

Compounds in English

Among the different processes of word-formation, compounding is probably one of the most productive, and, be it aware or as part of their linguistic competence, people make use of it pretty much all the time, because it allows them to create new meaning and to express themselves in a more specific manner. In the following essay, I would like to discuss the notion 'compound' in linguistic terms. I begin with a description of the basic structural properties of regular compounds.

Regular compounds have a **binary structure**, indicating that two constituents are combined in a **hierarchical** manner to form a new word. This binary branching can be done repeatedly which can result in complex multi-constituent words. This property is called **recursivity**. It is important to note, however, that only two constituents can be combined at a time, which is shown in (1).

(1) *the [[national [[football] league]] association]*

(2) *[over-the-fence gossip] (Plag 2003: 134)*

In (1), *foot* first combines with *ball* to form *football*, which then combines with *league* to form *football league*. It is then *football league* which combines with *national* to form *national football league*, which finally combines with *association*. Having a closer look at the constituents, it is visible that the first constituent can either be a root, as in *football*, a word, as in *football league*, or a phrase, as can be seen in *over-the-fence gossip* in (2). Interestingly, the second constituent can be a root, as in *football*, or a word, as in *national football league*, but not a phrase. If (2) would be *gossip over the fence*, the whole construction would be a nominal phrase and not a compound any longer which is one example why in a compound, the second constituent cannot be a phrase.

Another property of regular compounds is that they exhibit a **modifier-head structure**. The term **head** refers to the most important unit in a compound and in complex linguistic structures in general. It is also extensively used in the analysis of syntactic phrases, which shows the overlap between syntax and morphology. The difference between syntactic phrases and compounds lies in the fact, that in English, syntactic phrases are pronounced on the last word of the phrase, which is called the nuclear stress rule, whereas compounds are normally pronounced on the left-hand

constituent. The latter feature is the so-called **compound stress-rule**. Both are exemplified in (3).

(3) *gréenhouse* - *a green hóuse* (*Plag 2003: 138*)

(4) [*parks commisioner*]*Sg.* [*parks commissioners*]*Pl.* (*Plag 2003: 136*)

It is necessary to mention that there are exceptions to the compound stress rule. These are mainly restricted to certain well-defined meaning relationships as well as to certain types of combinations that are based on the analogy to existing compounds. It is only in these cases that the compound has rightward stress. These are, however, systematic exceptions that prove the rule. The application of the compound stress-rule, on the other hand, can be very beneficial, especially when dealing with complex multi-constituent compounds, where different hierarchical structures are possible. The stress assignment helps to clarify the binary structure of the compound, because according to the compound stress rule, the stressed unit of a compound has to be its left constituent.

In (4), another structural property of regular compounds is shown, namely, that a compound takes on the grammatical features of its head. Looking at the example, one can see that although *parks* is a noun in plural form, the whole compound remains in singular when *commisioner*, which is the head, remains in singular form. It is only when *commisioner* changes to its plural form that the compound becomes plural also. This inheritance of grammatical features from the head, as described above, is referred to as **feature percolation**.

So far, a compound can be defined as a complex word of specific types of constituents, which has a binarily branching hierarchical structure with a specific stress pattern attached to it. But there is more to it. In the following, I would like to expand this morpho-syntactic view of compounds by looking at the semantic characteristics of compounds, thus operating on the level of meaning.

Looking at the compound *tennis club* , one can say that it consists of a **modifier**, which is *tennis*, and a head, which is *club*. The head denotes a club and the modifier modifies this club by making it more specific. A *tennis club* is therefore a subclass of the referent of the head or, in other words, a hyponym of club. The **semantic head** of this compound is inside the compound, which is why these kind of compounds are referred to as **endocentric compounds**. Another traditional expression for compounds denoting a subset of the denotation of the head is **determinative compound**. Both notions can be used.

What makes endocentric compounds so productive is that they are relatively open in meaning. This creates an ambiguity and makes it possible that several different meanings of the same compound are plausible, as will be shown in the following examples:

(5) *cheese shirt* (*Olsen powerpoint: slide 10*)

(6) *power source*

(7) *milk steamer* (*Olsen powerpoint: slide 37*)

The compound *cheese shirt* in (5) is ambiguous, because it does not provide a clear-cut interpretation. Although *cheese* serves as a modifier of *shirt*, which creates a compound that is a subclass or a hyponym of the head, the meaning relation between these two words can be manifold: the compound could denote a shirt, which is yellow or a shirt which has a picture of cheese on it, but it could also denote a shirt which stinks or has holes in it or is worn when making cheese. As can be seen, the interpretation is relatively open when looking at the compound in isolation. This openness of interpretation is created by the lexical property of the noun *cheese*, which is a **sortal noun**. Sortal nouns are used for classifying entities.

It is worth noticing, however, that the semantic head of nominal compounds can not only be a sortal noun, as shown in (5), but a **relational noun** also. The compound in (6) is an example therefor. The meaning of *source* is very unspecific and requires some sort of specification, so that a person can make sense of the word. Semantically spoken, it triggers other words to specify its meaning. This is why the structure of these kind of compounds is called an **argument-head structure**. The property of a noun to be relational can be observed in syntax also, where in the noun-phrase *source of power*, *power* would be a necessary argument of *source*.

Looking at (7), we have an example of a **synthetic** or **verbal compound**. As opposed to the so-called **primary** or **root compound** where the head is a noun that is not derived from a verb, a synthetic compound's head is derived from a verb and therefore **deverbal**. The verb *steam* and the derivational suffix *-er* combine to form the noun *steamer*, which in turn is head of the compound *milk steamer*. The interesting thing here is how the first constituent of the compound, *milk*, is interpreted as an argument: the verb *steam* in the sense of *to place food over boiling water so that it cooks in the steam* is transitive and requires an **argument**. This feature is now passed on to the derivational noun *steamer*, so that the empty slot in the semantic representation can then be filled by the noun *milk*, which functions as an argument. This process of **argument-linking** can be very productive, especially within verbal or synthetic compounds, where the deverbal noun in head position can relate to its argument in many different ways. It needs to be added, however, that the openness of interpretation can be restricted when a specific lexical meaning has been assigned to it, as is the case in (7), or when the surrounding context or discourse conditions suggest a specific interpretation.

Apart from endocentric, determinative compounds and their variations, there is a second type of compound called **exocentric compound**. With regard to this type, the semantic head is not inside, but outside the compound. This defining feature will be illustrated in the following examples:

(8) *pickpocket* (Plag 2003: 145)

(9) *bonehead*

In (8), it is misleading to analyze the semantic relationship of the compound by means of endocentric analysis which is due to the fact that a pickpocket is not a hyponym of pocket. The semantic head lies rather outside the compound, denoting a third entity who in this case is *a person who steals money or other things from other people's pockets*. This idiomatic meaning is denoted by the word *pickpocket* and cannot wholly be derived by its constituents.

When an exocentric compound denotes a third entity outside the compound which is characterized by the properties of the constituents, as is the case in (9), this compound is also called a **possessive compound**. In (9), *bonehead* denotes a specific property of a person, namely the property of being stupid. This attribution of specific properties to a person as a whole is referred to as **metonymic shift** in meaning. Apart from this, the first member of possessive compounds are mainly adjectives.

A third type of compound is labeled **copulative compounds**. The entity denoted by copulative compounds is characterized to an equal degree by the properties named in the compound, as can be seen in (10), (11) and (12).

(10) *a scientist-explorer* (Plag 2003: 146)

(11) *the mind-body problem* (Plag 2003: 146)

(12) *an actor- songwriter-screenwriter* (Olsen powerpoint: slide 33)

Copulative compounds can further be subdivided into **appositional** and **coordinative compounds**. In (10), one entity is referred to that is both a scientist and an explorer. If one entity is described by the members of a copulative compound, it is said to be an appositional compound. A coordinative compound, on the other hand, denotes two entities that stand in a particular relationship with regard to the noun that follows. This is illustrated in (11), which denotes a problematic relationship between the mind and the body.

Interestingly, copulative compounds can be embedded inside determinative compounds. When looking at (11) again, it becomes obvious that the coordinative compound functions as modifier with regard to the head of the overall determinative compound, which is *problem*. In (12), however, the appositional compound consists of two determinative compounds, namely *songwriter* and *screenwriter*, showing that determinative compounds can also be embedded inside copulative compounds. Moreover, (12) shows that copulative compounds can be made up of more than two constituents, as was the case with determinative compounds.

It is important to remember, however, that the preceding classification of exocentric, possessive and copulative compound is a semantic one! On a formal level, however, things look different. Taking the exocentric compound in (9) as example, it can be said that the sentence “*There are many boneheads*” is possible, whereas the sentence “*There are many boneshead*” is wrong. The same holds for (10), where the plural marking also only occurs on the right-hand member of the

compound. This feature percolation indicates that at least in terms of their grammatical properties, exocentric, possessive and copulative compounds have indeed an endocentric head that is inside the compound. As for endocentric compounds, the semantic and the formal head are identical.

In addition to the semantic compound classification discussed above, compound patterns can also be established on a formal level. This approach is based on syntactic categories, namely the part of speech or word class of the compound head. It works as follows: in a first step, an inventory of compound patterns is created by finding complex words that appear to be compounds and whose constituents are a combination of the different word classes; in a second step, this matrix is analyzed to find out characteristic features as well as expressions that superficially appear to have the form of compounds, but are actually created by other word-formation processes. The result of this procedure is a typology of possible compounds.

With regard to **nominal compounds**, it can be shown that they are very productive in forming new words. This is mainly due to endocentric **noun-noun** compounds which, as already described above, can have a wide range of possible meanings. There are fewer **verb-noun** compounds such as *playground* or *spoilsport*, which in equal measures are endo- and exocentric. **Adjective-noun** compounds are highly lexicalized, which is seen in words such as *greenhouse*, *heavyweight* or *babyface*. Together with **preposition-noun** compounds such as *underarm* or *afterlife*, they often display a modifier-head structure.

Given the appropriate adjectival head, **noun- adjective** compounds as for instance *sugar-free* or *class-conscious* can have an argument-head structure. In other cases such as *knee-deep* or *blood-red*, the interpretation is less clear-cut, usually involving a comparison or a meaning of intensification. Apart from perhaps *fail-safe*, the category of verb-adjective compounds does not exist.

Adjective-adjective compounds are sometimes copulative compounds, being either appositional as in *sweet-sour* and *bitter-sweet* or coordinative as in *public-private (partnership)*. Moreover, the adjectival head can be a derived adjective as for instance in *awe-inspiring*. Preposition- adjective expressions such as *incoming* and *outgoing* only appear to be compounds, but are derived from phrasal verbs and thus the result of an **inversion** process. The same holds true for preposition- verb expressions as for instance *download* or *upgrade*.

Looking at noun-verb expressions such as *ghost-write* or *chain-smoke*, the verbal head can best be analyzed as the result of **backformation**. Backformation is a word-formation process, in which *a new lexeme is created by removing the suffix of the word*. As opposed to clipping, the word's part of speech or meaning may change. As a consequence of the strict verb-object order in the English language, *ghost-write* has to be a backformation of *ghost-writing*. With regard to **verb-verb** compounds, it seems that they are appositional copulative compounds, as can be seen in *drop-kick* or *stir-fry*. They are extremely rare, though. Adjective- verb expressions such as *to short-cut* or *to*

blindfold are so-called **conversions**. By means of conversion, a *new lexeme is created from an existing word without changing the form of the word*. Mostly, this is achieved when a word is assigned a different part of speech.

Prepositional “compounds” either do not exist, as is the case for nominal- preposition- and adjective-preposition expressions, or are the result of other processes. Verb- preposition expressions such as *push-up* or *start-up* can easily be unmasked as nominalizations of phrasal verbs and are thus the result of conversion. Preposition- preposition expression as for instance *into*, *onto* or *upon* are very limited in number and can be seen as **lexicalizations**, which is due to their grammatical nature as function words.

Apart from the preceding typology of possible compound patterns, which is based on the different word-classes, there are still some more compound patterns to be hinted at. Among these less frequent patterns are the so-called **phrasal compounds** such as the *over-the-fence gossip* discussed in (2), in which the first constituent takes on the form of a phrasal sentence. Talking about **identical constituent compounding**, the term already suggests that the two members of the compound are identical, as can be seen in the sentence “*I had an amazing night yesterday and by this I mean amazing-amazing*”. The emphasis of the compound lies in the fact that it qualifies the adjective as proper or prototypical. Another pattern worth mentioning is the use of **analogy**. Looking at words as for instance *landscape*, *cityscape*, *seascape* or *moonscape*, it becomes obvious that the building of analogies can be quite productive. As for **neoclassical compounds**, which can be seen as yet another process of word-formation, it can be said that they consist of either an initial combining form, a final combining form or both, as can be seen in *automobile*, *francophobe* or *astrology* respectively. The distinctive feature of **combining forms** is their ability to combine with each other which is not true for affixes or bound roots. Furthermore, they often display a different stress pattern to that of other compounds. Finally, neoclassical compounds sometimes include a **linking element -o-**, which is phonologically determined and needs to be analysed separately for each combining form.

Resuming what has been discussed so far, one can say that due to its nature as a linguistic sign, compounding can be analysed on both a semantic as well as on a formal level. It is therefore possible to classify compounds either according to their interpretation or according to their word-classes. Each approach leads to different results. If combined, both are useful tools in order to describe how complex compound patterns are formed and what their possible meaning could be. In day-to-day life, this knowledge can be used to deepen the understanding of the human language and, if needed, to actively create new words.