

English Exocentric Compounds[†]

Laurie Bauer

Victoria University of Wellington

Most compounds in English are endocentric, that is, one of the elements (typically the right-hand element) is the head of the construction. Headedness is shown most clearly by hyponymy: the compound as a whole is a hyponym of its head. For example, *traffic-light* is a hyponym of *light*, but not a hyponym of *traffic*. In some languages compounds take their declension class or gender from those of the head element, but in English this is not particularly important. An exocentric compound, by contrast, is one which is not a hyponym of one of its elements, and thus appears to lack a head or perhaps to have a head (or 'centre') external to the compound itself. English examples such as *redhead* 'a person with red hair', *flat-foot* 'policemen (slang)' and *egg-head* 'intellectual' abound. As far as I am aware, the first person to extend the notion of exocentricity from syntax to the morphological form of compounds was Bloomfield (1935: 235). Bloomfield himself links the term 'exocentric compound' with the earlier Sanskrit term *bahuvrihi*, thus establishing the equivalent use of the two terms in the subsequent technical literature. This is unfortunate. First, the original Sanskrit *bahuvrihi* compounds were adjectives not nouns. Examples are *bahuvrihi* itself, literally 'much rice' but meaning 'having much rice', or *gatāyus-* 'departed life = dead' (Gonda 1966: 83). Second, as will be shown, *bahuvrihis* represent a very small proportion of those compounds which may be termed exocentric. In this paper I shall consider some of the different kinds of exocentric compound in English, and suggest that some of them are not exocentric at all.

[†] This is the English version of the paper that was eventually published, in French, under the title 'Les composés exocentriques de l'anglais' in Dany Amiot (ed), *La composition dans une perspective typologique*. Arras: Artois Presses Université, pp. 35-47. The published version includes some editorial changes from this version.

1 The Romance type

There is a Romance type, found throughout the Romance languages and brought into English through contact with French after the Norman Conquest, in which the compound is made up of a verb, and a noun which functions as the direct object of that verb. The compound is not a hyponym of the final noun, and so the compound is exocentric. Some examples denoting people are given in (1) and some denoting objects are given in (2) (see Marchand 1969: 380-1).

- (1) cutthroat
killjoy
pickpocket
spoilsport
turncoat
- (2) breakfast
breakwater
dreadnought
scarecrow
wagtail

Many of the examples cited by Marchand are no longer current, and the type does not appear to be productive in modern English, though Bauer (1983: 205) cites the trade-name *Xpel-air* as a recent example. Modern usage would replace these with so-called synthetic compounds (for instance, *coat-turner*, *tail-wagger* – both found on the internet, but with meanings different from those of *turncoat* and *wagtail*) or with a compound headed by a converted noun (*waterbreak*).

There is a long history of discussion of the form of the verb in corresponding French constructions, since the verbal element in French has a form which could be taken to be an imperative. Although Marchand (1969: 381) seems ready to adopt this analysis, the question is rarely raised for English, where the verbal element looks like a stem. Although the imperative would be homophonous with the stem, there is no particular reason to assume any imperative function: to call people cutthroats is not to exhort them to cut more throats (Bauer 1980).

But if the form of the verb has generated some discussion, there seems to be general acceptance that these words are indeed exocentric. Since I have no reason to doubt this conclusion, the discussion can rest there for the moment.

2 The phrasal-verb type

The second type of exocentric compound discussed by Marchand (1969) is the type where a noun is made up of a verb + preposition/adverb/particle.

Marchand uses the example of *showoff*. Clearly there is no noun involved in *showoff*, yet the construction as a whole is a noun, so these must be exocentric.

Marchand has very few blind-spots, but the distinction between what Dokulil (1968) terms *Wortbildung* 'the process of forming words' and *Wortgebildetheit* 'the analysis of complex words' seems to be one of them. If we look at the synchronic structure of the English lexicon, forms like *showoff* appear to be compounds, in that they are made up of two independent lexemes.¹ If we look at the way in which this situation arose, we may come to a different conclusion about how we should view the structures.

To begin with, the two distinct elements of *showoff* are already present in the verb *to show off*. If there is a compounding process, we might argue, it operates in the formation of the verb, and not in the subsequent formation of the noun.² In the verbs, the process of composition (or formation, if we do not wish to be too specific) is probably endocentric: showing off might be deemed to be a type of showing, to clean up is a form of cleaning, and so on. In other cases, the semantic link is not so clear: to put up (at the inn) may not be a kind

¹ Saying this begs a number of questions concerning the nature of the lexeme, and whether function words belong to the class of lexemes. Since such questions do not lie at the heart of what it is that makes *showoff* a compound, I will not dwell on them here.

² Again, I am making assumptions, but this time they are not absolutely true. There is a whole series of forms on the pattern of *teach-in*, *love-in*, *sleep-in* for which there is no pre-existing verb. I assume that once the pattern of *showoff* exists, it can attract new members by analogy, even if the historical route to such forms is via the compound verb.

of putting, but at least *put* and *put up* are both verbs, and belong to the same conjugation class.

Moreover, the *showoff* type takes part in the same stress shift that many other verb to noun pairs show. Cruttenden (2001: 233-4) lists some thirty pairs of noun and verb where there is a stress difference between the two, the noun being stressed on the first syllable, the verb on the second. Most often, the stress difference is correlated with a difference in vowel quality as well. Examples are *'torment_N* vs. *tor^hment_V*, *'permit_N* vs. *per^hmit_V*, *'record_N* vs. *re^hcord_V* and so on. The difference between *show^hoff_V* and *'showoff_N* seems to follow the same rule (see also Selkirk 1982: 27).

Part of the reason that people seem to be reluctant to accept this analysis seems to be that the phrasal verbs are often considered to be syntactic constructions while the nouns are, perhaps for orthographic reasons, considered to be single lexemes. The difference seems to be terminological rather than substantive. Whether the verbs are lexemes or not, they are listemes (Di Sciullo & Williams 1987) or lexical items (Carstairs-McCarthy 2002: 13), just as the nouns are. This does not seem to be a suitable reason for suggesting that the path from one to the other cannot be a matter of word-formation.

3 The so-called possessive type

What are sometimes, rather misleadingly, called possessive compounds are also sometimes referred to as bahuvrihi compounds in a narrow sense, which means that the term 'bahuvrihi' is used ambiguously in the literature. I shall avoid using it any further here. I shall discuss three different sub-types that differ in the word-class of their left-hand element, but which clearly work in much the same way. A general discussion will follow the introduction of the three sub-types.

3.1 Adjective-noun

In this type we find an adjective (A) and a noun (N), combined with a single stress on the first element, used to denote an entity which has, possesses or is characterised by its A N. Examples are in (3).

- (3) flat-foot ('policeman', slang)
 greenback ('American dollar')
 hot-head ('impetuous person')
 redbreast ('European robin')
 red-cap (in Britain 'military policeman', in the US 'railway porter')
 red-eye (several meanings, including 'cheap whiskey').

The examples show several things.

- While descriptions of people are particularly common, names of animal and bird species are also frequent and even inanimates may be denoted by the type.
- The relationship between the unexpressed possessor or thing characterised and the expressed AN construction is rather vague.
- The existence of one of these with a particular meaning does not appear to block its existence with another, unrelated, meaning. This is important because lack of blocking is usually considered to be a sign of a productive process (Aronoff 1976: 45).

Jackendoff (1975: 657) suggests that all constructions of this type must be listed, because 'there is no way for a non-speaker of English to know that a *redhead* is a kind of person, but that a *blackhead* is a kind of pimple.' This seems to imply that the process by which these words are formed is not productive, yet we have just seen that there is contrary evidence. A solution for these words, which can solve this apparent discrepancy, will be presented below.

It should be noted that there are some expressions of this semantic type which do not meet the stress criterion. It is not clear whether stress is a sign of lexicalisation (see Bauer 2004) rather than a signal of the type, or whether the stress should be considered irrelevant in the definition of the type. Examples with phrasal stress are given in (4).

- (4) blue-bonnet ('Scotsman') (apparently with variable stress)
 green beret ('commando')
 red cross (medical organisation)

3.2 Noun-noun

In this sub-type, the first element of the expression is a noun, but otherwise the type works in the same way as the adjective-noun type illustrated above. Some examples are given in (5).

- (5) birdbrain ('idiot')
blockhead ('idiot')
egghead ('intellectual')
hatchback ('car whose boot and rear window form a single opening door')
paperback ('book in soft binding')
skinhead ('person — usually of a particular social type — with a shaven head')

This type seems to be fairly restricted, but nevertheless productive, as is witnessed by the name *Duckface* given to a character in the film *Four Weddings and a Funeral*.

There is a very similar type, which can be discussed alongside the examples in (5). It is illustrated by the example *butter-ball* 'fat person'. This is not a person who has a butter-ball, but one who resembles a butter-ball (whether *butter-ball* is taken literally as a lump of fatty material or as a plump game-bird, which is another meaning of the word). Although the semantic connection between the meaning of the compound and what it actually says is subtly different, the same points apply to examples like this as apply to those in (5).

3.3 Verb-noun

As is typically the case in English, it can be difficult to tell whether the first element in some compounds is a verb or a noun, given the predilection that English shows for conversion. Nevertheless, the examples in (6) appear to be similar to those given above, but with a verb in the first element.

- (6) scatterbrain ('person unable to concentrate')
Twinkletoes

3.4 Discussion

As others have pointed out (e.g. Plag 2003: 146), while these compounds are semantically exocentric in that a redcap is not a kind of cap nor a birdbrain a type of brain, they are nevertheless morphologically headed: in *redcap*, *red* modifies *cap*, and in *birdbrain*, *bird* modifies *brain*. Thus saying that these compounds are unheaded or have an external head is to that extent misleading.

It also misses an important point about the way in which language is used. Simple lexemes may also not denote the object to which they apparently refer, but instead denote an object connected with that ostensibly named. We call the phenomenon synecdoche. If we say, to use a classic example, *I saw a sail on the horizon* with the meaning 'I saw a ship on the horizon' we do not claim that *sail* is an exocentric lexeme, but rather that it is to be interpreted according to a well-known figure of speech. There seems no reason to treat the examples mentioned in this section in a different way. That being the case, all the examples cited here are endocentric compounds which happen to be interpreted figuratively.

This solution resolves the paradox from Jackendoff's observations. Figures of speech are productive, but some particular examples of the figure become established, when we might say that they become idioms or that the words involved become polysemous. Thus a syntactic tree is so-called by a metaphor, and *tree* becomes polysemous as a result of the establishment of the usage. So we can say that *redhead* and *blackhead* are idioms, but that the general figure of synecdoche is nevertheless productive.

4 Appositional compounds

The headedness of appositional compounds like *actor-manager* is in some doubt in the literature. For some authorities (e.g. Spencer 2004: 1264) they are double-headed. At the same time, Spencer lists these in a discussion of exocentric compounding, suggesting that exocentricity arises when it is difficult to take a decision on headedness. However, as Katamba (1993: 321) points out such compounds are syntactically headed, in that inflections are added to the right-hand edge of the word (**actors-managers*, *actor-managers*). This contrasts with the situation in, for example, Spanish as reported in Olsen

(2000: 912), where the plural of *actor-bailarín* 'actor-dancer' is *actores-bailarínes*, thus apparently confirming the dual-headedness of such constructions in Spanish. Unfortunately, the problem does not end there, since nouns like *cutthroat* are also inflected on the right-hand edge of the word, despite clearly being exocentric. The place of the inflection does not appear to be decisive, merely a default. The corresponding compounds in Dutch take their gender from the right-hand element (Booij 2002: 144), which seems like better evidence for headedness, but does not transfer neatly to English. The real question here is whether dual-headedness from a semantic point of view implies lack of headedness (or external headedness) from a syntactic one. In *actor-manager* there is no semantic external head as there is for *hatch-back* or *cutthroat* ('car' and 'person' respectively), and I would suggest that there is no need to analyse positionals as being exocentrics.

5 Some interestingly marginal cases

Bauer & Huddleston (2002) draw attention to some compounds which are usually classified as endocentric but which fail the hyponymy test. Although we may need to take care with the syntax for expressing these things, the fundamental point is that if something can be denoted by an endocentric compound of form XY it can also felicitously be denoted by Y alone. For example, a flourmill is a type of mill, grass-green is a type of green, to trickle-irrigate is to irrigate in a particular way. Examples in English where the relevant element of the compound is not the Y element are extremely rare (see Bauer & Renouf 2001), except in examples like *hanger-on*.

But while something that is grass-green is green, it is not true that something that is fat-free is free. So is *fat-free* an endocentric compound or an exocentric one? Neither is it true that a glow-worm is a worm, and the same question applies.

These two examples illustrate different problems. In *fat-free*, *free* can be used in the relevant sense, but only if it has a complement. The complement is normally present in the form of a prepositional phrase (*free of fat*), but can also be preposed in the form *fat-free*. Since adjectives with postposed complements cannot be used in pre-modifying position (**a free of fat meal*), preposing the complement is a necessary step if an attributive adjective is to be created (*a*

fat-free meal). The fact that this is a transformation of a usually syntactic structure into a morphological one suggests that *free* in *fat-free* should not be equated with *free* meaning 'gratis' or 'at liberty', the meanings *free* can have in isolation, but only with *free* in *free of fat*, and at that point it would seem that the construction is endocentric.

The *glow-worm* case is a matter of pragmatics and language history rather than a matter of hyponymy in the modern language. The meaning of the word *worm* has changed in the period since *glow-worm* was coined (*worm* meant 'insect' until long after that period). While *glow-worm* is no longer strictly endocentric, it was composed as an endocentric compound. An apparently similar case of German *Walfisch* 'whale fish = whale' is not actually parallel. There, the 'fish' element is simply a folk taxonomy which is scientifically incorrect.

It seems that many such cases can be treated similarly: they show lexicalisation; they show grammatical restructuring; they show a figurative interpretation. In none of these cases does it seem entirely appropriate to talk of these as being exocentric compounds.

6 Complex pre-modifiers

There are a number of exclusively pre-modifying constructions which look like compounds in that they are made up of two lexemes. These pre-modifiers show a number of different patterns, some of which are illustrated in (7) (examples from Bauer 1983, Bauer & Renouf 2001).

- (7) before-tax (profits)
 oestrogen-only (pill)
 pass-fail (test)
 quick-change (artiste)
 red-brick (university)
 roll-neck (sweater)
 wrap-around (skirt)

The question that such examples raise is whether they are to be analysed as compound adjectives (as is done in Bauer 1983), or whether some other analysis is possible. Note that none of these examples has an adjectival head, so that if they are compound adjectives they are all exocentric.

The reason that these items look like adjectives is that they occur in attributive position. However, to equate attributive position with adjectival status is to confuse form and function. These certainly function as pre-modifiers, but that does not imply that they are adjectives. These items are classifiers rather than epithets, and do not permit the usual range of sub-modifiers which occur with run-of-the-mill adjectives, nor are they usable as bases for deriving deadjectival nouns, verbs, etc. Examples of adverbs, prepositions and verbs in pre-modifying position are also found, as shown in (8), underlining the point that function and form do not necessarily match (examples from Klinge 2005: 343).

- (8) an inside toilet
 the must-see show of the season
 an off day
 the then king

An alternative analysis of some of the examples in (7) presents them not as adjectives but as pieces of syntax which have been captured as units and rank-shifted to be used as words. This option seems to be independently required for examples like *an if-you-really-want-to-know sneer*. Such an analysis seems to work well for *before-tax*, for example, where we could say *I paid \$3000 before tax*, but is less convincing for *roll-neck*, where there is no syntactic construction *roll neck* to be captured by the morphology. Thus at least in some cases it seems that we must see these constructions as morphological rather than syntactic.

A third analysis is to see all these examples as particularly complicated compound-types. Given that a noun is generally said to be able to take any part of speech as its modifier in a compound (see e.g. Selkirk 1982: 16), we would expect to find instances in which complex forms of any word-class are found as the modifying element in compounds. There are problems here with items such as *before-tax* and *red-brick* which appear to have phrasal rather than lexical structure, though the notion that compounds and syntactic constructions may interact with each other is familiar from discussions such as Chomsky & Halle's (1968: 22) coverage of examples like *black board eraser* and *American history teacher*. It does seem to be important that DPs are not

permitted in pre-modifying position in such instances: a construction like **The [the red panda]_{DP} cage* is completely impossible.

Although we might say that *pass-fail test* is then a VN compound, equivalent to an example like *hovercraft*, we do have the problem that *pass-fail* does not seem to be a possible V in English (and neither is it a possible N if we should believe that *pass-fail test* is NN, as suggested by *The Oxford English Dictionary*). It is not clear how big a problem this is, in terms of the proportion of such constructions which show patterns which could not occur in isolation. *Oestrogen-only*, discussed at some length in Bauer & Renouf (2001), is another such example. However, it must be pointed out that the precise nature of what is and is not permitted in modifying position, and how this can be guaranteed by a grammar, is not well understood. Consider, for example, the instances in (9), where both surprisingly acceptable and surprisingly unacceptable compounds are listed.

- (9) *a completely/back in-group
 *a dark red-brick university (with full AP)
 These range from small, few-person placements to large-scale missions
 (<http://www.ausaid.gov.au/partner/ozgov.cfm>)
 *in-supersonic-flight service (with AP)
 many-person games (with a QP, not submodified)
 a pain-in-stomach gesture (not a syntactic constituent, nor an idiom)
 *a stably hovercraft (with AdvP)
 a two-man boat (singular noun with plural quantifier)

In Bauer & Renouf (2001), the conclusion is drawn that we need a series of exocentric adjectival compounds to account for the attested data here. That may be so, but in these cases there is the potential for alternative analyses, even if none of them in isolation appears totally convincing. This is an area where more research is needed.

7 Out-compounding the compounds

Bauer & Renouf (2001) cite the example of *OJ [Simpson] outsoaped the soaps*, and point out that this construction goes back at least as far as Shakespeare, since Hamlet says he will *out-Herod Herod*. This particular pattern is productive, at least in journalistic prose, and has a long history. *The Oxford*

English Dictionary gives many examples, both with proper nouns (*out-Auden Auden*) and with common nouns (*out-hawk the hawks*), with the cognate direct object as above and without it (*outsavour rosemary*).

If these forms are to be interpreted as compounds, they would seem to be exocentric ones. However, it is feasible to interpret *out-* in these words not as a particle but as a prefix which acts as the head of the verb, just as *un-* does in *unhorse* and *en-* does in *encage*. Again, therefore, there is an alternative analysis to the exocentric one.

8 Discussion and conclusion

The one really clear case of exocentric compounding we have in English, the Romance type, is no longer productive. In all the other cases there is some doubt as to whether an analysis as an exocentric compound is the best one or not. Some of the alternatives are rather more far-fetched than others. It is clear that the complex pre-modifiers cannot be dismissed out-of-hand and that more research is required before any firm conclusion can be drawn about them. But it is tempting to suggest that English speakers are reluctant to deal with exocentric compounds at all, and that this may be part of the reason why the Romance type like *pickpocket* did not remain productive in English, while it thrived in Romance. Certainly, I would suggest that I have raised enough queries about English exocentric compounds in this paper to indicate that analysts should be rather less ready than they have been to assign English compounds to various classes of exocentrics. I would like to be able to suggest a really clear-cut conclusion: speakers of English do not coin exocentric compounds as a method of *Wortbildung*; any exocentric compounds in English are analysed as such after the event, from the point of view of *Wortgebildetheit*. Unfortunately, I do not believe that we can be that certain at the moment because of the complex pre-modifiers. The best we can say is that speakers of English do not productively coin exocentric compounds as head nouns. Whether the things we find as dependents modifying head nouns have to be treated as exocentric compounds is something which needs to be considered seriously. Further research in that area would be extremely welcome, and might allow the wider conclusion to stand.

References

- Aronoff, Mark 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Bauer, Laurie 1980. Deux problèmes au sujet des noms composés comprenant un premier élément verbal en français moderne. *Le français moderne* 48: 219-224.
- Bauer, Laurie 1983. *English Word-formation*. Cambridge: Cambridge University Press.
- Bauer, Laurie 2004. Adjectives, compounds and words. In *Nordic Journal of English Studies* 3/1 (= *Worlds of Words: A tribute to Arne Zettersten*): 7-22.
- Bauer, Laurie & Rodney Huddleston 2002. Lexical word-formation. In Rodney Huddleston & Geoffrey K. Pullum, *The Cambridge Grammar of the English Language*, Cambridge: Cambridge University Press, 1621-1721.
- Bauer, Laurie & Antoinette Renouf 2001. A corpus-based study of compounding in English. *Journal of English Linguistics* 29: 101-123.
- Booij, Geert 2002. *The Morphology of Dutch*. Oxford: Oxford University Press.
- Bloomfield, Leonard 1935. *Language*. London: Allen & Unwin.
- Carstairs-McCarthy, Andrew 2002. *An Introduction to English Morphology*. Edinburgh: Edinburgh University Press.
- Chomsky, Noam & Morris Halle 1968. *The Sound Pattern of English*. New York: Harper and Row.
- Cruttenden, Alan 2001. *Gimson's Pronunciation of English*. 6th edn. London: Arnold.
- Di Sciullo, Anna Maria & Edwin Williams 1987. *On the Definition of Word*. Cambridge, MA: MIT Press.
- Dokulil, Miloš 1968. Zur Theorie der Wortbildung. *Wissenschaftliche Zeitschrift der Karl-Marx-Universität Leipzig, Gesellschafts- und sprachwissenschaftliche Reihe* 17, 203-211.
- Gonda, Jan 1966. *A Concise Elementary Grammar of the Sanskrit Language*. University AL: University of Alabama Press.

Jackendoff, Ray 1975. Morphological and semantic regularities in the lexicon. *Language* 51: 639-671.

Katamba, Francis 1993. *Morphology*. Basingstoke: Palgrave.

Klinge, Alex 2005. The Structure of English Nominals. Unpublished thesis presented for the degree of doctor linguae mercantilis from the Copenhagen Business School.

Marchand, Hans 1969. *The Categories and Types of Present-Day English Word-Formation*. 2nd edn. Munich: Beck.

Olsen, Susan 2000. Composition. In Geert Booij, Christian Lehmann & Joachim Mugdan (eds), *Morphologie/Morphology* Vol 1, Berlin and New York: de Gruyter. 897-916.

Plag, Ingo 2003. *Word-Formation in English*. Cambridge: Cambridge University Press.

Selkirk, Elisabeth O. 1982. *The Syntax of Words*. Cambridge, MA: MIT Press.

Spencer, Andrew 2004. English (Indo-European: Germanic). In Geert Booij, Christian Lehmann, Joachim Mugdan & Stavros Skopeteas (eds), *Morphologie/Morphology* Vol 2, Berlin and New York: de Gruyter. 1255-1267.

