An Adequate Theory of Morphological Blind Alley Developments

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

LSA 2025 | Philadelphia, PA

Background: **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)

Background: **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)

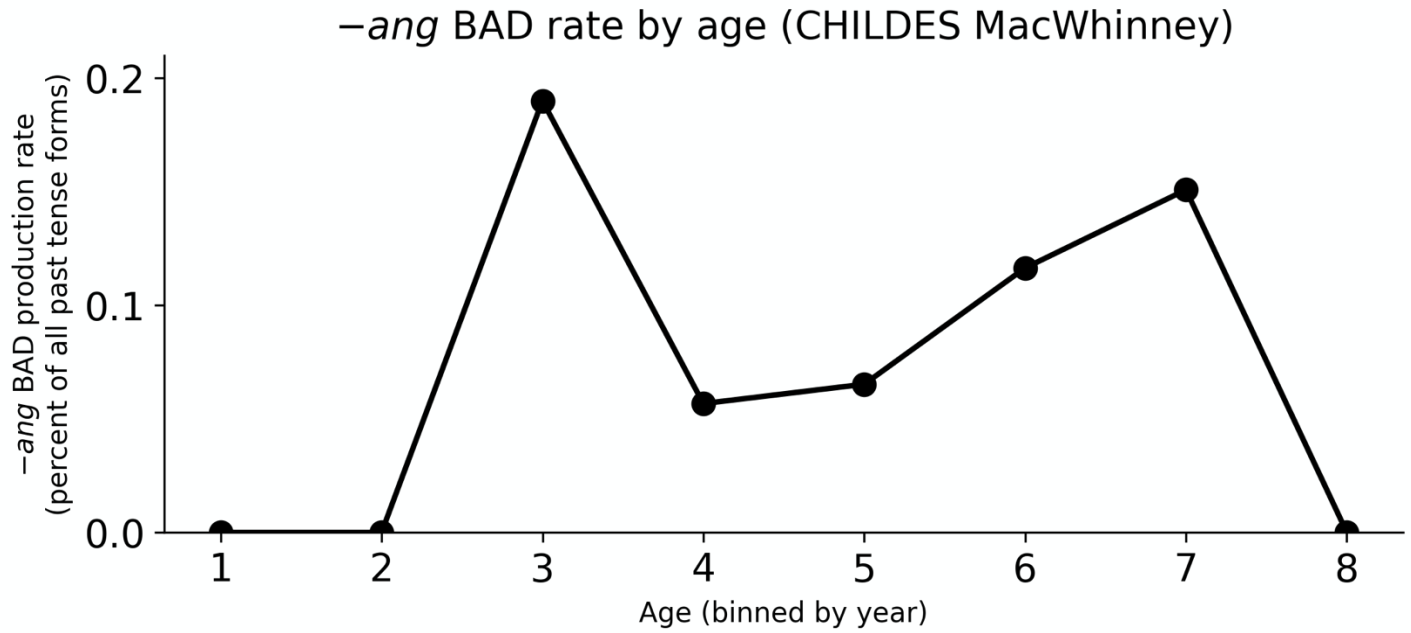
The Data: **English Past Tense ɪ → æ/__ŋ#**

- **ɪ → æ/__ŋ#** used for some past tense forms by adults
  - e.g., *sing-s**ang**, ring-r**ang***
- **Infrequent** and **unproductive**
  - New **–ɪŋ#** verbs take **productive –ed**
    - *Bing-Bing**ed**, bling-bling**ed***
- **Temporarily overapplied** by English-learning children
  - *bring-brought → bring-br**ang***
  - *fling-flung → fling-fl**ang***

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

The Data: **English Past Tense ɪ → æ/__ŋ#**

- **Temporarily overapplied** by English-learning children
  - *bring-brought → bring-br**ang***
  - *fling-flung → fling-fl**ang***



A Theory of BADs: **Desiderata**

- Explain their **ephemeral nature & timeline**
  - What causes children to **enter BADs**?
  - What causes them to **escape from them**?
- Explain the **patterns that children construct**
  - 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

Outline

- **Previous accounts** of Blind Alley Developments
  - Dressler et al. (2020) **Natural Morphology**
  - Payne & Yang (2023) **Tolerance Principle**
- **Proposal:** a mechanistic account of BADs
- **Applying the proposal: ɪ → æ/__ŋ#**
  - **Timeline**
  - **Content**
- A brief word on **strong BADs**
- **Conclusion & Open Questions**

Previous Work: **Natural Morphology**

- Morphological patterns that have no direct basis in the input must be explained by **cognitively-based universal preferences**
  - Saliency, frequency, etc.
  - **Natural morphology preferences:** iconicity, morphosyntactic transparency, morphosemantic transparency, bi-uniqueness, etc.
    - **"Tradeoffs" between these preferences**
- **NatMorph preferences** interact with **typology**
  - If patterns are **preferred based on universal principles**, they should occur in other languages
  - In the case of known BAD constructions, **this prediction is borne out!**
- Escape from BADs occurs because of **the opposing input**
  - **But the input is always opposing! Why then?**

Dressler et al. (2020)

Previous Work: **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)

Previous Work: **The Tolerance Principle**

- TP **evaluates** hypothesized process for productivity
  - Learner **enters** BAD when process **is productive over internal vocabulary**
  - Learner **exits** BAD when process **loses productivity over internal vocabulary**
- But how do we hypothesize the process 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**
  - **Natural Morphology preferences & trade-offs** may also play a role in **guiding children to some potential BAD constructions over others**

Payne & Yang (2023)

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	• 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**	• Accounts for children's escape in terms of **opposing input** • The **input is always opposing** – why does the child escape ***precisely when they do?*** • What is the **critical mass** of opposing input?

The Proposal

	The Tolerance Principle	Natural Morphology
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**	• Accounts for children's escape in terms of **opposing input** • The **input is always opposing** – why does the child escape ***precisely when they do?*** • What is the **critical mass** of opposing input?

The Proposal

	The Tolerance Principle	Natural Morphology
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**	• Accounts for children's escape in terms of **opposing input** • The **input is always opposing** – why does the child escape ***precisely when they do?*** • What is the **critical mass** of opposing input?
Patterns that Children Construct	• Does not predict which BADs are ***likely***, only which are ***theoretically possible*** given the child's **vocabulary**	

The Proposal

	The Tolerance Principle	Natural Morphology
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**	• Accounts for children's escape in terms of **opposing input** • The **input is always opposing** – why does the child escape ***precisely when they do?*** • What is the **critical mass** of opposing input?
Patterns that Children Construct	• Does not predict which BADs are ***likely***, only which are ***theoretically possible*** given the child's **vocabulary**	• Predicts the BADs that children are ***likely to construct*** in terms of **preferences & preference rankings**

The Proposal

	The Tolerance Principle	Natural Morphology
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**	• Accounts for children's escape in terms of **opposing input** • The **input is always opposing** – why does the child escape ***precisely when they do?*** • What is the **critical mass** of opposing input?
Patterns that Children Construct	• Does not predict which BADs are ***likely***, only which are ***theoretically possible*** given the child's **vocabulary**	• Predicts the BADs that children are ***likely to construct*** in terms of **preferences & preference rankings**

The Proposal

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**.

Weak BADs Under our Proposal

- TP sets **mechanistic bounds** on the **start and end** of the BAD
  - **Start:** the process is sufficiently dominant ***over the learner's vocabulary***
  - **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

Case Study: **Timeline of ɪ → æ / _ŋ# BAD**

- **Frequency & order of acquisition** correlated
  - **Model "typical" child** with strictly **frequency-based** order of acquisition
  - Frequencies calculated from **North American English CHILDES**

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

Case Study: **Timeline of ɪ → æ / _ŋ# BAD**

- When (frequency-based) vocab contains **200 verbs:**
  - **76 are irregular**, so **–ed** is not productive $(\boldsymbol{\theta_{200} = 37 < 76})$
  - **Subdivide, guided by NatMorph:**
    - **Bi-Uniqueness Preference: -ɪŋ#** verbs typically differentiate the simple past **(-æŋ#)** and past participle **(-ʌŋ#)**
      - Contrasts with the syncretism for other many verbs (e.g. ***thought***)
    - **3 -ɪŋ# verbs: *bring-brought, sing-sang, ring-rang***
    - **2/3 sufficient: ɪ → æ/_ŋ# productive** over internal vocabulary

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

Case Study: **Timeline of ɪ → æ / _ŋ# BAD**

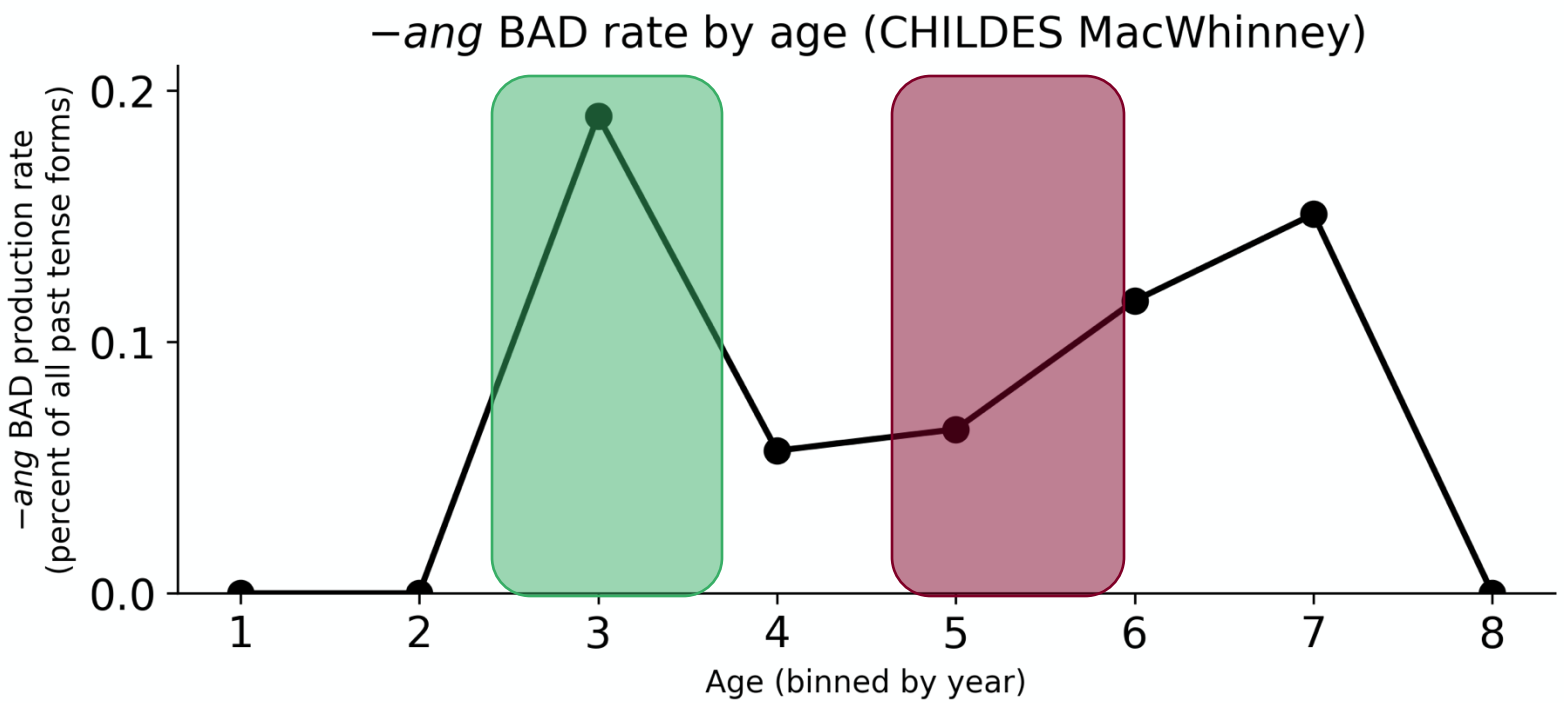
- When vocab contains **800 verbs:**
  - **8 -ɪŋ# verbs:** ***bring-brought, sing-sang, ring-rang, fling-flung, spring-sprang, sting-stung, swing-swung, wing-winged***
  - 3/8 not sufficient: **($\theta_8 = 3 < 5$)**
  - **ɪ → æ/_ŋ# cannot be supported anymore**

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

Case Study: **Timeline of ɪ → æ / _ŋ# BAD**

Verbs are about 25% of early vocabulary (Bornstein et al. 2004)

200 verbs ≈ 800 words **800 verbs ≈ 3200 words**



Fenson et al. (1994), Hart & Risley (1995), Biemiller (2005)

Case Study: **Content of ɪ → æ / _ŋ# BAD**

- **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#?

Case Study: **Content of ɪ → æ / _ŋ# BAD**

- **NatMorph: bi-uniqueness** favors –æŋ# BADs over -ɔt# BADs because of the syncretism between the past & past participle
- **Tolerance Principle:** even if an **-ɔt#** BAD were hypothesized, it will **never reach productivity**
  - Defining **the context** for **–æŋ# past tense forms:**
    - **Straightforward** (verbs ending in **-ɪŋ#** in the present)
    - **(Temporarily) supported by the input** (2/3 take **–æŋ#** in the past)

Case Study: **Content of ɪ → æ / _ŋ# BAD**

- Defining **the context** for **–ɔt# past tense forms**:
  - ***bring*** and ***think:*** **ɪŋ(k) → ɔt/ _ #**?
  - Verbs fitting **ɪ → æ/_ŋ#** also fit this rule description **(*sing, ring*)**
  - Verbs fitting **ɪ → æ/_ŋk#** also fit this rule description **(*drink, sink*)**
- When the vocab contains **300 verbs:**
  - **6 -ɪŋ(k)# :** ***think-thought, bring-brought, drink-drank, ring-rang, sing-sang, sting-stung***
  - 2/6 not sufficient $(\boldsymbol{\theta_6 = 3 < 4})$
  - Increasing vocabulary doesn't help **(*catch, teach*)**

Weak BADs: **Summary**

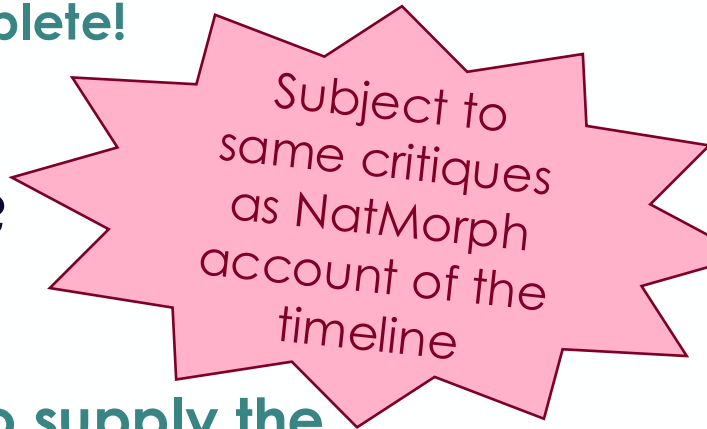
- Weak BADs persist only as long as they are productive ***over the learner's internal vocabulary,*** as measured by the TP
- **TP** delineates which possible BADs **may reach temporary productivity** and which may not
- **NatMorph** predicts some weak BADs to be more likely
  - Child is **unlikely to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

Strong BADs: **Overview**

- **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 bounds on the **timeline** of the strong BAD
- **NatMorph preferences** predict the **content** of the strong BAD
  - Since **the BAD process is not productive over the learner's internal vocabulary**, this is out of scope for the TP

Strong BADs: **Timeline**

- TP sets **mechanistic bounds** on **timeline** of the strong 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
    - If this process is the adult-like one, **the acquisition path is complete!**
    - If it is not, **a weak BAD is predicted**
- 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



Strong BADs: **Content**

- NatMorph preferences predict **what types of strong BAD constructions** children will build
  - Child knows **the category is marked**, but not how to mark it
  - e.g. Russian: reduplication used by children to indicate **iterativity**, but by adults for **intensification**
- Do we expect strong BADs involving a certain process in languages in which that **process is truly *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

Conclusions & Open Questions

- Marriage of **learning-theoretic TP** with **NatMorph preferences** gives a theory of acquisition with both:
  - A **formalization of the timeline** of BADs that maps well onto acquisition findings
  - Clear predictions about **which types of BADs** are **theoretically possible** & which are **likely or expected**
- **Open questions:**
  - How do children escape strong BADs when **no process ever reaches productivity**?
  - Are strong BADs attested that involve **patterns that *never* appear in the language?**

Thank you!!!

I am grateful to Wolfgang Dressler, 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*** ⇒ **± 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**)
  - tɕʉtʲ → tɕʉtʲ tɕʉtʲ (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**)
  - tɕʉtʲ → tɕʉtʲ tɕʉtʲ (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***$_{\text{EMPH}}$)
- 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