Linguistic differences and language design

Mark C. Baker

Department of Linguistics, Rutgers University, 18 Seminary Place, New Brunswick, NJ 08901, USA

A small number of discrete choices ('parameters') embedded within a system of otherwise universal principles create the extensive superficial differences between unrelated languages like English, Japanese, and Mohawk. Most current thinking about the evolution of language ignores or denies the existence of these parameters because it can see no rationale for them. That the human language faculty is organized in this way makes more sense if language is compared to a cipher or code. As such, it would have a purpose of concealing information from some at the same time as it communicates information to others.

Much linguistic research in the last 15 years has shown that the extensive superficial differences among languages can be attributed to a small number of differences (called parameters) in the otherwise universal principles that define them. Meanwhile, most theorizing about the evolution of language has ignored or denied the existence of these differences. This is an unfortunate disconnection between the two domains of inquiry. I claim that it is the side-effect of an unchallenged commitment to the idea that language developed exclusively as a medium for communication. This nearly universal presumption makes the existence of parameters seem more problematic than it needs to be. Instead, I argue that languages are rather like codes, which exist to conceal information as well as to communicate it.

Language as a finite system of discrete differences

The differences among historically unrelated languages have a peculiar quality. On the one hand, languages are not completely different. Not just any representation system that bears information can serve as a human language. Rather, all known languages obey abstract principles that are not logically necessary. Indeed, it is easy to find substantive similarities across languages, once one knows where to look. These similarities usually go unnoticed in descriptive grammars because they are so natural to us that we easily overlook them, even though artificial languages do not share those properties.

On the other hand, languages are not just slightly different. One might imagine that learning a new language would only be a matter of learning new words, but that is not the case. There are also important differences in sound patterns, in word orders, in how sentences can be transformed, and in how meanings are expressed. No aspect of language is immune to variation.

The curious situation, then, is that languages vary only slightly in the general principles that shape them, but they vary greatly in the actual sentence structures formed [1,2] (see Boxes 1 and 2). Most of the structure of language is contingent but universal; hence it is potentially innate to the human mind. This invariant machinery includes part-of-speech distinctions, notions of phrase and clause, the apparatus of subject, object, and indirect object, and much more. But the human language capacity is not entirely invariant. It allows for a few basic choices – known as parameters – within the universal machinery. These choices are a small subset of the options that could do the job in an information-theoretic sense, but they are more than a single person needs to get through life. Only a relatively small number of these parameters are required; current work suggests that the major syntactic distinctions among the world's 6000 or so languages can be explained with 10 to 20 parameters [2] (Fig. 1).

Most curiously of all, many of these parameters seem perfectly placed within the overall system so as to have the maximum impact on the superficial appearance of the language, without affecting its basic logical relations (see Boxes 1 and 2).

This situation is of inherent interest to cognitive science. In domains such as perceptual and motor systems, one assumes that the cognitive processes at work are essentially invariant across the human species. In other domains, such as social and cultural cognition, it has been standard to assume that the cognitive system can vary across populations in arbitrary ways. Language illustrates something in between: striking superficial diversity that can be reduced to a small number of discrete factors placed within a universal system. This cognitive architecture should be kept in mind as a possible model when studying other aspects of human culture.

Implications for the origins of language

This picture of linguistic variation poses an interesting puzzle for theorizing about language origins. After years of reticence, discussion of the evolution of language has exploded in recent years. One striking feature of this literature is how little it has to say about cross-linguistic variation. Most authors are completely silent on this point; they write as though only one language had ever existed [14–20]. This might make sense if linguistic variation...
were a minor phenomenon, or if it had no interesting structure. But the opposite is true.

The most obvious way to give an evolutionary account of our capacity to speak different languages is a non-starter. The cognitive systems of some animals leave parameters open for learning so that those animals can function in different environments. Bees, for example, can live at different latitudes because their solar navigation system leaves open a parameter for the angle of the sun [21]. But this is not a good model for understanding linguistic parameters, despite some suggestions along these lines [22–24]. There is nothing that makes Japanese word order more effective on Pacific islands and English word order better on Atlantic islands. Indeed, there is no ecological regularity in how the major linguistic types are distributed around the world [2,11]. In the absence of fitness advantages to speaking one language rather than another, mathematical studies show that having a parameterized language faculty is disfavored, because it makes language learning less reliable [23,24].

Another line of analysis is to say that our capacity for linguistic variation is not part of the adaptive design of the
human language faculty, but rather a kind of evolutionary accident, as suggested by Pinker and others [25–27]. The claim is that evolution has fixed many properties of human language, but not all of them. Once the innate endowment became rich enough to make learning a complex language reasonably reliable, there was little selectional pressure to keep elaborating it until all grammatical properties were determined. ‘Parameters’ are those grammatical points that are still left open by the innate endowment.

Although there might be no direct refutation of this view, I find it unsatisfying, given the quality of the linguistic differences we know about. The known parameters do not seem like minor quirks at the edges of language that were not worth worrying about. Rather, they seem like gadgets put close to the heart of language, exactly where they will have maximum impact on surface orders without changing underlying logical relationships. Given this, we should consider the possibility that parameters are a design feature of the language faculty after all.

Languages as codes

Most researchers take it for granted that the evolutionary purpose of language is to provide a way of communicating complex propositional information to kin and collaborators. From this point of view, the possibility of radically different languages is perplexing. But this is a theoretical assumption, not an established empirical result.

Suppose that the language faculty has a concealing function as well as a revealing function. Our language faculty could have the purpose of communicating complex propositional information to members of our group while concealing it from members of other groups. Many products of human engineering have been designed to fulfill exactly this purpose – namely codes and ciphers. Ciphers have a common structure: they consist of a publicly-known encryption algorithm that defines a family of ciphers and a secret ‘key’ that defines a particular cipher [28]. The innate endowment for language has a similar structure: the universal principles correspond to the general encryption algorithm, and the particular parameter settings correspond to the key. There are also intriguing correspondences of detail between particular ciphers and the specific ways languages differ from one another (see Box 3). This shift of perspective makes sense of the fact that many parameters seem perfectly designed to give languages very different ‘looks’ without affecting the fundamental content expressed. That is the nature of a good cipher.

We know that natural languages can make excellent codes; this was proven by the Navajo Code Talkers of World War II, whose native language was never cracked by cryptographers [29,30]. Could natural languages have functioned in a similar way in traditional societies, before radios and international transportation? I do not have the anthropological expertise to answer this question fully, but some broad outlines are clear. People throughout the world have lived in a state of more or less perpetual warfare, and where possible they prefer to fight against people from a different linguistic group. Furthermore, successful warfare in ‘primitive’ societies depended heavily on group coordination combined with secrecy and treachery [31]. Thus, the boundary conditions under which one can imagine an encoding function for human language being useful were present.
Box 3. Comparing language differences to 16th century cryptography

Claims that a biological system has a particular function are often reinforced by comparing it to products of human engineering that have that function. When the two have detailed structural similarities, the claim that they have similar functions gains support, a classic example being the comparison between the vertebrate eye and a camera. The ways human languages differ can be compared to cryptographic techniques of the 16th century (see Table I). Sixteenth century cryptographers used a variety of techniques: they both replaced and rearranged symbols in systematic ways, and they performed these transformations both at the level of letters and at the level of words and phrases [26]. This layered complexity evolved over time with the explicit purpose of defeating particular code-breaking strategies (such as frequency analysis). Natural languages also differ from one another in ways that show layered complexity, using substitution and rearrangement at multiple levels. Many of the specific tricks of early cryptographers have striking analogies in natural language. This gives credence to the notion that natural languages have the same concealing function as man-made ciphers.

Table I. Similarities between cryptographic methods and natural languages

<table>
<thead>
<tr>
<th>Cryptographic technique</th>
<th>Natural language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steganography (hiding the message)</td>
<td>Use of phonetic distinctions that are imperceptible to speakers of other languages</td>
</tr>
<tr>
<td>Letter for letter substitution ciphers</td>
<td>Use of different sound systems</td>
</tr>
<tr>
<td>Code expressions that stand for entire words and phrases</td>
<td>Saussurean arbitrariness (use of arbitrary sounds to express a concept)</td>
</tr>
<tr>
<td>Use of homophones (different code symbols for the same letter)</td>
<td>Allophonic variation (different pronunciations with no associated meaning difference)</td>
</tr>
<tr>
<td>Use of nulls (symbols added that do not stand for any letter)</td>
<td>Use of functional particles with no truth-conditional meaning</td>
</tr>
<tr>
<td>Removal of spaces between words</td>
<td>Phonological joining of words by rules of ‘external sandhi’</td>
</tr>
<tr>
<td>Transposition ciphers (systematic scrambling of words and letters)</td>
<td>Word order parameters (see Boxes 1 and 2)</td>
</tr>
</tbody>
</table>

Fig. 1. The parameter hierarchy. This is a systematic representation of some major parameters that distinguish languages (shown in violet). Each branch point is labeled with the name of one or more parameters; below each point are lower-level parameters that depend on them. The Polysynthesis parameter at the top of the hierarchy is the parameter that distinguishes Mohawk from English in Box 2. The Head Directionality parameter one step down is the parameter that distinguishes Japanese from English. At the bottom of the diagram are some historically unrelated languages that are syntactically similar as a result of having the same settings for these parameters.
Some researchers dismiss the idea that linguistic differences have a function out of hand. But we should also be wary of evolutionary psychological reasoning if it leads us to downplay what is, because of unprovable assumptions about what must have been. The use of language in early human societies is not observable, and yet the reality of linguistic variation of a particular kind is, so we should focus on that. There are, after all, plenty of unsolved questions about human cognition: intentionality, free will, a priori knowledge, abductive reasoning, and so on [32,33]. Many of these are clearly related to language [2]. We therefore need to remain open to new insights from outside the familiar range of assumptions.

**Conclusion**

The extensive differences among human languages can be attributed to a small number of discrete differences in the underlying rule system, known as parameters. In this respect, languages are similar to artificial codes, suggesting that language developed to prevent communication as well as to permit communication. Further research is needed to complete the parametric account of how languages differ, and to develop theories of the origins of language that explain not only the existence of human languages, but also how they can differ so systematically. In particular, we need to explore alternatives to the simplistic assumption that language exists only for communication.

**References**