



Word, word-form, lexeme

21 October 2011

Word

- Morphology – the study of the structure of words
- Lexicology – the study of the stock of words (*lexis*, *lexicon*) in a given language.
- Word is of central importance to lexicology
- In morphology, the word is not sufficient in the analysis, because there are smaller units.
- What do we mean by the term *word* in lexicology?
- Not easy to define *word*. Why?

Word

- **Speech** is a continuous stream of sound without a clear division into units, but it can be analysed into meaningful elements which recur and combine according to rules.
- In writing, such an analysis is expressed through the division into words and sentences.
- The essence of **grammatical units** is that they are meaningful and combine with each other in systematic ways. We may distinguish a **hierarchy of units**:
- A **sentence** consists of clauses, a **clause** consists of one or more phrases, a **phrase** consists of one or more words, a **word** of one or more morphemes, a **morpheme** consists of one or more **phonemes**.

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.
- However, although it may be difficult to define *word*, even nonliterate speakers can divide the speech chain into words.

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.
- In speech these pauses do not exist. Speech is a phonetic continuum and breaks are done only between some larger syntactic units, such as phrases or clauses.

E.g. *School, household, in, fall out, waste paper basket, forget-me-not, runner-up.*

Value of these spaces?

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.

Working definition

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.

Word

- Another difficulty when treating words is the term *word* itself, which may be used in a number of different ways:
 1. It may refer to the **word form**, the physical unit or concrete realisation, either the orthographical word (the written form) or the phonological word (the uttered or transcribed form).
 2. It may refer to the **lexeme**, which is rather like a dictionary entry. A lexeme includes all inflected formes of a word. It is an abstraction and is indicated by small capitals:
e.g. WALK – *walk, walks, walked, walking*
RUN – *run, runs, ran, running*
SING – *sing, sings, sang, sung, singing.*

Lexeme

- Since the lexeme is abstract, it is conventional to choose one of the inflected forms to represent it, such as infinitive of the verb or the singular of the noun. The same word form may in fact represent different lexemes:
- a. A **homonym** is a single orthographic and phonological word standing for two lexemes, as *bear* is either the verb or the noun.
- b. A **homograph** is a single orthographic word (but separate phonological words) standing for two lexemes, as *lead* is either the noun /lɛd/ or the verb /li:d/.
- c. A **homophone** is a single phonological word (but separate orthographical words) standing for two lexemes, as /mi:t/ is either the noun *meat* or the verb *meet*.

Lexeme

- The same lexeme might also have quite distinct word forms, as in the case of the definite article *the*, represented by /ðɪ:/ or /ðə/, or the indefinite article *a/an*, represented by /eɪ/, /ə/, /ən/, or /æn/.
3. Finally, word may also refer to a morphosyntactic word (or **grammatical word**). A **morphosyntactic word** consists of a lexeme and associated grammatical meaning. For example, in:

I take the garbage out every week. (TAKE + present)

I took the garbage out yesterday. (TAKE + past)

I have taken the garbage out already. (TAKE + past participle)

Word-form

the different morphosyntactic words are represented by different word forms (*take, took, taken*).

But in

- *I put the garbage out every week. (PUT + present)*
- *I put the garbage out yesterday. (PUT + past)*
- *I have put the garbage out already. (PUT + past participle)*


the different morphosyntactic words are represented by the same word form (*put*).

Lexical words

- We shall now come back to the distinction often made between **lexical (content)**, **function** words and **inserts**.
- 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. They are the main carriers of meaning in a text.
- For example, *bottle* has an independent meaning; and so does the series *boy, break, bottle, stone*.
- They are the words that remain in the information-dense language of telegrams, lecture notes, headlines, etc.

Function words

- The English vocabulary has thousands of **lexical words**, and new items can always be added to the list.
- **Function words** bind the text together. They often have a wide range of meanings and serve two main roles: indicating relationships between lexical words or larger units, or indicating the way in which a lexical word or larger unit is to be interpreted.
- Function words are elements like **prepositions, determiners, pronouns, numerals, auxiliaries, adverbial particles, coordinators, subordinators, wh-words, the negator *not*, existential *there*, the infinitive marker *to*...**
- Such classes contain relatively few members and the addition of new members is rather rare. They constitute closed classes or closed sets.

- 
- However, although prepositions may be classed as grammatical/function 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.

Inserts

- Inserts are a relatively newly recognized category of word. They do not form an integral part of a syntactic structure, but are inserted rather freely in the text.
- They are often marked off by intonation, pauses, or by punctuation marks in writing.
- They characteristically carry emotional and interactional meanings and are especially frequent in spoken texts.

E.g. **Hm hm**, very good. (CONV)

Cheers man. (CONV)

Yeah, I will. **Bye**. (CONV)

Biber et al. (1999)

Multiple class membership

- Word classes have fuzzy borderlines:

Homonymy – a single form belongs to more than one word class.

E.g. **right** – can be a lexical word (noun, verb, adjective, or adverb), or a function word (preposition or subordinator).

A number of forms can belong to more than one function word class, e.g. **To** (preposition and infinitive marker), **for** (preposition and subordinator), etc.

Other examples:

Before – adverb, preposition, subordinator

Round – Noun, verb, adjective, adverb, preposition

Weekly – noun, adjective, adverb

Early – adjective, adverb


(Biber et al. 1999)


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*, *happi-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*.

- 
- 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.