The formation of compound nouns in English

0. Introduction

This paper deals with compound nouns, e.g. milk man, death notes, Groucho Marx joke.

Compounds are usually defined as words made up of two independent bases, one of which is a head (a flag pole is a type of pole, not a type of flag) (Bauer 2009: 343 and 348, Bauer and Huddleston 2002: 1630-1631, Katamba 1994: 264, Plag 2003: 5). Compounding is one means of word-formation, amongst others, namely derivation (affixation or suffixation), conversion (to smile – a smile) – or more marginal means ones such as initialism (CIA for Central Intelligence Agency), acronyms (NATO for North Atlantic Treaty Organization), clipping (lab for laboratory), etc.

There are several types of compound nouns. (See Bauer (1983: 202ff and 2009: 343-356) for a typology, which is slightly different from the one presented here.)

- Noun + noun, which is the largest subgroup (e.g. boyfriend, woman doctor, manservant).
- Verb + noun (hovercraft, playpit). Let us note that Bauer (1983: 205) includes cut-throat, kill-joy, pickpocket, spoilsport, breakfast in this pattern, but that these examples may also be treated as conversions (to cut a throat → a cuthroat, to kill joy → a kill joy, etc.).
- Adjective + noun (fast food, software). Blackbird is a compound (a kind of bird) but black bird is not (it is a bird which is black). Let us note that it can be difficult to decide whether a given adjective + noun sequence is a compound or not. One criterion to distinguish between the two is stress: there is one primary stress in a compound, whereas an adjective + noun sequence has two. The word blackbird is a compound (a kind of bird) but black bird is not (it is a bird which is black).
- Preposition + noun, such as in-crowd or off-islander.
- Adverb + noun. This is a pattern which is not very productive, as it only works with adverbs of time or place, e.g., now generation.
- Dephrasal compounds, such as lady-in-waiting, son-in-law, love-in-a-mist, Darby and Joan (in that Darby and Joan feeling for instance).

Here, I only deal with noun + noun compounds.

I exclude from this study blends, e.g. chunnel (channel + tunnel), as they do not lend themselves to a clear analysis into morphemes (Bauer 1983: 234), and learned (or ‘neo-classical’) compounds, such as biology. Although these so-called neo-classical formations may be treated as compounds on the grounds that both components have a clearly identifiable

---

1 It has to be acknowledged, however, that some compounds (such as singer-songwriter), have no obvious head (Bauer 2009 : 349).
semantic content (bio- means ‘life’ and -logy means ‘science’ or ‘discourse’), compounds are typically made up of two free (rather than bound) bases. For instance, in milkman, we recognize milk and man, which may occur independently (there is some milk left, there is a man round the corner).

I have included, on the other hand, ‘synthetic compounds’, e.g. bus driver, although there is some discussion as to their status (Bauer 2009: 353, Lieber 1983: 251-286, Lieber 2004: 48). Bus driver can be seen as the result of a derivation process (bus drive + -er), as driver is derived from a verb (drive) and bus is an argument of that verb, but it can also be seen as the result of compounding (bus + driver). In fact, it may be the case that both morphological analyses are valid. Bus driver is indeed derived from drive (a) bus through a process of suffixation; the result (bus driver) is then reinterpreted as being made up of two bases, and is accordingly felt as a compound.

I also take into account all semantic types of compounds (Bauer 1983: 30-31, Bauer 2009: 353, Bauer, Lieber and Plag 2013: 463-483, Dressler 2006: 23-44), i.e., endocentric (milkman) and exocentric (birdbrain) (Bauer 2008a), including those that do not have an obvious head (coordinate compounds such as murder-suicide or singer-songwriter). (These are sometimes called ‘dvandvas’. See Bauer (2008b) for a discussion of that term and the various cases it covers.)

Compounds encapsulate a vast number of relations between the two bases (Bauer 2009: 353, Downing 1977, Hatcher 1960, Levi 1978, Warren 1978), and all types are included in this study. For example, a sun cream blocks the sun’s rays, a face cream is applied to the face, a hormone cream contains hormones, and a rash cream is meant to cure a rash.

In fact, one compound may encapsulate several semantic relations, as illustrated in the following extract, where it is clear that hash bar synthesizes the following relations: bar where hash is legal, bar where you smoke hash, bar where you can buy hash, bar where you can sell hash.

JULES Okay, so, tell me about the hash bars.
VINCENT Okay. What you want to know?
JULES Hash is legal there, right?
VINCENT Yeah, it’s legal, but it ain’t a hundred per cent legal. I mean, you just can’t walk into a restaurant, roll a joint, and start puffin’ away. I mean, they want you to smoke in your home or certain designated places.
JULES Those are hash bars?
VINCENT Yeah, it breaks down like this: it’s legal to buy it, it’s legal to own it and, if you’re the proprietor of a hash bar, it’s legal to sell it. It’s legal to carry it, but, but, but that doesn’t matter ‘cause - get a load of this all right - if the cops stop you, it’s illegal for them to search you. I mean, that’s a right the cops in Amsterdam don’t have.
(Tarantino 1994: 13-14)

I also take into account both lexicalised (milkman, postman, bedside) and non-lexicalised compounds (Truman Capote look alike contest, death notes). Some compounds are indeed formed by a particular speaker, on a particular occasion. Compounding is an extremely productive means of word-formation in English (Benczes 2006).
Lastly, I include both early stress (e.g. *farm hand*) and late stress (e.g. *baby brother*) compounds. These two sub-groups are identified by e.g. Boisson (1980), Chomsky and Halle (1968), Halliday (1970: 44-47), Huart (1984 and 1989), Jespersen (1909: §5.22 and 1933: §8.51), Quirk et al. (1985: §17.104), Marchand (1955: 216-227), and Viel (1981: 97-98).

In fact, the main object of this paper is to revisit the distinction between early stress compounds and late stress ones, and to show that only the early stress ones should be treated as nouns. Indeed, early stress compounds most closely resemble prototypical nouns (simple nouns) because they form a unit that cannot be interrupted.

My examples come from a corpus that consists of 1718 compound nouns (1097 single stressed and 621 double stressed), all found in context (novels, plays, film scripts, comics).

I first look at late stress compounds and aim to show that they have much in common with adjective + noun syntactic sequences. Then I deal with early stress compounds and argue that they represent the endpoint of a nominalization process.

1. Late stress compounds

Late stress compounds dress (*silk scarf, kitchen sink, underfoot gravel, back door*) differ form ‘real’ compounds, as noted by Huart (2010: 184). Contrary to early stress compounds, they are not to be treated as nouns. They are sometimes called ‘pseudo-compounds’, and there is a number of arguments for maintaining that late stress compounds behave as adjective + noun constructions. Indeed, *a cotton dress* is not very different from *a pretty dress*. One of those arguments is the adjectival meaning of the first noun.

1.1 Adjectival meanings

The meanings of the first nouns that occur in late stress compounds can be compared to those of adjectives, which (in a cross-linguistic perspective) typically denote dimension (*big, small*, etc.), age (*new, young, old*, etc.), value (*good, bad, lovely, atrocious*, etc.) colour (*black, white, red*, etc.) (Dixon 1982, Dixon and Aikhenvald 2004: 3-4). These adjectival meanings can be found in the following compounds:

\[
\begin{align*}
\text{a ten-ton behemoth} & \text{ (dimension), to be compared to e.g. a heavy animal} \\
\text{half-hour commercials} & \text{ (dimension) commercials, to be compared to e.g. long commercials} \\
\text{child man, baby sister} & \text{ (age), to be compared to e.g. immature man, younger sister.} \\
\text{second rate mystery novels} & \text{ (value) to be compared to e.g. bad mystery novels.}
\end{align*}
\]

In addition to these typical meaning, in English, adjectives often denote shape, origin and material (Cotte 1996: 135). The first nouns of late stress compounds also convey these meanings:

\[
\begin{align*}
\text{handlebar moustaches, lace curtain lungs} & \text{ (shape), to be compared with e.g. styled mustache, perforated lungs.} \\
\text{the Disney adaptation} & \text{ (origin), to be compared with e.g. my early portraits, the French adaptation.} \\
\text{a glass jar, an iron door, a log cabin, a metal stand, a steel cube} & \text{ (material), to be compared with e.g. a wooden spoon. These are very common.}
\end{align*}
\]
Another characteristic of these noun + noun constructions is the possible predicative use of the first noun.

1.2 Predicative use of the first noun

The first element of a compound can be used in predicative position, in the same way that an adjective can. Jespersen (1914: §13.61 and 13.61) noted this characteristic, and quotes the following (attested) examples: *our cousin was extremely low church* (compare with a low church person), *Mr Benson seems the most eighteenth-century of our later candidates for Parnassus* (compare with an eighteenth century novel), *measures looked upon at one time as purely labour are being passed by the Liberal Government* (compare with purely Labour measures, a Labour government).

In the following list, all compounds can be paraphrased with a sentence in which the first element occurs in predicative use.

- a crystal-set the set is crystal
- a glass jar the jar is glass
- the iron door the door is iron
- a leather travelling case the travelling case is leather
- out-of-date atlases the atlases are out of date
- paper feathers the feathers are paper
- plastic poker chips the poker chips are plastic
- a second-or third-hand children’s book the books are second (or third) hand
- second-rate mystery novels the mystery novels are second rate
- the seventeen-hundred line poem that poem is seventeen hundred lines
- a slate roof the roof is slate
- that steel cube the cube is steel
- a twenty-one inch telly the telly is twenty one inches

We may note that in these cases, the first noun often denotes a material (metal desks) or a dimension (a twenty-one inch telly).

1.3 Co-occurrence of the first noun with an adjective

A third sign that the first noun is close to an adjective (i.e., that the construction is not really a compound) is the possible co-occurrence of the first noun with an adjective. Jespersen gives the following examples: *her Christian and family name, He got into money and other difficulties, a Boston young lady* (1933: §8.52), *postal and telephone services, the lonely, bare, stone houses, all national, state, county and municipal offices, his silk high hat* (1914: §13.31, 13.32 and 13.33.).

In my corpus the following examples are to be found.

- the five-ton carnivorous lizard
- a tailor-made fifty-shilling grin
- a steady jog-trot rhythm
- the late-night final fog
- my chronic Los Angeles nausea
- some gray metal desks
- cold metal tools
a radiant New York June
a hard plastic dress-up pocketbook
a minuscule sixth floor maid’s room
those same sixth floor windows
those long summer days
a beautiful summer morning
that mysterious, three-month journey

1.4 An adverb before the first noun

One last characteristic of late-stress compounds is the possibility to find an adverb before the first noun. (If the noun can be modified by an adverb, it behaves as an adjective). This is actually quite rare, but Jespersen (1933: §8.54) provides the following (attested) examples.

on merely business grounds
a division on strictly party lines
from a too exclusively London standpoint

After this description of late stress compounds, let us turn to early stress ones, and see what they can tell us about nouns as a part of speech. Indeed, early stress compounds form a unit, and contrary to late stress ones, can be treated as nouns.

The aim is to show that early stress compounds result from the syntactic and semantic transformation of a preceding (in most cases theoretical) construction in which the two bases of the compound are combined.
In some cases the relation between the two nouns is prepared in the preceding co-text. We look at these first.

2. Early stress compounds resulting form a preconstruction in the preceding co-text

In the following example, the compound death notes occurs (and is interpretable) because the preceding co-text includes elements that make the relation between the two components explicit.

More than five years earlier, shortly after moving into the apartment on Riverside Drive, he had translated a number of the fragments Mallarmé wrote at the bedside of his dying son, Anatole, in 1879. These were short notes of the greatest obscurity: notes for a poem that never came to be written. They were not even discovered until the late 1950’s. In 1974, A. had done rough translation drafts of thirty or forty of them and then had put the manuscript away. When he returned from Paris to his room on Varick Street (December 1979, exactly one hundred years after Mallarmé had scribbled those death notes to his son), he dug out the folder that contained the handwritten drafts and began to work up final versions of his translations. These were later published in the Paris Review, along with a photograph of Anatole in a sailor suit. From his prefatory note: « On October 6, 1879, Mallarmé’s only son, Anatole, died at the age of eight after a long illness. The disease, diagnosed as child’s rheumatism, had slowly spread from limb to limb and eventually overtaken the boy’s entire body. For several months

2 Emphasis mine.
Mallarmé and his wife had sat helplessly at Anatole’s bedside as doctors tried various remedies and administrated unsuccessful treatments.

(Auster, 1982: 109)

The compound death notes is preceded by the following elements of preconstruction: the fragments Mallarmé wrote at the bedside of his dying son, Anatole, in 1879. These were short notes of the greatest obscurity: notes for a poem that never came to be written. The relation between the two nouns of death notes is therefore established in the preceding co-text and the compound noun is posterior to this initial relation. The compound is an end result, it could not have occurred without what precedes it: *Mallarmé wrote some death notes.

This is by no means exceptional, as illustrated in the following examples.

And it was when these cops were chasing the crooks that we played some good tricks with the telly, because when one of them opened his big gob to spout about getting their man I’d turn the sound down and see his mouth move like a goldfish or mackerel or a minnow mimicking what they were supposed to be acting – it was so funny the whole family nearly went into fits on the brand-new carpet that hadn’t found its way to the bedroom. It was the best of all though when we did it to some Tory telling us about how good his government was going to be if we kept on voting for them – their slack chops rolling, opening and bumbling, hands lifting to twitch moustaches and touching their buttonholes to make sure the flower hadn’t wilted, so that you could see they didn’t mean a word they said, especially with not a murmur coming out because we’d cut off the sound. When the governors of the Borstal first talked to me I was reminded of those times so much that I nearly killed myself trying not to laugh. Yes, we played so many good stunts on the box of tricks that mam used to call us the Telly Boys, we got so clever at it. (Sillitoe 1959: 22-23)

In the following extract, the content of a joke is described.

ALVY ( . . . ) The-the other important joke for me is one that’s, uh, usually attributed to Groucho Marx, but I think it appears originally in Freud’s Wit and its Relation to the Unconscious. And it goes like this. I’m paraphrasing: Uh ... ‘I would never wanna belong to any club that would have me for a member.’ That’s the key joke of my adult of my adult life in terms of my relationships with women. (Allen 1983: 4)

The compound noun occurs after the content of the joke has been explained.

ALVY Oh, my God! (Then, to the camera) She’s right! Why did I turn off Allison Portchnik?! She was beautiful. She was willing. She was real ... intelligent. (Sighing) Is it the old Groucho Marx joke? That-that I-I just don’t wanna belong to any club that would have me for a member? (Allen 1983: 22)

In some cases, a compound follows an of-construction, even though in most cases there is a clear meaning difference between compounds and of noun phrases (a coffee cup and a cup of coffee).

In twenty minutes’ time Carl returned alone, changed into a fresh T-shirt, and hunkered down on the front porch with his head in his hands.

She came out with a cup of coffee in each hand and sat down next to him, on his right. ‘What did he want?’ she asked.

‘Nothing,’ answered Carl. ‘We had some things to talk about. Nothing much. No big deal.’

Susan Marie handed him a coffee cup. ‘It’s hot,’ she said. ‘Be careful.’

‘All right,’ said Carl. ‘Thanks.’

‘I made him some,’ said Susan Marie. ‘I thought he was going to stay.’
‘It was nothing,’ said Carl. ‘It's a long story.’
(Gutterson 1994: 258-259)

The compound coffee cup can occur because the cup of coffee has been mentioned before.

All compound nouns are obviously not preceded by explicit elements of preconstruction. However, it can be argued that all compound nouns are the final stage of a nominalization process. Some noun + noun compounds result from the semantic and syntactic transformation of a theoretical construction.

3. Early stress compounds resulting from the transformation of a theoretical construction

I first look at the semantic side of the transformation. In what sense can it be said that compounds semantically re-elaborate syntactic constructions? There are different (not mutually exclusive) cases, which I list. In a last sub-section, I look at the syntactic side of the transformation.

3.1 A theoretical / potential (backgrounded) process

Compound nouns such as repair man or race horse encapsulate a reference to potential processes, which are backgrounded. A repair man normally repairs things on a regular basis, and a race horse races regularly. The fact that the agent (man, horse) is involved in a number of similar processes (repairing, racing) contributes to the creation of a category, i.e. a type of man, a type of horse. This is why, even if the processes of repairing or racing are not valid at the time of speaking, a man can still be called a repair man, and a horse, a race horse. The underlying predicative relation (/a man - repair something/, /a horse - race/) is not mentioned for its own sake but is backgrounded. This holds for many occupational denominations, such as cleaning woman, garbage man, water man, baseball player, camera men, gas man.

In the examples given above, the entity determined by the process is a participant in that process (more specifically, the agent). But in other cases, it may be a circumstance, as in e.g. coffee table (the coffee is served on the table) or lunchtime (people eat at a certain time). In the last example, the recurrent, collective process of eating leads to the creation of a category (a type of moment). The situation of eating is not mentioned for its own sake but because it characterizes the moment when it is likely to occur.

3.2 Artists and their works

Some other compound nouns illustrate a shift from the expression of authorship to that of a category.

FEMALE REPORTER I think there are more people here to see the Maharishi than there were to see the Dylan concert.
(Allen 1983: 65)
The compound *Dylan concert* may be compared with a genitive construction, *Dylan’s concert*. The shift from genitive to compound is most likely to occur when the first noun is a proper name. The semantic difference between the two is that a *Dylan concert* denotes a concert that has some special characteristics (and this is most likely to happen when the proper name denotes a famous person), whereas *Dylan’s concert* does not have that connotation, it merely means that the concert was given by Bob Dylan. The compound therefore adds a layer of meaning to the genitive: 1) the concert is given by Dylan, and 1) for that reason it is in some way special.

### 3.3 Oppositions

A compound noun may result from an opposition between two entities. If similar entities (e.g. rooms) serve different purposes (e.g. sleeping or washing), it may be grounds for opposing them, and creating sub-categories. For instance, if you use rooms for different purposes, e.g. sleeping or washing, it creates types of rooms, i.e. bedrooms and bathrooms. Those compounds nouns therefore add a semantic component to syntactic expressions such as *someone sleeps in a room*, or *the room in which someone sleep*. The compound denotes a *type* of room, which has some typical characteristics, and which still may be called that even if there is no bed in it.

These oppositions may be apparent in the context, as in the following extract.

They can drop all the atom bombs they like for all I care: I'll never call it war and wear a soldier's uniform, because I'm in a different sort of war, that they think is child's play. The war they think is war is suicide, and those that go and get killed in war should be put in clink for attempted suicide because that’s the feeling in blokes’ minds when they rush to join up or let themselves be called up. I know, because I’ve thought how good it would be sometimes to do myself in and the easiest way to do it, it occurred to me, was to hope for a big war so’ I could join up and get killed. But I got past that when I knew I already was in a war of my own, that I was born into one, that I grew up hearing the sound of ‘old soldiers’ who’d been over the top at Dartmoor, half-killed at Lincoln, trapped in no-man’s land at Borstal, that sounded louder than any Jerry bombs. Government wars aren't my wars; they've got nowt to do with me, because my own war's all that I'll ever be bothered about.

(Sillitoe 1959: 16)

In this context, several types of wars are contrasted: those fought by governments, and the one the speaker sees at work in society (the ruling class vs. the working class). There is an explicit contrast between *government wars* and *my own war*. The compound *government wars* encapsulates the fact that governments wage wars, but add something to that, namely the fact that this is only one type of war, amongst others. There is again a (partial) semantic transformation of the syntactic expression; one layer of meaning is added.

### 3.4 Shift from a possession meaning to the representation of a type

A compound noun may also result from the shift from a possession meaning to the representation of a new, sometimes *ad hoc*, category.

I still don’t think giving her « *Bambi eyes* » is going to get you a flame thrower.

(Waterson 1990: 31)
The compound *Bambi eyes* includes a backgrounded reference to a possession relation, which could be expressed with a genitive construction, *Bambi’s eyes*. However, the compound does not literally denote the eyes of the animal called Bambi (in the famous cartoon), but a moving look - in this situation, the look of a little boy who wants to get something from his mother. The eyes belong not to Bambi but to the little boy, and are a type of eyes (or look). There is a shift of meaning form possession to categorization.

This probably accounts for the difference between *a sailor suit* (compound) an *a soldier’s uniform* (genitive). Both denote a type of clothes, but the difference between the two is that only soldiers wear soldiers’s uniforms, whereas anyone (for example children) can wear sailor suits. The possession meaning is backgrounded, and the word *sailor* is meant to indicate a sub-category of clothes (including a striped top, for example) rather than denote any actual possessor. Again there is a shift in meaning - from actual possession to the expression of a category.

### 3.5 Entities that are inherently linked

We have not accounted for such compounds as *raindrop* or *snowflake*, which we may compare, respectively, with *a drop of water*, and *a heap of snow*. Why do we have a compound in the first case, and an *of-*construction in the second one?

One major difference between a raindrop and a drop of rain, from a referential point of view, is that rain necessarily comes in the form of drops, whereas that is not the case for water: we may have glasses of water, buckets of water, drops of water, etc. These various quantities of water are denoted by *of* constructions, which indicate the separation of a bounded quantity from an unbounded entity (*a cup of tea, a glass of water*).

We have a compound such as *raindrop* when the operation of separation is not relevant. The shape assumed by rain is spontaneously that of drops, and the extraction process is therefore skipped.

### 3.6 A syntactic transformation

We have seen that in all cases, compounds represent the transformation of a pre-conceived relation, and add an element of meaning to the (sometimes potential) syntactic expression of that relation. We can therefore assume that, from a syntactic point of view, compounds re-elaborate preceding constructions, be they theoretical, in which the relation would be explicit.

Those preceding constructions may be e.g. genitives (*a sailor’s suit* → *a sailor suit*), i.e. constructions which may themselves be seen as the result of a transformation, more specifically, a nominalization (*the sailor has a suit* → *a sailor’s suit*). A compound therefore cannot be seen as the transformation of a clause, but as that of a noun phrase, which may or may not be the result of a nominalization (*a sailor’s suit* is the result of a nominalization, but *a drop of rain* is not).

The noun phrases that can be transformed into compounds are the following:

- Qualifying gentives, e.g. *a sailor’s suit* → *sailor suit*,
- Determinative gentives (*Bambi’s eyes* → *Bambi eyes*),
Double genitives (*A movie of Fellini’s* → *A Fellini movie*),
Noun phrases that include a relative clause (The time when someone / people have
their lunch → lunch time).
of-constructions (*a bag of nail* → *a nail bag*).

In the last case (the *of*-construction), the preceding construction may be theoretical and
necessarily skipped (the syntactic alternative is not viable), as in the case of *raindrop*, but that
is not necessarily the case, as can be seen with *nail bag* in the following extract.

I’m nowhere near puffed despite *that bag of nails* that rattles as much as ever, and I can still give a
big last leap like galeforce wind if I want to, but everything is under control and I know now that
there ain’t another long-distance running runner in England to touch my speed and style. (…)
And down the drive I went, carrying a heart blocked up like Boulder Dam across my arteries, the
*nail-bag* clamped down tighter and tighter as though in a woodwork vice, yet with my feet like
birdwings and arms like talons ready to fly across the field except that I didn’t want to give
anybody that much of a show, or win the race by accident.
(Sillitoe 1959: 49-50)

**Conclusion**

Noun compounds, at least noun + noun compounds, and perhaps other types as well, can be
seen as the syntactic and semantic transformation of a noun phrase, which may or may not be
found in the preceding co-text.

The reason for positing a syntactic transformation is that there is a semantic re-
elaboration, ³ i.e. a difference in meaning between the syntactic construction and the
compound. A compound encapsulates a relation that could be expressed in a syntactic
construction and adds something to the meaning of that construction. In *Bambi eyes*, for
example, we are reminded that Bambi has eyes, and the compound contains an additional
element of meaning, i.e. the fact that those eyes are special. This means that they can be
treated as a category, which can then characterize other entities than Bambi (a little boy, in
our context). Meaning can be viewed as a dynamic process; it can be constructed. In this
example as in others, the meaning of the compound is posterior to that of the syntactic
construction.

The major difference between compound nouns and their syntactic counterparts is that
compound nouns belong to the lexical category of nouns. Even if they are nonce compounds,
i.e. not part of the lexicon, they are treated as nouns, and, as such, serve a categorization
purpose. With nouns we move away from the representation of a specific situation to that of a
category. For instance, *lunch time* does not denote the moment when a specific person has
lunch, but the time when people in general (typically) have lunch. It covers more situations
than the syntactic expression would, and denotes a type of moment (a category). This is what
nouns do. The study of compound nouns therefore helps us understand the nominal class as a
whole.

³ On the notion of « re-elaboration », see Cotte (2000: 214).
References


References of the examples


\(^4\) The scripts used are *Annie Hall* (1-109) and *Manhattan* (179-276).