

# Possible and Probable Errors in Child Language

**Sarah Brogden Payne**

Stony Brook University

[sarah.payne@stonybrook.edu](mailto:sarah.payne@stonybrook.edu)



Stony Brook  
University

Poznań Linguistics Meeting  
15<sup>th</sup> September 2023



**iACS**  
INSTITUTE FOR ADVANCED  
COMPUTATIONAL SCIENCE

# Blind Alley Developments

Systematic **deviations** of **child productions** from the input

**Two main types of BADs:**

- **Weak BADs:** *mis-application* of a pattern **present in the input**
  - e.g. overapplication of **-e-** at the boundary of German **noun-noun compounds**
- **Strong BADs:** use of a pattern *never attested* marking the relevant category **in the input**
  - e.g. use of **reduplication** to express iterativity in **Russian**

Dressler et al. (2020)

# Blind Alley Developments

Systematic **deviations** of **child productions** from the input

**Two main types of BADs:**

- **Weak BADs:** *mis-application* of a pattern **present in the input**
  - e.g. overapplication of **-e-** at the boundary of German **noun-noun compounds**
- **Strong BADs:** use of a pattern *never attested* marking the relevant category **in the input**
  - e.g. use of **reduplication** to express iterativity in **Russian**

Dressler et al. (2020)

# The Data: German Noun-Noun Compounds

- **-e-** is used at the boundary of some compounds by adults
  - e.g., *Hund-e-würstel* ‘dog poop’
- **Infrequent** and **unproductive** under most conditions
- **Temporarily overapplied** by German-learning children
  - *Luft-ballon-e* → \**Luft-e-bon-e* ‘air balloons’
  - *Mineral-wasser* → \*(*Mine*)*ral-e-wasser* ‘mineral water’
  - *Kind-er-kassette* → \**Kinn-e-sette* ‘child cassette’

Korecky-Kröll (2017), Dressler et al. (2020)

# The Data: English Past Tense $\text{ɪ} \rightarrow \text{æ}/\_\eta\#$

- $\text{ɪ} \rightarrow \text{æ}/\_\eta\#$  used for some past tense forms by adults
  - e.g., *sing-s**ang***, *ring-r**ang***
- **Infrequent** and **unproductive**
  - New  $\text{-ɪ}\eta\#$  verbs take **productive -ed**
    - *Bing-Bing**ed***, *bling-bling**ed***
- **Temporarily overapplied** by English-learning children
  - *bring-brought*  $\rightarrow$  *bring-br**ang***
  - *fling-flung*  $\rightarrow$  *fling-fl**ang***

Xu & Pinker (1995), Payne & Yang (2023)

# A Theory of BADs in Acquisition: Desiderata

- Explain their **ephemeral nature & timeline**
  - What causes children to **enter BADs**?
  - What causes them to **escape from them**?
    - Why do they escape so **quickly**?
- Explain the **patterns that children construct**
  - Does **everything** go?
  - Which constructions can we **expect the child to produce during BADs**?
  - Which constructions **don't we expect**?

Learning-Theoretic  
Account:  
The Tolerance  
Principle

Natural  
Morphology  
+  
The Tolerance  
Principle

# Contribution

I propose an account of **Blind Alley Developments** that marries the **complementary approaches** of the **Tolerance Principle** and **Natural Morphology** to provide both a **formalization of the timeline** of BADs and predictions about **which types of BAD constructions are likely or expected.**

# Outline

- **Proposal:** a mechanistic account of BADs
- **Applying the proposal:**  $\mathfrak{r} \rightarrow \text{æ}/\_\_\eta\#$ 
  - **Timeline**
  - **Does everything go?**
- A brief word on **strong BADs**
- **Conclusion & Open Questions**



# The Tolerance Principle

- **Intuition:** linguistic process must “**earn**” **productivity**
  - Do so by being applicable to a **sufficiently large number of candidates**, calibrated over the *learner’s internal vocabulary*
- Learner calculates two values for a rule **R**:
  - **N**: number of items *in a learner’s internal vocabulary* fitting **R**’s description
  - **e**: number of these items to which **R does not apply**
- Given these values, **R** is **productive** iff:

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

Yang (2016)

# Applying the TP: Recursive Learning

- TP **evaluates** hypothesized rules for productivity
  - Need a **mechanism to hypothesize the rules** to begin with!
- **Previous work: recursive** learning with **subdivision**
  - **Payne (2022)**: learn **inflectional categories** by counting “**collisions**”
  - **Belth et al. (2021)**: learn **inflectional processes** by **frequency**
- **NatMorph preferences & trade-offs** may also play a role in driving subdivision, instead of heuristics like frequency
  - **NatMorph Preferences** can guide children to some BADs over others

# The Proposal

## The Tolerance Principle

### Ephemeral Nature & Timeline

- Places ***precise, mechanistic bounds*** on the timeline of BADs grounded in a **formal, quantitative account of language acquisition**
- Provides a quantitative theory of **the critical mass of opposing input**

# The Proposal

	The Tolerance Principle	Natural Morphology
Ephemeral Nature & Timeline	<ul style="list-style-type: none"><li>Places <b>precise, mechanistic bounds</b> on the timeline of BADs grounded in a <b>formal, quantitative account of language acquisition</b></li><li>Provides a quantitative theory of <b>the critical mass of opposing input</b></li></ul>	<ul style="list-style-type: none"><li>Accounts for children's escape in terms of <b>opposing input</b></li><li>The <b>input is always opposing</b> – why does the child escape <b>precisely when they do?</b></li><li>What is the <b>critical mass</b> of opposing input?</li></ul>

# The Proposal

	The Tolerance Principle	Natural Morphology
Ephemeral Nature & Timeline	<ul style="list-style-type: none"><li>Places <b>precise, mechanistic bounds</b> on the timeline of BADs grounded in a <b>formal, quantitative account of language acquisition</b></li><li>Provides a quantitative theory of <b>the critical mass of opposing input</b></li></ul>	<ul style="list-style-type: none"><li>Accounts for children's escape in terms of <b>opposing input</b></li><li>The <b>input is always opposing</b> – why does the child escape <b>precisely when they do?</b></li><li>What is the <b>critical mass</b> of opposing input?</li></ul>
Patterns that Children Construct	<ul style="list-style-type: none"><li>Has little to say about what BADs are <b>likely</b>, only which are <b>theoretically possible</b> given the child's <b>vocabulary</b></li></ul>	

# The Proposal

	The Tolerance Principle	Natural Morphology
Ephemeral Nature & Timeline	<ul style="list-style-type: none"> <li>Places <b>precise, mechanistic bounds</b> on the timeline of BADs grounded in a <b>formal, quantitative account of language acquisition</b></li> <li>Provides a quantitative theory of <b>the critical mass of opposing input</b></li> </ul>	<ul style="list-style-type: none"> <li>Accounts for children's escape in terms of <b>opposing input</b></li> <li>The <b>input is always opposing</b> – why does the child escape <b>precisely when they do?</b></li> <li>What is the <b>critical mass</b> of opposing input?</li> </ul>
Patterns that Children Construct	<ul style="list-style-type: none"> <li>Has little to say about what BADs are <b>likely</b>, only which are <b>theoretically possible</b> given the child's <b>vocabulary</b></li> </ul>	<ul style="list-style-type: none"> <li>Predicts the BADs that children are <b>likely to construct</b> in terms of <b>preferences &amp; preference rankings</b></li> </ul>

# The Proposal

	The Tolerance Principle	Natural Morphology
Ephemeral Nature & Timeline	<ul style="list-style-type: none"> <li>Places <b>precise, mechanistic bounds</b> on the timeline of BADs grounded in a <b>formal, quantitative account of language acquisition</b></li> <li>Provides a quantitative theory of <b>the critical mass of opposing input</b></li> </ul>	<ul style="list-style-type: none"> <li>Accounts for children's escape in terms of <b>opposing input</b></li> <li>The <b>input is always opposing</b> – why does the child escape <b>precisely when they do?</b></li> <li>What is the <b>critical mass</b> of opposing input?</li> </ul>
Patterns that Children Construct	<ul style="list-style-type: none"> <li>Has little to say about what BADs are <b>likely</b>, only which are <b>theoretically possible</b> given the child's <b>vocabulary</b></li> </ul>	<ul style="list-style-type: none"> <li>Predicts the BADs that children are <b>likely to construct</b> in terms of <b>preferences &amp; preference rankings</b></li> </ul>

# Weak BADs Under our Proposal

- TP sets **mechanistic bounds** on the **start and end** of the BAD
  - **Start:** child knows that the *category must be marked*
  - **End:** the BAD process is *no longer sufficiently dominant*
- TP sets **mechanistic bounds** on the **possible types** of weak BADs
  - BAD process *must be sufficiently dominant* over internal vocabulary
- **NatMorph** predicts some weak BADs to be **more likely than others**
  - Child is *unlikely to consider every process* that may be productive
  - **NatMorph preferences** can *guide the child to some weak BADs* over others



# Timeline of the Weak BAD: $\text{ɪ} \rightarrow \text{æ} / \_ \eta \#$

- **Frequency & order of acquisition** correlated
  - **Model “typical” child** with strictly **frequency-based** order of acquisition
  - Frequencies calculated from **North American English CHILDES**
- By hypothesis, child knows that **past tense must be marked**
  - **Payne (2022) model prediction:** learn past tense marked by **112 verbs**

Goodman et al. (2008), Yang (2016), Belth et al. (2021), Payne & Yang (2023)

# Timeline of the Weak BAD: $\text{ɪ} \rightarrow \text{æ} / \_ \eta \#$

- When (frequency-based) vocab contains **200 verbs**:
  - **76 are irregular**, so **-ed** is not productive ( $\theta_{200} = 37 < 76$ )
  - **Subdivide**:
    - **3 -ɪη# verbs**: *bring-brought, sing-sang, ring-rang*
    - **2/3 sufficient**:  $\text{ɪ} \rightarrow \text{æ} / \_ \eta \#$  **productive** over internal vocabulary
      - $\text{ɪ} \rightarrow \text{æ} / \_ \eta \#$  **BAD is entered**

Goodman et al. (2008), Yang (2016), Belth et al. (2021), Payne & Yang (2023)

# Timeline of the Weak BAD: $\text{ɪ} \rightarrow \text{æ} / \_ \eta \#$

- When vocab contains **800 verbs**:

- **8 -ɪŋ# verbs**:

*bring-brought, sing-sang, ring-rang, fling-flung, spring-sprang, sting-stung, swing-swung, wing-winged*

- 3/8 not sufficient: ( $\theta_8 = 3 < 5$ )

- $\text{ɪ} \rightarrow \text{æ} / \_ \eta \#$  cannot be supported anymore

- The BAD is **abandoned**

Yang (2016), Belth et al. (2021), Payne (2022), Payne & Yang (2023)

# Does Everything Go?: ɪ → æ / \_ŋ#

- ɪ → æ/\_ŋ# BADs are **relatively common**
  - *bring-brang* is widely attested
  - *swing-swang* and *fling-flang* also attested
- -ɔt# BADs are **entirely unattested**
  - e.g. *stink-stought* from *think-thought*

Why ɪ → æ/\_ŋ# but not -ɔt#?

# Does Everything Go?: $\text{ɪ} \rightarrow \text{æ} / \_ \eta \#$

- **Tolerance Principle:** an  $-\text{ɔt}\#$  rule will **never reach productivity**, even if it is hypothesized
  - Defining **the context** for  $-\text{æ}\eta\#$  **past tense forms:**
    - **Straightforward** (verbs ending in  $-\text{ɪ}\eta\#$  in the present)
    - **(Temporarily) supported by the input** (2/3 take  $-\text{æ}\eta\#$  in the past)

# Does Everything Go?: $\text{ɪ} \rightarrow \text{æ} / \_ \eta \#$

- Defining **the context** for  **$-\text{ɔt} \#$  past tense forms**:
  - **bring** and **think**:  $\text{ɪ}\eta(\text{k}) \rightarrow \text{ɔt} / \_ \#$ ?
  - Verbs fitting  $\text{ɪ} \rightarrow \text{æ} / \_ \eta \#$  also fit this rule description (**sing, ring**)
  - Verbs fitting  $\text{ɪ} \rightarrow \text{æ} / \_ \eta \text{k} \#$  also fit this rule description (**drink, sink**)
- When the vocab contains **250 verbs**:
  - **6  $-\text{ɪ}\eta(\text{k}) \#$**  :  
*think-thought, bring-brought, drink-drank, ring-rang, sing-sang, sting-stung*
  - 2/6 not sufficient ( $\theta_6 = 3 < 4$ )

# Does Everything Go?: $\text{ɪ} \rightarrow \text{æ} / \_ \eta \#$

- Increased vocabulary & subdivision **don't make an -ɔt# BAD possible:**
  - Introduce *catch*:  $-\widehat{\text{æ}tʃ} \rightarrow \text{ɔt}/\_ \#?$ 
    - *scratch*, *hatch*, and *attach* are similarly frequent
  - Introduce *teach*:  $-\widehat{\text{itʃ}} \rightarrow \text{ɔt}/\_ \#?$ 
    - *reach* is similarly frequent
- Thus, **an -ɔt# BAD will never reach productivity**

# Does Everything Go?: ɪ → æ / \_ŋ#

- **Natural Morphology:** some BADs more expected than others
  - **Bi-uniqueness**
    - All verbs that take **-æŋ#** in the past take **-ŋ#** in the past participle
    - Many other verbs have a **syncretism between the simple past & past participle**
      - **This is true of most -ɔt# past tense forms**



# Weak BADs: Summary

- Weak BADs persist only as long as they are productive **over the learner's internal vocabulary**, as measured by the TP
- **TP** delineates which possible BADs **may reach temporary productivity** and which may not
- **NatMorph** predicts some weak BADs to be more likely
  - Child is **unlikely to consider every potential BAD**
  - NatMorph preferences can **guide the child to some weak BADs** over others, rather than considering all possible BADs

# A Brief Word on Strong BADs

- **Weak BADs:** *some process* is productive over the learner's vocabulary, albeit **not the adult-like one**
  - **Strong BAD:** *no process is productive*
- TP sets **mechanistic bounds** on **timeline** of the BAD
  - **Start:** child knows that the **category must be marked**, but not how
  - **End:** some process in the input becomes **sufficiently dominant** over the child's vocabulary

# A Brief Word on Strong BADs

- Why do children construct the strong BADs they do?
  - Child knows *that the category must be marked*, but **has no productive process** with which to mark it
  - **NatMorph preferences**: predict what types of constructions children are likely to build
    - **e.g. iconicity**: reduplication to express iterativity in Russian

# Strong BADs: Summary

- Strong BADs persist only as long as there is no productive process **over the learner's internal vocabulary**
  - **Open question:** how does the child escape the BAD if no process becomes productive? (e.g. **defectivity**)
- NatMorph preferences predict the **types of constructions we expect children to make** during their strong BADs

# Conclusions & Next Steps

- Marriage of **learning-theoretic TP** with **NatMorph preferences** gives a theory of acquisition with both:
  - A **formalization of the timeline** of BADs
  - Predictions about **which types of BADs** are **possible** & which are **likely or expected**
- Open questions:
  - How do children escape strong BADs when **no process ever reaches productivity?**

# Thank you!!!

I am particularly grateful to Prof. Dressler for his feedback and collaboration on this work, and the opportunity to present it here!

I am also grateful to Charles Yang, Jordan Kodner, Mark Aronoff, Jeff Heinz, Salam Khalifa, and Scott Nelson for comments.

This work was supported by the Institute for Advanced Computational Science Graduate Research Fellowship and the National Science Foundation Graduate Research Fellowship.



# Applying the TP: Recursive Learning

## Payne 2022: Inflectional Categories

- **Collision:** word appears in two categories in different forms
  - e.g. **walk-walked**  $\Rightarrow$   $\pm$  PAST
- Given **N** items, do enough have a collision between categories **A** and **B**?
  - If yes, **learn contrast and recurse**
  - Otherwise, **continue to take in input**
- **Terminate** when:
  - No more productive contrasts available

## Belth et al 2021: Inflectional Processes

- Given **N** items, do enough realize inflectional process **R**?
  - If yes, **learn productive rule**
  - If not, **subdivide based on the most frequent features** and recurse on each resulting set
- **Terminate** when:
  - Productive rule discovered
  - No more subdivisions possible

# Blind Alley Developments

Two main types of **Blind Alley Developments (BADs)**:

- **Strong BADs**: use of a pattern *never attested* marking the relevant category

## Root reduplication in Russian

- Reduplication is present **as a formal pattern in Russian** (e.g. used to express **intensification**)
  - тѣѣтѣ → тѣѣтѣ тѣѣтѣ (few → very few)
- **Iterativity** is marked in Russian with **imperfective verbs** or **secondary means**
- Reduplication used by children studied by Dressler et al. 2020 to mark **iterativity**:
  - *njam* → **njam-njam** 'I'm eating'
  - *prygat* → **pik-pik** (repeated jumping)



# Blind Alley Developments

Two main types of **Blind Alley Developments (BADs)**:

- **Strong BADs**: use of a pattern *never attested* marking the relevant category

## Root reduplication in Russian

- Reduplication is present **as a formal pattern in Russian** (e.g. used to express **intensification**)
  - тѣмѣ́ → тѣмѣ́ тѣмѣ́ (few → very few)
- **Iterativity** is marked in Russian with **imperfective verbs** or **secondary means**
- Reduplication used by children studied by Dressler et al. 2020 to mark **iterativity**:
  - *njam* → ***njam-njam*** 'I'm eating'
  - *prygat* → ***pik-pik*** (repeated jumping)

## Reduplication & Vowel Lengthening in Greek

- **Reduplication** and **vowel lengthening** appear as formal patterns in Greek
  - πρωί → πρωί-πρωί (morning → early morning)
  - πρωί → πρ:ωί (morning → ***morning*<sub>EMPH</sub>**)
- The **subjunctive** is marked in Greek with an **unstressed proclitic**
- Both patterns used by child studied by Dressler et al. 2020 to **mark the subjunctive**:
  - *káni* → ***ká:ni, ka+káni***

# Entering the Strong BAD: Greek Subjunctive

- Child learns that the **subjunctive must be marked** in Greek
  - **Payne (2022)**: Spanish subjunctive begins to emerge at just over 100 stems
  - Predictions for Greek subjunctive = **open question**
- **Fails to learn a productive process** to mark it
  - **TP**: **no sufficiently dominant process** over internal vocabulary
  - **NatMorph**: proclitics are challenging to take **from the input into the uptake**
- Uses iconic processes to **systematically differentiate the subjunctive**
  - **Vowel lengthening**: expresses **iconically marked categories** under NatMorph
  - **Reduplication**: **more iconic** than vowel lengthening under NatMorph

# Abandoning the Strong BAD: Greek Subjunctive

- Some process will eventually become **sufficiently dominant** over the child's vocabulary
  - If this process is the adult-like one, **the acquisition path is complete!**
  - If it is not, **a weak BAD is predicted** under our account
- What if **no process** becomes sufficiently dominant?
  - **Defectivity & memorization**
  - As the child's internal vocabulary grows, **greater ability to supply the necessary memorized forms**
  - **Lack of productivity** of the strong BAD likely causes child to give it up

# Does Everything Go?: The Greek Subjunctive

- **NatMorph preferences:** the child will use an **iconic pattern** to realize a category that **must be marked**
  - We don't expect e.g. **omission of marking** as a possible BAD for a category that is obligatorily marked
- Do we expect strong BADs involving a certain process in languages in which that **process is never attested?**
  - e.g. do we expect children to construct reduplication in a language with **no reduplication** in the input?
  - **NatMorph prediction:** yes, because innate preference will point to reduplication either way