Chronicling the Prehistory of Generative Grammar:
Problems of Origination and Attribution

Geoffrey K. Pullum
School of Philosophy, Psychology and Language Sciences
University of Edinburgh

Generative grammars, a.k.a. rewriting systems, are devices for defining sets of sentences by enumeration. This idea is generally thought to have been invented by Noam Chomsky in the 1950s. (For his first use of the verb 'generate' for the relation between a formal system and a set of strings, see Chomsky 1951: 3.) It has totally dominated theoretical linguistics for six decades. But Chomsky was not the first to conceive of such devices.

Pāṇini (India, 4th century BCE) developed what are now seen as rewriting systems 2,500 years ago, but his work was entirely unknown to rediscoverers in the West, and was not their inspiration; it will not be my topic here (but see Kadvany 2016 for an important recent contribution).

Albert Sechehaye (1870–1946) hints informally at the idea of grammars 'generating' sentences (Seuren 2018: 131). He called the Subject + Predicate pattern "the generative principle ['principe générateur], the central organ of the entire grammatical mechanism" (Sechehaye 1908: 30), and claims that "The spirit of syntax … must be constructive and architectural" rather than "an enormous learned compilation of superficially classified facts" (Sechehaye 1916: 76). This certainly resembles Chomsky's critique of "taxonomic" linguistics. But Sechehaye said nothing about the mechanisms now called generative grammars.

A generative grammar is a formal system comprising [i] a finite set of one or more given strings (or other objects such as trees) and [ii] a finite set of operations for making new ones from those. Two examples:

Initial string:	$\{0\}$	Initial strings:	$\{a, b, aa, bb\}$
Operations:	$X \Longrightarrow 1X$; $X \Longrightarrow 0X$	Operations:	$X \Longrightarrow aXa$; $X \Longrightarrow bXb$
Defined set:	Even numbers in binary.	Defined set:	Palindromes over $\{a, b\}$.

Categorial grammars are clearly generative in this sense. Bar-Hillel (1953) cites the 1935 revival by **Kazimierz Ajdukiewicz** (1890–1962) of a seminal 1929 paper by **Stanisław Leśniewski** (1886–1939). This antedates Chomsky's work by a quarter of a century, but still is not the earliest 20thC work on generative grammars.

The 20th-century development of generative grammars emerged from the formalist program in logic. *Principia Mathematica*, published by Whitehead and Russell (W&R) in 1910–1913, was surprisingly informal in its logic. Corresponding to Modus Ponens (**MP**) W&R had: "Anything implied by a true elementary proposition is true" (Vol. 1, p. 94). This muddles together syntax and semantics: it fails to separate **derivability** of symbol strings from **truth** of propositions. Urquhart (2009) notes that W&R "fail to make the basic distinction between axioms and rules of inference … they are lumped together under the heading of 'Primitive Propositions.' "

Clarence Irving Lewis (1883–1964) sketched a way of avoiding confusion between form and meaning (Lewis 1918: 344ff). Start with a list of strings (axioms, but you don't need to know that they are), and some operations for adding new strings to the list (i.e., deriving theorems — but you don't need to know that either). Assuming an alphabet $\{\sim, \supset, (,), \vee, p, p', p'', \ldots\}$, **MP** should say something like this (modified from Lewis 1918: 357):

Modus Ponens: Find a string w_1 on the list that begins with '(', ends with ')', and contains '$\supset$'. Find another string w_2 on the list that is identical with the bit of w_1 between the '(' and the '$\supset$'. Take the bit of w_1 that follows the '$\supset$', remove its final ')', and add the remainder of w_1 to the list.

Emil Leon Post (1897–1954), as a young mathematics graduate student, was inspired by work like Lewis's to try and turn the intuitively conceived operations of logic into pure math. His PhD project in mathematics at Columbia under the philosopher-mathematician **Cassius Jackson Keyser** (1862–1947) was ambitious: to prove the consistency and semantic completeness of W&R's informally assumed propositional logic. It required:

[1] a rigorous truth-table method for showing tautologousness of propositional formulae;

[2] an algorithmic procedure reducing proof to deriving new symbol strings from already available ones with no reference to meaning;

[3] a proof that from W&R's axioms (represented as symbol strings) the syntactic proof method **[2]** can produce a string x if and only if the truth-table method **[1]** establishes that x is a tautology.

Post's mathematicization of inference rules makes them purely mechanical operations on uninterpreted strings. **MP** says that if something matching $(X_1 \supset X_2)$ on the list (where X_1 and X_2 can match anything at all), and the string that matched X_1 is also on the list, then whatever matched X_2 can be added to the list. In postdoctoral work, Post generalized, placing no upper bound on the length or number of strings in a production, thus yielding a rather daunting schema:

Post's general 'canonical' form for productions:

$$\begin{array}{cccccccc} g_{1,0} & X_{1,1} & g_{1,1} & X_{1,2} & \dots & X_{1,n_1} & g_{1,n_1} \\ g_{2,0} & X_{2,1} & g_{2,1} & X_{2,2} & \dots & X_{2,n_2} & g_{2,n_2} \\ & & \vdots & & \dots & \vdots & \\ g_{k,0} & X_{k,1} & g_{k,1} & X_{k,2} & \dots & X_{k,n_k} & g_{k,n_k} \end{array}$$

produce

$$h_1 \quad X_{r_1,s_1} \quad h_2 \quad X_{r_2,s_2} \quad \dots \quad h_j \quad X_{r_j,s_j} \quad h_{j+1}$$

In an actual production, each of the g_i and h_i is specified string; each X_i is a free variable over substrings; there are k strings preceding the word 'produce'; each string has a fixed length n_i, seting an upper bound to the number of g's and X's; and $1 \leq r_i \leq k$ and $0 \leq s_i \leq n_{r_i}$, which means the X variables in the last line all have to be present somewhere in the earlier lines. Thus a production may specify that the value of the i^{th} variable of premise number j is to be inserted at some point in the last line, but you can't just insert arbitrary random material.

A generated set in Post's terms is any set of strings definable by some finite set of productions. Post sought a DECISION procedure for generated sets — a finite method for determining WHETHER OR NOT a string could be generated by a given production system (hence, in a logic, whether or not a given string was a theorem). His program was destined to collapse — but in a way that was itself a fascinating discovery.

Normal form systems are production systems in a radically simplified form: they have only a single 1-symbol initial string, and the productions are restricted to saying '$g_1X \Longrightarrow Xg_2$' (= 'erase g_1 from the beginning and add g_2 on the end'). Post proved that if you allow additional auxiliary symbols that can never appear in generated strings, then **any set generated by a production system can be generated by one in normal form**.

Post's strategy for his research on logic was to encode W&R's logic using formulae in a format as simple as normal form production systems, and then develop a decision procedure for validity. But early in his postdoc at Princeton, in the fall of 1921, he realized that this was impossible. Even extremely simple production formats elude decidability. As an example, consider these rules (due to Liesbeth De Mol):

$$\{\, axX \Longrightarrow Xbc; \qquad bxX \Longrightarrow Xa; \qquad cxX \Longrightarrow Xaaa \,\}$$

If x is any arbitrary single letter from $\{a, b, c\}$ and X can cover any string over those letters, what will these rules do to a string of a's, b's, and c's? What seems to be true, from experimenting, is that the rules will reduce any input over the vocabulary $\{a, b, c\}$ to a single a, after a seemingly chaotic series of lengthening and shortening steps. But no one has ever managed to prove that this holds for all strings — or to find a counterexample. (The problem is related to the notorious Collatz Conjecture.)

Post had glimpsed by 1921 several deep and intimately related mathematical truths of huge importance: that there are necessarily incomplete logics (including every finitary symbolic logic capable of expressing statements of arithmetic), and sets that cannot be mechanically enumerated, and therefore problems in mathematics that are absolutely unsolvable. But he did not publish these claims, which meant that the insights would later be attributed entirely to others:

— **Kurt Gödel** (1906–1978): W&R's predicate logic is inherently incomplete (Gödel 1931).

— **Alonzo Church** (1903–1995): lambda-calculus equivalence is incomputable (Church 1936).

— **Alan Turing** (1912–1954): there are uncomputable real numbers, and absolutely unsolvable computational problems (Turing 1937).

Tragic personal reasons lie behind Post's failure to publish his earthshaking findings in the 1920s. But more than two decades later, Post published two fundamentally important papers about production systems (= generative grammars) and their expressive power:

— **Post 1943** (henceforth '*Reductions*'): production systems in normal form can generate any set that is generable at all. [Result obtained by 1922.]

— **Post 1947** (henceforth '*Unsolvability*'): production systems limited to the format '$X_1\ g_1\ X_2\ \Rightarrow\ X_1\ g_2\ X_2$' (later called Type 0 by Chomsky) also allow every generated set. [Result obtained in 1946, answering a question of Axel Thue (1914).]

An address to the American Mathematical Society (Post 1944, henceforth '*Integers*'), also appeared in the early 1940s. The paper essentially founded modern computability theory, but it contained hardly any details about production systems. And another paper in the mid-1940s proved the unsolvability of an misleadingly simple problem about paired lists of strings: Post (1946), henceforth '*Variant*'.

Paul Rosenbloom (1950) published a textbook on mathematical logic, three years after *Unsolvability*, with detailed coverage of Post's production systems (but curiously inscrutable bibliographical notes).

Chomsky mentions Post on 8 occasions to my knowledge, mostly with no citation. He attributes the term 'generate' to Post, citing *Integers* on two occasions (1959: 137n; 1961: 7). He also cites *Variant* once (1963: 382, following Bar-Hillel et al. (1961)). And in three early works he cites Rosenbloom, but not in connection with production systems. He has NEVER cited either *Reductions* (1943) and *Unsolvability* (1947) — the two most crucial papers.

What explains this omission? Apathy (lack of interest in historical antecedents)? Myopia (failure to see the papers' relevance)? Amnesia (forgetting that he had read them)? Dishonesty (deliberate concealment of an intellectual debt)? Ignorance (simply not knowing the relevant papers)?
I conclude by giving an argument for what I think is the right explanation.

References

Ajdukiewicz, Kazimierz. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1:1–27. English translation published in Storrs McCall (ed.), *Polish Logic 1920–1939*, 207–231, Oxford University Press.

Bar-Hillel, Yehoshua. 1953. A quasi-arithmetical notation for syntactic description. *Language* 29:47–58.

Bar-Hillel, Yehoshua, Micha Perles, and Eliyahu Shamir. 1961. On formal properties of simple phrase structure grammars. *Zeitschrift für Phonetik, Sprachwissenschaft, und Kommunikationsforschung* 14:143–172.

Chomsky, Noam. 1951. *Morphophonemics of Modern Hebrew*. Typescript of a radical revision of Chomsky's MA thesis, dated December 1951; retyped and published by Garland, New York, in 1979.

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

Chomsky, Noam. 1961. On the notion 'rule of grammar'. In *Proceedings of the Twelfth Symposium in Applied Mathematics*, 6–24. Providence, RI: American Mathematical Society.

Chomsky, Noam. 1963. Formal properties of grammars. In R. Duncan Luce, Robert R. Bush, and Eugene Galanter, eds., *Handbook of Mathematical Psychology*, Volume II, 323–418. New York: Wiley.

Church, Alonzo. 1936. An unsolvable problem of elementary number theory. *Amer. J. Math.* 58:345–363.

Davis, Martin, ed. 1958. *Computability and Unsolvability*. McGraw-Hill: New York, NY.

Davis, Martin, ed. 1994. *Solvability, Provability, Definability: The Collected Works of Emil L. Post*. Boston, MA: Birkhäuser.

Gödel, Kurt. 1931. Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I. *Monatschefte für Mathematik und Physik* 38:349–360.

Jakobson, Roman. 1969. Linguistics in its relation to other sciences. In *Proceedings of the 10th International Congress of Linguists*, 75–111. Bucharest: Éditions de l'Academie de la République Socialiste de Roumanie. Reprinted in *Selected Writings, II: Word and Language* (The Hague: Mouton, 1971), 655–708.

Kadvany, John. 2016. Pāṇini's grammar and modern computation. *Hist. & Phil. of Logic* 37(4):325–346.

Leśniewski, Stanisław. 1929. Grundzüge eines neuen Systems der Grundlagen der Mathematik. *Fundamenta Mathematicae* 14:1–81.

Lewis, C. I. 1918. *A Survey of Symbolic Logic*. Berkeley, CA: University of California Press, first ed.

Post, Emil L. 1943. Formal reductions of the general combinatory decision problem. *American Journal of Mathematics* 65(2):197–215. Also in Davis 1994b: 442-460.

Post, Emil L. 1944. Recursively enumerable sets of positive integers and their decision problems. *Bulletin of the American Mathematical Society* 50:284–316. Also in Davis 1965: 305-337 and 1994b: 461-494.

Post, Emil L. 1946. A variant of a recursively unsolvable problem. *Bulletin of the American Mathematical Society* 52(4):264–268. Also in Davis 1994b: 495-500.

Post, Emil L. 1947. Recursive unsolvability of a problem of Thue. *Journal of Symbolic Logic* 12:1–11. Also in Davis 1994: 503–512.

Rosenbloom, Paul. 1950. *The Elements of Mathematical Logic*. New York: Dover.

Sechehaye, Albert. 1908. *Programme et méthodes de la linguistique théorique: psychologie du langage*. Paris: Champion.

Sechehaye, Albert. 1916. La méthode constructive en syntaxe. *Revue des langues romanes* 59(1/2):44–76.

Seuren, Pieter A. M. 2018. *Saussure and Sechehaye: Myth and Genius; A Study in the History of Linguistics and the Foundations of Language*. Leiden: Brill.

Thue, Axel. 1914. Probleme über Veränderungen von Zeichenreihen nach gegebenen Regeln. In *Skrifter utgit av Videnskapsselskapet i Kristiana, I*. Oslo: Norske Videnskaps-Akademi.

Turing, Alan M. 1937. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* series 2, 42, issue 1:230–265.

Urquhart, Alasdair. 2009. Emil Post. In Dov Gabbay and John Woods, eds., *The Handbook of the History of Logic, Volume 5*, 617–666. Amsterdam: Elsevier.

Whitehead, Alfred North and Bertrand Russell. 1910–1913. *Principia Mathematica*. Cambridge: CUP.