

Century of history  
David Lightfoot

**Introduction\***

The decades following the publication of *On the Origin of Species* were the final flourish of what can reasonably be viewed as the century of history. The grand ideas of the nineteenth century built on what had been developed in the eighteenth and focused on documenting and understanding changes in species, languages and political systems, changes through time - history. People working in fields that we now call evolutionary biology, linguistics and political science read each other, talked to each other and converged to develop common ideas on the nature of change, which led to common successes and common failures. The central commonality was that Darwin, Marx and the philologists, mostly German, thought that there were general principles of history to be discovered. That belief brought some successes but turned out to be a chimera; the new fields failed to achieve the explanations that had been sought and twentieth century work largely turned away from seeking to explain change.

Now, almost two centuries later, there is a new cross-disciplinary focus on change through the lens of complexity science, focusing on phase transitions and emergent phenomena in disparate domains. This belongs to a family of ideas that includes earlier work on catastrophes and chaos. Again evolutionary biologists, linguists, social scientists and now physicists and chemists are reading each other. Where complexity *science* moves beyond the level of metaphor and develops meaningful principles that might constitute a complexity *theory*, then it may usher in a new century of history that will achieve greater explanatory success, offering the prospect of cross-fertilization with complexity theories across the domains of physics, biology, linguistics and the social sciences.

Here I will look at the triumphs and failures of nineteenth-century linguistics, suggesting that they have analogs in biology and political science, and show how the current century's focus on tipping points and emergent phenomena opens new prospects if we construe the language faculty as a biological entity. A key difference is that the early work took languages to be social entities in the world outside, while we take languages to be mental constructs growing in the minds of individual children.

**The comparative method**

It was clear that God did not place the languages of the modern world where they are now spoken, and the historical record shows that they evolved from earlier forms. This view, articulated clearly in the eighteenth century, has never attracted the attention of the "creationists," who have engaged evolutionary biologists. People thought that we might be able to find precisely how languages developed from one another. A central idea was that if one could understand the sound changes that transform words as they are transmitted from generation to generation, so one might understand the historical relationships among languages and how a language descended from some earlier language. The upshot was a century of research, which discovered much about sound changes affecting the Indo-European languages.<sup>1</sup>

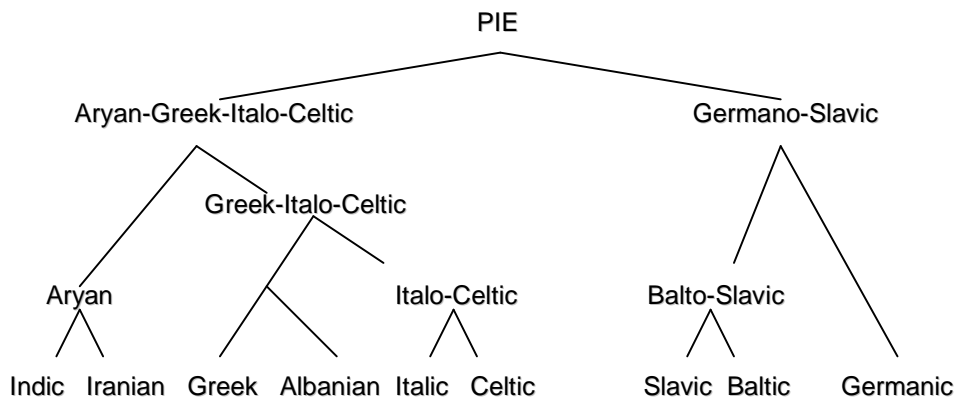
The word for ‘father’ in the Romance languages is French *père*, Spanish *padre*, Italian *padre*, Sardinian *patre*, Catalan *pare*, Portuguese *pai*, all transmogrifications of Latin *pater*. The Germanic languages have similar, but different words: English *father*, Dutch *vader*, German *Vater*, Danish *fader*, Gothic *fadar*. Compare very different words for ‘father’ in other languages, where the words do not have a common ancestor (as far as we can tell): Chinese *fuqin*, Japanese *titi-oya*, Basque *aita*, Finnish *isä* and Korean *apeci*. The degree of similarity between the Germanic and Romance forms reflects the fact that the Germanic languages are historically related to each other and are also historically related to the Romance languages, though less directly. And much less directly to Chinese, Japanese, Korean, Basque and Finnish.

Such observations had led Sir William Jones, a British judge who had worked in Calcutta with the East India Company, to speculate famously in an after-dinner speech in 1786 that Greek, Latin and Sanskrit are similar in ways that indicate that they descended from a common source, which might no longer exist. The key idea here was that the source language was not Hebrew, Sanskrit or Latin but rather a hypothetical language that no longer exists. A critical, corollary axiom was that languages change over time and that new languages are constantly emerging: modern French and Romanian evolved from forms of Latin, Hindi and Urdu developed from forms of Sanskrit, and Latin and Sanskrit were both outgrowths of an earlier source language for which we have no records.

That idea was taken up passionately in Germany, where scholars deduced many properties of the hypothetical source language that we now know as Proto-Indo-European (PIE). The PIE word for ‘father’ was a two-syllable word and the initial segment was a labial consonant, probably *p* (/p/ in Romance, /f/ in Germanic), followed by an open vowel and an alveolar consonant, then some kind of vocalic *r*.

Words are transmitted from one generation to the next, and they may change their form over time. By examining CORRESPONDENCES, as in the Romance words for ‘father,’ linguists developed the comparative method and postulated that languages are historically related to each other to greater or lesser degrees, and that they cluster in families. English and Dutch have more cognate words, and the cognate words are more similar to each other, than English and Spanish, so English and Dutch are more closely related, even though they have plenty of words that are not cognate: *bike* and Dutch *fiets* refer to the same object but do not have a common origin. This reasoning was the comparative method, the only real basis for reconstructing the properties of a hypothetical proto-language.

In 1863 August Schleicher published a short book on Darwinian theory (Schleicher 1863) and he postulated what we now call cladograms (or trees), indicating the genetic relationship among language families, based on a rough quantification of cognate words, shared derived characters or “synapomorphies.” Such language cladograms are comparable to classifications of botanical species and genera in the Linnaean system and reflected the cross-disciplinary concerns of the late nineteenth century.



This is the cladogram for the earliest stages of the Indo-European languages proposed by Schleicher. I have anglicized and modernized some of his labels. The tree is incomplete, of course: many languages that I have not specified fall under Germanic, and Celtic was subdivided into two groups, Brythonic and Goidelic, the former consisting of Cornish (which died out in the eighteenth century), Breton and Welsh, the latter embracing Manx (which died out in the nineteenth century), Irish and Scots Gaelic. The tree expresses the idea that, say, the Celtic languages are more closely related to the Latin-derived "Italic" languages (which we now call "Romance": Italian, French, Sardinian, Spanish, Galician, Catalan, Rumanian, etc.) than the Slavic languages (Polish, Russian, etc.) are to either; and that the Baltic languages (Latvian, Lithuanian) are more closely related to Germanic than to Celtic, but not as closely related to Germanic as they are to the Slavic languages. And so on. This was the first tree proposed for the Indo-European languages or, in fact, for any language family. We have no records for any of the parent languages and we now doubt that Albanian and Greek represent any kind of unity, nor Italo-Celtic. On the other hand, some features of Schleicher's cladogram remain undisputed and many relationships not specified here have come to be established convincingly.

These cladograms idealize away from the fact that languages do not split sharply and suddenly emerge in their full individuality. The splitting process is more gradual and complex, and is initiated by relatively minor divergences. Languages are dynamic systems and not organic entities: there is no linguistic distinction between a dialect and a language, so there is some arbitrariness in what we call a distinct language. We might say that the first change that affected, say, Latin and not any of the other languages is the bifurcation point, the point at which Latin suddenly splits. But that is not enough. Saying that French and Italian are descended from Latin glosses over the fact that they descended from different forms of Latin, and that "Latin" is a cover term for many different forms of speech. As a result, the conventional tree models of historical linguists would require vast elaboration to be equivalent to modern, biological cladograms, which are usually based strictly on the molecular structure of organisms.

Not only do these cladograms treat languages in the aggregate as if they were organic entities in the outside world but they capture only homologies, features “inherited” from a common ancestor, and not analogies, features arising independently through common responses to environmental similarities. Humans share hair and a warm-blooded physiology with chimpanzees and mice as a result of evolutionary history, or homology. Birds and bats, on the other hand, both fly by analogy; they have a very different evolutionary history, bats being mammals. Similarly languages share features because of a common ancestry AND because of common developments.

There may be common innovations that have nothing to do with a common history. To take a trivial example, many historically unrelated languages have a word for ‘television’ that sounds much like English *television*. During the period of the Scandinavian settlements, English drew many common words directly from the language of the settlers: *bait, bull, egg, fellow, give, hit, husband, law, low, loose, meek, oar, sister, skin, sky, take, wrong*. English took much from French during the period of the Norman occupation, musical terms from Italian, *pundit, thug* and *calico* from India, and *moccasin, toboggan* and *tomahawk* from American Indian languages. This kind of commonality, due to factors other than a common history, is not expressed by cladograms, linguistic or biological.

Nineteenth-century linguists developed rich hypotheses about the words of PIE but much is speculative; the further back we go, there is more conjecture. The Indo-European and Semitic languages probably descended from a common source, which predated PIE, but it is difficult to be confident in what the words for ‘father’, ‘mother’, etc. were. Indeed, it is likely that human language evolved just once, in East Africa, and that all the languages of the world derive from that single evolutionary step and are historically related. We believe this because the human language faculty appears to be uniform across the species; if it had evolved at different times in different places, one would expect to see different faculties in different groups, as one sees different visual systems that have different evolutionary histories.

However, trying to reconstruct very ancient superfamilies, like a common ancestor for Indo-European and Semitic, is precarious. Some of the reasoning is based on the genetic affinities of speakers, and this is quite dubious. South Indians are closely related to certain African groups genetically but this tells us nothing about the languages, because linguistic and biological affiliations do not necessarily correlate. English remains a northwest European language, even though it is spoken by people from southeast Asia and Australia; and it would remain a northwest European language even if a cataclysm were to eliminate English-speaking communities outside the southern hemisphere. Modern Hebrew is a Semitic language even though it has been influenced greatly by Indo-European languages.

The Nostratic superfamily has captured people’s imagination and has been treated in long articles in popular magazines. The term is due to Holger Pedersen but the idea of a superfamily goes back to the nineteenth century. Henry Sweet (1900) argued that the Indo-European family came from the same source as “Ugrian” (Finno-Ugric), “Altaic” (which included Turkic, Mongolian, Tungusic and Japanese) and Sumerian. Nostratic

has changed over the years. A recent taxonomy (Bomhard 1990) includes Indo-European, Kartvelian (south Caucasus), Afro-Asiatic, Uralic-Yukaghir, Elamo-Dravidian and perhaps Sumerian. Others have added a Korean-Japanese family and a new Chukchi-Eskimo group. From there it is a small step to "ProtoWorld," which has also been advocated by more imaginative colleagues. Proponents of Proto-World assume monogenesis for all the languages of the world AND assume that it can be demonstrated through the surviving properties of recorded languages. The first assumption is probably right but the possibilities for reconstruction are too limited for us to know peculiarities from so far back (Ringe 1995).

There has always been a tension in reconstruction work, some linguists claiming to be reconstructing prehistory and others taking reconstructions to be abstractions that express relationships among existing languages (see the debate between Lightfoot 2002a,b and Campbell & Harris 2002). That tension was highest when language study was entirely historical, as in the nineteenth century, and the only question to ask about a language was how it got to be the way it is. At the end of the century Hermann Paul (1880) pontificated that 'it has been objected that there is another scientific view of language possible besides the historical; I must contradict this.'

Work on the history of languages first became central in Germany, and it grew not only out of Sir William Jones's insight but also from a general intellectual movement of the late eighteenth to mid-nineteenth century that we call "Romanticism." The Romantics focused on ethnic and cultural origins of various phenomena. Since race, language and culture were seen as closely related, the reconstruction of the prehistory of Germanic was attractive to the Romantic temperament. These links were quite clear in the work of the linguists Herder and Grimm.

### Laws of history

Nineteenth-century linguists studied similarities among cognate words, words derived from the same historical source; this was the basis for establishing historical relationships, and then for establishing the sound changes that derived one form from another historically. As the century progressed, they formulated historical "laws" with ever greater precision. To get a taste of the enterprise, we can track one matter in some detail: the shift in the Germanic consonant system, which became famous as "Grimm's Law."

In 1822 Jacob Grimm, one of the brothers who collected fairy stories, and who was also a philologist, revised his *Deutsche Grammatik* by adding a 595-page account of the phonology of 15 languages. He explicated the Germanic consonant shift, formulating a historical cycle (*Kreislauf* 'rotation'). He observed that the ancient languages showed a voiceless stop (the sounds represented roughly by the letters *p*, *t*, *k*) where Germanic languages such as Gothic and English showed a corresponding fricative (*f*, *th*, *h*).

	Sanskrit	Greek	Latin	Gothic	English
<i>p</i>		<i>pod-</i>	<i>ped</i>	<i>fotus</i>	<i>foot</i>
<i>t</i>	<i>trayas</i>	<i>treis</i>	<i>tres</i>	<i>threis</i>	<i>three</i>

*k*                      *kardia kor*                      *hairto heart*

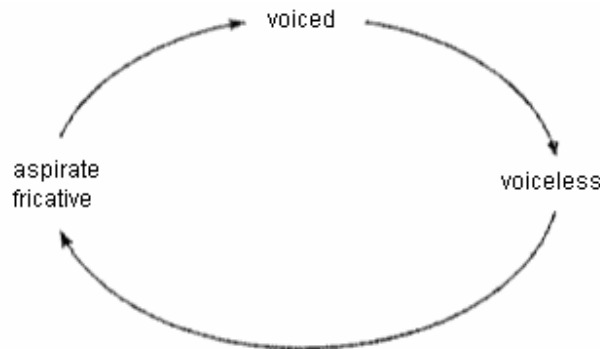
Similarly, where the ancient languages showed a voiced stop, Germanic showed a voiceless stop.

*b*                      *turbe turba*                      *thaurp thorp*  
*d*    *daśa*            *deka decem*                      *ten*  
*g*                      *agros ager*                      *akrs acre*

And where the ancient languages showed an aspirate (a stop pronounced with a puff of air and written *b<sup>h</sup>*, etc.), Germanic showed an unaspirated voiced stop.

*b<sup>h</sup>*    *b<sup>h</sup>arāmi*        *p<sup>h</sup>ero fero*                      *baira bear*  
*d<sup>h</sup>*    *d<sup>h</sup>ā-*              *tit<sup>h</sup>ēmi facio*                      *do*  
*g<sup>h</sup>*    *stig<sup>h</sup>-*              *steik<sup>h</sup>o*                      *steiga go*

Grimm took the ancient languages to manifest the consonants of the hypothetical parent language, PIE, more or less directly. The manifestation was not always direct: the PIE voiced aspirates *d<sup>h</sup>* and *g<sup>h</sup>* were realized as voiceless *t<sup>h</sup>* (written θ, theta) and *k<sup>h</sup>* (χ, chi) in Greek, the aspirates *b<sup>h</sup>* and *d<sup>h</sup>* as a voiceless fricative *f* in Latin. This meant that there were some changes between PIE and the ancient languages. Grimm was interested mostly in the changes between PIE and early Germanic, viewing them as a cycle.



There were exceptions, cases in which the hypothesized correspondences did not hold, but he showed no interest in them. Others were more interested, however. Two generations of scholars sought to address the exceptions systematically, and succeeded with a discovery by Verner fifty years later.

There were three classes of exceptions. A voiceless stop did not change to an expected fricative if preceded by a fricative; so Germanic shows *ist* (rather than *isth* with the expected *th*) where the ancient languages have *est*, *esti* and *asti* for 'is.' Second, the mathematician Hermann Grassmann showed ingeniously that PIE must have had aspirates where the ancient languages showed voiced stops, as in the initial segments of the Sanskrit words for 'daughter' and 'offer,' *duhitā* and *bod<sup>h</sup>āmi*, so one is not surprised to find Gothic *dauhtor* and *biudan*. Third, in 1876 Karl Verner observed that certain voiceless stops in the ancient languages did not become voiced stops in Germanic, as Grimm's Law would lead us to expect, but became voiced fricatives. So Sanskrit *pitār*, Greek *patēr* and Latin *pater* show two voiceless stops, indicating that the consonants for

'father' in PIE were *p-t-r*. The first of these stops behaved according to Grimm's Law and became a fricative *f* in Germanic: Gothic *fadar*, English *father*, etc. However, the second stop, the *t*, did not become a voiceless fricative as in English *thin*; unexpectedly, it became a voiced stop *d* in Gothic. On the other hand, the word for 'brother' worked as expected: Sanskrit shows a medial *t* (*bhrātā*), which corresponds to a voiceless fricative in Gothic (*brōþar*). Verner showed that the different histories of the medial *t* in 'father' and 'brother' were a function of the phonetics of the words: in one case the ancient accent preceded the *t* (*b<sup>h</sup>rātā*) and in the other case it followed (*pitar*). This observation, not surprisingly, entered the canon alongside Grassmann's Law, as Verner's Law.

Verner's triumph in rendering Grimm's Law exceptionless made 1876 the *annus mirabilis*, yielding the two theses of the Neogrammarians: that sound change is regular and exceptionless and that it is phonetically conditioned. The Neogrammarians represented the conceptual culmination of the nineteenth century and they had a profound influence on the twentieth: American structuralism essentially translated the two theses about historical change into principles for synchronic analyses, whereby abstractions were limited to surface, phonetic factors.

### **Explanation**

The triumph of nineteenth-century linguists lay in describing phonetic changes that words undergo from generation to generation. There were problems with the descriptions and there were changes that were not phonetically conditioned – they were assigned to a different category of change, loosely defined as “analogy” – and some changes, like those described in Grimm's Law and the Great Vowel Shift of Middle English, were systematic, properties not just of individual sounds but of the whole system. Also, languages were, in effect, taken to be only collections of words and there was no work done on syntax in anything like the sense we understand it today. Finally, referring to individual sound changes as “laws” was a misnomer, because they were not general laws like Boyle's Law and explanations were needed for why they applied when and where they did. Therein lay the principal failure: by the end of the century there were impressive compilations of changes that had occurred but for no apparent reason.

Contemporary linguists struggled with this failure and paid much attention to matters of explanation. There were, at the time, two models of explanation: Newtonian mechanics and Darwinian selection. Newton saw phenomena as describable by deterministic laws of force and motion, such that all future states were, in principle, predictable in a “straight-line,” linear fashion from a complete knowledge of the present state. This inspired the notion of sound laws to describe the history of changes and Franz Bopp offered a mechanical explanation of vowel changes by invoking a “law of gravity” and postulating that syllables had different “weights.” Darwin was inspired by work on language history and he, in turn, inspired linguists like Schleicher to treat languages like natural organisms, plants and animals, identifying forces that would make languages fitter. Languages, like species, compete in a struggle for survival and there were inexorable laws of change to be discovered. Darwin himself thought that languages tended to change in the direction of having shorter, “easier” forms and that this could be explained by natural selection (Darwin 1874).

Nineteenth-century linguists knew that language reflected psychological properties, but there was a strict demarcation between the work of linguists and that of psychologists and contemporary ideas were problematic (Lightfoot 1999: ch2). Grimm, for example, adopted a mystical belief in a Hegelian *Sprachegeist*, which had some existence above and beyond individuals (Grimm 1848). He explained his law of consonant shifts as

Connected with the Germans' mighty progress and struggle for freedom ... the invincible German race was becoming ever more vividly aware of the unstoppable advance into all parts of Europe ... How could such a forceful mobilization of the race have failed to stir up its language at the same time, jolting it out of its traditional rut and exalting it? Does there not lie a certain courage and pride in the strengthening of voiced stop into voiceless stop and voiceless stop into fricative?

Apart from this kind of thing, linguists generally did not appeal to psychology to explain historical changes. Instead, as thinking in the "century of history" demanded, there were independent laws of history to be found. Grimm's, Grassmann's and Verner's laws operated on the sounds of languages and were manifested in the relationship between corresponding words in different, historically related languages. They required a deeper explanation and changes were taken to be *directional*, as in biology, where the replacement of one species by another was taken to result from a mutation that yielded an organism more successful in the struggle for survival in a particular environment.

There was consensus that language change followed fixed developmental laws and that there must be a direction to change but there was active disagreement about what that direction was. Rasmus Rask held that languages became simpler; Schleicher identified a progression from isolating to agglutinating to inflectional types, although this was said to hold for preliterate societies, whereas Rask's drive to simplicity held for literate societies.

By the early twentieth century the data of linguistics comprised an inventory of sound changes occurring for no good reason and tending in no agreed direction. The historical approach had not brought a scientific, Newtonian-style analysis of language, of the kind that had been hoped for; there was no predictability to changes. The psychological moves could not provide the necessary underpinning. Consequently the program was not viable and there was no science of language history that met nineteenth-century demands. The main problem was that the demands were too ambitious.

For all the talk of directionality, nineteenth-century linguists seem not to have been at ease with it; certainly their analyses allowed for particular, contingent factors. After all, under certain circumstances some forms of spoken Latin developed into some form of French, and under other circumstances other forms of Latin developed into Spanish and Sardinian; there was nothing intrinsic to Latin that made it develop into French.

As we have noted, the deterministic view of history, the idea that there are laws determining the way that history proceeds, is a hallmark of the nineteenth century. We have seen how it guided the study of language and it played a role in the development of Darwinian ideas and in the domain of political history.

Darwin read the linguists and vice versa and Marx dedicated *Das Kapital* to Darwin. Marx too had an interesting theory of change, whereby ideas are socially embedded and

are amended through conflict, through the clash of theses and antitheses. He understood social change and revolution in terms of small insults to a system building up until the system breaks. However, he built on eighteenth-century ideas that there is a political *science*, a science of the relationships of human beings to each other and to their environment, and was very much a nineteenth-century thinker, caught up in notions of predestiny and determinism. He developed historical laws predicting that a feudal society must necessarily develop into a mercantilist society, a mercantilist into a capitalist society, capitalism into socialism, and socialism into communism. For Marx, the real task of economics was to explain how society evolved over time. At his funeral, Engels eulogized him: ‘Just as Darwin discovered the law of evolution in organic matter, so Marx discovered the law of evolution in human history.’

The nineteenth-century historicist paradigm – the notion that there are principles of history to be discovered, which would account for a language’s development – was largely abandoned in the 1920’s. Indeed, there was a virulent anti-historicism in the writing of structuralists like Franz Boas, Leonard Bloomfield and Edward Sapir. They worked on language change, showing that the comparative method could be applied to the unwritten, indigenous languages of North America; Bloomfield worked on the reconstruction of proto-Algonquian for most of his career. They also perpetuated many of the analytical procedures of historical linguists in their own synchronic work. However, they abandoned *historicism* and with it the earlier program of seeking to discover how languages came to be the way they are. The historicist program wasn’t really refuted or shown to be seriously inadequate; rather it was abandoned as yielding diminishing returns. The paradigm had turned to psychology to avoid a built-in circularity and then collapsed because of the inadequacy of the psychology invoked.

The synchronic work of the twentieth century developed ideas about the structures of language, going beyond ideas that languages were essentially just collections of words. The second half of the century developed new approaches to syntax, distinguishing properties that are intrinsic to the organism from those that arise as a result of environmental influence, and distinguishing between external language in the world outside and internal systems that develop as part of an individual’s biological make-up. This, in turn, has opened new approaches to change over time.

### **Poverty-of-stimulus reasoning and the linguistic genotype**

This year we celebrate Darwin’s 200<sup>th</sup> birthday and the 150<sup>th</sup> anniversary of *On the Origin of Species*, but we are also celebrating the 150<sup>th</sup> anniversary of the major publication of Gregor Mendel’s work on pea plants and the idea of genes. Syntacticians follow Mendel’s reasoning from the poverty of the stimulus in identifying factors that must be built into the organism in advance of experience and cannot be derived from experience. This entails that we postulate properties that hold of all humans intrinsically and that affects the way we can address Darwinian questions about diachronic change.

When children develop language, they do not acquire English or Japanese but rather a private system that enables them to communicate. English is not a biological entity and its sentences do not constitute a recursively enumerable set. This can be seen most easily by considering a sentence like *John might could drive* and asking whether it is a sentence of English; the answer is that it is in Alabama but not in Alaska.

Children do not just imitate what they hear but develop a system that is far richer than the fragmentary and limited speech that they happen to encounter in their first few years. For example, they hear a finite number of utterances and the system they develop must be finite but it ranges over infinity; children develop an internal system that generates an infinite range of expressions. The system is *recursive* and that property cannot be derived purely from experience.

All human language systems have three recursive, looping devices that permit structures of indefinite length:

RELATIVIZATION: *This is the cow that kicked the dog that chased the cat that killed the rat that caught the mouse that nibbled the cheese that lay in the house that Jack built.*

COORDINATION: *Ray and Kay went to the movie and Jay and Fay to the store, while Gay and May and Clay worked where Shay and Jack were watching, but Zach and Mack and Shaq slept.*

COMPLEMENTATION: *Ray said that Kay said that Jay thought that Fay said that Gay told me that ...*

No child ever heard a sentence of indefinite length - they all ended - but the linguistic system that every child acquires, with the three looping devices, has the capacity for indefiniteness. Children understand and use novel sentences all the time and they do that by virtue of having an open-ended system.

Furthermore, virtually every generalization breaks down in ways that are unlearnable by children and the limits to generalizations illustrate the poverty-of-stimulus reasoning used by Mendel, which in turn illuminates intrinsic properties. For example, children hear expressions with *is* in its full form (1a) or reduced (1b), and therefore may deduce that there is an operation reducing *is* to 's. However, in 2a.b *is* is not reducible and nobody would say *\*Kim's taller than Jim's* or *\*I wonder what the problem's with him* (non-occurring, hypothetical forms are marked \*). The limits to the generalization are not learnable, because children have no evidence for the non-occurrence of the latter cases. Mere non-occurrence is not evidence for children, because they say many things they have not heard. However, experimental work shows also that children do not try out such forms in the way that they use *go-ed* or *foots*; somehow at the earliest stages when they can be tested, they just know. This is an example of what is meant by the poverty of the stimulus: what children hear is not rich enough to determine their eventual behavior.<sup>2</sup> That behavior is shaped in part by an internal genetic component, much as the properties of Mendel's pea plants were subject to the requirements of internal factors, what we came to call genes.

1.a. Kim is taller

b. Kim's taller

2.a. Kim's taller than Jim is

b. I wonder what the problem is with him

Another poverty-of-stimulus problem. Alongside 3, children hear forms without the clause introducer *that* (4) and may deduce that there is an operation deleting *that*.

However, (5a,b) have no corresponding forms with a deleted *that*: \**The car arrived yesterday Kay drove* or \**Kay drove was obvious to all of us*. Again, nobody tells children that these forms do not occur and children have no direct evidence.

3.a. Peter said [that Kay drove]

b. The car [that Kay drove]

c. It was obvious [that Kay drove]

4.a. Peter said [Kay drove]

b. The car [Kay drove]

c. It was obvious [Kay drove]

5.a. The car arrived yesterday [that Kay drove]

b. [That Kay drove] was obvious to all of us

Following Mendel, one might solve this poverty-of-stimulus problem by postulating information that is built into the organism, perhaps a condition on deletion (6).<sup>3</sup>

6. Something may be deleted, if it is (in) the complement of an adjacent, overt word.

*That Kay drove* is the complement of the adjacent word in (3) but not in (5); in (5a) it does not complete the meaning of the adjacent *yesterday* and *that* may not delete. In (5b) *that Kay drove* isn't adjacent to anything and *that* may not delete. In this way the general principle of (6) interacts with a general and learnable operation deleting clause introducers to distinguish what we say (e.g. 4) from what does not occur: \**The car arrived yesterday Kay drove*, \**Kay drove was obvious to all of us*.

Similarly (6) interacts with a general, learned property of English speakers: they allow the second of two identical verbs to be "gapped" (*Fay saw Ray and Jim Tim*) and (6) distinguishes (7a) from the non-occurring (7b) without *that*. In (7b) *that Kim stayed* is the complement of an adjacent verb that is not pronounced or "overt" and *that* does not delete.

7.a. Fay said that Ray left and Jim <sub>ve</sub> [that Kim stayed]

b. \*Fay said that Ray left and Jim <sub>ve</sub> [Kim stayed]

Postulating (6) as a general principle also allows us to understand the difference between (2) and the non-occurring *Kim's taller than Tim's* and *I wonder what the problem's with him* (8).

8.a. Kim's taller than Tim's ~~tall~~

b. I wonder what the problem's ~~what~~ with him

In (8a) the understood (deleted) *tall* is not the complement of anything after *is* has been incorporated into the preceding word, nor the understood (deleted) *what* in (8b); therefore the deletions may not take place and forms like (8) do not occur.

Consider one final operation that speakers of English learn and use over a wide domain: whole verb phrases may be deleted, as in (9).

- 9.a. Max left for Rio on Wednesday and Mary did <sub>VP</sub> as well
- b. Max left for Rio, although Mary didn't <sub>VP</sub>
- c. Although Max couldn't <sub>VP</sub>, Mary left for Rio
- d. Susan went to Rio. Yes, but Jane didn't <sub>VP</sub>
- e. The man who left for Rio knows the woman who didn't <sub>VP</sub>
- f. Don't <sub>VP</sub>!

In many different structural configurations the ellipsed verb phrase is interpreted as *leave for Rio*. In all these cases, the deleted VP is the complement of an adjacent overt verb (*did*, *didn't*, *couldn't*, *don't*). However, principle (6) distinguishes (10) from the non-occurring (11).

- 10.a. They denied reading it, although they all had <sub>VP</sub>
- b. They denied reading it, although they often/certainly had <sub>VP</sub>
- c. I haven't seen that movie but John has <sub>VP</sub>
- 11.a. \*They denied reading it, although they had all <sub>VP</sub>
- b. \*They denied reading it, although they had often/certainly <sub>VP</sub>
- c. \*I haven't seen that movie but John's <sub>VP</sub>

In (10) the understood VP is the complement of and adjacent to *had/has* and in (11) it is not.

There is much more to be said about these analyses, but Mendelian poverty-of-stimulus reasoning enables us to postulate four simple, learnable operations for which children have evidence in their everyday experience. Hearing a simple sentence like *Peter said Kay left* at the appropriate stage of development enables children to learn an operation like (12a). Children in London and New York have such experiences and children in Amsterdam and Toulouse do not; therefore Dutch and French children do not acquire an operation comparable to (12a), because they have no relevant triggering experience.

- 12.a. delete *that*
- b. empty V
- c. reduce *is* to *'s*
- d. ellipse VP

Similarly hearing *Jay saw Ray and Jim Kim* would trigger (12b), *Kim's happy* triggers (12c) and *Mary didn't* triggers (12d). Children experience nothing to indicate the limits to these generalizations and, under the view sketched here, they LEARN nothing more elaborate that blocks non-occurring forms. Rather, the interaction of the learned (12) with the general, pre-existing (6) yields the right distinctions.

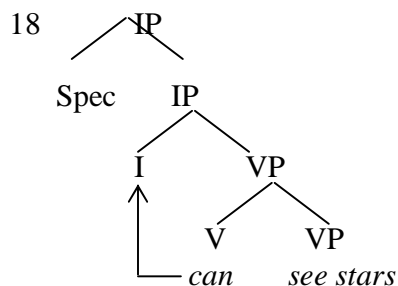
### **Language change**

So a person's internal language capacity is a complex system that depends on an interaction between learned operations and principles that are expressed by the genetic material. It grows in children in response to the external language that they encounter and becomes part of their biology. If language growth in young children is viewed in this way, then we can understand language change over generations of speakers differently, in terms of the dynamics of these complex systems. In particular we can understand how languages shift in bursts, in a kind of punctuated equilibrium.<sup>4</sup> Consider two structural shifts that English has undergone.

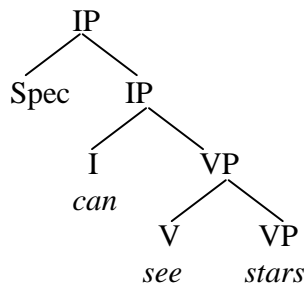
Modern English has forms like (13-17a) but not (13-17b).

- 13.a. He has understood chapter 4
  - b. \*He has could understand chapter 4
- 14.a. Understanding chapter 4, ...
  - b. \*Canning understand chapter 4, ...
- 15.a. He wanted to understand
  - b. \*He wanted to can understand
- 16.a. He will try to understand
  - b. \*He will can understand
- 17.a. He understands music
  - b. \*He can music

However, earlier forms of English had the b forms, which were used by speakers up to Sir Thomas More in the early sixteenth century. More used all the forms of (13-17) and the b forms occur in nobody's writing after him. There is good reason to believe that there was a single change in people's internal systems such that words like *can*, *could*, *must*, *may*, *might*, *will*, *would*, *shall*, *should* and *do* were once categorized as more or less normal verbs but then they were recategorized as Inflectional elements in all known grammars of English speakers after the time of More. Before More, verbs like *can* moved to a higher Inflection position, as in (18), and after More they were generated directly as Inflectional elements and occurred in structures like (19), a single shift in the system, which was manifested by the simultaneous loss of the phenomena in (13-17b). The singularity of the change accounts for the parallelism in the loss of phenomena. Syntacticians have shown that sentences like (13-17b) are not compatible with a system with structures like (19).



19



And we know why this change happened. Early English had very complex morphological properties. For example, we find *fremme*, *fremst*, *fremþ*, *fremmaþ* in the present tense and *fremed*, *fremedest*, *fremede*, *fremedon* in the past tense of ‘do;’ *sēo*, *siehist*, *siehp*, *sēop* in the present tense for ‘see;’ *rīde*, *rītst*, *rīt*, *rīdaþ* for the present tense of ‘ride,’ and *rād*, *ride*, *rād* and *ridon* for the past tense. There was a massive loss of verbal morphology in Middle English, beginning in the north of England and due to intimate contact with Scandinavian speakers. Again I skip interesting details but external language changed in such a way that the modern modal auxiliaries like *can*, *shall*, etc came to be morphologically distinct from other verbs, because as the members of the small preterite-present class, they lacked the one surviving feature of present tense verb morphology, the *-s* ending of the third person singular. The evidence indicates that they were recategorized in people’s internal systems, as we have discussed. So we see domino effects: changes in what children heard, the newly reduced verb morphology, led to a different categorization of certain verbs, which yielded systems that were compatible with (13-17a) but not (13-17b).

A later major change was that English lost (20-22a). Such forms occurred frequently in texts up through the seventeenth century, although diminishing over a long period in favor of the *do* forms of (20-22b).<sup>5</sup>

20.a. \*Understands Kim chapter 4?

b. Does Kim understand chapter 4?

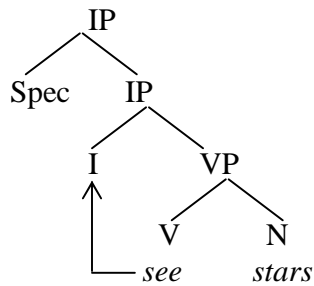
21.a. \*Kim understands not chapter 4

b. Kim does not understand chapter 4

22.a. \*Kim reads always the newspapers

b. Kim always reads newspapers

Again we can understand the parallelism of the three changes in terms of a single change in the abstract system, namely the loss of the operation moving verbs to a higher Inflection position (23). This is another change peculiar to English and not affecting any of the other European languages, whose systems have all retained the verb movement operation. In present-day English verbs do not move to the higher position and therefore cannot move to clause-initial position (20a), to the left of a negative (21a), or to the left of an adverb (22a). The equivalent movements continue to occur in French, Dutch and German systems.



This shift was due to two prior changes and we see another domino effect. The first was the recategorization of modal verbs that we just discussed, and the second was the emergence, first in the Westcountry, of “periphrastic” *do* forms as an alternative option for expressing past tense: *John did leave*, *John did not leave*, etc., instead of *John left* and *John left not*. As a result of these changes affecting what children heard in external language, the Inflection position was occupied by modal auxiliaries and by *do* in internal systems and was not available as a target for verb movement in those instances. Thus, lexical verbs did not occur in that position as often as before the days of periphrastic *do* and before modal auxiliaries were no longer verbs, and as a result, the  $I[V]$  structure fell below the threshold that had permitted its acquisition by children. Again this is too brief an account but I hope to have made clear that two prior changes in external language had the effect of reducing enormously children’s evidence for the  $I[V]$  structure, triggering a new internal system, and that three simultaneous but apparently unrelated changes were a function of a single change in the abstract system.<sup>6</sup>

### Conclusion

Amorphous external language and internal *systems* are different in kind and the modern distinction between external and internal language is crucial (Lightfoot 2006a). In addition, there is interaction between nature and nurture, and also within the system, syntax interacting with the articulatory and semantic systems, providing pairings between sound and meaning over an infinite range.

We have seen that the language capacity does not consist just of a set of words but is a complex adaptive system. Children are exposed to speech and their biological endowment, a kind of toolbox, enables them to interact with their external linguistic experience, thereby growing a private, internal system that defines their linguistic capacity.

Since the systems are complex and adaptive, they involve particular abstractions, categories and operations and these, not the behaviors themselves, constitute the real points of variation and change. Phenomena do not change in isolation but they cluster, depending on the abstract categories involved. As a result, change is bumpy and takes place in a kind of punctuated equilibrium. We understand the bumps, the clusters of changes, in terms of changes in the abstract system, as we illustrated in the two structural shifts in the history of English. If we get the abstractions right, we understand why phenomena cluster in the way they do.

Everybody’s experience varies and people’s internal systems may vary, but not linearly. Also they change over time, and sometimes variation in experience crosses thresholds

and triggers the development of a different internal system. Children are sensitive to variation in initial conditions, in the terminology of chaos theory. In general, we understand change in internal systems through the acquisition process, by virtue of children being exposed to different experiences. We explain changes where we can identify changes in the external language that children are exposed to such that the new experiences trigger different internal systems with different categories and operations. For example, after the comprehensive morphological changes of Middle English, young children had different experiences that led them to categorize words like *may* and *must* differently from verbs like *run* and *talk*. Assigning these words to a different category, Inflection, enables us to understand why (13-17b) all disappeared in parallel.

Under this approach, change is contingent, dependent on particular circumstances, and we are not surprised to see English undergoing two changes that other European languages have not undergone. English had peculiar morphological properties that were affected in peculiar ways by contact with Scandinavian speakers and that led to the new categorization. Other European languages were not affected in that way. If change is contingent like this, then there is no general direction to change and there is no reason to believe that languages all tend to become simpler or more efficient or less anything. There are no general principles of history of the kind that nineteenth-century thinkers sought and explanations are local.

Also, if change is sensitive to variation in initial conditions, then we can understand why the nineteenth-century enterprise of reconstructing prehistoric proto-languages met major obstacles, particularly in the structural, systematic aspects of language.

Consequently we can achieve better explanations for linguistic change than were possible in the nineteenth century and we assimilate the study of language change into study of the dynamics of complex systems in other domains, such as changes in species, physical environment, social organizations, economic systems, etc. We expect to find phase transitions, “catastrophes,” where many phenomena change in parallel: that happens when external language changes to reach a tipping point where it triggers a different internal system in young children. And we expect to find “emergent phenomena,” new things emerging that are not determined directly by the initial conditions of language acquisition but follow from the properties of the complex, abstract system being acquired.

The nineteenth century may have been a “century of history” across several scientific domains, but it did not yield satisfactory explanations for language change. Cross-disciplinary work on complex adaptive systems is leading to better explanations and the twenty-first century may prove to be a more successful century of history, drawing linguists, evolutionary biologists, economists, chemists, political scientists, environmentalists and many other scientists together more effectively even than in the remarkable convergence of the late nineteenth century.

## Notes

\* With thanks for comments on an earlier version from Betty Tuller, Tineke Scholten and Sharon Klein.

<sup>1</sup> There is no better account of nineteenth-century linguistics than Davies (1998) and Pedersen (1931) provides useful portraits of the major figures.

<sup>2</sup> *The Linguistic Review* devoted a whole issue to discussion of poverty-of-stimulus reasoning but the lead article restricted its discussion to defective data and excluded discussion of cases where there was an *absence* of relevant data (Pullum & Scholtz 2002: 14-17). The cases discussed here all involve absence of relevant data.

<sup>3</sup> For more details, see Lightfoot 2006b.

<sup>4</sup> Niyogi & Berwick 2009 provide a formal model of this approach.

<sup>5</sup> One might ask why this change appears to have taken place more slowly than the category change affecting the modal auxiliaries. Shakespeare and other writers alternated easily between the coexisting old and new systems. Perhaps it is easier for systems to coexist when they differ in terms of movement operations than it is for systems that differ in the categorization of certain words.

<sup>6</sup> For a cue-based approach to language acquisition by children, as distinct from the conventional grammar evaluation approaches, see Lightfoot 1999, 2006a.

## References

Bomhard, A. (1990) A survey of the comparative phonology of the so-called 'Nostratic' languages. In P. Baldi, ed. *Linguistic change and reconstruction methodology*. Berlin: Mouton de Gruyter.

Campbell, L. & A. Harris (2002) Syntactic reconstruction and de-mythologizing "Myths and the prehistory of grammars." *Journal of Linguistics* 38: 599-618.

Darwin, C. (1874) *The descent of man and selection in relation to sex*. (2<sup>nd</sup> ed. 1889) New York: D. Appleton, 88-92.

Davies, A. M. (1998) *Nineteenth-century linguistics*. London: Longman.

Grimm, J. (1848) *Geschichte der deutsche Sprachen*, vol. 1. Leipzig: Weidmannsche Buchhandlung.

Lightfoot, D. W. (1999) *The development of language: Acquisition, change and evolution*. Blackwell: Oxford.

\_\_\_\_ (2002a) Myths and the prehistory of grammars. *Journal of Linguistics* 38: 113-136.

\_\_\_\_ (2002b) More myths. *Journal of Linguistics* 38: 619-626.

\_\_\_\_ (2006a) *How new languages emerge*. Cambridge: Cambridge University Press.

\_\_\_\_ (2006b) Minimizing government: Deletion as cliticization. *The Linguistic Review* 23: 97-126.

Niyogi, P. & R. Berwick (2009) The proper treatment of language acquisition and change in a population setting. *Proceedings of the National Academy of Sciences* ??.

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

Pedersen, H. (1931) *The discovery of language: Linguistic science in the nineteenth century*. Bloomington: Indiana University Press.

Pullum, G. & B. Scholtz (2002) Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19: 9-50.

Ringe, D. A. (1995) 'Nostratic' and the factor of chance. *Diachronica* 12.1: 55-74.

Schleicher, A. (1861-2) *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*. Weimar: Hermann Böhlau.

\_\_\_ (1863) *Die darwinische Theorie und die Sprachwissenschaft*. Weimar: Hermann Böhlau.

Sweet, H. (1900) *The history of languages*. London: Dent.