba***	*trabajía*	***amaré***
2SG	*tomas*	***amaste***	*mirabas*	*mirarías*	*esperás*
3SG	***ama***	*esperó*	***amaba***	*espería*	*tomará*
1PL	***amamos***	*miramos*	*mirabamos*	*tomaríamos*	*miraremos*
2PL	*tratáis*				
3PL	*esperan*	*miraron*	*entraban*	*tratarían*	*entrarán*

(Chan 2008, Lignos & Yang 2016)

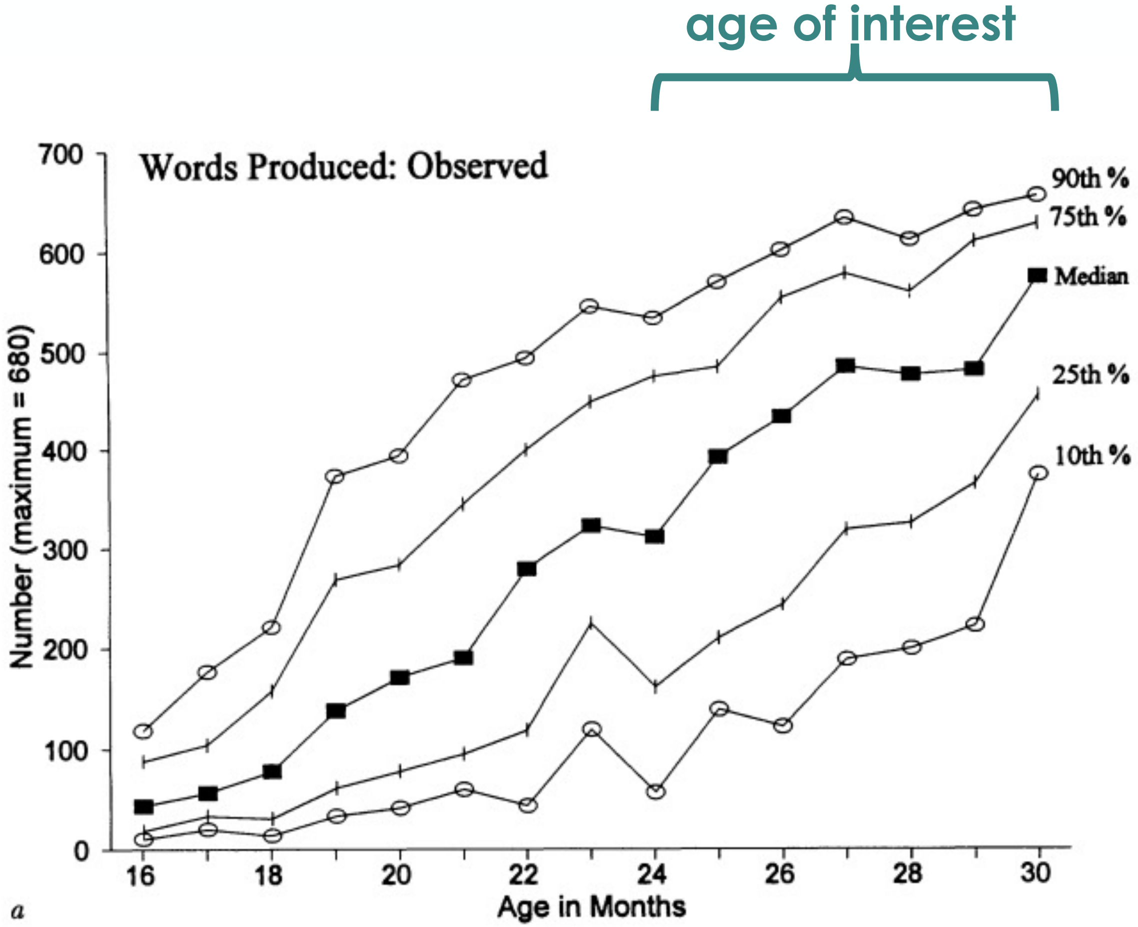
Input Sparsity: Paradigm Saturation



(data from Payne et al 2021, Belth et al 2021, Payne 2022, and Payne 2023)

Input Sparsity: Early Vocabulary

- **At 2;0: 200-500 words** cross-linguistically
- **At 3;0: <1000 words** cross-linguistically
- Early vocabulary makeup:
  - ~50% **nouns**
  - ~25% **verbs**



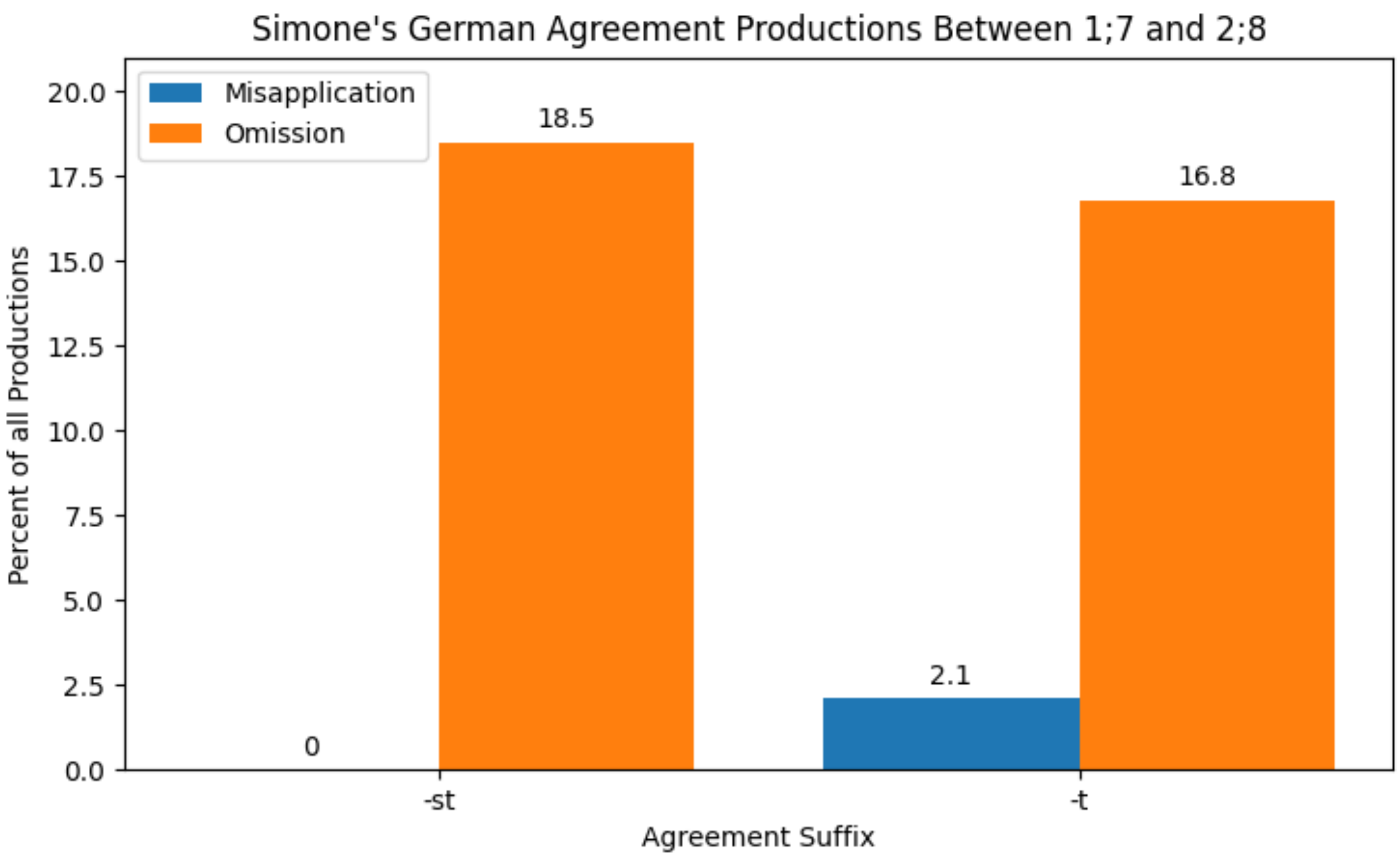
(from Fenson et al 1994)

What makes a plausible model?

Developmental Findings

Background: Production Errors

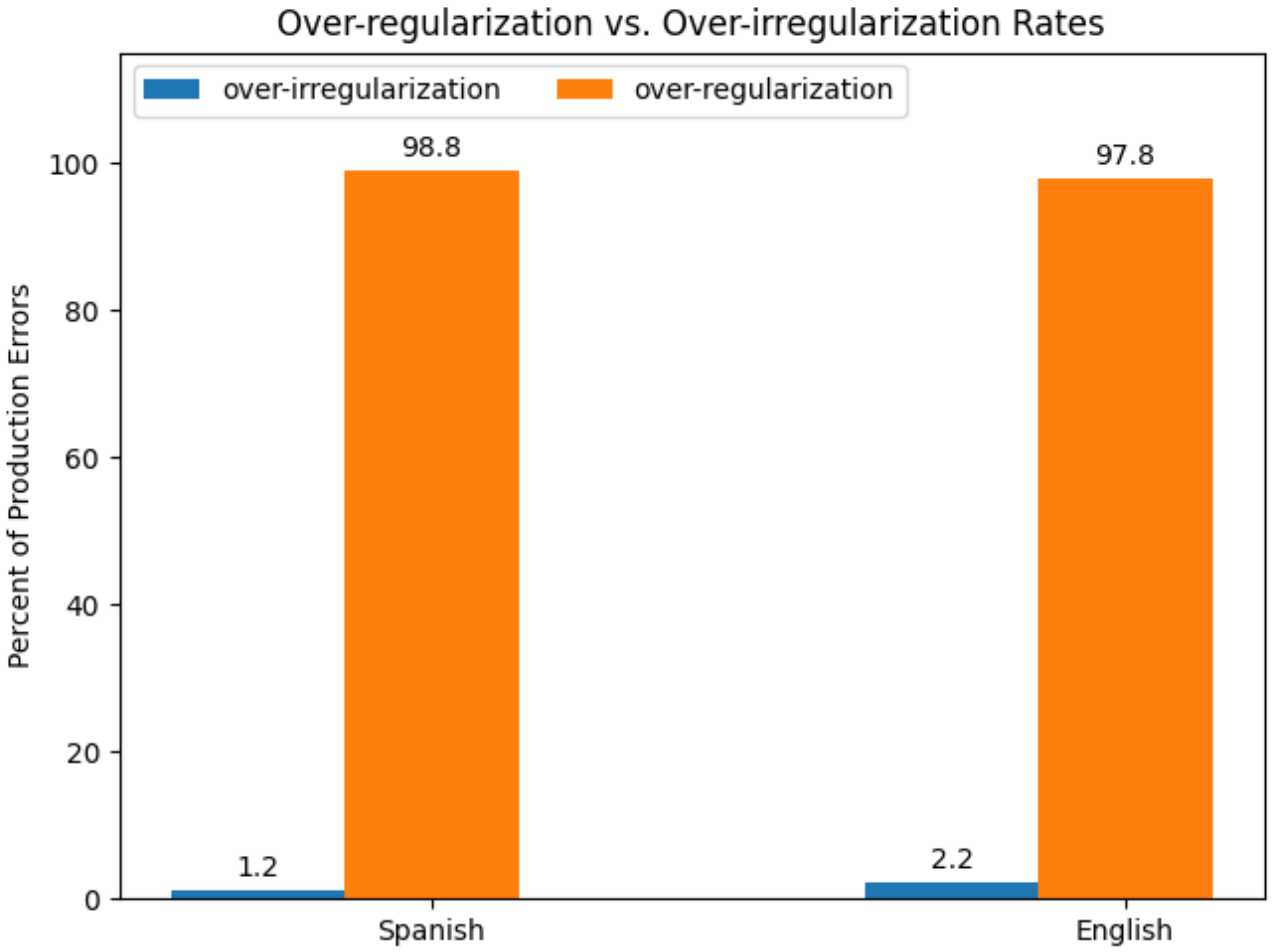
- **Omissions:** ***Root Infinitives***
  - e.g. "Papa have it"
- **Substitutions:** incorrect overt affix
  - e.g. "I has it"



(Clahsen & Penke 1992, Philips 1995, Legate & Yang 2007)

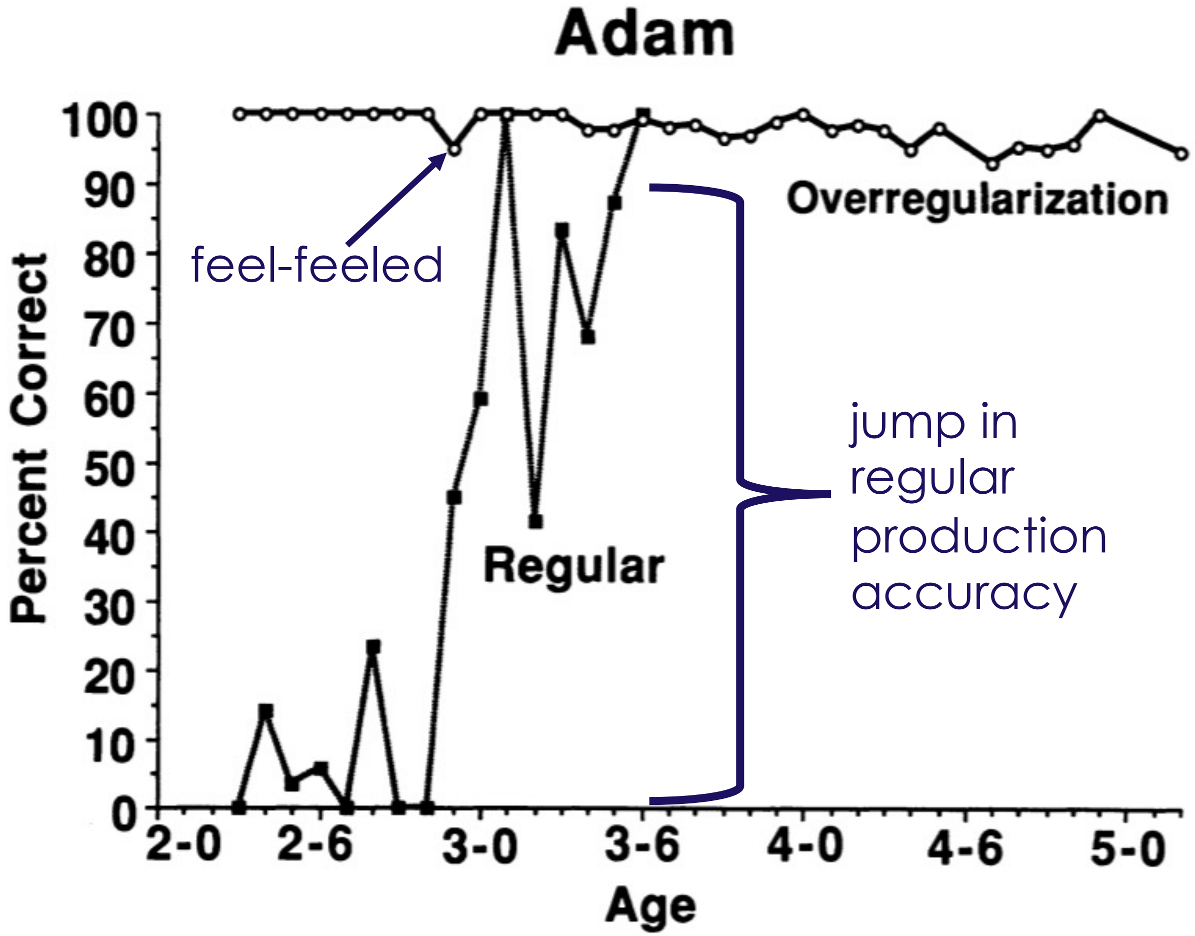
Background: Production Errors

- **Over-regularization**
  - e.g. ***feel-feeled***
- **Over-irregularization**
  - e.g. ***bite-bote***



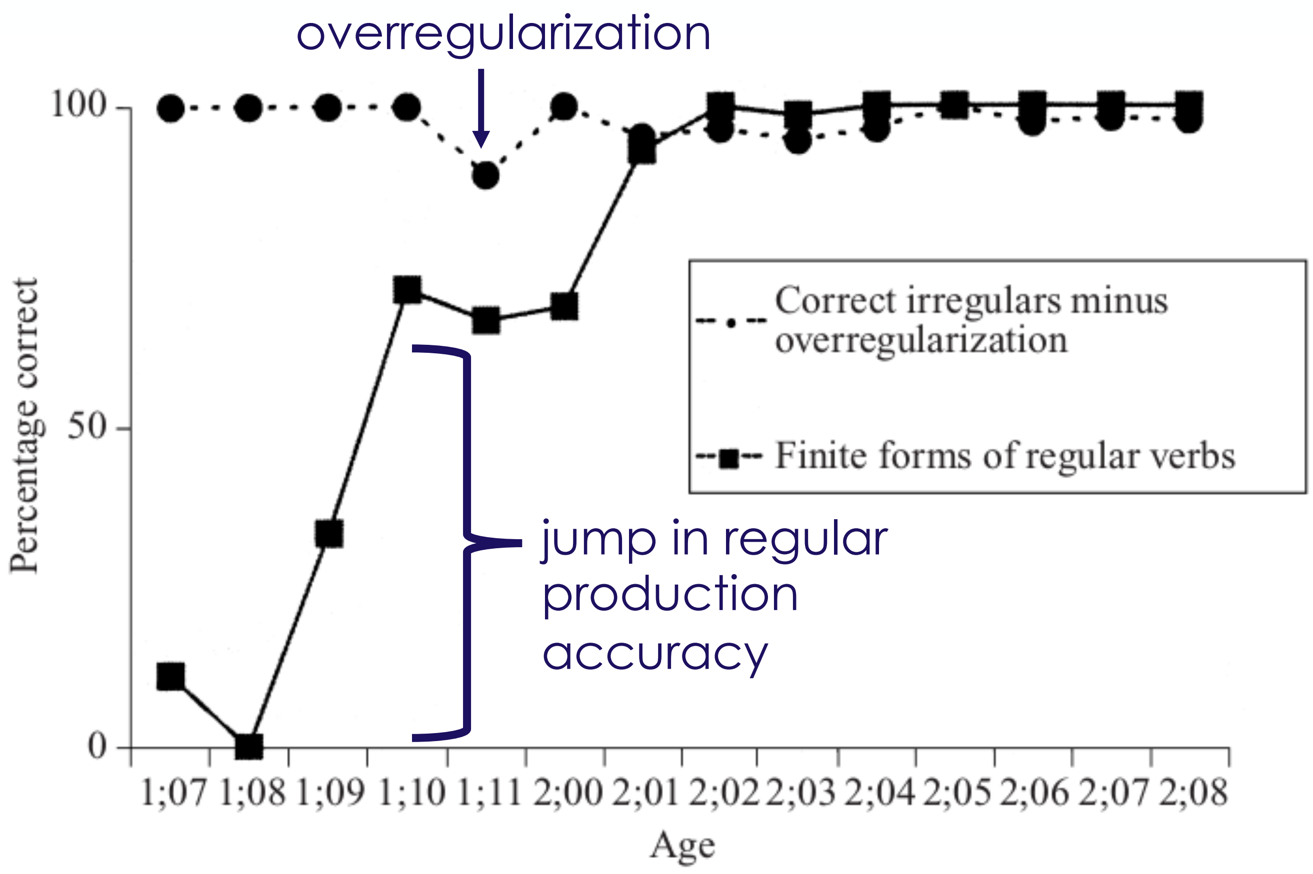
(Maslen et al 2004, Xu & Pinker 1995, Clahsen et al 2002)

Background: Developmental Regression



(from Marcus et al 1992)

Background: Developmental Regression



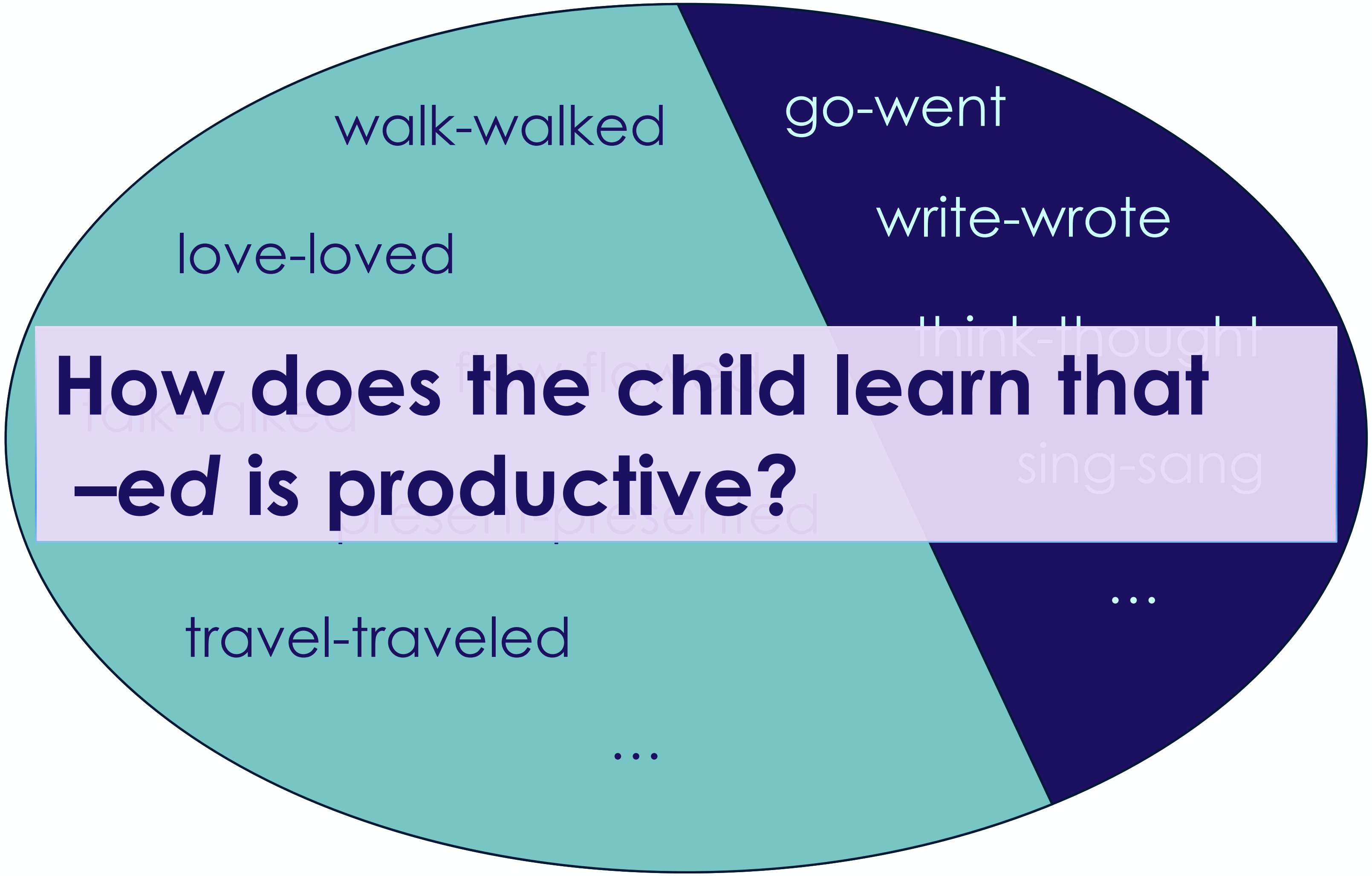
(from Clahsen, Aveledo, and Roca 2002)

Summary: What Makes a Plausible Model?

- **Learn from:**
  - ***Small*** vocabulary
  - ***Sparse*** paradigms
- **Errors:**
  - ***Omissions***, not substitutions
  - ***Over-regularizations***, not over-irregularizations
  - Developmental ***regression***

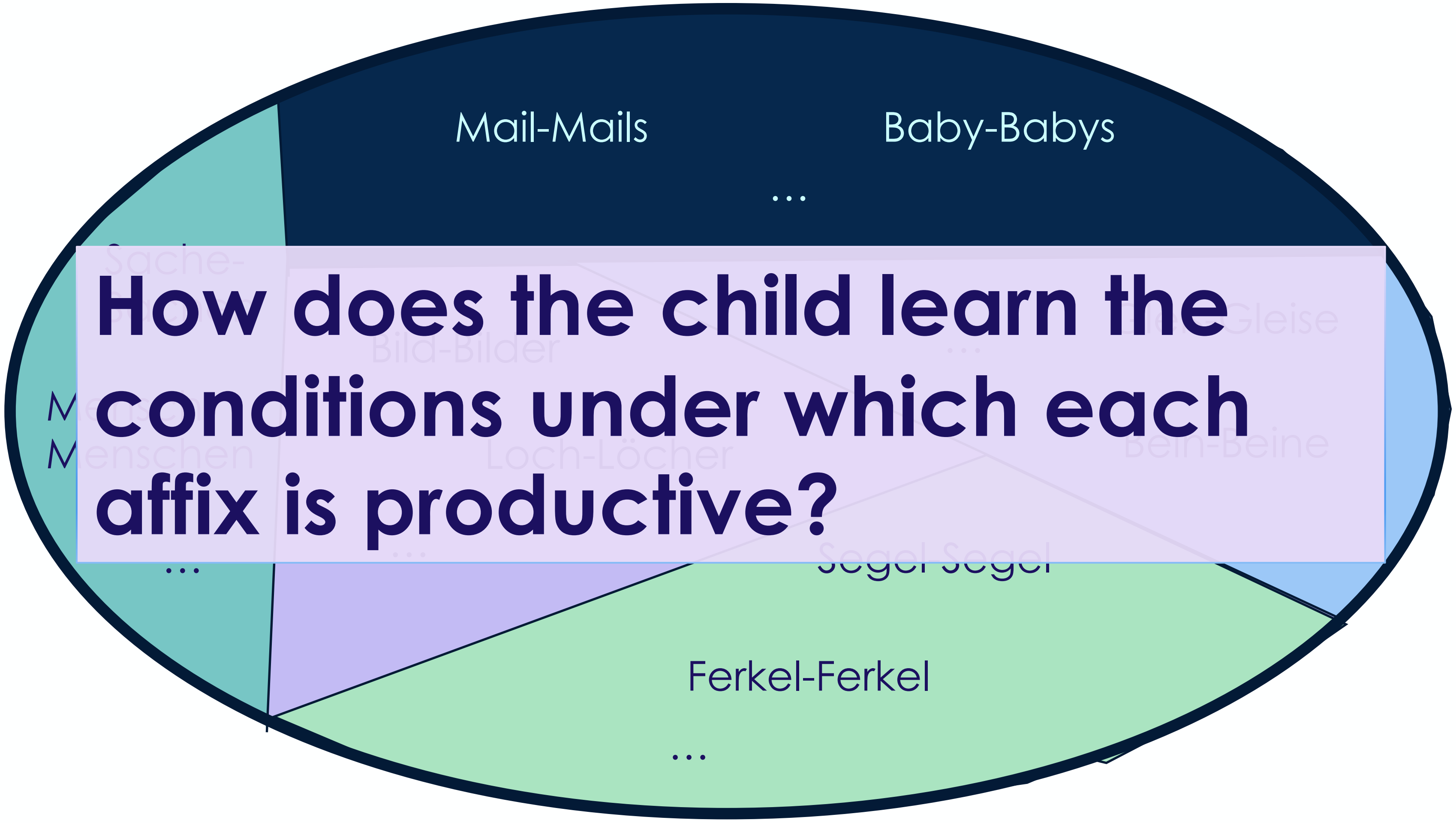
Previous Work: The Past Tense Debate, Rounds 1 and 2

Handling Exceptionality



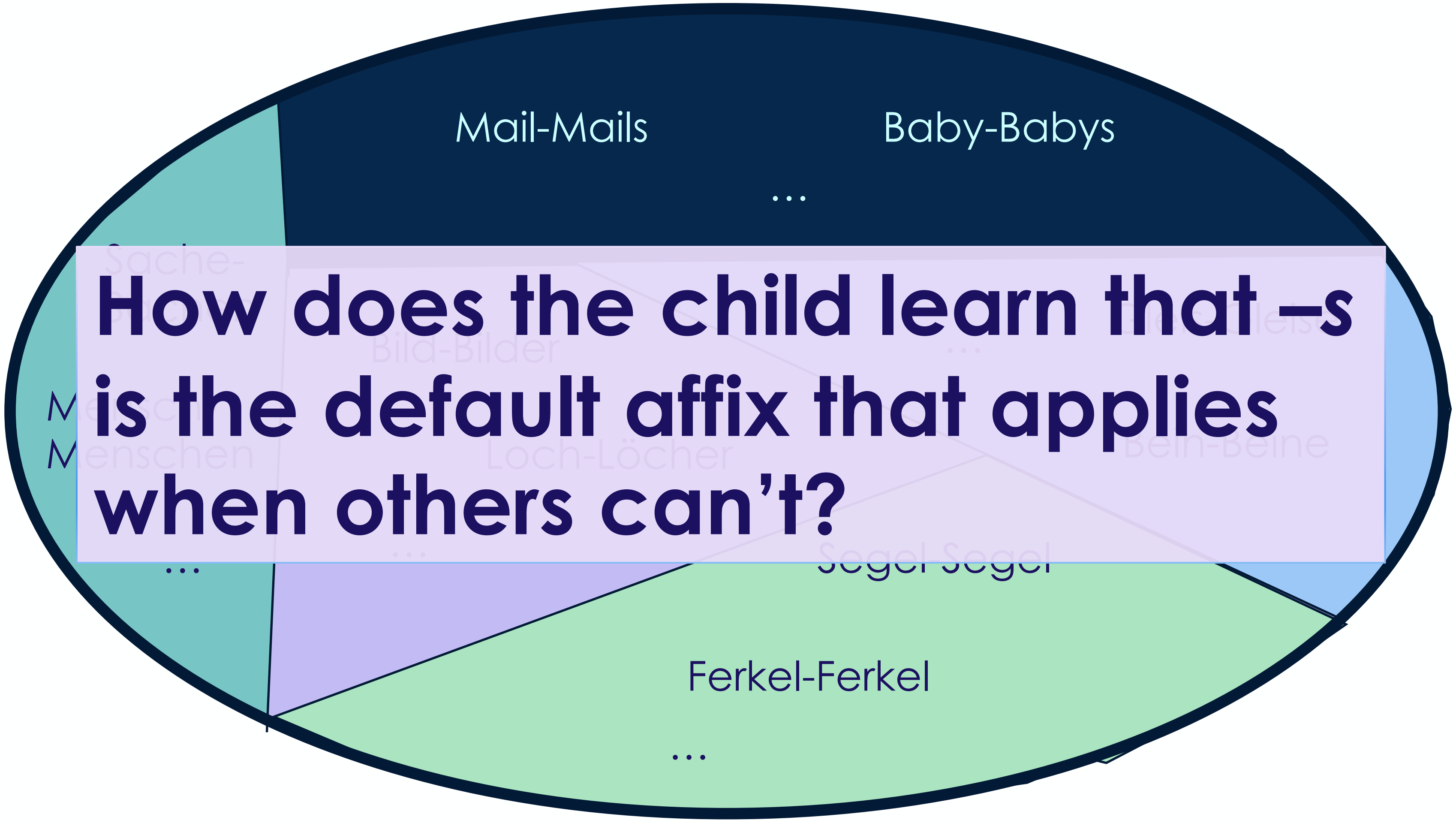
How does the child learn that –*ed* is productive?

Handling Allomorphic Productivity



How does the child learn the conditions under which each affix is productive?

Handling Allomorphic Productivity



How does the child learn that –s is the default affix that applies when others can't?

Background: The Past Tense Debate

- **Rumelhart & McClelland (1986):** ***single-route, connectionist*** model can:
  - Exhibit ***developmental regression***
  - Exhibit ***overregularization***

  ∴ **Rule-like behavior**

- **Pinker & Prince (1988):** actually…

  Developmental regression = artifact of training data
  - First trained on ***80% irregulars***
  - Then trained on ***80% regulars***

  Exhibits ***over-irregularization***
  - ***sip-sept, type-typeded, mail-membled***

  ∴ **No rule-like behavior**

Background: The Past Tense Debate Revisited

Recurrent Neural Networks in Linguistic Theory: Revisiting Pinker and Prince (1988) and the Past Tense Debate

Christo Kirov
Johns Hopkins University
Baltimore, MD
ckirov1@jhu.edu

Ryan Cotterell
Johns Hopkins University
Baltimore, MD
ryan.cotterell@jhu.edu

Background: The Past Tense Debate Revisited

- **Kirov & Cotterell (2018):** encoder-decoder RNNs can overcome empirical limitations
  - Near ***100% test accuracy***
  - Learn ***several classes at once***
  - Trained on ***developmentally-representative*** data
  - Main errors = ***overregularizations***

- **Corkery et al (2019):** ED model **still fails empirically!**
  - Predictions **don't match well with humans** on nonce English past tense forms
    - **Still over-irregularizes!**
  - Massive **variability in model rankings** between seeds
    - **Correlation with human ratings** also varies massively

Background: The Past Tense Debate Revisited

- **Kirov & Cotterell (2018):** encoder-decoder RNNs can overcome empirical limitations
  - Near ***100% test accuracy***
  - Learn ***several classes at once***
  - Trained on ***developmentally-representative*** data
  - Main errors = ***overregularizations***

No ***developmental regression!***

Trained on ***>3500 verbs in their full paradigm***

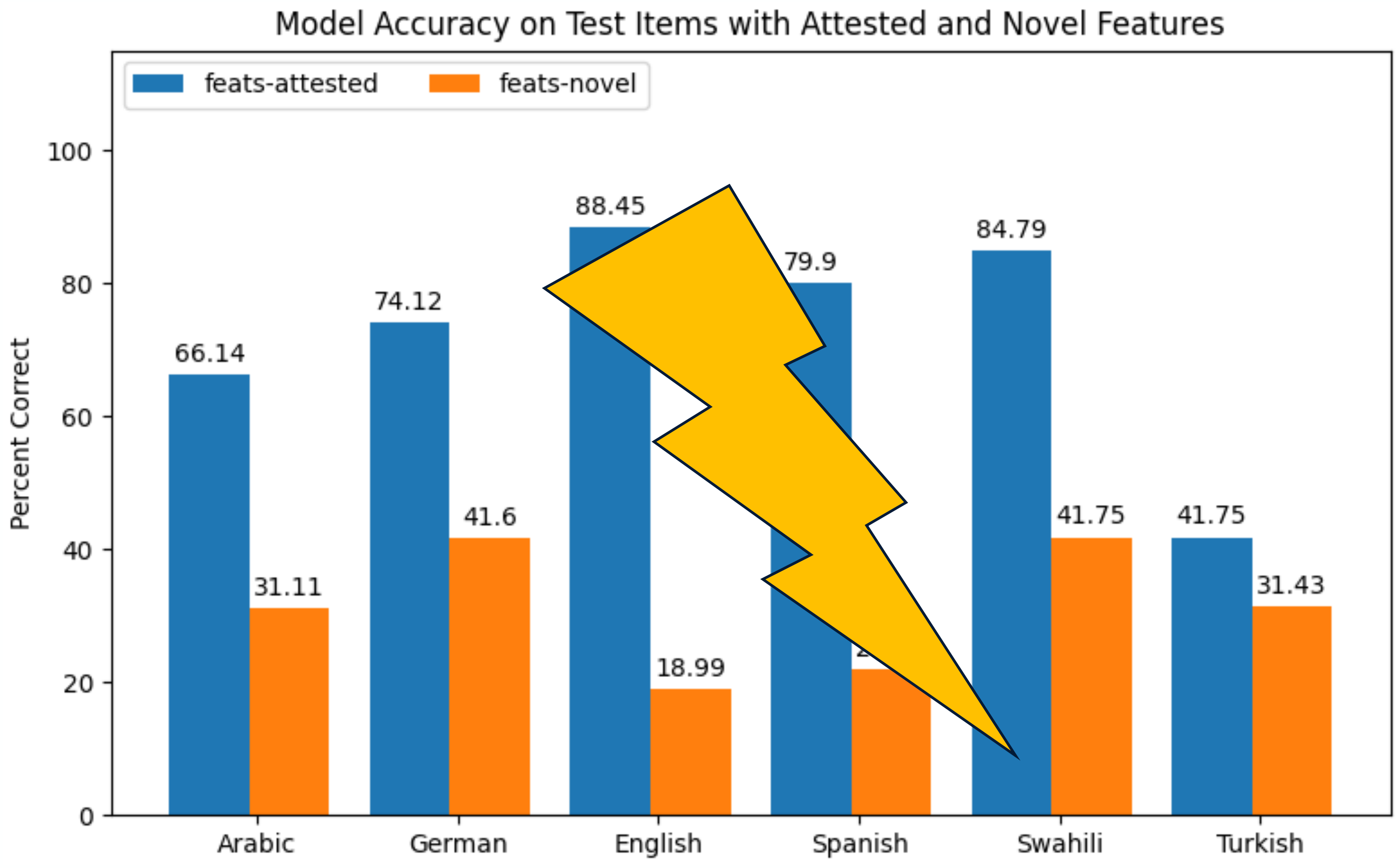
- Children know **< 350 verbs** at 3;0
- Would need to see **> 15k lemmas** to see 3,500 in complete paradigm

German Noun Plurals: We *really* aren't there

- **Marcus et al (1995):** NNs overapply the ***most common process*** rather than the ***default***
  - **German:** most common ≠ default
- **McCurdy et al (2020a):** Train on German noun plurals & **test on nonce words**
  - Model predictions ***don't match well with human predictions***
  - ***Overproduction of frequent*** affixes rather than default
- **McCurdy et al (2020b):** Model **uses gender** as main cue, humans **use phonology**

Generalization(?) by Neural Models

- Children **generalize** to novel feature combinations
  - e.g. Spanish:
    - See ***amarian, ama, aman, amaba***
    - Generalize: ***amaban***
- Can NNs do the same?
  - Evaluate 3 models on feature sets **attested vs. unattested** in train



(ongoing work with Salam Khalifa, Jordan Kodner, and Zoey Liu)

Summary: The Past Tense Debate(s)

- What have we gotten from ~30 years of NN research?
  - Better accuracy
  - More developed architecture

  } Good for NLP(?)

- What haven't we gotten?
  - Still ***overproduce irregulars***
  - Still ***no developmental regression***
  - Still ***data-hungry***:
    - Too much, too saturated

  } Persistence of issues ⇒ **fundamental difference between connectionist models & language faculty**

Proposal: Recursive, Rule-based Learning

Models: Making Sense of Errors

- Children **over-regularize** & don't **over-irregularize**
  - Account for this with **rule-based mappings**:
    - Learn rule like **PAST → *-ed***
    - Apply rule when no exception known
      - **Over-regularization** when exception not yet learned
      - **Developmental regression** when rule first learned
  - **Abduction of Tolerable Productivity (ATP):** recursively learn productive rules & their exceptions

Models: Making Sense of Errors

- Children **omit** inflectional affixes, but don't **substitute** them
  - Account for this with **initially-underspecified** inflectional categories:
    - Must learn e.g. that English contrasts **+3SG vs. -3SG**
    - Underspecified category can't be productively mapped to form, so **omit inflection**
  - **Sufficient Contrast Learner (SCL):** recursively learn inflectional categories

Preliminaries: The TSP

Intuitions: given a set of *N* items:

- If ***most*** do X, then all do X **(generalization)**
- If ***few*** do X, memorize those that do **(lexicalization)**

Tolerance of exceptions

Generalize a rule applying to *N* items with e exceptions iff:

$$e \leq \theta_N = \frac{N}{\ln N}$$

Sufficient positive evidence

Generalize a rule applying to *N* items and seen applying to *M* iff:

$$\underbrace{N - M}_{\text{worst-case } e} \leq \theta_N = \frac{N}{\ln N}$$

(Yang 2016)

Preliminaries: Training Data

- Children learn **frequent forms earlier** (Goodman et al 2008)
  - Use ***most frequent forms from CHILDES***
- Children use of **distributional cues** to learn meaning
  - Intersect CHILDES with ***UniMorph features as a proxy for these cues***
- Input: **(lemma, inflected, features)**

Language	Lemma	Inflected	Features
English	*walk*	*walked*	{V, PAST, 3, SG}
Spanish	*amar*	*amaban*	{V, 3, PL, PAST, IMPFV}
German	*Sache*	*Sachen*	{N, FEM, PL}

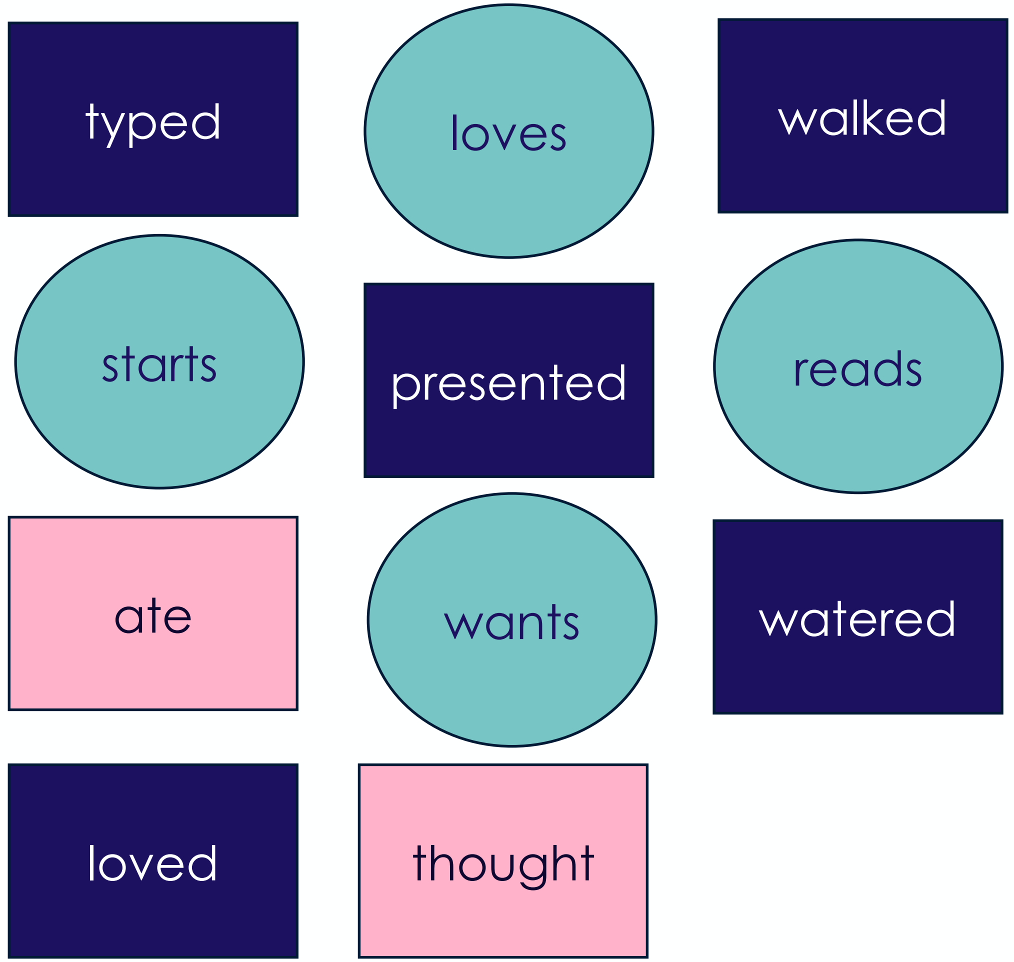
Mapping Features to Form: Abduction of Tolerable Productivity

(Belth et al 2021)

ATP Model: Recursive Subdivision

- Apply TSP **recursively**
  - Given ***N*** items, do **enough** of them take **-x affix**?
    - If yes, **productive rule learnt!**
    - If not, **subdivide** into disjoint subsets & **recurse**
- **Terminate** when:
  - Productive rule found **(generalization)**
  - No more subdivisions possible **(lexicalization)**
- Apply to **English past tense** and **German noun plurals**

ATP Model: Toy Example



- 11 items: 4 **-s**, 5 **-ed**, 2 **other**
- **Generalize** most frequent?
  $N - M = 11 - 5 = 6 > \theta_{11} = 4.5$
- **Subdivide!** Hypothesize a rule:

ATP Model: Toy Example

typed

walked

presented

ate

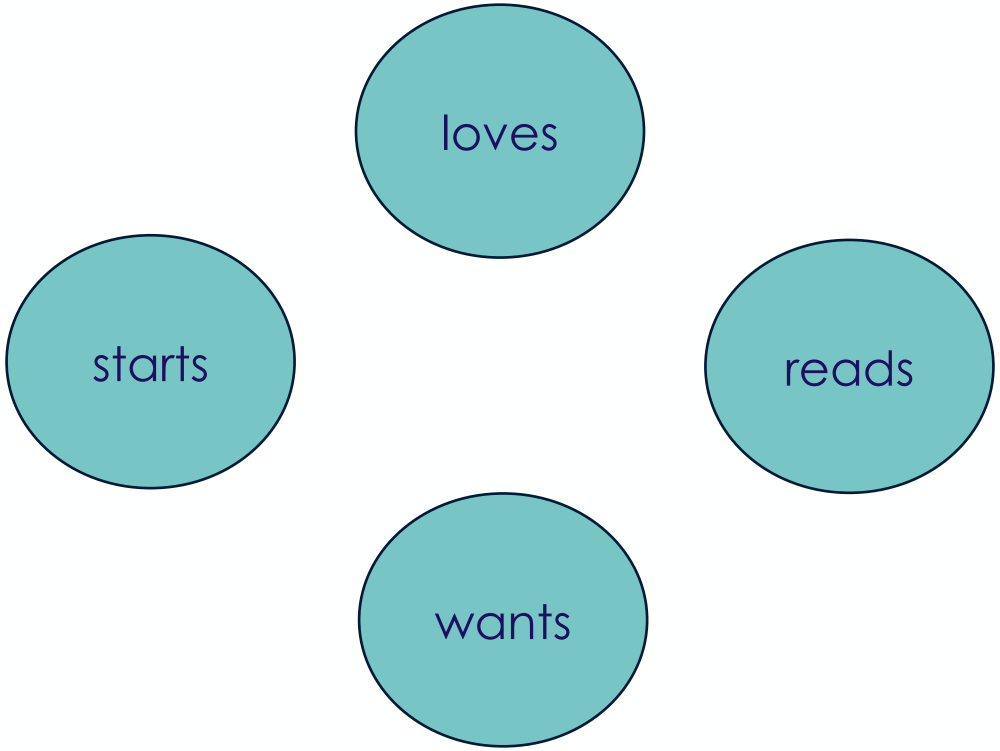
watered

loved

thought

- 11 items: 4 **-s**, 5 ***-ed***, 2 **other**
- **Generalize** most frequent?
  $\boldsymbol{N - M = 11 - 5 = 6 > \theta_{11} = 4.5}$
- **Subdivide!** Hypothesize a rule:
  - PAST → ***-ed***
- **Test** the rule:
  - $\boldsymbol{N - M = 2 < \theta_7 = 3.5}$ ✅
- R1 productive! PAST → ***-ed***
  - Memorize ***ate*** and ***thought***

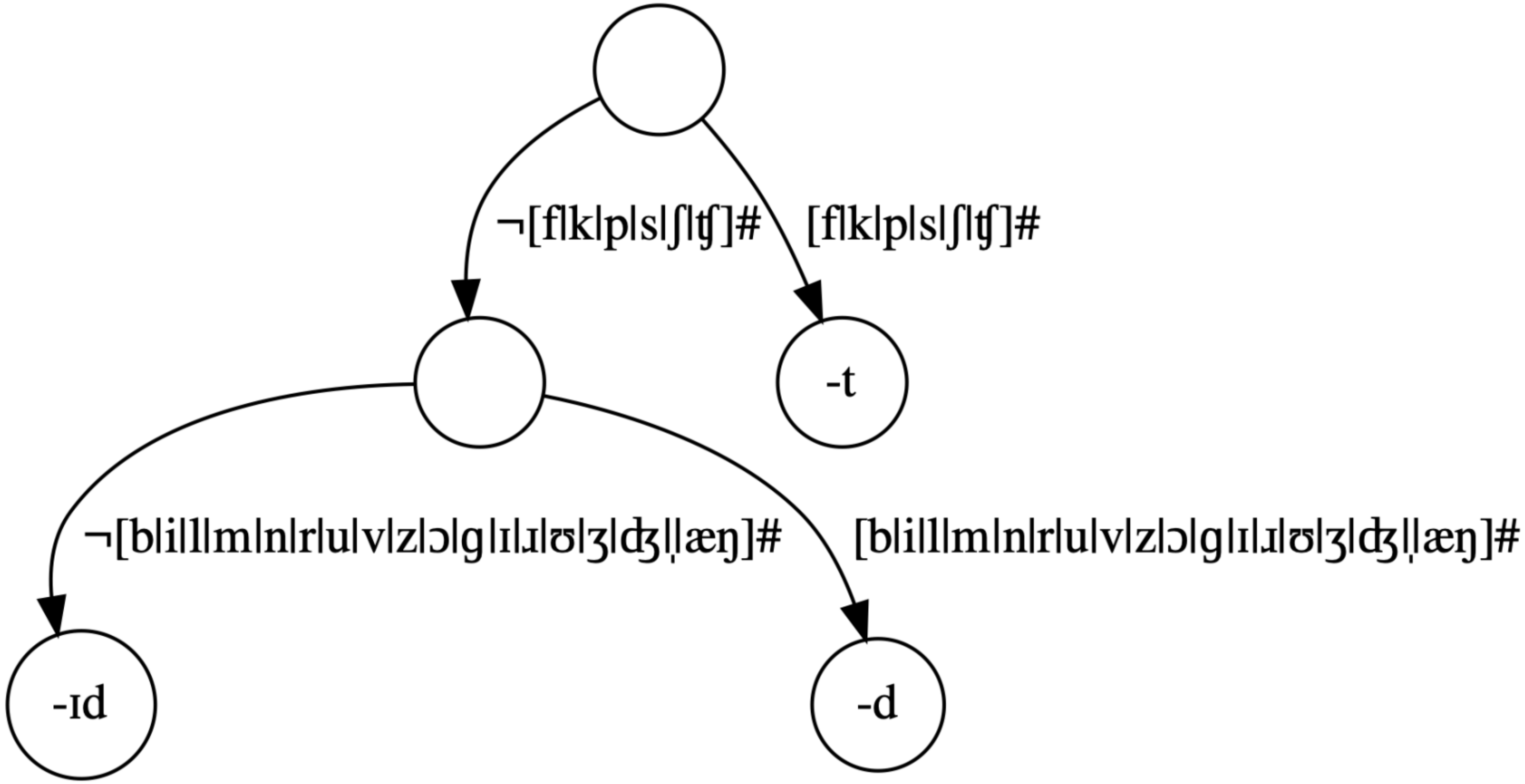
ATP Model: Toy Example



- 11 items: 4 **-s**, 5 **-ed**, 2 **other**
- **Generalize** most frequent?
  - $N - M = 11 - 5 = 6 > \theta_{11} = 4.5$
- **Subdivide!** Hypothesize a rule:
  - PAST → ***-ed***
- **Test** the rule:
  - $N - M = 2 < \theta_7 = 3.5$ ✅
- R1 productive! PAST → ***-ed***
  - Memorize ***ate*** and ***thought***
- **Recurse:** PRES,3,SG → **-s**

ATP Model: Sample learning trace

English past tense: morphophonological conditioning



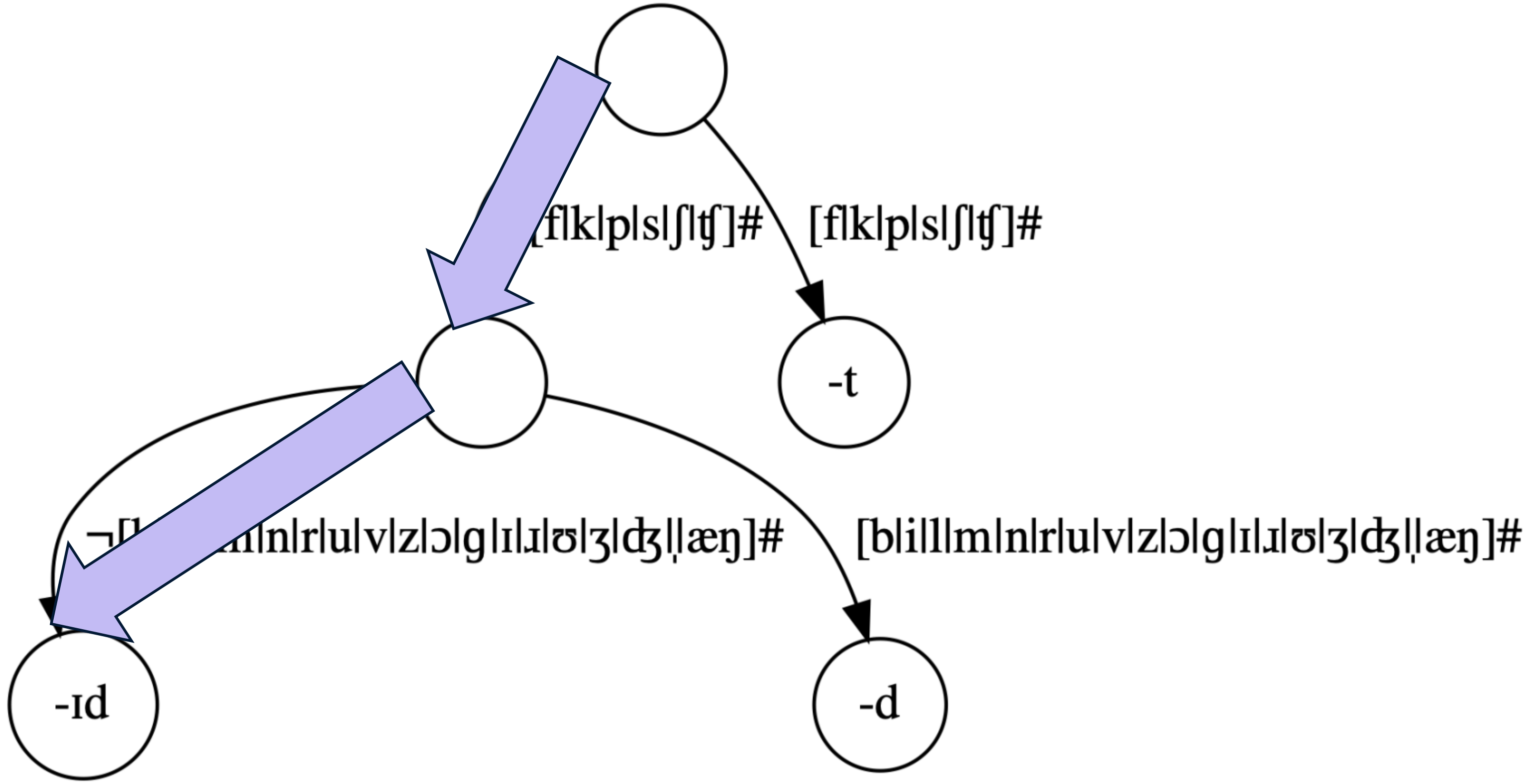
ATP Model: Inflection and Generation

- During test, given **novel forms & features** to inflect
- Traverse decision tree to correct node
  - If node has ***productive rule***, apply the rule
  - If no ***productive rule***, either:
    - Produce unmarked form
    - **Analogize** to a known form at this node

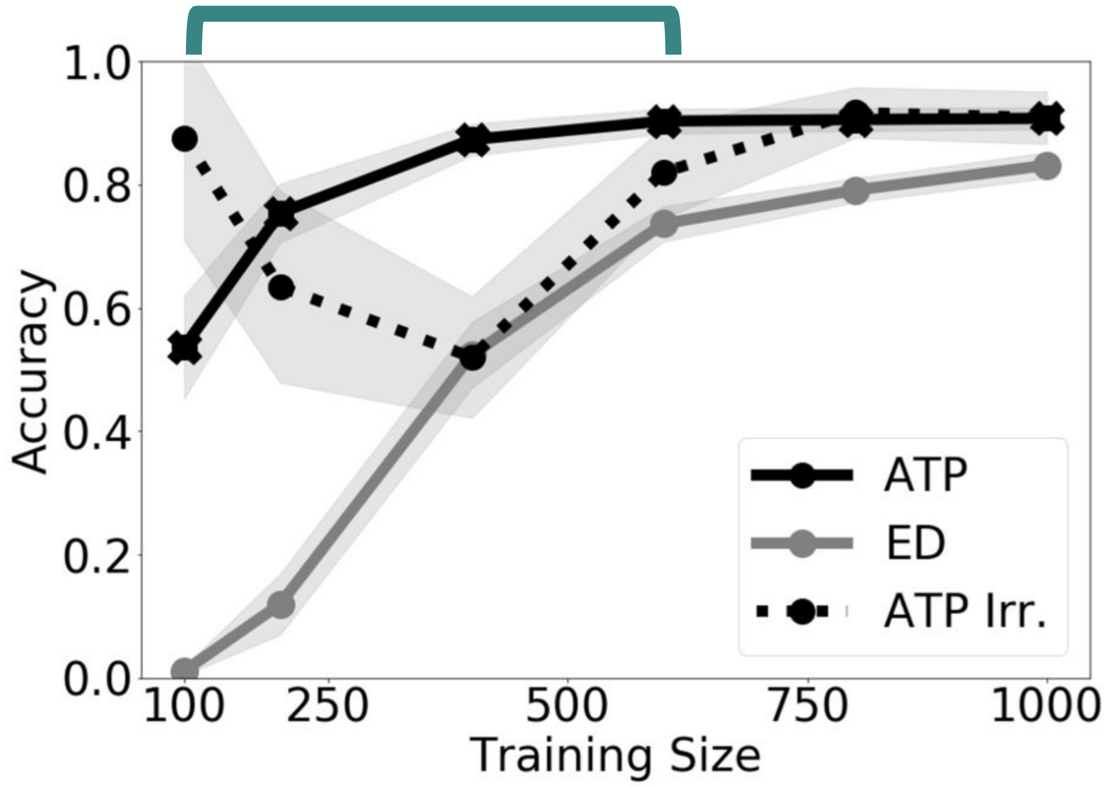
ATP Model: Sample learning trace

English past tense: inflect /want/

/wantɪd/



ATP: English Results

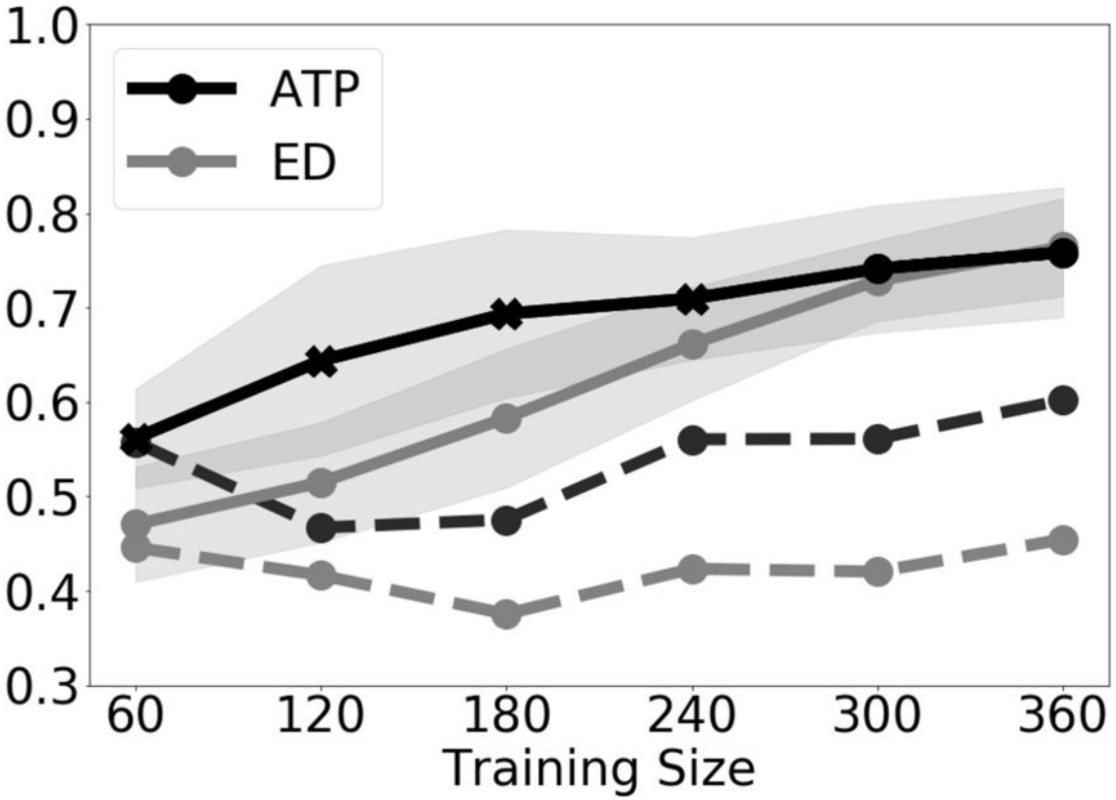


(a) English

- **Developmental regression** and **overregularization**
- Trained on **plausible vocabulary**
  - **1000** inflected forms
- Mechanistic account of developmental regression

ATP: German Results

- Relies less on gender than K&C
  - **Solid lines** = gender info given at test
  - **Dashed lines** = gender info not given at test
- Trained on **plausible vocabulary**
  - **400** inflected forms



(b) German

ATP: Summary

- Children **overregularize**
  - ***So does ATP!***
- Children show **developmental regression** when learning some paradigms
  - ***So does ATP!***
- Children learn from **extremely sparse, skewed input**
  - ***So does ATP!***

ATP gives mechanistic account of *why* these errors occur and what the resulting grammar looks like

Learning which Features are Marked: Sufficient Contrast Learner

(Payne 2022, 2023)

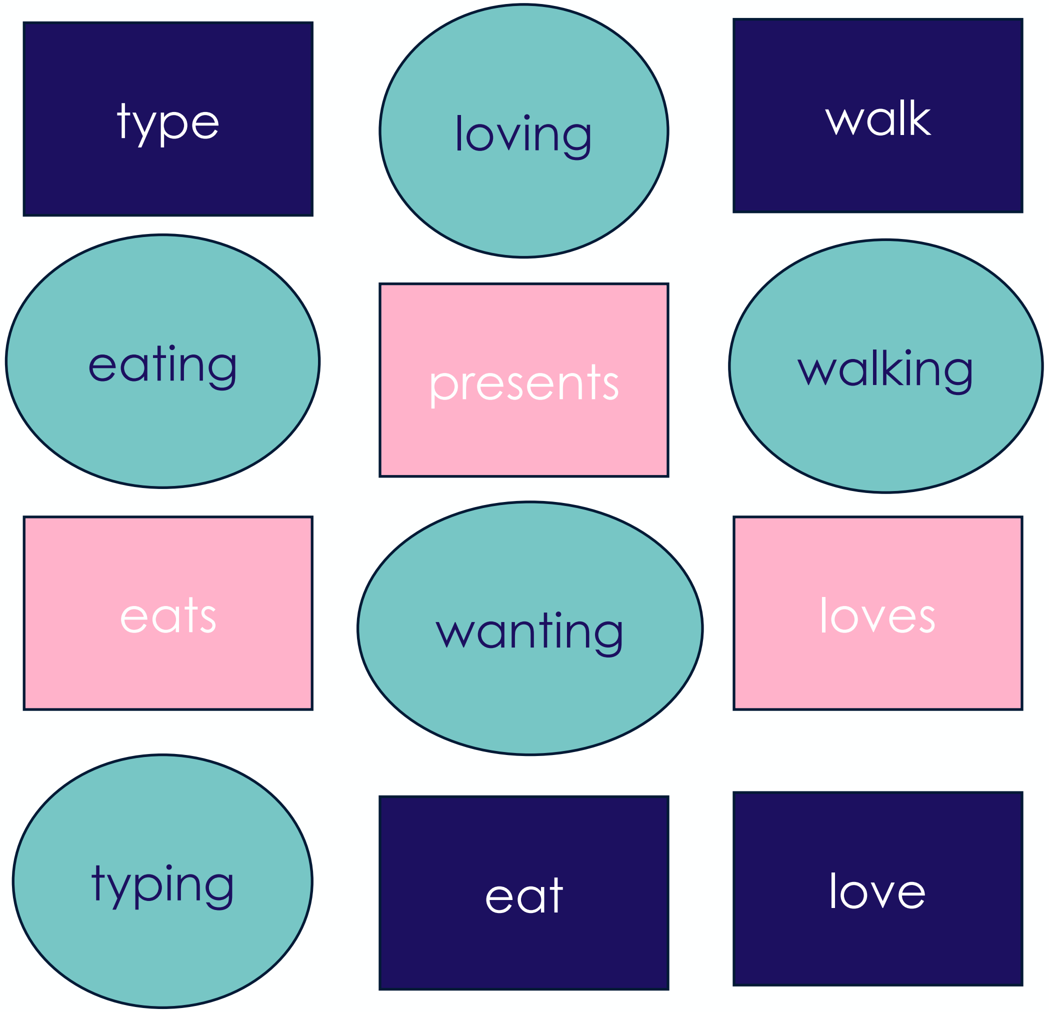
SCL Model: Preliminaries

- **Principle of Contrast: distinct forms ⇒ distinct meanings**
  - e.g. ***walk*** and ***walked*** must mean something different
- **Collisions:** one lemma in multiple inflected forms
  - e.g. ***walk-walked*** tells us that **+/-PAST** is marked
- **TSP:** when are there enough collisions to learn marking?
  - e.g. don't learn from ***am-are*** that **1 vs. 2** marked in English
  - Learn from ***walk-walked, sing-sang***, etc. that **+/-PAST** marked

SCL Model: Collisions

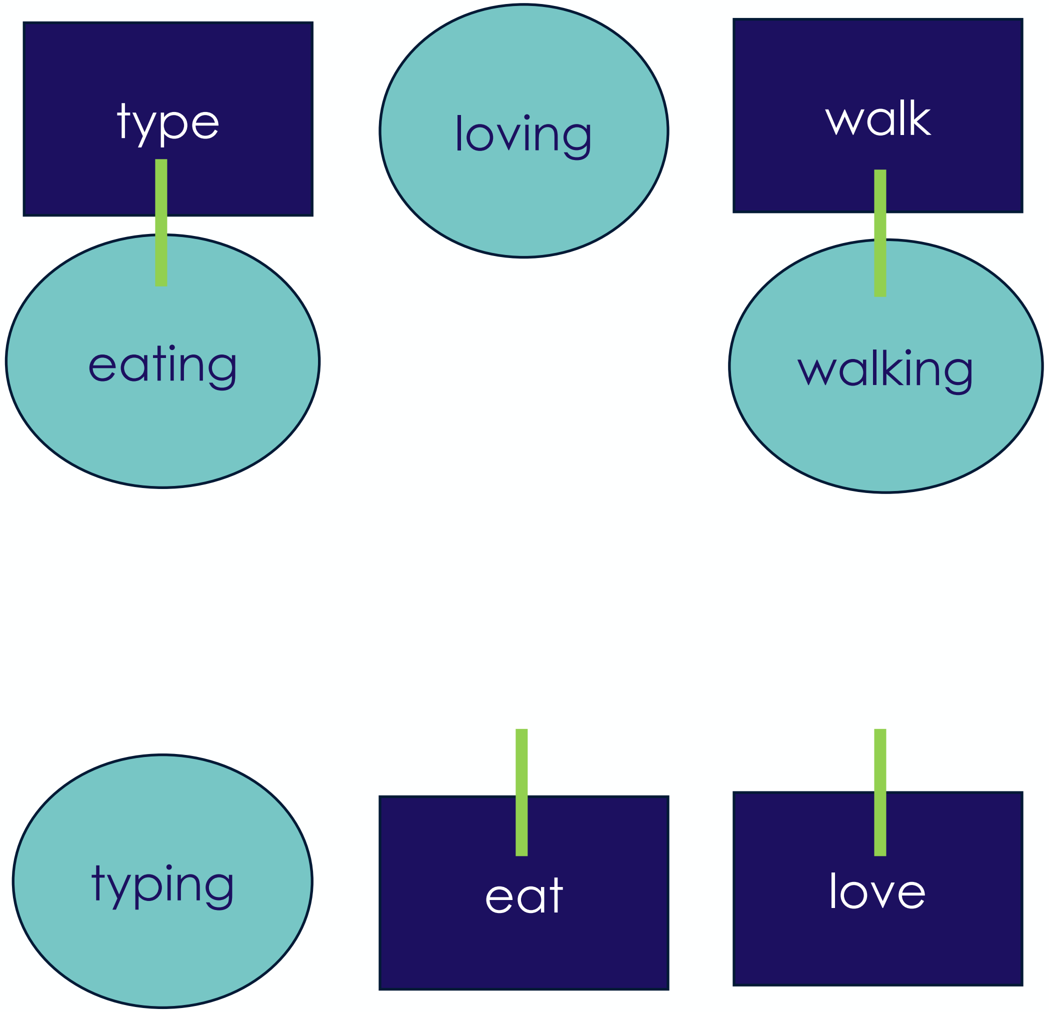
- Apply TSP **recursively** again!
  - Input taken in **incrementally**
  - When ***j^{th}*** **input** encountered, is there a collision?
  - If so, do enough forms appearing in **inflection A** also appear in **inflection B** in a different form?
    - If yes, **productive contrast learnt!** Subdivide and recurse
    - If no, **continue** to take in input
- Apply to **English verbs, German noun plurals, Spanish verbs**, and **Hebrew verbs**

SCL Model: Toy Example



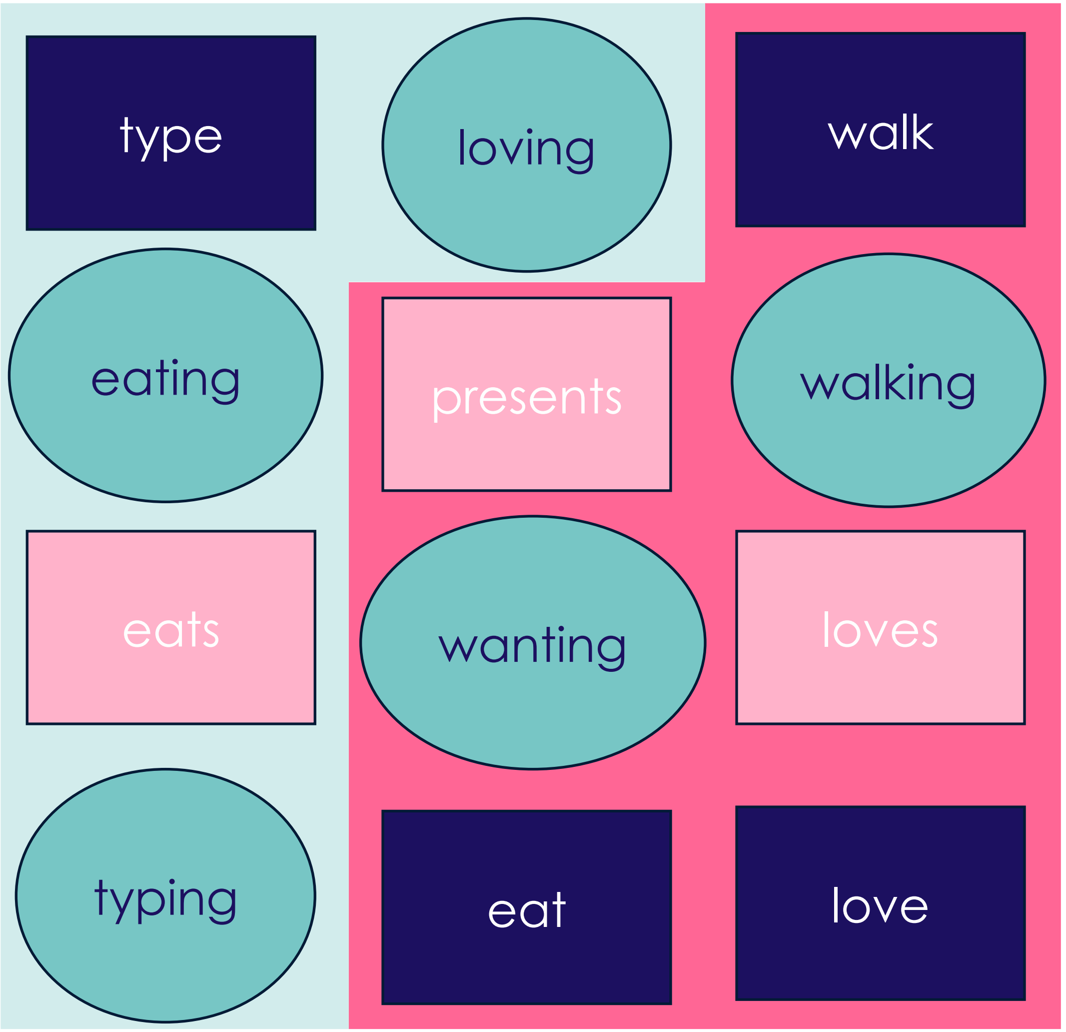
- Collision: ***walk*-*walking***
- **+/-PARTICIPLE** marked?

SCL Model: Toy Example



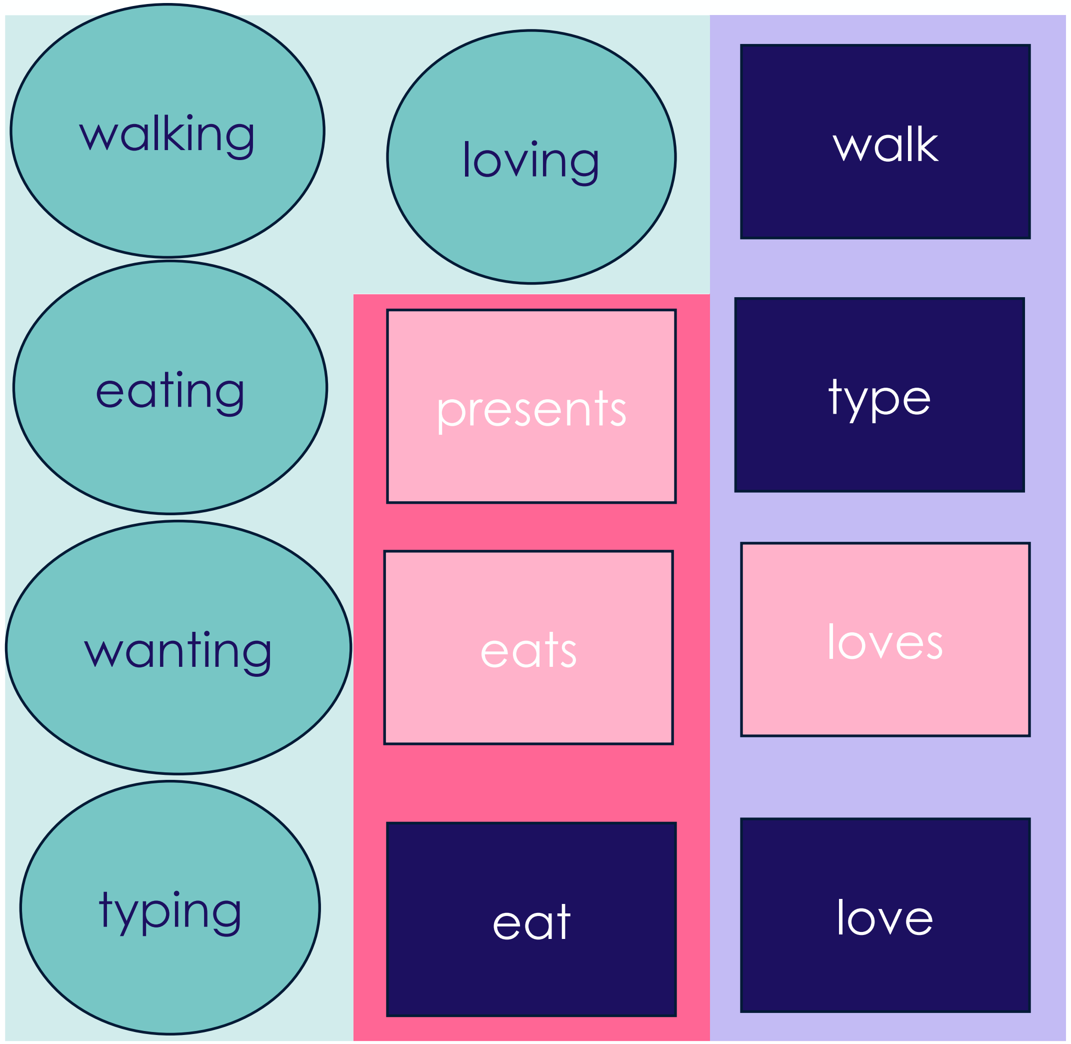
- Collision: ***walk-walking***
- **+/-PARTICIPLE** marked?
    - 5 participles, 4 collisions (not ***wanting***)
    - $N - M = 1 < \theta_5 = 3$ ✅
- Contrast 1 productive! **+/PARTICIPLE marked**

SCL Model: Toy Example



- Collision: ***walk-walking***
- **+/-PARTICIPLE** marked?
  - 5 participles, 4 collisions (not ***wanting***)
  - $N - M = 1 < \theta_5 = 3$ ✅
- Contrast 1 productive! **+/PARTICIPLE marked**
- **Subdivide** into **+PARTICIPLE** and **–PARTICIPLE** forms

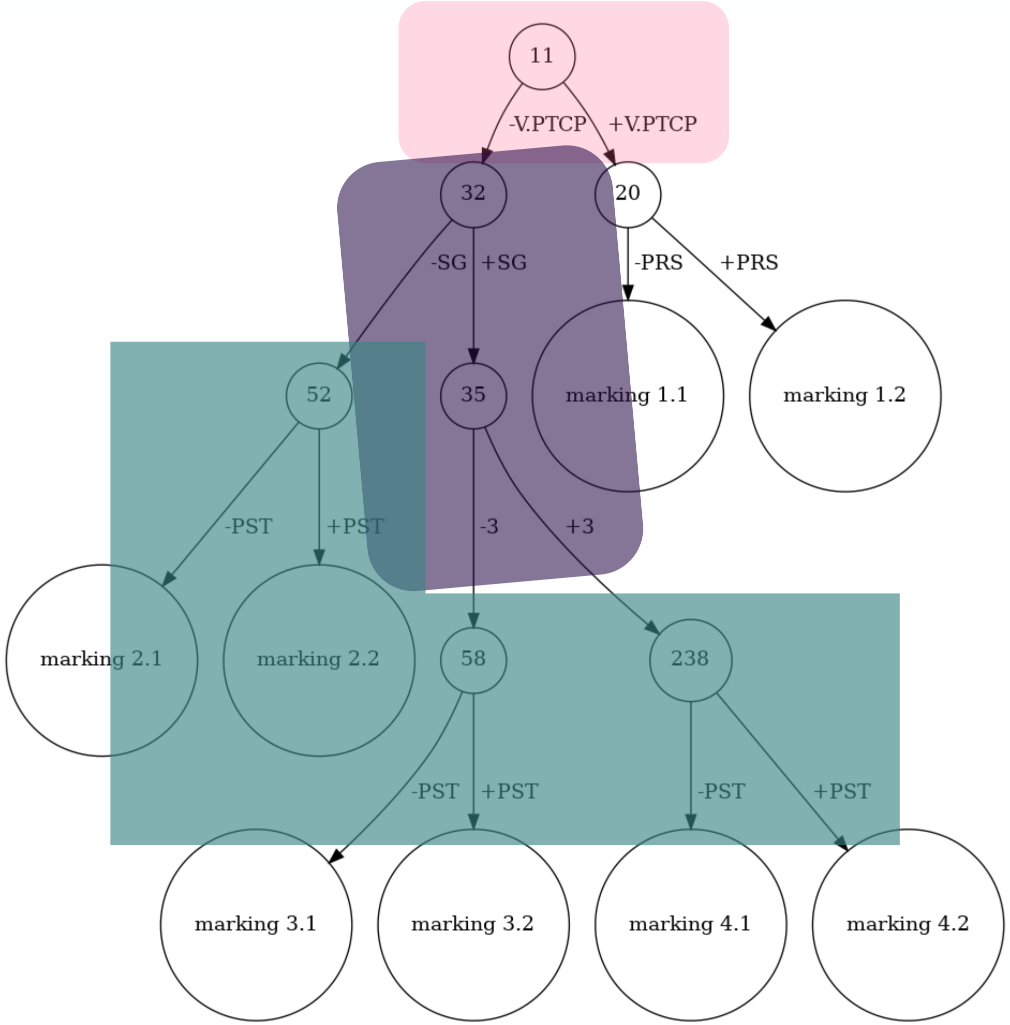
SCL Model: Toy Example



- Collision: ***walk*-*walking***
- **+/-PARTICIPLE** marked?
    - 5 participles, 4 collisions (not ***wanting***)
    - $N - M = 1 < \theta_5 = 3$ ✅
- Contrast 1 productive! **+/PARTICIPLE marked**
- **Subdivide** into **+PARTICIPLE** and **–PARTICIPLE** forms
- Recursively learn that **+/-3SG** marked

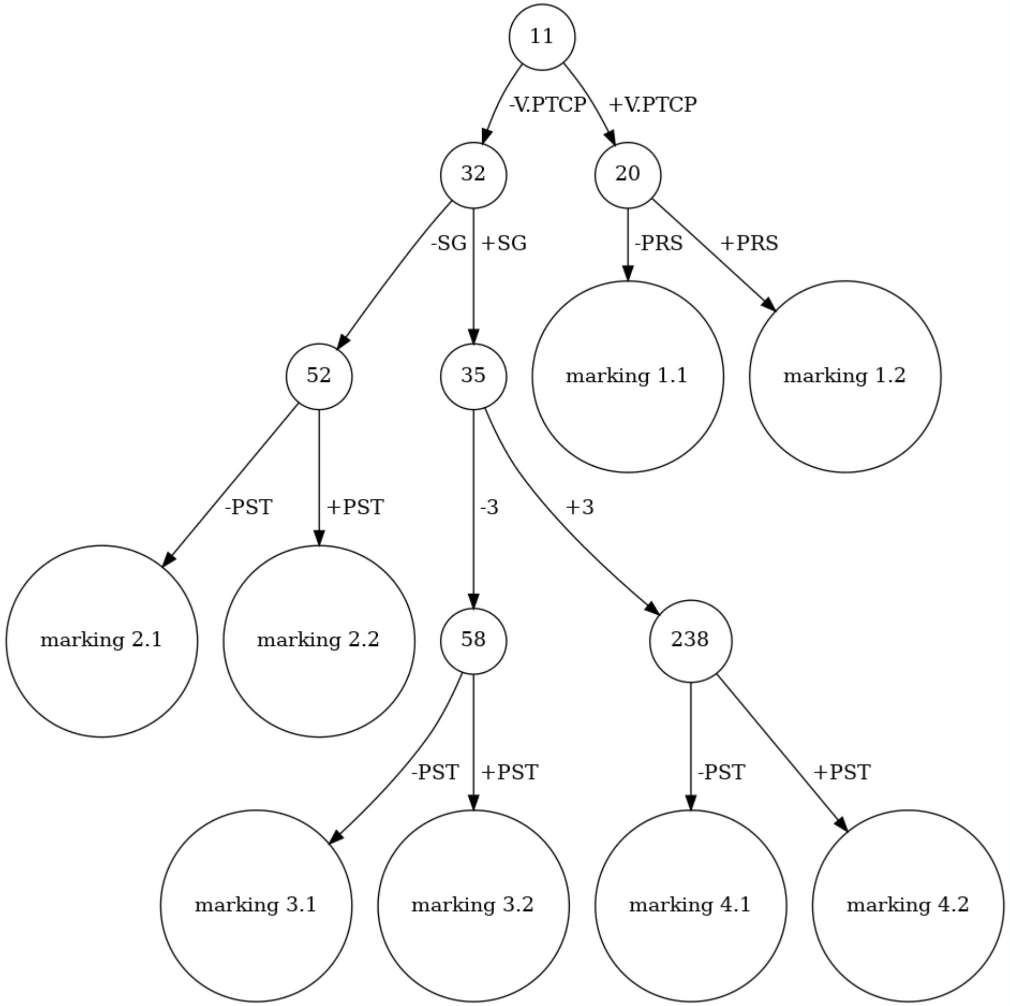
SCL: English Results

- Plausible **order of acquisition**
  1. **Participle** (*–ing*)
  2. **3,SG** (*-s*)
  3. **Past** (*-ed*)
- Plausible vocabulary size:
  - **112** lemmas
  - **238** inflected forms
- Learning past tense separately for each agreement?



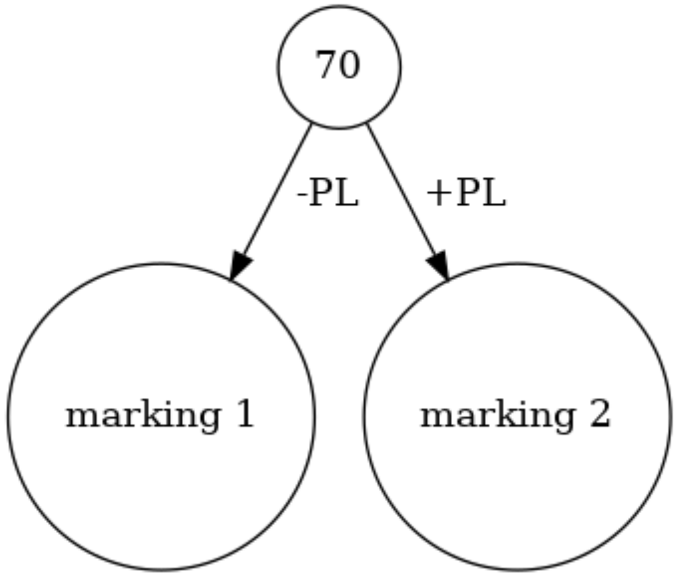
SCL: English Results

- Learning past tense separately for each agreement?
  - **Yang, Ellman, and Legate (2015):** past tense acquired later for learners of AAE
  - **Difference in input** = agreement, not tense
  - TSP tolerates ***relatively fewer exceptions*** for larger *N*



SCL: German results

- Plausible vocabulary size:
  - **66** lemmas
  - **70** inflected forms
- Well under vocab size at which plural affix overapplication begins



SCL: Spanish Results

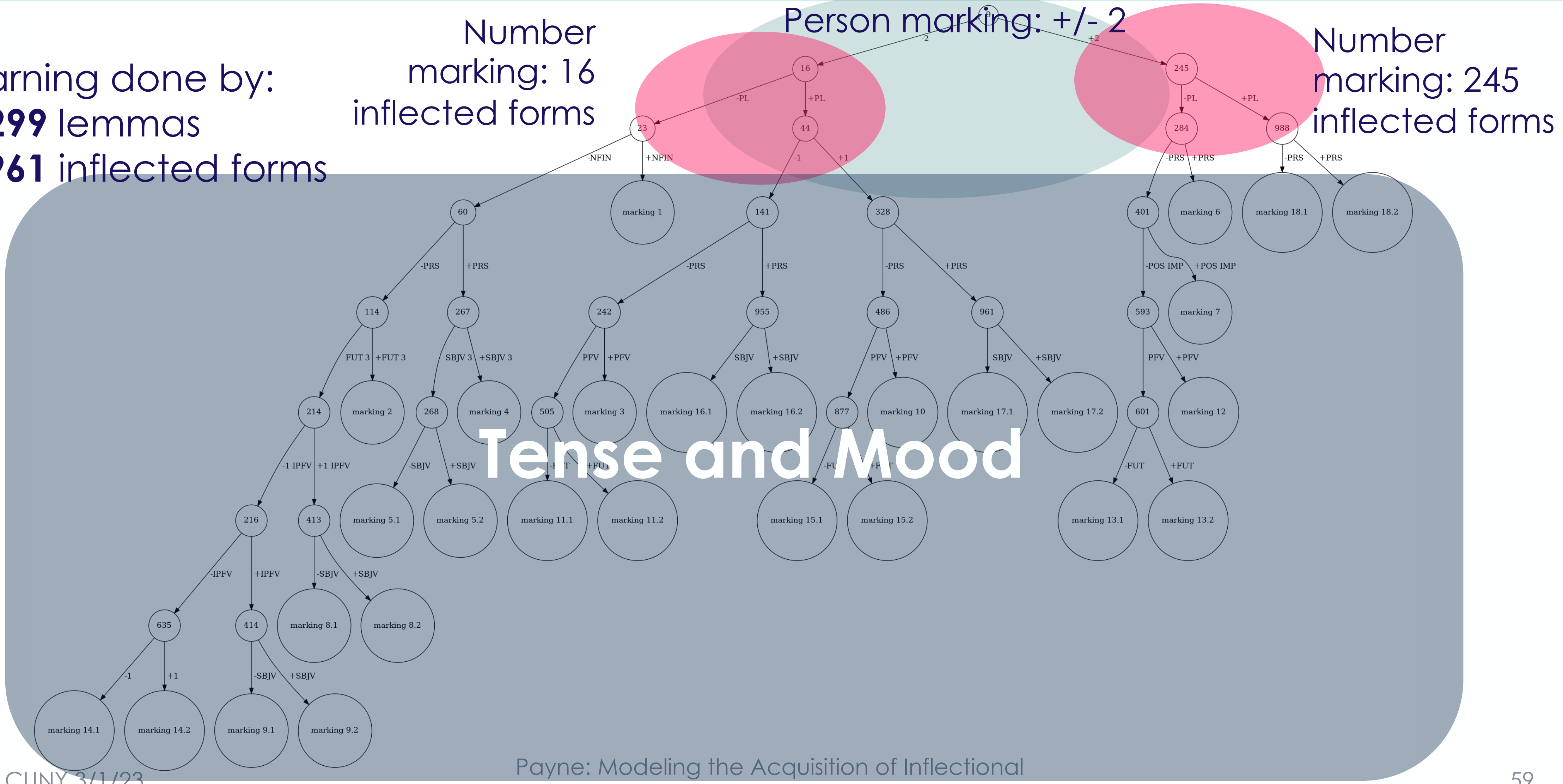
- Spanish order of acquisition:
  - **Finiteness & person marking**: 1;7
  - **Number marking**: 1;7-2;0
    - Second plural emerges later than other agreements in many learners
  - **Tense**: 2;0-2;2
  - **Mood**: 1;7-2;2

(Montrul 2004)

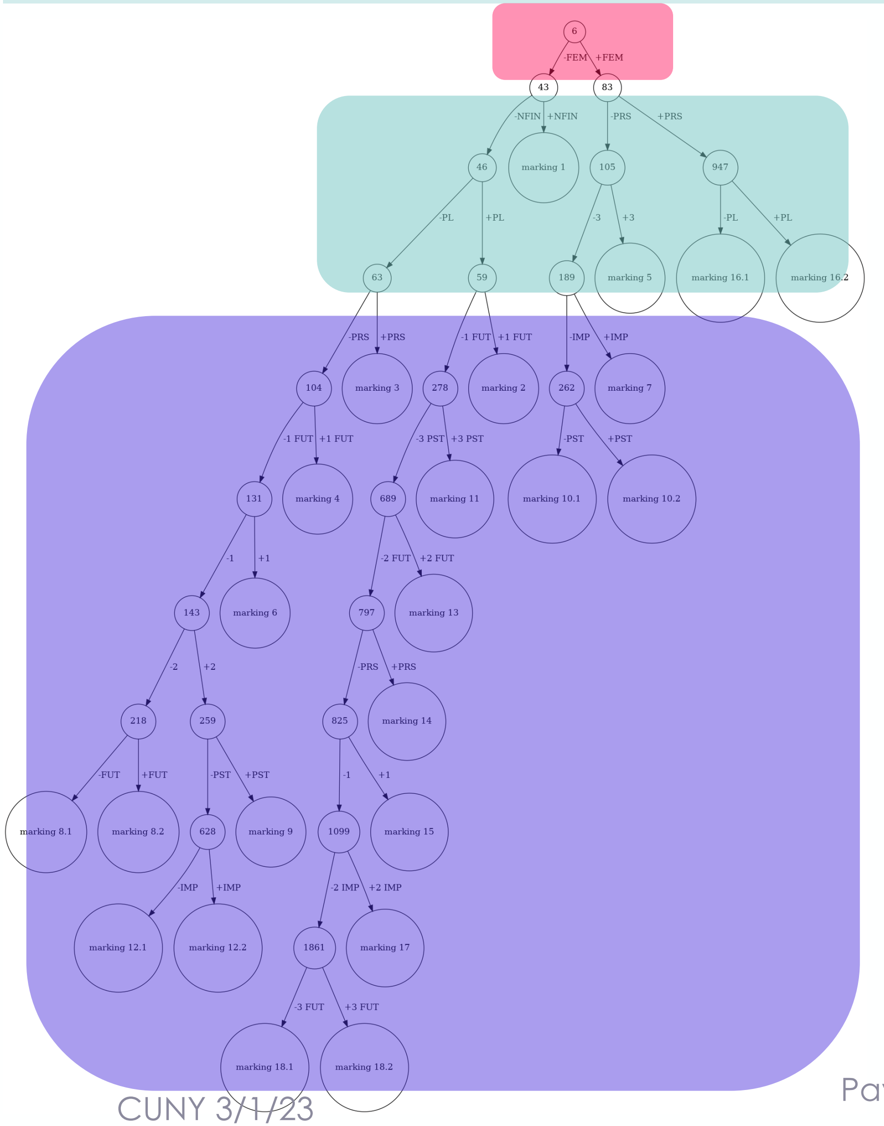
SCL: Spanish Results

Learning done by:

- **299** lemmas
- **961** inflected forms



SCL: Hebrew Results



- Hebrew order of acquisition:
  - **Person, number, gender** before tense
  - **Person vs. number** varies
  - **Gender** appears before or at the same time as **number**
- Our model:
  - Order of acquisition:
    - **Gender**, **person & number**, **tense**
  - Vocab size:
    - **323** lemmas
    - **1861** inflected forms

SCL Implications: Root Infinitives

- **Root Infinitive (RI) Stage:** stage of omission errors
- Cross-linguistically, **"richer" morphology ⇒ shorter RI stage**
- Richer morphology also means **more subdivision**
  - TSP tolerates more exceptions for **smaller *N***
  - More subdivision ⇒ **smaller *N***
  - Smaller *N* ⇒ **quicker learning** of inflectional categories
- SCL gives a **mechanistic account** of cross-linguistic differences

(Philips 1995, Legate & Yang 2007)

SCL: Summary

- Children **omit affixes**
  - ***SCL gives an account for why!***
- Children show clear **order of acquisition effects**
  - ***So does SCL!***
- Children learn from **extremely sparse, skewed input**
  - ***So does SCL!***

SCL gives mechanistic account of order of acquisition, omission errors and cross-linguistic differences in acquisition

Conclusion & Future Directions

Conclusion: Getting the Right Stuff Wrong

	NNs	ATP	SCL
Learn from plausible data	❌	✅	✅
Account for over-regularization and developmental regression	❌	✅	---
Account for omission and the RI stage	❌	---	✅
Interpretability	❌	✅	✅

SCL & ATP

- **SCL** learns the features required by **ATP**
  - **Combination** of these two learning strategies
  - First learn the inflectional classes, then map them to form
- **Expand ATP:**
  - Handle **templatic & agglutinative** morphology
- **Expand SCL:**
  - Explore model **subdivision predictions**
  - Learn features from **distributional information**

Thank you!!

I am grateful to Charles Yang, Jordan Kodner, Jeff Heinz, JulieAnne Legate, Mark Aronoff, Bob Berwick and his lab, Zoey Liu, Caleb Belth, Salam Khalifa, and attendees of the Stony Brook University Brown Bag for comments, feedback, and collaboration on the work presented here.

I am grateful for funding by the Instituted for Advanced Computational Science Graduate Research Fellowship and the National Science Foundation Graduate Research Fellowship.

References

Belth, Caleb, Sarah Payne, Deniz Beser, Jordan Kodner, and Charles Yang. 2021. "The Greedy and Recursive Search for Morphological Productivity." *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society*.

Chan, Erwin. 2008. *Structures and Distributions in Morphology Learning*. PhD dissertation, University of Pennsylvania.

Clahsen, Harald, Fraibet Aveledo, and Iggy Roca. 2002. "The Development of Regular and Irregular Verb Inflection in Spanish Child Language." *Journal of Child Language*, 591-622.

Clahsen, Harald, and Martina Penke. 1992. "The Acquisition of Agreement Morphology and its Syntactic Consequences: New Evidence on German Child Language from the Simone Corpus." *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Acquisition*, 181-223.

Corkery, Maria, Yevgen Matusevych, and Sharon Goldwater. 2019. "Are we there yet? Encoder-Decoder Neural Networks as Cognitive Models of English Past Tense Inflection." *Proceedings of the 57th Conference of the Association for Computational Linguistics*, 3868-3877.

Fenson, Larry, Philip S Dale, Steven Reznick, Elizabeth Bates, Donna Thal, Stephen Pethick. 1994. *Variability in Early Communicative Development*. Monographs of the Society for Research in Child Development, 1-185.

Goodman, Judith, Dale Philip, and Ping Li. 2008. "Does Frequency Count? Parental Input and the Acquisition of Vocabulary." *Journal of Child Language*, 515-531.

Johnson, Valerie, Jill de Villiers, and Harry Seymour. "Agreement Without Understanding? The case of third person singular –s." *First Language*, 317-330.

Kim, Yun Jung and Megha Sundara. 2021. "6-month-olds are sensitive to English morphology." *Developmental Science*.

Kirov, Christo and Ryan Cotterell. 2018. "Recurrent Neural Networks in Linguistic Theory: Revisiting Pinker and Prince (1988) and the Past Tense Debate." *Transactions of the Association for Computational Linguistics*, 651-665.

Legate, Julie Anne & Charles Yang. 2007. "Morphosyntactic Learning and the Development of Tense." *Language Acquisition*, 315-344.

Lignos, Constantine, and Charles Yang. 2016. "Morphology and Language Acquisition." *The Cambridge Handbook of Morphology*, 743-764.

MacWhinney, Brian. 2000. The CHILDES Project: Tools for Analyzing Talk, Transcription Format, and Programs. *Psychology Press*

Marcus, Gary, Steven Pinker, Michael Ullman, Michelle Hollander, T John Rosen, Fei Xu, and Harald Clahsen. 1992. "Overregularization in Language Acquisition." *Monographs of the Society for Research in Child Development*.

Marcus, Gary, Ursula Brinkmann, Harald Clahsen, Richard Wiese, and Steven Pinker. 1995. "German Inflection: The Exception that Proves the Rule." *Cognitive Psychology*, 189-256.

Marquis, Alexandra and Rushen Shi. 2012. "Initial Morphological Learning in Preverbal Infants." *Cognition*, 61-66.

Maslen, Robert, Anna Theakston, Elena Lieven, and Michael Tomasello. 2004. "A Dense Corpus Study of Past Tense and Plural Overregularization in English." *Journal of Speech, Language, an Hearing Research*, 1319-1333.

McCurdy, Kate, Sharon Goldwater, and Adam Lopez. 2020a. "Inflecting When There's No Majority: Limitations of Encoder-Decoder Neural Networks as Cognitive Models of German Noun Plurals." *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 1745-1756.

McCurdy, Kate, Adam Lopez, and Sharon Goldwater. 2020b. "Conditioning, but on Which Distribution? Grammatical Gender in German Plural Inflection." *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, 59-65.

Montrul, Silvina. 2004. *The Acquisition of Spanish: Morphosyntactic Development in Monolingual and Bilingual L1 Acquisition and Adult L2 Acquisition*. John Benjamins Publishing.

Payne, Sarah, Jordan Kodner, and Charles Yang. 2021. "Learning Morphological Productivity as Meaning-Form Mappings." *Proceedings of SCiL*.

Payne, Sarah. 2022. *When Collisions are a Good Thing: The Acquisition of Morphological Marking*. Bachelor's Thesis, University of Pennsylvania.

Payne, Sarah. 2023. "Contrast, Sufficiency, and the Acquisition of Morphological Marking." *Proceedings of the 47th Boston University Conference on Language Development*.

Pinker, Steven, and Alan Prince. 1988. "On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition." *Cognition*, 73-193.

Phillips, Colin. 1995. "Syntax at Age Two: Cross-Linguistic Differences." *Language Acquisition*, 70-120.

Rumelhart, David and James McClelland. 1986. "On Learning the Past Tenses of English Verbs." In *Psycholinguistics: Critical Concepts in Psychology*.

Soderstrom, Melanie, Kenneth Wexler, and Peter Jusczyk. 2002. "English-learning Toddlers' Sensitivity to Agreement Morphology in Receptive Grammar." *Proceedings of the 26th Annual Boston University Conference on Language Development*.

Xu, Fei, and Steven Pinker. 1995. "Weird Past Tense Forms." *Journal of Child Language, 531-556.*

Yang, Charles. 2016. *The Price of Linguistic Productivity: How Children Learn to Break the Rules of Language*. MIT Press.

Zipf, George Kingsley. 1936. *The Psycho-Biology of Language: An Introduction to Dynamic Philology*. Routledge.