

Computational Limits on Natural Language Suppletion

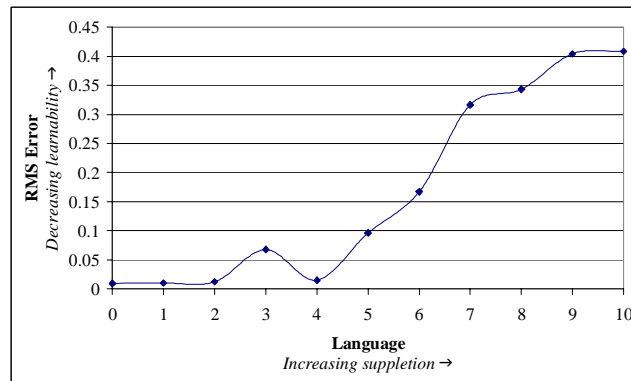
While most natural languages tend to contain suppletive pairs, suppletion is vastly overshadowed in all languages by regular form-to-form mappings. What enforces the cross-linguistically low level of suppletion? This work makes the intuitive argument that suppletive mappings are kept to a minimum for a very simple reason: they are harder to learn. In what follows, this point is illustrated through an examination of suppletive and regular (uniform) verbal paradigms.

Most contemporary theories of morphology offer no way to constrain the amount of suppletion that occurs in a language. In inferential-realizational theories (Stump, 2001) for instance, a verbal paradigm can be realized using either rules that enforce uniformity, or rules that allow suppletion, as in the following examples from Spanish:

		<i>hablar</i> 'to speak'				<i>ser</i> 'to be'	
		Sg	Pl			Sg	Pl
1		SPEAK<1Sg>habl → habl-o	SPEAK<1Pl>habl → habl-amos	1		BE<1Sg> → soy	BE<1Pl> → somos
2		SPEAK<2Sg>habl → habl-as	SPEAK<2Pl>habl → habl-áis	2		BE<2Sg> → eres	BE<2Pl> → sois
3		SPEAK<3Sg>habl → habl-a	SPEAK<3Pl>habl → habl-an	3		BE<3Sg> → es	BE<3Pl> → son

Repetition of the root *habl-* in each cell of the paradigm for *hablar* gives rise to its uniform nature. In contrast, no identifiable root exists to anchor the forms that make up *ser*'s paradigm. As a result, the relationship between any two members of the paradigm is suppletive. The problem here is that theories that make use of these kinds of rules offer no reason to favor the class of rules that realizes a paradigm uniformly over the class that realizes a paradigm suppletively. This lack of constraint erroneously predicts that a language could contain more suppletive than uniform paradigms, or even be composed solely of suppletive paradigms.

The fact that the grammar does not provide a way to limit suppletion is not problematic, however, if we adopt the position that grammars are embedded within a biological system that has limited computational resources. In order to demonstrate the validity of such an approach, I devised a set of 11 "languages," each containing a different number of suppletive verbal paradigms, ranging from no suppletion, to a language in which all paradigms are suppletive. These languages were then presented to a neural network designed with a standard feedforward architecture, and running the backpropagation-of-error learning algorithm (Rumelhart & McClelland, 1986). The results show that, as the number of suppletive paradigms the network is asked to master increases, learnability decreases:



Further, there is an upward limit, or threshold, on the number of suppletive paradigms that can be learned without significantly affecting network performance. In effect, the model predicts that suppletion in natural language will be tolerated, so long as it is kept to a minimum.

Although this work focuses on the way in which performance limitations can supplement inferential-realizational theories of morphology to provide constraints on suppletion, it can be applied to other morphological theories as well, most of which (if not all) also fail to put limits on whatever mechanism they use to account for suppletion.