

Introduction: What is a word? Formal domains

Alana Johns and Keren Rice
University of Toronto

As Marianne Mithun discusses in her book *The Languages of Native North America*, one of the most striking features of these languages to the early Europeans was the enormous complexity of the words. The languages are not of the isolating, or analytic type well-known to the early explorers and missionaries, but rather show tremendous morphological complexity. This morphological complexity has fascinated scholars over the generations, and is the theme of the 2000 Workshop on the Structure and Constituency of the Languages of the Americas held at the University of Toronto.

In her discussion, Mithun raises the important question of what a word is. She begins with the following example, from Yupik:

- (1) kaipiallulliniuk
 kaig-piar-llru-llini-u-k
 be hungry-really-past-apparently-indicative-they 2
 'The two of them were apparently really hungry.' (page 38)

How do we know this is a single word? Mithun suggests that the best criterion is usually the judgment of the native speaker. She goes on to present several additional criteria. For instance, speakers are usually able to pause between words, but seldom pause within a word. In addition, speakers are aware of meanings of words, but often not of meanings of individual morphemes. Phonological criteria might be relevant as well; for instance the word may serve as the domain of stress.

In the workshop, many issues relating to the nature of a word were addressed, and we summarize our presentation here.¹

We begin by pointing out some of the morphosyntactic issues about the nature of the word. Here the most fundamental question that occurs to us is the following – 'Do complex words differ from sentences, and, if so, how?'

A major issue that follows from this question concerns the relation between the word and the lexicon. While Mithun takes the traditional position that the word is formed in the lexicon, this has been challenged in work on many of the language families of the Americas. From the perspective of syntax, the traditional word has been argued to be syntactically complex, being a phrase or phrases rather than a word. The richness of syntactic structure allows for the existence of both the analytic structures and the morphologically complex structures as syntactic rather than lexical objects. The possibility of syntactic status also opens the door to many questions usually considered to fall under the domain of syntax, questions about issues such as configurationality, movement, and argument structure. We ask then 'What is the relation between the word and the lexicon?'

While there are many syntactic properties to these complex words, there remain ways in which they appear to differ from syntactic objects. An often noted difference involves the order of morphemes within a word. Morphemes within a morphologically complex word are generally strictly ordered. This strict ordering has lead linguists to propose templates for these languages. In recent work, researchers have sought to rely on principles to determine morpheme ordering. While the fixed ordering of the template looks quite different from syntax, where there seems to be visible rearrangement of words, the step towards finding principles that underlie the ordering of morphemes within a word is reminiscent of syntax. Thus, we can ask 'Is the fixed position of morphemes within a word a superficial property or does it represent something fundamental about the word?'

Words also differ from sentences in generally lacking properties such as iteration and intensification. For instance, in English, one can say 'it is really really hot today' or 'it is really very hot today.' Leaving

¹ We do not include here the extensive linguistic literature which addresses the questions raised in this article.

aside reduplication, this property does not seem to be present in many of the morphologically complex languages. Is this then a true way in which the word is different from the phrase?

Another way in which words and sentences appear to differ is by what holds the units together. Functional morphology is, in a sense, the glue that holds words together in sentences. However, it is phonology that is the glue within a word. Are these equivalent in some sense, or is there a fundamental difference between them?

Another question concerns how words are related to one other. A common perspective within the generative literature is that words are related to each other through derivation. However, derivational models have been challenged, and it has been argued that words stand in relationship to words, without derivation. A set of questions can be raised: 'Are words necessarily related to other words? If so, how is this relationship established? More particularly, do inflectional paradigms relate to each other through words?' Related to these questions are those about access to word-internal structure: 'Is there internal structure to words? Or are words simply the spell-out of part of the computational component of the grammar? Is it possible to predict what couldn't be a word?

Questions of lexicalization and phraseologization also are very important in considering these morphologically complex items. In terms of lexicalization, it is often the case in morphologically complex languages that words take on conventionalized meanings rather than being clearly compositional. Does the conventionalization of meaning have consequences for the lexicon? Even within utterances, individual words may have idiosyncratic properties. How does idiosyncrasy bear on the nature of the lexicon and what a word is?

Turning now to the phonological side of words, phonological evidence is often of prime importance in arguing that the objects of study in the languages of the Americas are words. Considerable evidence has been amassed over the years that the phonological word is not the equivalent of the morphological word. The languages of the Americas have been important to the study of the internal prosodic structure of words. One often finds that morphemes within a word are not all equally related to each other phonologically – a word itself may have internal phonological domains. One can ask what these domains say about the nature of the word. On the one hand, they could be used as evidence that the word is formed on a number of lexical levels as, say, in lexical phonology or as a number of distinct prosodic domains. Alternatively, internal phonological structure might be used to argue for the word having internal syntactic structure which is mirrored in the phonology. We ask again, this time from a phonological perspective, 'What is the relationship between the word and the lexicon?'

Phonology contributes to our understanding of words in other ways. While in some of the languages of the Americas it is straightforward to identify morphemes from surface forms, in other cases, the languages are highly fusional. The existence of such forms has, perhaps, implications for lexical insertion. A relevant question is: 'What does the phonology tell us about lexical insertion? Does the phonology bear on whether words are simply the spell out of part of the computational component?

Important issues also arise when one looks beyond the word. Clitics continue to be a source of debate. Their prosodic status is of interest, as is their host. One can question what types of units clitics can attach to, and also whether the attachment of a clitic creates a word, or some higher level unit. In addition, it is sometimes observed that proclitics and enclitics differ in their patterning. The question continues to be important: 'What are the properties of a clitic and what do they tell us about the word?'

In this brief introduction, we hope to have set out some of the issues that motivated us to focus on the nature of the word as the theme for WSCLA 2000.

References

Mithun, Marianne. 1999. *The Languages of Native North America*. Cambridge, U.K.: Cambridge University Press.