"Generative Perspectives of Language Change and Acquisition" *in* Diachrony and Language Evolution

Jordan Kodner

Department of Linguistics
Institute for Advanced Computational Science
Stony Brook University
jordan.kodner@stonybrook.edu
+1 (262) 573-2812

Key Points

- Generative approaches to language change primarily focus on changes to the grammar. Theory and diachrony are seen as mutually informative.
- Child language acquisition is seen as the locus of grammar change.
- Corpus, computational, and experimental approaches play core roles in the field's understanding of grammar change and the relationship between grammar change and child language acquisition.

This chapter looks at the relationship between language change and child language acquisition from a broadly generative perspective, with a focus on syntactic change. Generative approaches to language change focus primarily on changes to the grammar and have typically seen language acquisition as the crucial locus for change. They address the problem both from theoretical perspectives, drawing insight from generative syntax, formal language theory, and learning theory on one hand, and from empirical approaches, including corpus linguistics and experiments with children, on the other hand. Taken together, theoretical linguistics, historical linguistics, and research in child language acquisition prove to be mutually informative in developing our understanding of language change.

1 Language Acquisition and Language Change

The idea that child language acquisition plays a crucial role in language change is one of the oldest ideas in linguistics. Hermann Paul famously championed this idea in his *Prinzipien* (Paul, 1880). He also argued that language should properly be thought of as something seated in the mind of each individual rather than just something that exists nebulously out in the community. The language of each individual changes the most during childhood as children are learning it, and the language of the community is made up of the language of individuals. So, as the reasoning went, acquisition is the point at which most language change (at least of the type that Paul was interested in) must occur.

Perhaps the clearest visualization of the idea that language changes primarily through child language acquisition, is the Z-model of Andersen (1973). Language change was conceived of as an abductive process of learning and mis-learning as individuals sometimes unfaithfully reconstruct the grammars of their elders from sample outputs from those grammars. This is visualized in Figure 1, with the diagonal line labeled "acquisition" responsible for change.

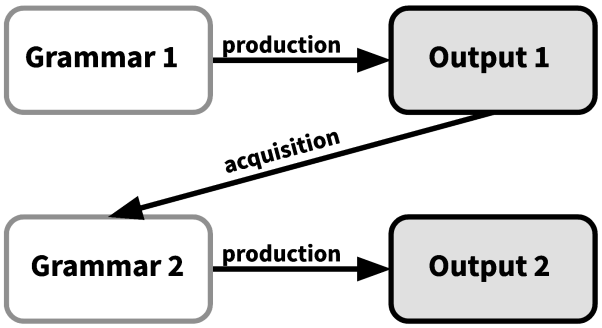

Figure 1: Visualizing the Z-Model of language acquisition and change

Paul's late 19th century perspective on the language of the individual presaged the cognitive revolution in the mid-20th century. This revolution gave birth to generative linguistics, which is largely concerned with the *grammars*, or the formal aspects of the mental underpinnings of language, within individuals. Since generative linguistics is primarily concerned with understanding the grammar, generative approaches to language change and language acquisition are primarily concerned with grammar change and grammar learning respectively. Paul's approach also presaged the development of variationist sociolinguistics (see the discussion in Weinreich et al. 1968), which views language variation and change as phenomena emerging from the interaction of individuals within speech communities. While child language acquisition is not a primary focus of variationist sociolinguistics, it is nevertheless well-understood that language acquisition is the primary way that language is transmitted across generations and thus plays a unique role in language change (see the discussion of language *transmission* and *diffusion* in Labov 2007).

Extensive research in child language acquisition has bolstered a relationship between grammar learning and grammar change. A critical period for language development (Lenneberg, 1967) limits much (but not all) of individual change to childhood. And documented child learner innovations and frequently parallel changes to the grammar suggest a causal relationship in some cases. The most famous example of this type is probably analogical leveling in morphology, which closely parallels over-regularization, which becomes the most common error (or innovation) which children produce during language development.

Beyond analogical change, child language acquisition is understood to play a crucial role in changes to all aspects of the grammar from phonetics and phonology (e.g., Kiparsky, 1965; Hyman, 2008; Richter, 2021) to morphology (e.g., Anderson, 2015; Björnsdóttir et al., 2020; Kodner, 2023), syntax (e.g., Lightfoot, 1979; Roberts, 2003; Biberauer and Walkden, 2015), and semantics (e.g., Lightfoot, 1979; Cournane, 2014). This chapter looks at the relationship between language change and child language acquisition from a broadly generative perspective, with a focus on syntactic change. First, Section 2 discusses the generative interest in grammar change. Second, Section 3 discusses the role that child language acquisition plays in generative theories of grammar change. Third, Section 4 discusses empirical methods that relate language acquisition to language change. While this division between 'theoretical' and 'empirical' approaches is useful for exposition, it is largely spurious, as all approaches in this chapter combine theory with empirical work to a substantial degree.

2 A Focus on Grammar Change

Generative approaches to theoretical linguistics are primarily interested in characterizing the grammars of individual speakers (Chomsky, 1957, 1965). The *grammar* in this sense is a mental construct that *generates* (hence *Generative* linguistics) all and only the utterances that belong to the speaker's language. This puts

generative linguistics squarely within cognitive science, which is interested in understanding the processes and representations that underlie human cognition. Formal language theory, a branch of computational theory within computer science and mathematics, is closely related to generative linguistics and often intersects with it – Chomsky himself contributed to early work in formal language theory (e.g., Chomsky, 1959; Chomsky and Schützenberger, 1963). It is also worth noting that any approach to linguistics which aims to understand the cognitive representations and processes which generate human language are technically "generative," though by convention, the term is generally reserved for approaches that remain explicitly aligned to some degree with Chomsky's classic work. That is the narrow definition that I adhere to in this chapter.

Theoretical work in generative linguistics targets grammatical competence and largely abstracts away from factors of variation, processing, and acquisition. This methodological idealization is sometimes characterized as the *ideal speaker-listener* (Chomsky, 1965). However, generative work beyond theory relaxes this idealization as appropriate. Language processing is focused on the interplay between competence and performance, so it cannot abstract away from performance factors. Language acquisition is studies how learners become full-fledged competent speakers of their native languages, so it cannot be assumed that all speakers know their language perfectly. Finally, language variation and change are concerned with heterogeneous populations of speakers whose grammars are different in some ways.

In sum, given that generative linguistics is primarily concerned with the grammar, generative approaches to language change are concerned with how the grammar changes over time and how the grammar underpins observable changes in language use. Care is taken to consider both children who have not yet achieved adult-like competence in their native languages as well as language communities which contain heterogeneous expressions of language. The next two sections discuss influential theoretical and empirical approaches to the study of grammar change which invoke child language acquisition.

3 Theoretical Approaches

This section discusses three influential theoretical approaches to syntactic change which invoke child language acquisition. Most work of this type is focused on the innovation of grammar changes, though there is also research on incrementation and population-level modeling. I begin with two classic perspectives on grammar change within theoretical syntax before discussing a closely related perspective influenced by computational learning theory.

In his work on syntactic change, Lightfoot discusses how small perturbations in the input to children could *trigger* more substantial changes to the grammars that they eventually acquire, which is tantamount to the innovation of syntactic changes (Lightfoot, 1979, 1999). In these cases, an understanding of the grammar reveals how surface-disparate changes may actually be causally linked – something that would not be possible with only a surface perspective on language use. The loss of V-to-I movement in English, which manifests on the surface as several distinct changes to word order (e.g., **Alice paints always pictures.* → *Alice always paints pictures.*), can be accounted for in this way. Early earlier changes in the history of English: the advent of modals as a new inflectional category, the loss of V2, the loss of rich subject-verb agreement, and the ongoing emergence of *do*-support (e.g., **Alice painted not the pictures.* → *Alice did not paint the pictures.*), obscured evidence to learners that the verb should raise in the syntactic structure. A learner who failed to acquire a grammar with this movement has, from a diachronic perspective, created a syntactic innovation which yields all the observed changes in language use. This then had the opportunity to spread through the population.

Van Gelderen (2011) invokes Economy Principles employed by children as they build their language-specific grammars as a way of operationalizing observed cycles of language change. Consider the Subject Cycle as a simple example: pronouns often shift diachronically from emphatics to full pronouns to clitics to agreement morphemes. Two Economy Principles together can account for this. The Head Preference

Principle states that words should be evaluated as heads when possible rather than phrases, and the Late Merge Principle prefers late insertion to early insertion and movement in the syntactic structure. Taken together then, a learner should prefer to treat pronouns as heads rather than specifiers (something else is the head of the phrase) when possible. Furthermore, they should be base generated in that position rather than moved there after being merged lower in the tree. Full pronouns need to appear both low and high in the tree during a derivation, hence early merge and movement, but later steps in the subject cycle do not. These principles can further be implemented within the mechanics of the syntax in terms of feature change and loss, thus providing an operationalization of the descriptively observed subject cycle in the grammar.

Formal language theory and computational learning theory are branches of computer science which share a notion of formal grammar with generative theoretical linguistics. The highly mathematical nature of this perspective allows practitioners to write formal proofs, an opportunity that is not available in other theoretical approaches to language. Work in this area also supports a causal role for learning in language change. One example of this approach is Niyogi and Berwick (1997), which employed a Trigger Learning Algorithm (TLA) and a parametric view of the grammar to understand change. Since the input is finite, and much of it may be ambiguous between two given alternative grammars, learners may misconverge with some probability which can be directly calculated from the formalism. This misconvergence, from the perspective of language change, is innovation. While conceptually similar to Lightfoot's work, this approach is much more abstract and mathematical, and it is embedded in a population-level model: language change is an observation about distribution of language(s) in populations or speech communities of individuals. Niyogi and Berwick's approach allows them to work out the population-level dynamics induced by different parametrization of the grammar including, the emergence of S-curves. The work also draws strong parallels to formal models of population genetics (see the discussion in Niyogi 2006).

To summarize, the important takeaways from theoretical approaches to grammar change extend beyond the specific theories that they are implemented under. Child language acquisition is the locus of innovation for grammar changes, it relates disparate changes within a language through their impact on the grammar, and it connects the language of individuals to the population as a whole. Combined with empirical methods such as those discussed in the next section, they contribute to full-fledged explanations for grammar change.

4 Empirical approaches

While most empirical work in historical linguistics is corpus-based, some recent work instead adopt experimental methods. Cournane (2014, 2017) presents such an experimental approach for investigating the relationship between language acquisition and change in the semantics of modal verbs. Children are known to persist in non-adult-like interpretations of model verbs well past school age (Papafragou, 1998), offering ample time to influence their peers. Cournane shows that children's developmental trajectories across target languages closely parallel the diachronic modal cycle, where lexical modals become functional modals, for which root meanings (e.g., deontic *must* "required to") initially dominate, while epistemic meanings (e.g., *must* "known to be true") gain ground over time. The meanings that are acquired earlier in development occur earlier in the diachronic pathway. Furthermore, children may over-generalize root modals to epistemic meanings, providing a means for a continuation of the cycle. As part of an overall defense of a causal role for child language acquisition in language change, she invokes peer-to-peer interaction as an important link between individual innovation and community-level change, reflecting a convergent perspective with Niyogi and Berwick's mathematical work.

Corpus linguistics plays a central role in the study of grammar change. This has been facilitated by a substantial effort to develop syntactically annotated corpora of historical texts (e.g., Kroch and Taylor, 2000; Taylor, 2007; Wallenberg et al., 2011). These corpora, which provide phrase structure parses, part-of-speech tagging, and sometimes morphological annotation and orthographic normalization, share a common origin

with the phrase-structure-parsed syntactic corpora that enabled early statistical natural language processing, with overlapping annotation standards and even shared personnel (Marcus et al., 1993). Historical annotated corpora provide direct information about language use over the centuries, from which we can infer the timeline for the innovation of grammar changes as well as their incrementation.

While some changes unfold over centuries, others spread in the course of a single generation. Individual writers are frequently observed to use both conservative and innovative grammatical variants, which has been conceptualized in terms of *competing grammars* within individual speakers (Kroch, 1989). Thus, discrete changes to the grammar manifest gradiently in both individual language use and community uptake, providing time for individual innovations to snowball into population-level change. These findings are consistent with basic principles of variationist sociolinguistics, which has been described as the investigation of gradient distributions of discrete phenomena (Sankoff, 1988). While triggering accounts may provide an explanation for the innovation of *do*-support, corpus work such as Kroch (1989) provides accounts for its incrementation. Furthermore, the observable rise of *do*-support provides evidence for the structure of the grammar, through the Constant Rate Hypothesis (Kroch, 1989). *Do*-support rose in use in different sentential contexts following S-curves at different times, and Kroch shows that these S-curves share the same slope, arguing that all instances of *do*-support derive from a single change to the grammar, with their rates of use further mediated by grammar-external factors. Thus theories of grammar and theories of change can mutually inform one another.

Since historical corpora are similar to child directed speech in their distribution of many linguistic units (Kodner, 2019), they can often be substituted for child directed speech as the input to model acquisition on languages of the past. Some approaches take advantage of this insight to synthesize the empirical corpus-based methodology with computational models of acquisition to evaluate hypotheses on something more closely approximating real data. Two learning models which were originally developed to explain synchronic learning trajectories have seen the most use in the study of grammar change: the Variational Learner (Yang, 2002) and the Tolerance Principle (Yang, 2016).

The Variational Learner differs from previous mathematical models like the TLA in that it incorporates competing grammars and reinforcement learning. Children assign probabilities to hypothesized grammars and "reward" grammars that parse input utterances by increasing their probabilities relative to other hypotheses. Over time, a child will settle on some distribution of competing grammars dictated by the proportion of utterances in the input that are unambiguously parsed by each grammar. Heycock et al. (2013) used the Variational Learner to model the gradual historical shift from V-to-T to V-in-situ in North Germanic languages (cf. the loss of V-to-I in English). They observe that V-raising or lack thereof interacts with optional embedded V2 in these languages, creating a complex distribution of evidence for a learner. Most constructions are ambiguous as to which grammar generated them, with limited exceptions in subordinate clauses. Using the IcePaHC corpus of historical Icelandic (Wallenberg et al., 2011) and a corpus of modern Swedish to investigate the empirical proportion of such clauses, they find a slight advantage for V-in-situ. Thus the Variational Learner predicts that children acquiring these languages should slightly prefer such a grammar to V-to-T and slowly increase its usage over time. This is consistent with the attested diachronic trajectory. Additionally, the quantitative approach provides a more elaborated and testable mechanism than trigger-based or cue-based approaches alone.

The Tolerance Principle (TP) is a model of productivity learning which has been successful at accounting for several phenomena in language acquisition. Acknowledging the ubiquity of exceptions to linguistic rules across all levels of the grammar, the TP provides a quantitative description for how a learner determines how many exceptions a hypothesized productive rule can tolerate and still enter into the grammar, while a rule with few enough exceptions can be learned and its exceptions memorized. A rule with too many exceptions should be abandoned. Experimental support for this tolerance threshold has emerged from a range of psycholinguistic studies (Schuler, 2017; Emond and Shi, 2020; Li and Schuler, 2023).

Several historical changes ranging from the development of the modern English stress system (Dresher

and Lahiri, 2022), changes in the distribution of past participle forms in Latin and Romance (Kodner, 2023), the "dative sickness" epidemic in Icelandic (Nowenstein et al., 2023), and more have been accounted for quantitatively with historical corpus linguistics and the Tolerance Principle. Trips and Rainsford (2022) modeled the shift from object experiencer verbs (patterning like modern *amuse*) to subject experiencer verbs (patterning like *admire*) with the TP. Calculated over a corpus of Middle English, they show that *admire*-type verbs, together with verbs borrowed from Old French used reflexively and non-reflexive constructions with omitted stimulus could have provided sufficient evidence to a learner to acquire a productive "if V is a psych verb, then it can be used with a subject experiencer" rule. This predicts that learners would overgeneralize subject experiencers at the expense of object experiencers as is observed diachronically in the language. Combined with a model for incrementation, this would account for the attested gradual shift in favor of subject experiencers in the history of English.

5 Conclusion

In conclusion, generative approaches to grammar change draw heavily from generative theoretical linguistics but are inseparable from both corpus-based and experimental methodologies. Together with an understanding of child language acquisition as both the locus and driver of innovations, they account for how and why the grammars of languages change over time. In combining the theoretical and empirical, they reveal both descriptive principles of grammar change as well as their underlying mechanisms.

6 Bibliography

Andersen, H. (1973). Abductive and deductive change. *Language*, pages 765–793.

Anderson, S. R. (2015). Morphological change should be book chapter? *The Routledge handbook of historical linguistics*, pages 264–285.

Biberauer, T. and Walkden, G. (2015). *Syntax over time: lexical, morphological, and information-structural interactions*, volume 15. Oxford Studies in Diachronic a.

Björnsdóttir, S. M., Westergaard, M., and Lohndal, T. (2020). The effects of attrition on grammatical gender: A view from north american icelandic. *Heritage Language Journal*, 17(3):332–354.

Chomsky, N. (1957). *Syntactic structures*. Mouton de Gruyter.

Chomsky, N. (1959). On certain formal properties of grammars. *Information and control*, 2(2):137–167.

Chomsky, N. (1965). Aspects of the theory of syntax. page 3.

Chomsky, N. and Schützenberger, M. (1963). The algebraic theory of context-free languages, computer programming and formal systems. volume 9, page 1968.

Cournane, A. (2014). In search of L1 evidence for diachronic reanalysis: Mapping modal verbs. *Language Acquisition*, 21(1):103–117.

Cournane, A. (2017). In defense of the child innovator. In Mathieu, E. and Truswell, R., editors, *Micro Change and Macro Change in Diachronic Syntax*, pages 10–24. Oxford University Press, Oxford.

Dresher, B. E. and Lahiri, A. (2022). The foot in the history of english: Challenges to metrical coherence. *Patrick Honeybone, Graeme Trousdale, Bettelou Los, & Clare Cowie (ed.), English Historical Linguistics: Change in Structure and Meaning*, pages 42–59.

Emond, E. and Shi, R. (2020). Infants' rule generalization is governed by the Tolerance Principle. In *45th Annual Boston University Conference on Language Development*.

Heycock, C., Sorace, A., Hansen, Z. S., and Wilson, F. (2013). Acquisition in variation (and vice versa): V-to-t in faroese children. *Language Acquisition*, 20(1):5–22.

Hyman, L. M. (2008). Enlarging the scope of phonologization. *UC Berkeley PhonLab Annual Report*, 4(4).

Kiparsky, P. (1965). *Phonological change*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Kodner, J. (2019). Estimating child linguistic experience from historical corpora. *Glossa: A Journal of General Linguistics*, 4(1):122.

Kodner, J. (2023). What Learning Latin Verbal Morphology Tells us about Morphological Theory. *Natural Language & Linguistic Theory*, 41:733–792.

Kroch, A. and Taylor, A. (2000). The penn-helsinki parsed corpus of middle English (ppcme2). department of linguistics, university of pennsylvania. cd-rom. *Department of Linguistics, University of Pennsylvania, CD-ROM*, second edition, release 4 ((http://www.ling.upenn.edu/ppche-release-2016/PPCME2-RELEASE-4).

Kroch, A. S. (1989). Reflexes of grammar in patterns of language change. *Language variation and change*, 1(3):199–244.

Labov, W. (2007). Transmission and diffusion. *Language*, 83(2):344–387.

Lenneberg, E. H. (1967). Biological foundations of language.

Li, D. and Schuler, K. D. (2023). Acquiring recursive structures through distributional learning. *Language Acquisition*, 30(3-4):323–336.

Lightfoot, D. (1999). *The development of language: Acquisition, change, and evolution*. Wiley-Blackwell, Hoboken, NJ.

Lightfoot, D. W. (1979). *Principles of diachronic syntax*. Cambridge University Press, Cambridge.

Marcus, M., Santorini, B., and Marcinkiewicz, M. A. (1993). Building a large annotated corpus of english: The penn treebank. *Computational linguistics*, 19(2):313–330.

Niyogi, P. (2006). *The computational nature of language learning and evolution*. MIT press Cambridge, MA.

Niyogi, P. and Berwick, R. C. (1997). A dynamical systems model for language change. *Complex Systems*, 11(3):161–204.

Nowenstein, I. E. et al. (2023). *Building yourself a variable case system: The acquisition of Icelandic datives*. PhD thesis, University of Iceland.

Papafragou, A. (1998). The acquisition of modality: Implications for theories of semantic representation. *Mind & language*, 13(3):370–399.

Paul, H. (1880). *Prinzipien der Sprachgeschichte*. Tubingen. Niemeyer.

Richter, C. (2021). *Alternation-sensitive phoneme learning: Implications for children's development and language change*. University of Pennsylvania.

Roberts, I. (2003). *Syntactic change: a minimalist approach to grammaticalization*. Cambridge University Press.

Sankoff, D. (1988). Variable rules. *Sociolinguistics: An international handbook of the science of language and society*, 2:984–997.

Schuler, K. D. (2017). *The acquisition of productive rules in child and adult language learners*. PhD thesis, Georgetown University.

Taylor, A. (2007). The york—toronto—helsinki parsed corpus of old English prose. In *Creating and Digitizing Language Corpora*, pages 196–227. Springer.

Trips, C. and Rainsford, T. (2022). Tolerating subject-experiencers? Yang's Tolerance Principle applied to psych verbs under contact in Middle English. *Journal of Historical Syntax*, 6(12):1–43.

van Gelderen, E. (2011). *The linguistic cycle: Language change and the language faculty*. Oxford University Press.

Wallenberg, J. C., Ingason, A. K., Sigurðsson, E. F., and Rögnvaldsson, E. (2011). Icelandic Parsed Historical Corpus (IcePaHC) version 0.9.

Weinreich, U., Labov, W., and Herzog, M. (1968). Empirical foundations for a theory of language change. In Lehmann, W. and Malkiel, Y., editors, *Directions for historical linguistics*, pages 95–195. University of Texas Press.

Yang, C. (2002). *Knowledge and learning in natural language*. Oxford University Press, Oxford.

Yang, C. (2016). *The Price of Linguistic Productivity*. MIT Press, Cambridge, MA.