Syntactic reanalysis as internal change: exploring the dangers of intake

Brian Gravely— University of Arizona
Diachronic Generative Syntax 23 (DiGS23)
Workshop: *Child learners in syntactic change: theory and models*
8 June, 2022— NYU

1. INTRODUCTION

If information that was present in the signal fails to make it into the learner's perceptual intake, then that information cannot impact their interpretation of the input, and as a consequence children may make inferences about the lexicon or grammar that are appropriate for that perceptual intake but not for the actual input.
(Omaki & Lidz 2015:180)

- There have historically been two primary points of view with respect to *how* change occurs in grammar:

  i. CHILD INNOVATOR APPROACH: Children are the movers and shakers of change; more specifically, it is their experience with the input presented to them that serves as the catalyst for change (Cournane 2014, 2015, 2017; Hale 1998, 2007; Lightfoot 1979, 1999, 2020; Pearl 2007; Tagliamonte & D'Arcy 2007; *i.a.*)
  ii. 'INVISIBLE HAND OF CHANGE': Grammar change is a natural process of language that happens as a result of a domain-general, pre-linguistic phenomenon (Bybee 2003a,b; Heine & Kuteva 2002, 2012; Hoefler & Smith 2009; Traugott & Dasher 2001; *i.a.*)

- A different question related to the CIA is *what* causes change:

  i. EXOGENOUS CHANGE: Change that is provoked by external forces (e.g., a certain pattern in the input)
  ii. ENDOGENOUS CHANGE: Change that results as a novel creation with no plausible source of said change from the input the child receives

⇒ I contend that all syntactic change is, in fact, endogenous and internal to the language acquisition device (LAD)
  - This desideratum should hold regardless of the *type* of syntactic change

- Many of the theories behind diachronic change are mere *descriptions* of how change might have occurred, but there have been little to no advancements as to what *explanations* we may give the recorded changes
  - That is, what about the language acquisition (LA) process might have caused the changes we know to have happened?

⇒ I claim that we must look to recent developments and discoveries in the LA literature regarding what we know about language acquisition
  - The pinpointing of the place in which change must takes place in the language acquisition device (LAD) allows us to identify more accurately the influence the LA process has on language and potential change more generally

⇒ I present diachronic data from Galician and discuss the ways in which different syntactic changes may be re-created using LA models (and, in turn, different types of change and how what we know about language acquisition may provide insight into explaining change)

⇒ I assert that approaching the explanation of syntactic change from the purview of LA immediately allows us
  - to dispense with the ideas that Universal Grammar should ever be accredited with/blamed for a given change (*pace* van Gelderen 2004, Roberts 2007)
  - to view the simple attribution of any type of diachronic reanalysis to "weak evidence in the PLD" as untenable (*pace* Lowell Sluckin 2016)

2. THE INNERWORKINGS OF THE LAD

Do we truly know what the language acquisition device consists of?

There is good reason to believe we do!

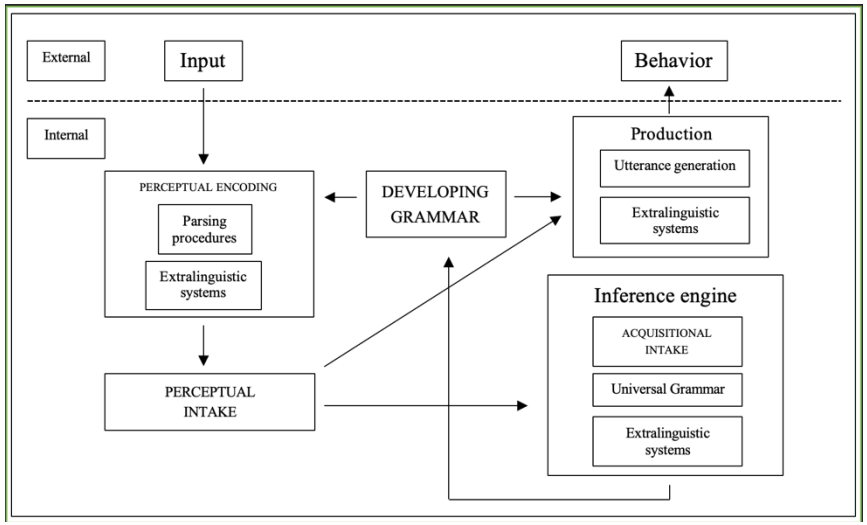


Figure 1. The language acquisition device (modified from Pearl 2021:4)[1]

Key components with respect to potential dangers for change:

PERCEPTUAL ENCODING

- The input signal perceived at any given stage of development
    - i.e., dependent on linguistic and extralinguistic modalities alike (e.g., developing grammar vs. constraints on memory)
- Presumably where structural syntactic modeling takes place (i.e., tree building)

PERCEPTUAL INTAKE

- The result of what has been perceived in the PERCEPTUAL ENCODING mechanism

ACQUISITIONAL INTAKE

- Most likely only a portion of the PERCEPTUAL INTAKE
    - What is the child currently trying to learn? What is relevant to the child's growing linguistic experience?
- The assessment made in the PERCEPTUAL INTAKE either positively or negatively affects the inferences and statistics taken on any given phenomenon
    - It is here that the idea of 'multiple grammars' would presumably be entertained

DEVELOPING GRAMMAR

- Consistently feeds multiple components related to perception, learning, and production
- All updates may immediately affect both perception and production

A few notes on this model:

- UG plays a minimal role
    - "… it can filter the perceptual intake down to the relevant pieces by providing both constraints on possible hypotheses and attentional filters." (Pearl 2021:5)
    - "… we show that UG does not by itself define a learning mechanism, but rather must be paired with an inference that links the representations provided by UG with the data of the experience." (Lidz & Gagliardi 2015:337)
- The importance of perceptual encoding is crucial
    - Based on where the child is developmentally and how advanced her developing grammar is, her ability to encode incoming strings (familiar or not) will undoubtedly dictate what is captured in the perceptual intake and pushed through to the acquisitional intake

[1] This model was originally formulated by Omaki & Lidz (2015) and further elaborated on by Lidz & Gagliardi (2015).

⇒ Even though it is impossible to track the specific processing issues of a phenomenon that developed 500 years ago, we can (and should!) look to what we now know about the individual components found in the LAD in order to explain certain changes (e.g., multi-word reanalysis)
  - Not every type of syntactic change is as dependent on these processes (e.g., loss of V2 in English; Biberauer & Roberts 2016, 2017)

If the child is fully reliant on parsing and aspects related to perceptual encoding, what are the difficulties she may face along the way?

3. PARSING, CATEGORIZATION, AND WHEN THINGS GO WRONG

Let's take a look at what we know about parsing…

Three parsing outcomes (Phillips & Ehrenhofer 2015:416-417)

i) A learner may fail to assign a parse to an arriving string, either because it is too complex or it arrives too quickly;

ii) A learner may systematically assign an incorrect parse (per the established adult grammar) to a particular sentence type due to parsing biases or reanalysis failure;

iii) A learner that parses an input string successfully so that it may reach the perceptual intake mechanism of the LAD may extract more or less information from said input string depending on her ability to predict the contents of the sentence as it unfolds.

- It is important to remember that scenario (ii) is a real potential outcome even when the input is optimal from a communication standpoint (Omaki 2010; Omaki & Lidz 2015)
  - Thus, we should not claim that the "input is the problem"

In Gravely (2021), I claimed that it is plausible that some hypotheses posited by the child are simply not present based on the evidence in her input
  - This may have a considerable amount to do with the expectations that children have regarding how their input *should* be structured (Lidz, Gleitman & Gleitman 2003; Lidz & Gagliardi 2015)
  - It may also be related to internal computational processes based on the biological development of the child (Belletti 2017)
  - This does not entail probabilistic learning (*pace* Ferreira & Patson 2007)

Where does categorization fall in all of this?

- Work by Cassani (2019) and Cassani et al. (2018, 2019) highlights the three main factors that play a part in the child's determining to what category a particular lexical item belongs

Three factors of category learning in order of importance

i) Contextual diversity

ii) Frequency

iii) Predictability

- Contextual diversity has been shown to be crucial for LA models that rely on framing (Chemla et al. 2009; Mintz 2003, 2006; Naigles 1996; St. Claire et al. 2014; Weisleder & Waxman 2010; *i.a.*)
  - Category acquisition via framing has been shown to aid in early language processing on multiple fronts (Bar-Sever & Pearl 2016)
- From the perspective of categorization (and learning more generally), a word that appears in fewer syntactically diverse contexts should be harder to learn

⇒ In Gravely (2021), I claimed that this is a plausible source for reanalysis
  - Imagining the child has a syntactically limited exposure to a particular string with minimal variation, acquisitional deviance may ensue

⇒ I claim that multi-word reanalysis may be the result of the inability to correctly categorize (i.e., encode) a particular string
  - Moreover, subtle morphosyntactic variation of a given structure may aid in reanalysis (Hudson Kam & Newport 2005)

- In Gravely (2021), I showed that this is precisely what we find in the engenderment of the root-clause presentative complementizers *velaí* and *velaquí* in Galician
  - From the declarative *ve-l(o) aí*/*aquí* 'You see it there/here'

<u>Stage 1: Biclausal structure</u>

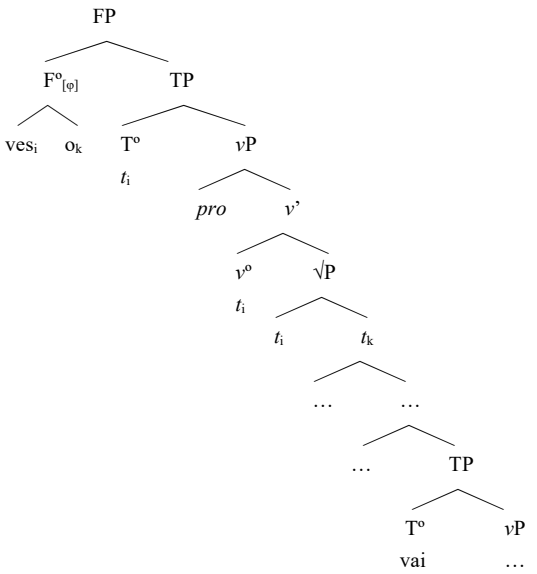
(1) Ve-lo aí vai Xan
see.PRS.2SG-CL there go.PRS.3SG Xan
'You see him, there goes Xan.'

<u>Stage 2: Monoclausal structure; reanalyzed verbal structure</u>

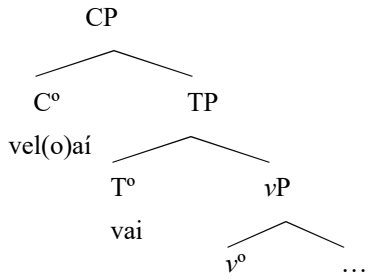
(2) Vel(o)aí vai Xan
behold go.PRS.3SG Xan
'There goes Xan.'

- There was considerable variation between the GENDER and NUMBER of the clitic pronoun
- There were only a handful of recorded situations in which this phrase was embedded, plausibly due to the illocutionary force and pragmatic use associated with it

(3) *Stage 1 structure*



(4) *Stage 2 structure*



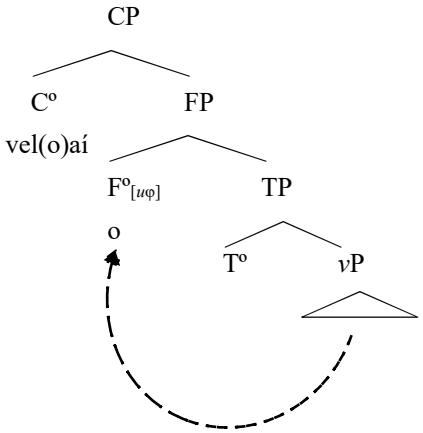
- Considering the overwhelming sentence-initial frequency of this string, we may imagine that the syntactic diversity was relatively scarce
  ⇒ Prime case for multi- to single-word reanalysis

But we should be careful about assuming the clitic was simply lost in the mix…

Stage 3: Monoclausal, root-clause complementizer agreement structure

(5) Vel(o)aí o vai Xan
behold CL go.PRS.3SG Xan
'There goes Xan.'

(6) *Stage 3 structure*



- Early on after the reanalysis of *velaí/velaquí*, we see the accusative clitic behaving as an agreement marker
  - The dependencies of this construction are unexpected due to the lack of C-AGR paradigms in Romance
  - However, there was always a pragmatic connection in Stage 1, as the clitic always agreed with the subject of the subordinating clause in GENDER and NUMBER (occasionally in a topicalized position; cf. 7)
  - There are recent investigations tying in pragmatics and syntax as two modules of grammar that may keep one another in check during the acquisition process (e.g., Hacquard & Lidz 2019)

(7) [As mociñas de Boado]$_i$ ve-las$_i$ acolá che van
the girls of Boado see.PRS.2SG-CL$_{F.PL}$ there CL$_{DAT}$ go.PRS.3PL
'The girls from Boado, you see them, there they go.'

- This engenderment of a subject agreement marker is akin to the type of endogenous structure mentioned in Belletti (2017)

(8) *Deviant structures in Italian children* (Belletti 2017:2; from Adele age 4,9)
a. Che cosa succede ai miei amici,
what thing happen.PRS.3SG to-the my friends
il pinguino e la mucca?
the penguin and the cow
'What is happening to my friends, the penguin and the cow?'
b. Il coniglio a i'pinguino lo tocca
the rabbit DOM the-penguin CL touch.PRS.3SG
'The rabbit touches the penguin' (lit. 'The rabbit to the penguin (he) touches him')

- Neither differential marking nor pre-verbal object fronting in CLLD contexts are properties of the Tuscan variety these children are exposed to
  - However, these are precisely the situations that lead to change, particularly when such a high percentage of children converge on a divergent structure

Can we truly argue that this type of endogenous change is the same as the modified structures compared to the data in the PLD?

⇒ If we consider…
  - the conjunction of aspects (linguistic and non-linguistic alike) present in the processing procedures and encoding mechanisms of the LAD
  - the fact that not only learning but accurate parsing and computation are reliant on biological development (cf. Kidd et al. 2012, 2014; Perfors et al. 2011), making internal reanalysis a real possibility as the child develops cognitively
  - variation and deviance can arise when the PLD is optimal

⇒ … then the answer is yes!

4. Conclusive observations

What about each step in the learning process is so important and how does it piece together with the other components?

Prerequisites for accurate transmission (Phillips 2012:282)

i) The learner must be exposed to relevant linguistic input (Input I)

ii) The learner must experience that input in suitably informative contexts (Input II)

iii) The learner must attend to these informative situations (Uptake I)

iv) The learner must appropriately encode these informative situations (Uptake II)

v) The learner must store relevant examples and perform appropriate computations over them in order to form grammatical hypotheses (Computation I)

vi) The learner's computation may be guided by a constrained hypothesis space (Computation II)

- It is clear that (i)-(iii) have been at the heart of work related to variation and change
- However, I suggest that we must look to (iv) and (v) as sources for change
- This may only be done when we pair findings and novel advancements in LA with the diachronic records we have

Acknowledgements

First and foremost, I would like to thank Lisa Pearl for the lengthy discussions we have had on this topic and her unending support in guiding me through many of the complexities presented here. I would also like to thank Giovanni Cassani for his helpful comments related to the link between his work and my data cited here, as well as Jeff Lidz for insightful commentary on various aspects of this work. Finally, many thanks to Tim Gupton for his innumerable comments and criticisms on these proposals.

Feel free to reach out to me!

💻 www.brian-gravely.com
✉ brian.gravely@icloud.com

References

Bar-Sever, Galia and Lisa Pearl. 2016. Syntactic categories from frequent frames benefit early language acquisition processing in English and ASL. *Proceedings of the 40th Annual Boston University Conference on Child Language Development*, 32-46.

Belletti, Adriana. 2017. Internal grammar and children's grammatical creativity against poor inputs. *Frontiers in Psychology* 8:1-7.

Biberauer, Theresa & Ian Roberts. 2016. Parameter typology from a diachronic perspective: the case of Conditional Inversion. In Ermenegildo Bidese, Federica Cognola & Manuela Caterina Moroni (eds.) *Theoretical Approaches to Linguistic Variation*, 259-291. Amsterdam: John Benjamins.

Biberauer, Theresa & Ian Roberts. 2017. Conditional inversion and types of parametric change. In Bettelou Los & Pieter de Haan (eds.) *Word Order Change in Acquisition and Language Contact: Essays in honour of Ans van Kemenade*, 57-77. Amsterdam: John Benjamins.

Bybee, Joan. 2003a. *Phonology and language use.* Cambridge: Cambridge University Press.

Bybee, Joan. 2003b. 19 Mechanisms of Change in Grammaticalization: The Role of Frequency. In Brian Joseph & Richard Janda (eds.) *The Handbook of Historical Linguistics*, 602-623. New York: John Wiley & Sons.

Cassani, Giovanni. 2019. *The Bootstrapping Toolbox: Which Cues are More Useful to Learn Lexical Categories and Why*. Doctoral dissertation, University of Antwerp. Antwerp, Belgium.

Cassani, Giovanni, Robert Grimm, Walter Daelemans and Steven Gillis. 2018. Lexical category acquisition is facilitated by uncertainty in distributional co-occurrences. *PLoS ONE* 13(12):1-36.

Cassani, Giovanni, Robert Grimm, Steven Gillis and Walter Daelemans. 2019. What makes a distributional context useful? Lexical diversity is more important than frequency. Ms.

Chemla, Emmanuel, Toben Mintz, Savita Bernal and Anne Christophe. 2009. Computational principles of language acquisition: Categorizing words using 'frequent frames": what cross-linguistic analyses reveal about distributional acquisition strategies. *Developmental Science* 12(3):396-406.

Cournane, Ailís. 2014. In Search of L1 Evidence for Diachronic Reanalysis: Mapping Modal Verbs. *Language Acquisition* 21(1):103-117.

Cournane, Ailís. 2015. *Modal Development: Input-Divergent L1 Acquisition in the Direction of Diachronic Reanalysis*. Doctoral dissertation, University of Toronto. Toronto, Canada.

Cournane, Ailís. 2017. In defence of the child innovator. In Éric Mathieu & Robert Truswell (eds.) *Micro-change and Macro-change in Diachronic Syntax*, 10-24. Oxford: Oxford University Press.

Ferreira, Fernanda and Nikole Patson. 2007. The 'good enough' approach to language comprehension. *Language and Linguistics Compass* 1(1-2):71-83.

Gagliardi, Annie and Jeffrey Lidz. 2014. Statistical insensitivity and the acquisition of Tsez noun classes. *Language* 90(1):58-89.

van Gelderen, Elly. 2004. *Grammaticalization as economy*. Amsterdam: John Benjamins.

Gravely, Brian. 2021. *Language acquisition and endogenous grammar change: the rise of Galician complementizer agreement*. Doctoral dissertation, University of Georgia. Athens, GA, USA.

Hacquard, Valentina and Jeffrey Lidz. 2019. Children's attitude problems: Bootstrapping verb meaning from syntax and pragmatics. *Mind & Language* 34(1):73-96.

Hale, Mark. 1998. Diachronic syntax. *Syntax* 1(1):1-18.

Hale, Mark. 2007. *Historical Linguistics: Theory and methos*. Oxford: Wiley-Blackwell.

Heine, Bernd and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambrudge University Press.

Heine, Bernd and Tania Kuteva. 2012. Grammaticalization theory as a tool for reconstructing language evolution. In Maggie Tallerman & Kathleen Gibson (eds.) *The Oxford Handbook of Language Evolution*, 511-527. Oxford: Oxford University Press.

Hoefler, Stefan H. & Andrew D. M. Smith 2009. The pre-linguistic basis of grammaticalization: A unified approach to metaphor and reanalysis. *Studies in Language* 33(4):886-909.

Hudson Kam, Carla and Elissa Newport. 2005. Regularizing unpredictable variation: The roles of adult and children learners in language formation and change. *Language learning and development* 1(2): 151-195.

Kidd, Celeste, Steven Piantadosi and Richard Aslin. 2012. The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE* 7(5):e36399.

Kidd, Celeste, Steven Piantadosi and Richard Aslin. 2014. The Goldilocks Effect in Infant Auditory Attention. *Child Development* 85(5):1795-1804.

Lidz, Jeffrey and Annie Gagliardi. 2015. How nature meets nuture: Universal Grammar and statistical learning. *Annual Review of Linguistics* 1, 333-353.

Lidz, Jeffrey, Henry Gleitman and Lila Gleitman. 2003. Understanding how input matters: verb learning and the footprint of universal grammar. *Cognition* 87(3):151-178.

Lightfoot, David. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.

Lightfoot, David. 1999. *The development of language: Acquisition, change, and evolution*. Oxford: Wiley-Blackwell.

Lightfoot, David. 2020. *Born to parse: How children select their languages*. Cambridge, M.A.: MIT Press.

Lowell Sluckin, Benjamin. 2016. *Exaptation across grammatical domains: a generative account of source contexts, mechanisms and outcomes*. Master's thesis, University of Cambridge. Cambridge, England.

Mintz, Toben. 2003. Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* 90(1):91-117.

Mintz, Toben 2006. Finding the verbs: Distributional cues to categories available to young learners. In Kathryn Hirsh-Pasek & Roberta Golinkoff (eds.) *Action meets word: How children learn verbs*, 31-63.

Naigles, Letitia. 1996. The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition* 58(2):221-251.

Omaki, Akira. 2010. *Commitment and flexibility in the developing parser*. Doctoral dissertation, University of Maryland. College Park, MD, USA.

Omaki, Akira and Jeffrey Lidz. 2015. Linking Parser Development to Acquisition Syntactic Knowledge. *Language Acquisition* 22:158-192.

Pearl, Lisa. 2021. Modeling syntactic acquisition. *The Oxford Handbook of Experimental Syntax*. To appear. lingbuzz/003416.

Perfors, Amy, Joshua B. Tenenbaum and Terry Regier. 2011. The learnability of abstract syntactic principles. *Cognition* 188:306-338.

Phillips, Collin. 2012. Individual variation and constraints on language learning. *Linguistic Approaches to Bilingualism* 2(3):281-286.

Phillips, Collin and Lara Ehrenhofer. 2015. The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism* 5(4):409-453.

Roberts, Ian. 2007. *Diachronic Syntax*. Oxford: Oxford University Press.

St. Clair, Michelle, Padraic Monaghan and Morten Christiansen. 2010. Learning grammatical categories from distributional cues: Flexible frames for language acquisition. *Cognition* 116(3):341-360.

Tagliamonte, Sali A. and Alexandra D'Arcy. 2007. Frequency and variation in the community grammar: tracking a new change through the generations. *Language Variation and Change* 19:199-217.

Traugott, Elizabeth Closs and Richard B. Dasher. 2001. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Walkden, George. 2011. Abduction or Inertia? The logic of syntactic change. *Proceedings of the Sixth Cambridge Postgraduate Conference in Langauge Research*, 230-239.

Walkden, George. 2012. Against inertia. *Lingua* 122:891-901.

Weisleder, Adriana and Sandra Waxman. 2010. What's in the input? Frequent frames in child-directed speech offer distributional cues to grammatical categories in Spanish and English. *Journal of child language* 37(5):1089-1108.