WORD, WORD-FORM, LEXEME

English Lexicology
14 November 2007
Word

- Lexicology – the study of the stock of words in a given language
- Word is of central importance
- What do we mean by the term word?
- It is not easy to define word. What are the difficulties?
The definition of word

- The term word is used to designate an intermediate structure smaller than a whole phrase and yet larger than a single sound segment.
- It can be defined depending on whether we focus on its representation, the thought which it expresses, or purely formal criteria.
First definition

- This definition relies mainly on writing traditions that separate by spaces sequences of letters or characters. These separations do not always correspond to functional realities.
- E.g. a new waste *paper basket*
- The first two spaces do not have the same value as the last two.
- Consequently, a definition based on writing traditions alone cannot be entirely satisfactory.
Second definition

- The second type of definition considers the indivisible unit of thought as the most essential criterion. The main problem faced by this view is the delimitation which offers us three possible alternatives:
  - A) the word as represented in writing represents a thought unit or a psychological unit, e.g. table, house, courage, faith, intelligence, tall, short, sleep, eat...
  - B) The word forms one block but includes two units of thought: e.g. farmer, rethink, spoonful.
  - C) The psychological unit exceeds the limit of the graphological unit and spreads over several words, which is then a more complex unit: e.g. all of a sudden, as usual, coconut.
Third definition

- By L. Bloomfield, who suggested a formal definition of word. He contrasted it with other significant units, the morpheme or minimal meaningful unit, and the syntagma or structure, consisting potentially of more than one word.

- For Bloomfield, a minimal form is morpheme. A form which may occur alone is free, and the one which cannot occur alone is bound: F book, man; B -ing, -er.

- Word is a minimal free form, which can occur in isolation and have meaning but which cannot be analysed into elements which can all occur alone and also have meaning.
The word defined

- Most fluent speakers think they know what a word is.
- The word is an uninterruptible unit of structure consisting of one or more morphemes and which typically occurs in the structure of phrases.
- The morphemes are the ultimate grammatical constituents, the minimal meaningful units of language.
The word defined

- For example, the different forms of the verb *speak*, i.e. *speak, speaks, speaking, spoke*, and *spoken* are separate words grammatically. Equally, *grave* (noun) and *grave* (adjective) are different grammatical words. So are the plural, the plural possessive, and the possessive of the word *boy*, all represented by the pronunciation /bɔɪz/, but spelt *boys*, *boys'*, and *boy's* respectively.
Lexeme

- If we consider word to be an abstract unit, that belongs to grammatical or primary articulation of language, then we shall call it a LEXEME.
- Lexeme is a lexical unit and is entered in dictionaries as the fundamental element in the lexicon of a language.
- To summarise the distinction between the orthographic word, which belongs to secondary articulation, we will say that dies, died, dying and die are forms of the lexeme DIE, that man, men are the singular and plural forms of lexeme MAN, that lexeme MAN is a noun, but DIE a verb.
WORD-FORM

- Since there is another distinction to be done, we will refer to ‘orthographic’ or ‘phonological’ word as WORD-FORM.
- This term is introduced because we have to make distinction between WORD-FORM, LEXEME, and WORD.
- The last term will be used for the third sense of ‘word’.
- E.g. 1. a trying day; 2. They are trying hard.
- Here we have two lexemes, but we still think that these words are the same. The first one is a form of TRYING, while the other is the form of TRY, and these are homonymous. These lexemes are formally identical.
- Therefore, we will say that the word-form tried is the form of the word which we call ‘the Past Participle’ / ‘the Past Tense’ of TRY.
Summary

- E.g. *That is no country for old men.* (W.B. Yeats)
- Here we have six monosyllabic word-forms: *this, is, no, for, old,* and *men.* The remaining word-form is *country,* which is a disyllable.
We shall now come back to the distinction often made between lexical and grammatical words. In most general terms, lexical words are nouns, verbs, adjectives, and adverbs. They have fairly independent meanings and may be meaningful even in isolation or in a series. For example, bottle has an independent meaning; and so does the series boy, break, bottle, stone. On the other hand, a word such as a, with, or a series such as a, the, to, with does not automatically suggest an identifiable meaning. Classes of lexical words contain hundreds or even thousands of members, and they form open classes.
For example, the English vocabulary has thousands of nouns, and new items can always be added to the list. Grammatical words are elements like prepositions, articles, conjunctions, forms indicating number or tense, and so on. Such classes contain relatively few members and the addition of new members is rather rare. They constitute closed classes or closed sets.
However, the distinction between grammatical and lexical words must be handled with care. There is no clear-cut dividing line between the two types of word. We should, in fact, speak of a continuum ranging from words with semantic content such as exam, students, to words devoid of semantic content such as it and that in a sentence like «It is obvious that some students will pass this exam». 
However, although prepositions may be classed as grammatical words, they are not completely empty of semantic content. The sentence «The book is on the table» has quite a different meaning when on is replaced by under, near, off, etc. Similarly, the co-ordinators and, or, and but are not mutually interchangeable, because they are not synonymous.
Characteristics of words

- First, the word is an uninterruptible unit. When elements are added to a word to modify its meaning, they are never included within that word. They respect the internal stability of the word and are added either at the beginning as prefixes of the word or at the end as suffixes. For example, the prefix *un-* and the suffix *–able* may be added to the words *aware* and *drink* and give *unaware* and *drinkable* respectively.
Note, however, that an affix may also occur not at the beginning or at the end, but simultaneously with the word; we then speak of a suprafix. Compare for example the words 'export' (noun) and ex'port (verb); they differ only in the position of the primary stress represented by the symbol ('). The stress pattern may be referred to as a suprafix. The word to which affixes are added and which carries the basic meaning of the resulting complex word is known as 'the stem', which may consist of one or more morphemes. The label 'root' is used to refer to a stem consisting of a single morpheme.
Secondly, the word may consist of one or more morphemes. When it consists of one morpheme only, then it cannot be broken down into smaller meaningful units, e.g. *dog, hand, man, out, work*. These are called 'simple' words, which are typically 'minimum free forms', in the sense that they may stand by themselves and yet act as minimally complete utterances, e.g. in answer to a question.
When words consist of more than one morpheme, they may be either complex or compound. Complex words may be broken down into one free form and one or more bound forms: e.g. dog-s, happy-ly, quick-er, work-ing, whereas compound words consist of more than one free form: e.g. birth+day, black+bird, candle+stick, coat+hanger. We also need to mention cases which incorporate the characteristics of both complex and compound words: e.g. gentle-man-ly consists of the compound word gentle+man and the suffix -ly; wind+shield+wipe-er consists of the compound word wind+shield and the complex word wip-er.
Thirdly, the word occurs typically in the structure of phrases. According to the hierarchy adopted here, morphemes are used to build words, words to build phrases, phrases to build clauses, and clauses to build sentences.
Finally, it is also an important characteristic of each word that it should belong to a specific word class or part of speech. Where the same form appears in more than one class, as frequently happens in English, we regard the various occurrences as separate words (for example, smoke (verb) as distinct from smoke (noun)). It may even be suggested that a word is defined by two factors: its semantic 'nucleus' and the class to which it belongs.