

A few notes on definition of word

Petra Jesenská

Matej Bel University, Banská Bystrica, Slovakia

0 Introduction

The definition of word itself has been an issue of some controversy also because it is quite a difficult task to define the word without knowing what should be at 'least one word' (Cruse, 1991:35). That is why one must examine what it entails. A great deal of scholarly discussion has been centred on the linguistic status of the word. *The Oxford Companion to the English Language (OC)* describes word from various points of view but basically it characterizes word as 'an element of speech or a combination of vocal sounds, or one such sound used in a language to communicate an idea' (OC, 1996:1026). Alena Štulajterová (2005:89), the Slovak structural linguist, does not problematize the issue and understands word as a 'lexical unit' within a language word-stock. Bertrand Russell, however, distinguishes two kinds of word, so-called *object word* and *dictionary word*¹ (Palmer, 1979:23-24). The English grammarian Henry Sweet made distinction between *full* and *form words* which are known as lexical (content) and grammatical (empty) words in nowadays' linguistics.

Word, however, has numerous connotations. After clicking on the Internet a lot of information can be found about a computer programme Word. One must search *word definition* to find description of word in the linguistic sense. What is available on the Internet are characterizations of word as:

- 'a part of speech',

¹ Russell asserts: 'Object words are learnt ostensibly, i.e. by pointing at objects, while dictionary words have to be defined in terms of the objects words. The object words thus have OSTENSIVE DEFINITIONS.' (Palmer, 1979:24)

- 'a unit of language that native speakers can identify'²,
- 'a brief statement' (ibid.), for example *he didn't say a word about it*, etc.
- 'the Word was God' - set of sacred rules given to people from god, etc.

One could continue in searching like this and digress from the major topic – *word*.

Now I would like to draw readers' attention to two general and constant characteristics of *word* across a wide range of languages.

The first is that a word is typically the smallest element of a sentence and the second is that a word is the largest unit. A word is the smallest and largest unit at the same time. We are going to explain into more detail.

1 Word – the smallest and largest unit at the same time

A word is typically the smallest element of a sentence which has positional mobility – that is, the smallest element that can be moved around without destroying the grammaticality of the sentence (ignoring any semantic effects, of course): John saw Bill.

Bill saw John.

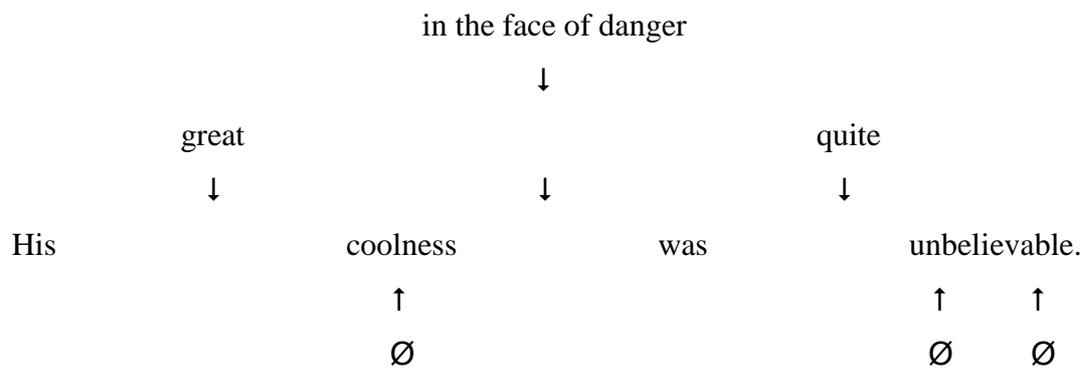
Bill, John saw.³

As one can see by no means all words are equally mobile in this sentence, but with very few exceptions, the smallest mobile units are words. The morphemes constituting a single word have, however, a rigidly fixed sequential order: we find *unwillingly*, but not *lywillingun* or *unlywilling* (ibid.), etc.

² Resource: <http://wordnet.princeton.edu/perl/webwn?s=word>

³ All three examples taken from Cruse, 1991:35.

According to Cruse (1991:36) 'the second major characteristics of words is that they are typically the largest units which resist 'interruption' by the insertion of new material between their constituent parts. Consider the following sentence, and observe where extra material can be inserted: *His coolness was unbelievable.*'



The possible insertion points clearly represent word-boundaries.

However, this characteristic of words may seem less important in other languages (Chinese or Japanese). We have chosen Turkish as a representative of an Indo-European language and thus it can serve a good example because it is a kind of language, in which words composed of a relatively large number of morphemes are very common. For example, the word *öldü* consisting of two morphemes. This means 'he / she / it died', the final *-dü* indicating third person singular, past tense. Quite a lot can be inserted between the root *öl-* and the past tense element, as the following word shows: *öl-dür-ül-e-mi-yecek-ti*. This can be translated as 'He would be unable to be killed': the final *-ti* represents the same morpheme as *-dü*. However, there is a marked difference in the degree of interruptibility between words and phrases. In the Turkish example, although several grammatical elements can be inserted within the word, they are strictly determinate in number and identity; whereas between words, if one takes into account coordinations and parenthetical insertions, the possibilities are infinite. This can be illustrated from English; Turkish is no different in principle:

HIS great courage and imperturbable COOLNESS in the face of what must at times have seemed to him to be insuperable odds WAS, I must confess – although I do not really like him – quite UNBELIEVABLE⁴.

2 Word in the context of structural linguistics

There are too many definitions of word but most common it is described as the fundamental or primary unit of the lexical level of the language: as a sequence of phonemes or its representation in writing (graphemes), that expresses a meaning. It is obvious that a word cannot be divided into smaller units of independent use, although, a linguistic analysis may uncover in it more than one morpheme. Also, a word cannot accept any insertion of further material, it is a relatively stable unit which can be arranged in structures with other such units to form sentences of the language. This description is more or less identical with Cruse's understanding of a word.

Cognates of 'word' with Afrikaans is *woord*, Danish *ord*, Dutch *woord*, French *mot*, German *Wort* (word), Greek *weírein* (speak), Latin *verbum* (word), Sanskrit *vrátam* (command), Spanish *palabra*, and Swedish *glosa* or *ord*.

In linguistics there are eight kinds of word distinguished when certain features are fundamental to English (the following distinction is taken from *OC*):

- | | |
|---------------------------|-----------------------------|
| 1 the orthographic word, | 5 the grammatical word, |
| 2 the phonological word, | 6 the onomastic word, |
| 3 the morphological word, | 7 the lexicographical word, |
| 4 the lexical word | 8 the statistical word. |

⁴Cruse, 1991:36

1 The orthographic word is a kind of word described in terms of alphabetic or syllabic writing systems: a graphic (visual) sign with an empty space around it. It may or may not have a canonical form: according to *OC* in the 14th Century, before print encouraged standardization, *merry* was also spelled in various ways: *myry*, *myrie*, *murie*, and *mery*.

2 The phonological word is the word understood in terms of sound: a spoken signal that occurs more commonly as part of a longer utterance than in isolation and is subject to rhythm. Traditional spoken English is a series of stressed and unstressed syllables which behave in more or less predictable ways: where an experienced listener hears *It's no good at all* being pronounced in a relaxed, informal way, a foreigner may hear *Snow good a tall*⁵. In the flow of speech, words do not have such distinct shapes as on paper, and syllable boundaries do not necessarily reflect grammatical boundaries: the phrases *a notion* and *an ocean* are usually homophonic and only context establishes which has in fact been said.

3 The morphological word is the word in terms of form that lies behind both the orthographic and the phonological word. For example, word *big* has a spelt-out realization *b-i-g* and a spoken realization /bɪg/, but is independent of both, because it can be expressed in either medium and also in sign language. This entity is capable of realization in different 'substances'; it is distinct from such spelt-out variants as *colour* and *color* as well as from the innumerable ways in which African, American, Australian, Caribbean, English, Irish, Scottish or other people may say 'colo(u)r'. However, all such users have it in common and it is the basis of such further forms as *colourful* and *discoloured*.

⁵ Both instances taken from *The Oxford Companion to the English Language*

4 The lexical word (*content word, lexeme, lexical item*) means the word in terms of content relates to things, actions, and states in the world. It is usually realized by one or more morphological words, as when *do, does, doing, did, done* are taken to be five 'versions' of the one verb *do*. Lexical words are generally fitted into the flow of language through such mechanisms as *affixation, suppletion*⁶, *stress shift*, and *vowel change*, all of which have morphological and other effects. The set of such words is always open to new members, and in English embraces nouns, verbs, and adjectives, and other parts of speech when they behave like nouns, verbs, and adjectives, as in 'But me no buts'. Lexical words may be simple in structure (*cat, mouse*), or composite (*cold-bloodedness, incomprehensible, teapot, blackbird, Commonwealth, stamp collector, put up with* etc).

5 The grammatical word (also called a *form word, function word or structure word*) is the word that in terms of syntactic function contrasts with the lexical word and is an element in the structural system of a language. It serves to link lexical words. In English, conjunctions, determiners, interjections, particles, and pronouns are grammatical words. They occur frequently and have their own semantic systems, as which such particles as *up* and *down*, which relate to position, direction, space, and time. In principle, such words are a closed set to which new items are seldom added. As lubricants, grammatical words are like affixes: the *out* in *throw out* is like the prefix *e-* in *eject*; the *before* in *before the war* means the same as *pre-* in *pre-war*. They can also function like affixes, as in *he-man* and *yes-man*.

6 The onomastic word – the word in terms of naming establishes special, often unique reference: the difference between *Napoleon* and *emperor*. It may be simple like *Smith* or complex like *Smithsonian*. Names may be motivated, like *Sitting Bull* (a Sioux name derived from

⁶ 13th C: from Latin *suppletio/suppletionis* making complete – a term in linguistics for a situation in which a form in a grammatical paradigm bears no family resemblance to the base form: for example, *went* in *go/goes/going/went/gone* and *better* and *best* in *good/better/best*. Suppletive forms are common in irregular usage in many languages.

an omen involving a bull buffalo) or conventional, like *Smith* today (though not in the Middle Ages, when the name was occupation-based). Although such words are lexical, they are not usually listed in dictionaries and may or may not be relevant in encyclopaedias. They are often regarded as apart from normal vocabulary, though they also ought to be learned.

7 The lexicographical word – the word in terms of dictionaries is usually presented in an alphabetic setting. Many dictionaries have an entry, for example *did* as the past of *do*, an entry *them* as the object form of *they*, and so on, with cross-references to the representative form. There are therefore two kinds of entry: anything the compilers think anyone might look up, and the *citation forms* under which definition proceeds. The conventional citation form for nouns is the singular (unless a word is always plural) and for the verb is the bare infinitive (unless the verb only occurs as a participle, or is a modal verb).

8 The statistical word – the word in terms of occurrences in texts is embodied in such instructions as ‘Count all the words on the page’: that is, count each letter or group of letters preceded and followed by a white place. This instruction may or may not include numbers, codes, names, and abbreviations, all of which are not necessarily part of the everyday conception of ‘word’. Whatever routine is followed, the counter deals in tokens or instances and as the count is being made the emerging list turns tokens into types: for example, there could be 42 tokens of the type *the* on a page, and 4 tokens of the type *dog*. Both tokens and the types, however, are unreflectingly spoken of as words.

3 Conclusion

There are different points of what a *word* is (or means). Two kinds of definitions have been presented in this article. The first one comes from Cruse who understands the *word* as the smallest unit of a sentence (viewpoint of syntax) and at the same time as the largest inseparable unit where no other material can be inserted (morphological and lexicological viewpoint). There are also other eight definitions understood in terms of various kinds of points of view in structural linguistics: orthography, phonology, morphology, lexicology, grammar, onomastics, lexicography and statistics. The main aim of the article was to illuminate definition of word in regard to its semantics.

The article is a compilation of various characteristics and definitions of the *word*. The aim of an author was not to bring new ideas, but to summarise existing and accepted ideas about the *word* in terms of structural linguistics. The greatest help were publications recorded in *bibliography* below this article.

Bibliography

- CRUSE, D.A.1991. *Lexical semantics*. Cambridge: Cambridge University Press, 1991, pp310. ISBN 0-521-27643-8
- CRYSTAL, David. 1996. *The Cambridge Encyclopedia of the English Language*. Cambridge: Cambridge University Press, 1996, pp 490. ISBN 0-521-40179-8
- The Handbook of English Linguistics*. 2006. Malden, Oxford, Carlton: Blackwell, 2006, pp 808. ISBN 1-4051-1382-0
- The Oxford Companion to the English Language*. Abridged Edition. (In the text as *OC*). 1996. Oxford, New York: Oxford University Press, 1996, pp1054. ISBN 0-19-863136-7
- PALMER, F.R.1979. *Semantics*. Cambridge, London, New York, Melbourne: Cambridge University Press, 1979, pp 23 – 24, pp 37 – 42; 164pp. ISBN 0 521 09999 4 paperback
- ŠTULAJTEROVÁ, Alena. 2005. K problematike adaptácie nových anglicizmov v systéme slovenského jazyka. In: *Teória a prax prípravy učiteľov anglického jazyka 3. Zborník katedry anglistiky a amerikanistiky FHV UMB*. Banská Bystrica : UMB, 2003, pp 89 – 106. ISBN 80-8083-148-3.

The Internet Resources

Definitions of word: <http://wordnet.princeton.edu/perl/webwn?s=word>
(November 14, 2007)

Definition of word: <http://www.brainyquote.com/words/wo/word240447.html>
(November 20, 2007)