בס"ד

Sequence to Sequence Rewriting Rules: Unification of Morphophonology and Morphosyntax

Sasson Margaliot
Jerusalem, Israel

DRAFT, February 2017

Abstract This article argues for the novel understanding of minimalist Merge as a special case of extended Rewriting Rules which allow more than one elementary symbol on the left side of the arrow. This makes it possible to extend the use of Rewriting Rules to Distributed Morphology and, possibly, even to phonology.

The idea behind the minimalist approach is that a single "rule" should be able to accomplish as much as possible. An illustration of this is Merge, which indeed accounts for a large number of facts. It was originally introduced to describe syntax, but then the framework of the "single generative engine" was introduced, with its syntax-all-the-way-down, extending the domain of Merge to also cover the structure of words. The same procedure of Merge that combines words into sentences also builds the words from morphemes.

But this article argues for the application of Merge to phonology as well. Traditionally, syntax and phonology are described in completely different terms, but it may simply be a result of an historical accident. After all, why would the ordered sequences be handled by two different generative engines?

In fact, the version of Merge introduced here is suited equally well to work in both phonology and syntax. I would suggest that every instance of Merge must be licensed by a Rewriting Rule. A regular Rewriting Rule consists of two parts separated by an arrow: a symbol (called *label*) to the left of the arrow, and a sequence of labels representing syntactic objects to the right of the arrow. Merge, then, is combining together a number of syntactic objects (possibly generated by previous applications of Merge), and the resulting new syntactic object is assigned the label of the Rewriting Rule it was licensed by.

For example, using outdated terminology:

(1) VP → V NP

The existence of a Rewriting Rule (1) is licensing merging syntactic objects labeled V and NP together, creating a new syntactic object labeled as VP. This is the traditional use of Rewriting Rules in regular syntax.

In order to extend the use of Rewriting Rules to cover Distributed Morphology, I suggest that more than one label be allowed on the left side of the arrow. Also, morphosyntactic features in *List A* will be considered elementary symbols alongside the regular syntactic labels. Then Vocabulary Items in *List B* are represented each by a Rewriting Rule as in (2) and (3):

(2) Root[dog] +count +animate → /d/ /o/ /g/

(3) Root[do] +past → /d/ /o/

In theoretical linguistics various processes are traditionally described in entirely different terms and notations, but many of them can easily be represented in the language of sequence-to-sequence Rewriting Rules.

For example

(4) /d/ /e/ /l/ /e/ → /d/ /u/

The notation of Rewriting Rules can be helpful for an easier introduction of theoretical linguistics for specialists in the cognitive sciences and artificial intelligence. There is no need to teach them an overwhelming multitude of various notations, when in fact they are all just special cases of sequence-to-sequence rewriting.

Since the language faculty is responsible not only for the generation of utterances (as in 'speaking'), but also for their parsing (as in 'listening'), the universal notation should be equally appropriate for both generation and parsing. This is exactly the case with the sequence-to-sequence Rewriting Rules, since changing the direction of all arrows in Generation Grammar produces another set of Rewriting Rule feasible as a grammar for parsing.