

COMPLEXITY IN LANGUAGE AND IN LAW

GEOFFREY SAMPSON
University of South Africa
tsampsgr@unisa.ac.za

ABSTRACT

Two kinds of objection are made to the claim that some languages are simpler than others. Many linguists have asserted that, as a matter of empirical observation, all languages are roughly equal in complexity, but very little evidence has been cited. A more weighty objection is that the claim is meaningless because the complexity of different languages is incommensurable. It may not be numerically quantifiable; but a comparison with the evolution of legal systems shows that that does not make claims of differential complexity meaningless.

KEYWORDS: Complexity; law; Common Law; creoles; precedent.

Complexity invariance among natural languages was for many years treated as a linguistic axiom. According to Benjamin Fortson (2010: 4), for instance, “A central finding of linguistics has been that all languages [...] are equally complex in their structure”. This axiom has recently been denied or questioned. A number of linguists have begun to suggest that languages may often differ in overall complexity (see e.g. Miestamo et al. 2008; Sampson 2008; Sampson, Gil, and Trudgill 2009).

That suggestion is met in practice by two types of objection. The purpose of this paper is to distinguish explicitly between these objections, and to argue that even the apparently weightier one is less serious than it first appears.

The two types of objection, briefly, are: (i) as a matter of observation, languages are in fact about equal in complexity; or alternatively (ii) languages are incommensurable in complexity, so that it is meaningless to assert either that their overall complexity varies or that it does not vary.

The former of these seems to be the position which corresponds to what many twentieth-century linguists saw as axiomatic (cf. the Fortson quotation

above). But it is not easy to take seriously. Fortson wrote as if there were some well-known body of “finding[s]” which establish point (i), but I am not aware of anything of that kind in the literature of linguistics (Fortson gives no reference). In practice, if one were to ask a linguist who agreed with (i) what grounds we have for believing it, he would probably quote a well-known but quite brief passage in Charles Hockett’s 1958 *Course in modern linguistics*, a book which served to introduce the subject to my generation of students. Hockett wrote (1958: 180–181) that the American Indian language Fox has a more complex morphology than English, therefore it “ought to have” a simpler syntax, and indeed (Hockett told us) it has.

This idea that complexity-differences in various areas of language structure balance one another was far from new when Hockett wrote. As long ago as 1899 Henry Sweet had written “If a language is very regular and simple in one department, we may expect it to be irregular and complex in another” (Sweet 1899: 68).¹ But neither writer offers us much reason to accept the idea. Hockett justified his use of the word “ought” by saying “all languages have about equally complex jobs to do”. This remark seems to assume that the job done by a natural language is in some sense definable without reference to the properties of the language, which is surely at least questionable – one might alternatively feel that a language defines the range of tasks it is capable of executing, and that many of those tasks could not be identified independently of linguistic structure. (Cf. Sampson and Babarczy 2013.) But, leaving that point aside, a one-line comparison between just two languages is clearly very thin evidence indeed on which to establish a profound axiom about relationships among all the world’s languages. It is vulnerable to any pair of languages in which morphological and syntactic complexity differences pull in the same direction rather than balancing each other. At the Poznań conference, Ilona Koutny commented that her native language, Hungarian, has more complex morphology than English and a syntax which is at least no simpler than that of English. That brief observation is already enough to counterbalance Hockett’s claim, and I believe similar remarks could be made about very many language-pairs. John McWhorter (2001) has argued that creole languages as a class possess “the world’s simplest grammars”. One can debate that, as a number of linguists have done, but McWhorter’s generalization is at least backed up with a great deal more concrete evidence than Hockett gave in his comparison of Fox and English. Whether it is ultimately true or false, point (i) is certainly not axiomatic.

¹ I am grateful to Kees Hengeveld for drawing my attention to this passage.

If the linguistic community was for many years inclined to accept an axiom of complexity invariance, this was not, I believe, because of the force of Sweet's, Hockett's, or any other linguist's arguments; it has to be explained in ideological terms. Much of the motivation for the study of synchronic linguistics sprang from the desire to show that there are no such things as "primitive languages" – as Edward Sapir put it ([1921] 1963: 22), "the delicate provision for the formal expression of all manner of relations [...] meets us rigidly perfected and systematized in every language known to us". In linguists' minds it was often a short step from the political principle that all human societies are entitled to equal respect to the scientific idea that all human languages are equal in complexity. But scientific assertions cannot be demonstrated or refuted by reference to political or ethical principles.

(Since Hockett's book, one idea has emerged within the discipline which, if correct, would provide a separate reason for believing in complexity invariance: namely the idea that most of the detailed grammatical structure of human language is determined by our biological inheritance and hence cannot vary between languages. But, although this idea has won many converts, there seems to be no real evidence supporting it and a great deal of evidence against it – Johansson 2005; Sampson 2005; Evans and Levinson 2009. I shall not consider it further here.)

Turning to objection (ii), that it is meaningless to compare languages in terms of their relative complexity: this was powerfully argued for instance by Guy Deutscher (2009). Although it seems to have been voiced much less often down the years than objection (i), I see it as in principle a more weighty objection than (i). To claim that languages may vary in complexity seems to imply that we have some kind of unit or metric in terms of which we can assign a numerical value to the overall complexity of a language, so that we might be able to say "language *A* is 1.07 times (or 5.8 times, or ...) as complex as language *B*". Few linguists have claimed to be able to define any such metric even in crude, provisional terms.

I certainly have not done that, and shall not do so here. For the sake of argument I am happy to concede that no such metric might in principle be available. (There are linguists who have been trying to develop metrics of that sort; I do not know whether their aim is a reasonable one, but for present purposes I am willing to accept that it might not be.) Nevertheless, it does not follow that natural languages cannot be said to differ in overall complexity.

In order to show that this does not follow, I shall draw an analogy between language and another aspect of human culture with which language has much in

common (though one which is not often discussed by linguists), namely law. The evolution of law offers a clear example of a process which yields a system that is indisputably much more complex at later periods than it was when it was young, but where nobody claims to be able to put figures on that difference and it is quite unlikely than anyone ever could.

As is well known, the legal systems of various Western countries fall into two families: many Continental European nations have so-called Civil Law systems, which trace their ancestry to Roman law, whereas England, the USA, and many other English-speaking countries have Common Law systems, which have evolved ultimately out of the customs of the Germanic tribes who immigrated into southern Britain over the centuries immediately preceding and after the collapse of the Roman Empire. The highly codified nature of Civil Law makes it less relevant for our present purposes; I shall develop my analogy by reference to English Common Law (which was classically expounded by Sir William Blackstone in the 1760s – Blackstone 1768–1769; for an up-to-date layman’s introduction, see e.g. Elliott and Quinn 2008).

The “Common Law” as such was established over the period between the Norman Conquest of 1066 and about 1250, during which time the sometimes divergent legal customs of the tribes who inhabited different parts of England were merged into a single nationwide “Common” system, and the fundamental principle of *stare decisis* (follow precedents) was adopted. From the thirteenth century onwards, the body of English law grew through the accumulation of precedents set in individual cases. Human life is so complex that no finite set of rules will ever explicitly cover all the issues which fall to be resolved; when a judge confronts an open issue, he makes a decision that harmonizes as well as possible with the existing body of law, and that precedent then becomes a fixed rule binding on future cases. (This considerably simplifies the reality, of course; for instance, higher courts are entitled to overrule precedents set by lower courts. But my purpose here is to describe the general complexion of the system rather than to go into detail.)

Judges are not supposed to make new law on their own initiative. The traditional theory was that judges respond to novel questions by “discovering” rules which were somehow implicit in the existing body of law but had not previously been spelled out. However, this was a fiction. In reality, precedents might often have been settled differently, but, whichever way they were settled, those decisions then became part of English law.

From the thirteenth to the nineteenth century this was close to a complete description of the process of English legal development. The picture has

changed over the past hundred years or so, because of an explosion of statute law. Parliament can enact a statute on any topic it pleases, which will override any common-law rules with which it conflicts; by the present day probably most issues that arise are governed by some statute (though, since no form of words is ever unambiguous, the mechanism of precedent-setting comes into play again in determining what a given statute means in practice). But the large current role of statute law is a new thing (and the impact of the rather different type of law emanating from the European Union is of course much newer still). Until the late nineteenth century, Acts of Parliament were relatively few, and tended to relate to specialized purposes that did not affect the population as a whole. (In the eighteenth century, for instance, divorces were individual Acts of Parliament.) As late as about 1910, a lawyer was able to write:

Parliament is continually pouring forth a stream of legislation: but that legislation rarely touches the law as it affects people in their ordinary everyday life. I venture to say deliberately that the law relating to business and trade, the law as it affects the household, the family man, the shopkeeper and the merchant, the employer and the workman, the mistress and her maid, has not changed appreciably since Queen Anne died [that is, over the preceding two centuries].

(Anon. 1910?: 2; emphases in original removed)²

Another hundred years later, the same could no longer be said. To keep things simple, let us consider just the 600-year development of the legal system from the thirteenth to the nineteenth century.

My point is that, on the one hand, English law at the end of that period was a far more complex system than at the beginning. Anyone knowledgeable about legal history would, I believe, see that as too obviously true to be worth discussing. The statement just quoted that the law had not “changed” meant that the principles round which precedents accumulated had not been altered by statute, but that accumulation certainly made the overall body of law more complicated. (The paragraph following the one quoted referred to the law having been “built up by succeeding generations of judges”.) The standard cliché, originally coined by the poet Tennyson (1842), describes how the law “broadens slowly down / From precedent to precedent”, where we are entitled to take “broaden” as

² The publication quoted is undated, but internal evidence shows that it was published not earlier than 1908 and not later than 1913. Lawyers’ professional etiquette of the period required it to be published anonymously; on the title page it is attributed to “A Barrister-at-Law”.

amounting to “grow more complex”.³ But on the other hand, there is no way of putting figures on that complexity difference.

One might imagine that, if nobody has attempted to define a metric in this area, that is merely because doing so is not something that would interest legal scholars. But there is more to it than that: quantification is not possible. One could not merely count precedents, for instance. Some precedents are simple, others are individually complex; but also, most legal decisions do not constitute precedents (established law is adequate to resolve the cases), and there is no algorithm for deciding which particular cases extend the law and hence count as precedents. (One might say that a case counts as a precedent if it is reported in one of the standard series of law reports, but it takes human judgement to decide whether a case merits reporting.) Furthermore, new precedents do not invariably add complexity to the existing body of law; on occasion they simplify it, for instance by treating a series of previous precedents as special cases of a more general rule, or by deciding that two apparently-relevant precedents cannot be reconciled with one another in the circumstances of a new case and must be replaced by a more straightforward rule. (For how precedent-based legal evolution works in practice, see Manchester and Salter 2000.)

I believe that knowledgeable people would not see it as controversial to maintain both that English law in the nineteenth century was (much) more complex than English law in the thirteenth century, and also that this difference is not open to formal quantification. Thus, it can make sense to say that alternative systems in some area of human culture may vary in complexity although their respective degrees of complexity are incommensurable. In other words, objection (ii) is not unanswerable.

The existence of a clear but unquantifiable difference in level of complexity between two legal systems which are in a “mother–daughter” relationship obviously does not entail the existence of such differences between pairs of languages which may be related as “cousins” or not related at all; but it does establish that asserting the existence of such differences is not simply meaningless.

Some might accept that this is true in principle, but yet argue that languages are so different in kind from legal systems that objection (ii) remains valid in practice. Much twentieth-century linguistics was wedded to an assumption that human languages are definable in terms of clearcut sets of formal grammatical

3 Tennyson’s poem actually says that *Freedom* “broadens slowly down [...]”, but this is poetic compression. “Precedent” is inherently a legal rather than political concept, and Tennyson’s lines are regularly understood as meaning that English freedom is underpinned by a system of laws which grow increasingly complex through accumulation of precedents.

rules, akin to the Backus–Naur notation which defines the well-formed code sequences of a computer programming language.⁴ If so, this would make languages very different kinds of animal from legal systems. However, that assumption looks very threadbare, now that it is widely recognized after decades of effort that “No-one has ever successfully produced a comprehensive and accurate grammar of any language” (Graddol 2004).

Human languages are not formally-definable systems (Sampson and Babarczy 2012, 2013). Languages, like the Common Law, are systems which evolve through usage rather than in accordance with clearcut plans. In the writings just cited, I have explained the patterns of grammatical frequency that emerge from analysis of real-life usage via an analogy between the grammar of a human language and the range of paths available in a prairie inhabited by a society lacking developed systems of property rights. There will be a few broad, well-beaten tracks (corresponding to the most basic sentence structures, such as, in English, Subject–Verb–Object), and many other lesser routes, ranging smoothly down to barely-detectable disturbances of the grass where only a handful of people have passed, corresponding to grammatical patterns which have been used less frequently or scarcely ever. It makes no more sense to partition the set of possible sentence patterns into a group of “grammatical” sheep and a complementary group of “ungrammatical” goats than it would to offer a comprehensive routemap of the prairie. The most one could aspire to is to map the more heavily-used tracks (in the prairie case) or the more usual grammatical patterns (in the case of a language), down to some cut-off which will be essentially arbitrary in either case, without pretending that these are “all and only” the available paths or grammatical patterns. (Before the rise of generative linguistics in the late twentieth century, that was all that grammarians did claim to do.) Describing the current state of a human language is a task much more like specifying the current state of some area of Common Law, which is generally acknowledged as a task that cannot be done perfectly because the unavoidable vagueness and messiness of law will always escape precise definition, than like specifying what counts as well-formed code in a programming language (which is routinely done by a compiler – if code compiles, it is well-formed, and if not, not).

In sum: of the two standard objections to the idea that human languages vary in overall complexity, objection (i) has never been seriously justified and seems intuitively implausible, while objection (ii) may look plausible at first

⁴ For Backus–Naur notation see <foldoc.org/Backus-Naur>, downloaded 14 Nov 2013, or the Wikipedia article “Backus–Naur Form”.

sight but embodies a fallacy. In reality, there is nothing illogical in the assertion that languages differ in complexity although the differences are unquantifiable; this assertion, I believe, is true. In this respect, languages resemble many other components of human cultures.

REFERENCES

- Anon. 1910? *Cassell's lawyer and business man's legal handbook: a popular exposition of the civil, commercial, and ecclesiastical law of Great Britain, with such criminal law as affects the trader*. (3 vols.). London: Cassell.
- Blackstone, W. 1768–1769. *Commentaries on the laws of England*. (4 vols., 3rd ed.) Oxford: Clarendon Press.
- Deutscher, G. 2009. “‘Overall complexity’: a wild goose chase?” In: Sampson, G.R. et al. (eds.). 243–251.
- Elliott, C. and F. Quinn. 2008. *English legal system*. (9th ed.) Harlow, Essex: Pearson Longman.
- Evans, N., and S.C. Levinson. 2009. “The myth of language universals: language diversity and its importance for cognitive science”. *Behavioral and Brain Sciences* 32. 429–492.
- Fortson, B.W. 2010. *Indo-European language and culture: an introduction*. (2nd ed.) Oxford: Wiley-Blackwell.
- Graddol, D. 2004. “The future of language”. *Science* 303. 1329–1331.
- Hockett, C.F. 1958. *A course in modern linguistics*. New York: Macmillan.
- Johansson, S. 2005. *Origins of language: constraints on hypotheses*. Amsterdam: John Benjamins.
- McWhorter, J. 2001. “The world’s simplest grammars are creole grammars”. *Linguistic Typology* 6.125–66. [Reprinted in McWhorter, *Defining creole*, Oxford: Oxford University Press, 2005.]
- Manchester, C. and D. Salter. 2000. *Exploring the law: the dynamics of precedent and statutory interpretation*. (3rd ed.) London: Sweet & Maxwell.
- Miestamo, M., K. Sinnemäki and F. Karlsson (eds.). 2008. *Language complexity: typology, contact, change*. Amsterdam: John Benjamins.
- Sampson, G.R. 2005. *The “language instinct” debate*. (2nd ed.) London and New York: Continuum.
- Sampson, G.R. 2008. “A linguistic axiom challenged”. <www.grsampson.net/ALac.html>, uploaded 30 Mar 2008.
- Sampson, G.R. and A. Babarczy. 2012. “Introduction to Sampson and Babarczy 2013”. <www.grsampson.net/BGwgCh1.htm>, uploaded 18 Dec 2012.
- Sampson, G.R. and A. Babarczy. 2013. *Grammar without grammaticality*. Berlin: De Gruyter.
- Sampson, G.R., D. Gil and P. Trudgill (eds.). 2009. *Language complexity as an evolving variable*. Oxford: Oxford University Press.

- Sapir, E. 1921. *Language: an introduction to the study of speech*. Reprinted by Hart-Davis, London, 1963.
- Sweet, H. 1899. *The practical study of languages: a guide for teachers and learners*. London: Dent.
- Tennyson, A. 1842. "You ask me why, tho' ill at ease". In: Tennyson, A., *Poems* (vol. i). London: Edward Moxon. 116.

Address correspondence to:

Geoffrey Sampson
University of South Africa
P O Box 392
Unisa 0003
South Africa
tsampmgr@unisa.ac.za