Introduction

Diego Gabriel Krivochen
University of Oxford

The origins of generative grammar are inextricably linked to the birth in the 40's and 50's of formal language theory as we know it. This is not to say that there was no 'formal grammar' in the linguistic sense before that (the reader is referred, e.g., to Ajdukiewicz, 1935), but rather that the revolution that gave rise to post-structuralism in the US, and took the works of Bloomfield and Harris to new heights, was heavily based on the mathematical framework that became available through the works of Alan Turing (1936), Alonzo Church (1936), and Emil Post (1943), to mention but some of the main figures. A foundational assumption in generative grammar is the idea that languages are (infinite) sets of (finite) strings (Chomsky, 1959: 137; 2015: 156) whose structural descriptions are enumerated by a recursive procedure (Chomsky & Miller, 1963: 283; Langendoen & Postal, 1984: 18, ff.). The aim of a generative grammar of a language, as defined in the foundational works, is to enumerate structural descriptions for *all and only* the well-formed sentences of that language. This procedure has the form of a function that relates a structural description to each well-formed sentence in a language. In this sense, a language is a *generated* set insofar as an adequate grammar can recursively enumerate all and only well-formed sequences of basic building blocks.

In Chomsky's words,

> *We must require of such a linguistic theory* [a theory of generative grammars] *that it provide for:*
>
> i. *an enumeration of the class* $S_1, S_2, \ldots$ *of possible sentences*
> ii. *an enumeration of the class* $SD_1, SD_2, \ldots$ *of possible structural descriptions*
> iii. *an enumeration of the class* $G_1, G_2, \ldots$ *of possible generative grammars*
> iv. *specification of a function f such that* $SD_{f(i,\, j)}$ *is the structural description assigned to sentence* S_i*, by grammar* G_j*, for arbitrary i,j*
> v. *specification of a function m such that m(z) is an integer associated with the grammar G, as its value (with, let us say, lower value indicated by higher number)* (Chomsky, 1965: 31)

https://doi.org/10.1075/elt.00030.kri
Evolutionary Linguistic Theory 3:2 (2021), pp. 123–128. ISSN 2589-1588 | E-ISSN 2589-1596

The structural descriptions recursively enumerated by a generative grammar make the dependencies between elements in a string explicit in terms of a number of binary predicates: *is-a, daughter-of, mother-of*, etc. Grammars and the languages they generate are organised in an inclusive hierarchy known as the Chomsky Hierarchy (CH):

> *Theorem 1:*
> *for both grammars and languages, Type 0 ⊇ Type 1 ⊇ Type 2 ⊇ Type 3*
> (Chomsky, 1959: 143)

where Type 0 = unrestricted; Type 1 = Context-Sensitive; Type 2 = Context-Free; Type 3 = regular (Chomsky, 1956).

Each of these levels imposes restrictions on an unrestricted system. If we look at grammars in terms of rewriting rules, these restrictions pertain to what can appear in the left- and right-hand sides of rules of the form X → Y; we can present the CH in terms of the rule format that each level allows for.

Let capital Roman letters be variables over non-terminal symbols, lower-case letters be variables over terminals, and Greek letters be variables over sequences (of terminals or non-terminals). Then,

		Rule types	**Comments**
Type 3	Finite-state	A → Bc	B may be empty
Type 2	Context-free	A → αb	α is a possibly empty sequence of nonterminals
Type 1	Context-sensitive	αAβ → αγβ	α and β may be empty, γ cannot
Type 0	Unrestricted	φ → ψ	No restrictions on φ or ψ

The mathematical results obtained in the early days of generative grammar opened up the possibility of evaluating the *generative capacity* of a grammar. In Chomsky's words (1965: 60),

> *Suppose that the linguistic theory T provides the class of grammars G_1, G_2,… , where G_i weakly generates the language L_i, and strongly generates the system of structural descriptions Σ_i. Then the class {L_1, L_2, …} constitutes the weak generative capacity of T and the class {Σ_1, Σ_2, … } constitutes the strong generative capacity of T.*

Restricting the strong generative power of the grammar (that is: adequately characterising the classes of structural descriptions that the grammar could generate to avoid both *undergeneration* and *overgeneration*) was a central concern for generative linguists of this era (e.g., Chomsky, 1965: 99), and the ties between formal language theory and grammatical research were strong. In that period, the rules proposed to account for specific phenomena were subject to close formal scrutiny: in addition to Chomsky's (1956, 1959) works on the (local) inadequacy of Markov models and pure Context-Free grammars to generate all and only grammatical English sentences, we may cite McCawley's (1968) discussion of the generative power of the base component of an *Aspects*-style grammar (Chomsky, 1965) and the consequences of allowing for unordered rules, Reich's (1969) arguments for abandoning the CF / CS analysis of natural languages in favour of a FS analysis, Lyons' (1968) discussion of the (in)adequacy of strictly CF grammars for accounting for English subject-verb agreement, Langendoen's (1975) case for FS parsing of finite CF languages, Peters & Ritchie's (1973) highly influential formal results about the generative power of an *Aspects*-style grammar which allows for unbounded deletion, to mention but a few. Outside the generative camp, research on categorial grammars and their expansion in Montague (1973) also became the object of discussion of their generative power: Partee (1975) and Hamblin (1973) argue that Ajdukiewicz-style CG has the generative power of a CFG, but that Montague's quantification rules (S11-S14) '*introduce an element of context-sensitivity*' (Hamblin, 1973: 43).

During the late 70s and 80s, formal syntacticians seemed to split into two schools. In one, scholars following the lead of pursued the idea that natural language grammars were strictly context-free; notable works include Gazdar (1981); Pollard (1984), and Gazdar et al. (1985); see also Maling & Zaenen (1982) for an analysis of extraction phenomena under CF assumptions. Blevins & Sag (2013) provide an excellent overview of this research. Starting in the mid-70s, on the other hand, Aravind Joshi and colleagues explored the possibility that natural language grammars were in fact mildly-context sensitive, allowing at most for limited crossing dependencies between *two sets* of elements (Joshi et al., 1975; Joshi, 1985; Kroch & Joshi, 1985); these insights were formalised by means of Tree Adjoining Grammars (TAGs) (see however Swanson et al., 2013 for a modification of a TAG in which restricted adjunction permits a limitation to CF power).

All in all, the relations between FLT and linguistically-oriented- (as opposed to computationally-oriented) syntax and semantics '*have been on the wane since the mid-1990s*' (Joshi, 2011). The goal of this issue is to show that FLT has much to contribute to the development of empirically successful theories of syntax, semantics, and morpho-phonology. Furthermore, the contributions in this issue illustrate the idea that not all levels of linguistic analysis need to be modelled using the same formal tools.

Jane Chandlee's 'Nonderived environment blocking [NDEB] is input strictly local' argues that NDEB belongs in the most restrictive level in the sub-regular hierarchy (SRH; Heinz, 2018). This is in stark contrast to the models usually used in syntax, which tend to converge around mild context-sensitivity, under the traditional idea that languages are sets of strings. Chandlee presents a detailed approach to the computational complexity of phonological representations, where the idea that 'phonology is regular' is refined in terms of a hierarchy of subregular languages. She provides an account of NDEB in terms of input strictly local functions (ISL) within the SRH; these are processes that are locally evaluated and whose formulation does not use quantification, so that global evaluation of a formula is banned. Chandlee details the computational processes involved in NDEB in terms of strictly local input-output mappings which constitute the blocking processes.

Thomas Graf also refers to the SRH in 'The computational unity of Merge and Move', but in the context of an analysis of Merge and Move operations in Minimalist grammars (MGs). Graf combines considerations of computation and cognitive evolution, and argues that a system implementing Merge via a conjunction of constraints on feature satisfaction operations in derivation trees, such that feature requirements must be locally satisfied, supplemented with recursive adjunction (in the MG- rather than the TAG sense) gets Move 'for free'. Without adjunction, Merge is *strictly local* (SL): in SL-*k* dependencies, well-formedness is checked in terms of units of size *k*. Strict locality reduces cognitive computational cost, and also restricts the set of structures to be checked at any point. However, Graf argues, recursive adjunction creates unbounded dependencies that cannot be checked locally, and thus pushes the system beyond SL.

Robert Frank and Tim Hunter's 'Variation in mild context-sensitivity: Derivational state and structural monotonicity' presents a 2-dimensional typology of theories according to whether or not syntactic operations only apply at the root (±**Ext**) and whether or not the probing space is restricted (±**Fin**). TAGs are −Ext / +Fin systems (since trees may grow 'in the middle' by means of adjunction), whereas MGs are +Ext/+Fin systems. They demonstrate that head movement and/or remnant movement endows +Ext/+Fin systems with greater generative power and then test the respective empirical adequacy of the two system classes with an examination of Bulgarian multiple wh-movement, which creates centre embedding. Frank & Hunter show that despite the greater generative power of an MG (due to head movement and/or remnant movement), a TAG is capable of providing a linguistically revealing analysis of center-embedded multiple wh-questions that is unavailable for +Ext/+Fin systems.

The last paper argues against the standard assumption that syntactic structure is computationally uniform – that descriptive adequacy requires that a natural-language sentences may have both computationally flat and computationally hierarchical portions. The contribution of this paper is the idea that is indeed possible to formulate an explicit syntactic framework in which this computational variation using a version of lexicalised TAGs.

Finally, I would like to thank Jane Chandlee, Thomas Graf, Robert Frank, and Tim Hunter for their willingness to contribute to this special issue with their insightful papers, and the reviewers for their collaboration in making this special issue a reality. Also the Editorial Board of ELT for trusting me with this Special Issue. There is much that FLT can contribute to the work of grammarians of all theoretical persuasions, and I hope this special issue illustrates how much linguistic inquiry can be enriched by explicitly considering the formal issues that underpin syntactic, semantic, and morpho-phonological analysis.

References

Ajdukiewicz, K. (1935). Die syntaktische Konnexität. *Studia Philosophica* (Warszawa) *1*.1–28.

Blevins, J. & Sag, I. (2013). Phrase structure grammar. In M. den Dikken (ed.) *The Cambridge handbook of generative syntax*. Cambridge: CUP. 202–225. https://doi.org/10.1017/CBO9780511804571.010

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory* (2). 113–124.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

Chomsky, N. & Miller, G. (1963). Introduction to the Formal Analysis of Natural Languages. In R. Luce, R. Bush & E. Galanter (eds.), *Handbook of Mathematical Psychology* (pp. 269–321). New York: Wiley & Sons.

Chomsky, N. (2015). *The Minimalist Program* [2nd Edition]. Cambridge, Mass.: MIT Press.

Church, A. (1936). An unsolvable problem of elementary number theory. *American Journal of Mathematics*, *58*. 345–363. https://doi.org/10.2307/2371045

Gazdar, G. (1981). Unbounded dependencies and coordinate structure. *Linguistic Inquiry 12*. 155–184.

Gazdar, G., Klein, E., Pullum, G. & Sag, I. (1985). *Generalized Phrase Structure Grammar*. Cambridge, Mass.: Harvard University Press.

Hamblin, C.L. (1973). Questions in Montague English. *Foundations of Language 10* (1). 41–53.

Joshi, A.K. (1985) Tree adjoining grammars. In D. Dowty, L. Karttunen & A. Zwicky (eds.) *Natural Language Parsing* (pp. 206–250). Cambridge, Mass.: CUP. https://doi.org/10.1017/CBO9780511597855.007

Joshi, A.K. (2011). Review: Formal grammars in linguistics and psycholinguistics, vol. 1: An introduction to the theory of formal languages and automata by Willem J. M. Levelt. *Language 87*(2). 414–416.

Joshi, A.K., Levy, L. & Takahashi, M. (1975). Tree adjunct grammars. *Journal of the Computer and System Sciences 10*(1). 136–163.

Heinz, J. (2018). The computational nature of phonological generalizations. In L. Hyman & F. Plank (Eds.), *Phonological typology: Phonetics and phonology* (pp. 126–195). Berlin: Mouton De Gruyter.

Kroch, A. & Joshi, A.K. (1985). Linguistic relevance of tree adjoining grammars. Technical Report MS-CIS-85-18, Department of Computer and Information Science, University of Pennsylvania. Available online at: ftp://babel.ling.upenn.edu/papers/faculty/tony_kroch/papers/relevance3.pdf

Langendoen, D.T. (1975). Finite-State Parsing of Phrase-Structure Languages and the Status of Readjustment Rules in Grammar. *Linguistic Inquiry*, *6*(4). 533–554.
Langendoen, T. & Postal, P. (1984). *The Vastness of Natural Languages*. Oxford: Blackwell.
Lyons, J. (1968). *Introduction to theoretical linguistics*. Cambridge: Academic Press.
Maling, J. & Zaenen, A. (1982). A Phrase Structure Account of Scandinavian Extraction Phenomena. In P. Jacobson & G. Pullum (eds.) *The nature of syntactic representation*. Dordrecht: Reidel. 229–282. https://doi.org/10.1007/978-94-009-7707-5_7
Montague, R. (1973). The proper treatment of quantification in ordinary English. In J. Hintikka, J. Moravcsik & P. Suppes (eds.), *Approaches to natural language* (pp. 221–242). Dordrecht: Reidel.
Partee, B. (1975). Montague Grammar and Transformational Grammar. *Linguistic Inquiry* *6*(2). 203–300.
Peters, S. & Ritchie, R.W. (1973). On the generative power of transformational grammars. *Information Sciences 6*. 49–83. https://doi.org/10.1016/0020-0255(73)90027-3
Pollard, C. (1984). Generalized Context Free Grammars, Head Grammars, and Natural Languages. Ph.D. thesis, Stanford University.
Post, E. (1943). Formal Reductions of the General Combinatorial Decision Problem. *American Journal of Mathematics*, *65*(2). 197–215. https://doi.org/10.2307/2371809
Reich, P. (1969). The finiteness of natural language. *Language 45*(4). 831–843.
Swanson, B., Yamangil, E., Charniak, E. & Shieber, S. (2013). A Context Free TAG Variant. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, 302–310.
Turing, A. (1936). On Computable Numbers, with an Application to the Entscheidungsproblem, *Proc. London Math. Soc. 42*(2). 230–265.

Address for correspondence

Diego Gabriel Krivochen
Department of Linguistics, Philology and Phonetics
University of Oxford
Centre for Linguistics and Philology
Walton Street
Oxford OX1 2HG
England
diego.krivochen@ling-phil.ox.ac.uk