Distributional Properties of Affixes in English

RIA MUKHERJEE

Abstract. Language users have a remarkable ability to create, produce, and comprehend complex words. Words such as 'undercut' and 'bakery' appear to be composed of units, traditionally called morphemes, that recombine in rule-like ways to form other words, such as 'underline' and 'cannery'. However, morphological systems are quasi-regular; they are systematic and productive but admit many seemingly irregular forms. Thus, 'bakery' is related to bake and 'cannery' to can but what is the groce in 'grocery'? There is no liver in 'deliver', corn in 'corner'. Such words exhibit partial regularities concerning the correspondences between form and meaning, the treatment of which has important implications for linguistic and psycholinguistic theories. This article describes an approach to morphological phenomenon called the convergence theory, in which morphology is a graded, inter-level representation that reflects correlations among orthography, phonology, and semantics. Our goal in this opinion piece is to articulate an approach to think about complex words rather than exhaustively review the literature or propose a specific model and so some disclaimers are in order. The article emphasises the phenomena concerning derivational morphology in English; issues concerning other types of morphology and cross-linguistic variation are discussed only briefly.

Keywords : Distributional properties; convergence theory; connectionist theory; word recognition; cross-linguistic variation; lexical syntax.

Derivational morphology is the aspect of language concerned with the structure and formation of words. Language users' ability to create and understand new words also implicates this level of lexical structure. Most theories assume that complex words consist of discrete units that are recombined by rules. In this article I have briefly reviewed some of the limitations of this approach and have offered an alternative, inspired by connectionist theories of knowledge representation and learning, in which graded, non-discrete morphological structures emerge in course of learning relations among the spellings, sounds, and meanings of words. The classical treatment of word formation within structural linguistics was that words consist of primitives, called morpheme, that are minimal meaning-bearing units. Words are created by rules that combine these morphemic primitives, allowing the creation of new expressions. This approach can be called discrete morphology, because morphemes are discrete units, like beads on a string. For example, the assumption that morphemes are minimal meaning-bearing units was called into question by Aronoff who noted, "there are subword patterns that seem to function as morphemes in so far as they enter productively into word formation, but have

RIA MUKHERJEE

little or no meaning". (Aronoff 2016 : 49) In most human languages important components of linguistic structure are carried by affixes, or bound morphemes. The affixes in a language comprise a relatively small but frequently occurring set of forms that surface as parts of words, but never occur without a stem. While bound morphemes always occur as part of a larger word, they are viewed as having an independent status by virtue of the fact that they combine productively with stems and other grammatical elements in systematic and predictable ways. For example, any English verb root that is inflected with the suffix -ing and is preceded by a form of the auxiliary verb, be, results in a verb form that is marked with particular tense and aspect: present progressive (e.g. she is reading). Mastering the morphological system of a language thus involves acquiring the generalizations about the relationships between formal elements (e.g. auxiliary-be and -ing), as well as the semantic and functional properties of the language that are represented in the morphological system (e.g. mood, aspect, number, etc.). However, before learners can acquire morphological facts about their language, they must first identify the sub-lexical combinatorial units: they must identify the bound morphemes. The patterns, in which semantic and phonological properties are co-related, is quite common. There are also forms in which the meanings of the component morphemes are unrelated to the meaning of the word. For traditional theories the central question is whether such words are to be treated as morphologically complex or listed in the mental lexicon. It is widely assumed that there should be a principled basis for deciding the issue, but there has been little agreement about where to draw the line.

Word Recognition

Words could be recognised without using morphological structure at all because spelling and sound usually provide sufficient information for this purpose. Psycholinguistic research has therefore, focused on demonstrating that there is a level of morphological representation in lexical memory and that this information is used in processing. Like most of the studies in the literature, our discussion focuses on word reading, but similar issues arise in connection with other uses of language. In an influential study by Tuft and Forster, "subjects performed a lexical decision task, deciding if letter strings were words or not. The critical comparison was between pseudo-words such as Dejuvenate (analysed as containing the stem morpheme that occurs in Rejuvenate), and ones such as Depertoire (not incorporating a stem morpheme)". (Tuft and Forster 1979 : 638) Subjects found it more difficult to decide that the pseudowords containing morphological stems were not words. The findings from this and many subsequent studies suggested that recognition involves decomposing words into component morphemes. A second approach involves using frequency effects to diagnose the use of morphemes. The frequency with which a word is used affects how hard it is to process. This effect is standardly taken as evidence that frequency information is stored as part of

word-level representations in the mental lexicon. A third approach utilises the phenomenon of lexical priming, in which the processing of a word is affected by its similarity to a preceding word. They concluded that the effect was due to morphological overlap rather than formal (orthographic or phonological) overlap. This priming methodology has also been used in many subsequent studies. However, the usual strategy of designing a factorial experiment that manipulates all the relevant factors runs into difficulty when they are intrinsically highly co-related, as in morphology. Clearly the concerns about these data also bear on the validity of theoretical models developed to explain them.

The Convergence Theory

An alternative approach is provided by pursuing the parallels between derivational morphology and orthographic-phonological correspondences. As we have noted, both systems are quasi-regular; there are cases that can be described as rule-governed but also exceptions that deviate from the rules in differing degrees. Both are traditionally thought to require two mechanisms, a set of rules for the "regular" cases and a word-based mechanism to handle the exceptions. Seidenberg and McClelland's model showed how "both the rule-governed cases and exceptions could be processed within a single network employing distributed representations". (Seidenberg 1989 : 96) The theory did not embody a categorical distinction between rule-governed and rule-violating; the weights used in pronouncing all words encoded different degrees of consistency in the spelling-sound mapping. Both the rule-governed cases and exceptions could be processed within a single network employing distributed representations. The model did not embody a categorical distinction between rule-governed and rule-violating; the weights used in pronouncing all words encoded different degrees of consistency in the spelling-sound mapping. The main reason to pursue this approach is because it captures an essential characteristic of quasi-regular systems: the existence of partial regularities e.g have violates the pronunciation rules of English; however, have's the pronunciation of have is not arbitrary; it overlaps with "rule-governed" forms such, as had, has, and hive. In dual-mechanism theories, the procedure by which 'have' is pronounced is unrelated to the one used in pronouncing had, has, and hive.

The Connectionist Approach

Since in the connectionist approach, there is a single set of weights, what is learned about one word carries over to partially overlapping words. This characteristic of the model allows it to explain empirical phenomena such as consistency effects: the fact that a word such as wave (which has an irregular neighbour, have) takes longer to name than wade (whose neighbourhood is more consistent). Morphology can be construed in a similar manner. The basic idea is that it is a graded, inter-level representation that develops in the course of acquiring

RIA MUKHERJEE

lexical knowledge. The lexicon encodes information about the spellings, sounds, and meanings of words, the regularities within these codes and between codes, plus additional information derived from the contexts in which words occur. The problem of lexical learning is framed in terms of a connectionist network employing distributed representations of these types of information. The lexical network supports computations from orthography to phonology, phonology to meaning, meaning to phonology, and so on, which are utilised in performing tasks such as production and comprehension.

Current Understandings and Problems

I. Are Statistics Sufficient?

In the convergence theory, morphology is given a statistical interpretation. It corresponds to statistical regularities in the mappings among spelling, sound, and meaning. In traditional theories morphemes are defined structurally rather than statistically; people may encode facts such as how often morphemes are encountered or how similar they are to each other, but the status of a unit as a morpheme does not depend on these factors. The same holds for other grammatical structures as well. Thus the traditional view asserts that morphological structure has properties not predictable from mere statistics. It therefore, naturally suggests that if the statistical properties of words were equated, there would be residual effects due to morphological structure for each.

II. Lexical Syntax

Words consist of more than just spellings, sounds, and meanings; they also have syntactic properties that need to be addressed. Words fall into grammatical categories such as noun and verb, which derivational but not inflectional rules may change. These phenomena raise important issues about syntactic representation that clearly extend well beyond the scope of this article. Aside from acknowledging their importance, we would point out that progress might be made by embedding a lexical processing system such as the one we have described within a device that tracks the distributions of words in sentences.

III. Cross-Linguistic Differences ?

Languages vary considerably in morphological richness; English morphology is usually considered to be impoverished compared to languages such as Serbo-Croatian, Italian, or Hebrew. It therefore, cannot be assumed that what holds for English will necessarily apply to other languages. On the other hand, one of the principal claims of modern linguistic theory is that languages are similar in so far as they exhibit universal properties, with the degree of variation across

languages quite limited. Whether the convergence theory will extend to other languages is not known. At this point three observations seem relevant. First, experiments to date have not tended to yield radically different patterns of results across languages. To the contrary, what is striking is the extent to which the results have been similar. Secondly, the methodological issues that have arisen in studies of English also apply to studies in other languages. Thirdly, although there are differences between languages in terms of morphological complexity and types of morphological mechanisms, they exhibit deeper commonalities.

IV. Understanding vs. Production

A final issue is whether the same principles are at work in word 'recognition' and production. Intuitively the tasks seem intimately related; learning a word typically involves learning both how to produce and comprehend it. The meaning and grammatical properties of a word do not change as a function of whether it is spoken, written, or read. It would therefore, be odd if completely different principles underlay these different aspects of performance. Studies of production and comprehension have nonetheless tended to be pursued independently, with different models, experimental methods, and theoretical approaches. In a recent review Levelt summarised models of word production "that contrast witht he approach proposed here in several respects". (Levelt 1999 : 223) Most research on word production assumes multiple discrete levels of linguistic representation, including morphology.

Conclusion

The important point to emphasise at this early stage in the development of the convergence theory is the quasi-regular character of derivational morphology. Beyond distributional properties such as frequency and transitional probabilities, phonological factors could also influence infants' early representation of affixes. To the degree that affixes in a given language have phonotactic tendencies that infants can detect, once infants have segmented enough affixes to detect the patterns, they could use the tendencies as cues to guide further segmentation and the discovery of new affixes. This possibility raises a potential concern in this study. The main reason for pursuing the approach we have described is not because there is something intrinsically correct about multi-layer networks employing distributed representations or incorrect about other approaches. Rather, it is because the quasi-regular character of morphology seems to demand the kinds of graded representations that the convergence theory entails. This approach is a radical departure from linguistic tradition and it remains to be seen whether it can account for the mass of cross-linguistic data that have accumulated. The practical and conceptual obstacles in developing explanatory models of non-trivial scope are also considerable. The issues raised here are therefore, likely to play out for some time to come.

RIA MUKHERJEE

Ria Mukherjee
Research Scholar, Dept. of English
Ranchi University, Ranchi

Works Cited

Aronoff, M. *Word Formation in Generative Grammar*. New Delhi : MIT Press, 1976.
Levelt, W.J.M. *Models of word Production : Trends in Cognitive Sciences,* 1999.
Seidenberg, M.S. and J.L. McClelland, "A Distributed, Developmental Model of Word
        Recognition and Naming". *Psychological Review.*, 1989.
Taft, M. and K. Forester. "Lexical Storage and Retrieval of Prefixed Words". Jouranal of
        Verbal Learning and Verbal Behaviour, 1975.