ADJECTIVES, COMPOUNDS, AND WORDS

Laurie Bauer, Victoria University of Wellington, New Zealand

English spelling is, as everyone knows, full of peculiarities. Those peculiarities account for George Bernard Shaw's argument that 'fish' could be spelt *ghoti,* the absurd claim that *York-Los* appears as a 'word' in the larger construction *the New York-Los Angeles flight,* and the amusement provided by sentences like:

The dough-faced ploughman coughed and hiccoughed his rough way through Scarborough.

Linguists take as given the primacy of the spoken word and the derivative nature of written language. We are aware of absurdities like those mentioned above, but see them as artifacts of the spelling system and dismiss them as being of marginal relevance to the structure of English. Yet at other times we appear to have difficulty in discarding the idea that English orthography tells us something important about the language. In this article, I should like to discuss one such instance.[1]

However misleading English spelling may be on occasions, there is one place where it seems to match our intuitions perfectly. The description *a black bird* has *black* and *bird* in two orthographic words, whereas the naming function illustrated by a compound form such as *a blackbird* is a single orthographic word, and no longer a series of two. We can find these intuitions justified in the literature. *Black* in *blackbird* is no longer available for syntactic or morphological modification (we cannot have **a rather black-bird,* nor *a blackerbird). This indicates that it is not a full word in its own right. *Blackbird* carries stress on the left-hand element of the compound. This is sometimes called 'compound stress' in the literature, but for reasons which will become clear, it is, in this article, referred to as 'first-element stress'. This points out the difference between the word and the

phrase (where nuclear stress tends to fall on the rightmost element in what we can, for the purpose of this discussion, term 'second-element stress'). The meaning of *black bird* can be deduced from the meaning of its elements and the meaning of the construction, while the meaning of *blackbird* cannot be entirely predicted from the meaning of the elements (if that were possible, a sentence such as *I saw a brown blackbird this morning* would be nonsensical, which is not the case). This means that *blackbird* must be a dictionary entry, and in that sense is a lexical item (a term used in this article in preference to the alternatives 'listeme' and 'dictionary word'[2]). We thus have a strong set of coincidences, which match our intuitions, as set out in Table 1. Indeed, the whole pattern of Table 1 appears so convincing that it may seem odd to bring up the matter at all in this context.

Table 1

The evidence for blackbird *as a word*

black bird	blackbird
Second-element stress	First-element stress
Independent elements, each of which can be inflected	First element dependent, inflection belongs to the unit as a whole
Meaning predictable from the elements	Meaning not entirely predictable from the elements, so must be listed
Each element is a separate lexical item	The unit as a whole is a lexical item
Sequence of two orthographic words	Single orthographic word
Conclusion: a phrase	*Conclusion*: a word

If all examples were like this, there would be no problem; a problem does arise, though, with the notion that this example is in some way typical of English. To show this, we need to see how general or how limited the pattern illustrated in Table 1 is.

We can start with the observation that the number of adjectives that work in the way that *black* does in our *exemple-type* seems to be very restricted. If we require exactly parallel conclusions to those laid out in Table 1, we find the kind of adjectives set out in Table 2. Whatever these adjectives may have in common, they are not a random sample of words labeled 'adjective' in our dictionaries.

Table 2

Adjectives which produce words like blackbird

Some colour adjectives: *black, blue, brown, green, grey, red, white*	*blackboard, blue-tit, brownstone, greenfly, greyhound, redfish, whiteboard*
Grand in words of family relationships	*grandfather*
A miscellaneous set of monosyllabic gradable adjectives of which only a few are illustrated here: *broad, dry, free, hard, hot, mad, small, sweet*	*broadcloth, dry-cell, freepost, hardboard, hotbed, madman, small-arm, sweetcorn*
A small set of non-gradable monosyllabic adjectives: *blind, dumb, first, quick* (= 'alive'), *square, whole*	*blindside, dumbcluck, (†)first-day, quicksand, squaresail, wholestitch*
A very small number of disyllabic adjectives: *bitter, narrow* and possibly *silly*[3]	*bitter-cress, narrow-boat, silly-season*

Before we move on, some comments need to be made about the examples in Table 2. First, consider the colour adjectives listed. While there is no implicit claim that the list given in Table 2 is absolutely exhaustive, some of the omissions here might need as much explanation as the inclusions. Only endocentric compounds have been considered, and not bahuvrihis such as *blackcap* (a type of bird) or *blackjack* (a game which involves black jacks), and only adjectives which appear in compound nouns have been listed. *Yellow* may never occur in first-element stressed compounds

of the right type: *yellow pages* (where both stress patterns are heard) is presumably a bahuvrihi when spoken with first-element stress - in any case, it is always written as two separate words; even *Yellowstone* is a bahuvrihi. The lack of *pink* is interesting in view of discussions about basic colour terms in English, but could be an accidental gap. The second point about the examples in Table 2 is that endocentric compounds with these adjectives are rarer than, for instance, bahuvrihis with the same adjectives, where the first-element stress seems to be better established. This is true not only of lexical items with the colour adjectives, but also of lexical items with the gradable adjectives. Third, it is clear that the set of relevant examples is not fixed. *The Chambers Dictionary* (1994) has *old 'boy* and *old 'girl* (of a school) where I would have *'old-boy*, and *'old-girl*. *Chambers* has *loose box* (presumably with phrasal stress) where *The Hamlyn Encyclopedic World Dictionary* (1971) has *'loosebox*. Fourth, some of the omissions in the gradable adjectives in Table 2 look as though they may be significant: no examples were discovered in *The Chambers Dictionary* (1994) with *big, deep, loud, mild, tall, thin, warm,* and *young*. There is no apparent influence of the unmarked term in any pair of gradable antonyms, since *coldstore* and *hothouse* are both found, as are *dry-cell* and *wetland, sour-dough* and *sweetcorn*. It seems that the gaps are largely accidental, though it is difficult to be sure of this.

Now we need to consider some of the criteria that led us to believe that *blackbird* was a single word, and show that these criteria do not necessarily coincide.

First we can consider the coincidence of first-element stress and writing as a single word: the orthographic and phonological criteria. This coincidence can be shown not to hold generally. We find lexical items written as a single word which *Chambers* lists as having second-element stress: *first-'aid, ill-'will*. *Chambers* does not mark stress on lexical items written as two words, so that we might be led to assume second-element stress on all of

these, were it not that some of them clearly have first-element stress: *funny business* and *little people* ['leprechauns'], for example, are both marked with first-element stress in the *Macmillan English Dictionary* (2002). The same may be true of some of the items with monosyllabic first elements like *long stop, wise guy*. Thus single orthographic words may have second-element stress, and sequences of two orthographic words may have first-element stress, as well as the patterns illustrated in Table 1.

Further, stress and orthography need not be consistent with grammatical isolation. Although we can find only established lexical items by reading dictionaries, we can see that these need not have compound stress and need not be written as a single word. Examples such as *black death, black ice, blue duck, brown rat, brown trout, green tea, red giant, red squirrel, white line, white meat,* and hundreds of others show that listed items need not be single orthographic or phonological words. In each of these examples the colour-adjective is as inaccessible to syntactic or morphological modification as it is in the *blackbird* type of example. The moment we discuss *blacker ice* or *a redder squirrel*, we are no longer using these as the names of the entities given in the dictionary definitions. Rather we are using them as descriptions, in the same way that we might use *black bird*. The same is true if we talk of *a very brown trout* or *rather white meat*. Thus what we define as lexical items on grammatical criteria need not have a single stress or be written as a single orthographic word.

Consider what would happen if we started with an example like *funny business*. We would probably say that this is a lexical item because its meaning is not entirely predictable from the meanings of its parts and because if *funny* is sub-modified in any way, the whole no longer retains its idiomatic meaning, but becomes compositional. However, in all other respects it meets the criteria for a phrasal construction. However, when this is so, we accept that orthography and stress are subsidiary criteria which do not need to be met for something to be a lexical item. Thus, implicitly,

we admit that orthography and stress are, if not irrelevant, then no more than supporting material in the discussion of *blackbird*. And at that point we should acknowledge that *blackbird* (and other words like it) just happen to have various criteria align but that this is not crucial, and that stress and orthography are not ways of defining lexical items.

At this point, though, we need to cast our net wider, because there are also first-element stressed adjective-noun constructions which are not covered in the above discussion, largely because they are never written as a single orthographic word. Some examples are given in Table 3.

Table 3

Examples of compound-stressed adjective-noun constructions with other adjectives

'feudal system, 'nervous system, 'solar system, cardio-'vascular system
inter'mediate school, 'normal school,⁴ 'primary school, 'secondary school
'classical period, ro'mantic period
'choral society, co-'operative society, dra'matic society, ope'ratic society
'cultural centre, 'cultural club, 'social club, 'social worker
'musical box
'floral arrangement

The items in Table 3 differ from the adjective-noun compounds illustrated in Table 2 in that they cannot be glossed as 'an N which is (stereotypically) A'. That is, while a *blackbird* is 'a bird which is stereotypically black', the *romantic period* is not 'a period which is stereotypically romantic'. While that factor does appear to distinguish some of the examples in Table 3 from examples like *classical 'music, primary 'colour, private 'school, public 'school, secret so'ciety*, it fails to explain the stress in examples like *cultural 'desert, primary edu'cation, social 'secretary, social se'curity, solar 'film,*

and *solar 'panel*. It should also be noted that this table clearly does not provide an exhaustive list of relevant examples (more keep turning up!), but without a strategy for finding examples it is difficult to elicit them. Some of the adjectives from Table 2 might conceivably fit in here. A *dumb-show*, for example, is not a show which is dumb, and a *stillbirth* is not a birth which is still. Similarly *easy-chair* and *happy hour* seem, in some ways, to fit better in Table 3 than in Table 2. We might hesitate about how to gloss *silly-season* in Table 2: is it 'a season which is silly' or 'a season in which silly things get reported' or 'a season in which the silly is done/reported'?

So we find first-element stress doing at least two different things. With the adjectives listed in Table 2, it indicates that the adjective is to be interpreted as non-gradable (as a classifier, in one terminology), while with the adjectives in Table 3, first-element stress indicates that the adjective is to be interpreted in its non-predicate meaning. However, and this is crucial, in neither case is the stress pattern a reliable marker of the function shown in the relevant table. In both instances, phrasal stress can have precisely the same reading, sometimes with precisely the same adjectives (as in *black bear* and *primary education*).

What, then, is the function of first-element stress? It seems as though its function is not to delimit a compound in any structural sense, which is why the label 'compound stress' has been avoided here.

To consider the type of construction illustrated in Table 3 in more detail, the patterns in which *school* appears will be considered more closely. What seems likely to be relevant in assigning stress to the first element of the constructions illustrated in Table 3 is a set of factors including the frequency of the particular collocations involved, contrasting patterns of premodification, and the collocations in which the particular adjectives are used. In order to elucidate these factors, the collocates of *school* in the one million words of the Wellington Corpus of Written New Zealand

English (WCWNZE) (Bauer 1993) were listed, and these are set out in Table 4. Where *school* occurred in a complex string of premodifiers, it was placed with its immediate constituent, thus *[Sunday School] floor* and *compulsory [school uniforms]*. The verb *school* and its form *schooling* were not included in the counts. No distinctions are drawn in Table 4 as to whether *school* or *schools* appeared in the text.

Table 4 shows that almost a third of the attestations of *school* are in a context where it is premodified by a word which says what kind of school we are dealing with. While some of the attested premodifiers are purely descriptive (*newer, previous, specific*) many of them, including the most frequent ones, name categories of school. Depending on how we count, perhaps one quarter of all uses of *school* have this kind of premodification. I must confess to always having been rather skeptical of Kingdon's (1958: 151) notion that *teacup* (for instance) is stressed on the first element because of 'an implied sense of contrast' with items such as *breakfast cup* and *coffee cup*. But here we do seem to have some evidence which would point to just such a conclusion: *school* appears so often with a modifier that it is the modifier which is more important than the head noun. This becomes even clearer when we look at some of the modifiers involved.

Table 4 - Collocates of school *in* WCWNZE.

Use of *school*	number	%
area school	2	
board school	1	
boarding school	6	
city school	1	
Correspondence School	12	
country school	2	
<other descriptive adjective> school (e.g. compulsory, existing)	4	
grade 0 school	1	
<gradable adjective> school	5	
grammar school	1	
high school	48	
household school	1	
independent school	3	
integrated school	1	
intermediate school	5	
local school	3	
<location name> school	9	
Maori school	1	
native school	1	
New Zealand school	4	
neighbouring school	1	
night school	2	
preparatory school	1	
primary school	33	
private school	8	
public school	1	
<religious or philosophical interest> school (e.g. Catholic, Rudolph Steiner, etc.)	7	
rural school	7	
secondary school	45	
shack school	1	
state school	1	
<subject> school (e.g. medical, journalism, etc.)	16	
summer school	1	
Sunday school	7	
town school	1	
training school	6	
tribal school	1	
400-pupil school	1	
Total premodified school	251	32%
AFTER school	7	
GO to school	17	
IN / AT (the) school	46	
LEAVE school	13	
OUT OF school	2	
Total special PP / VP	85	11%
Titles of schools not pre-empted by the categories above	35	4%
Figurative uses (school of thought, school of fish)	25	3%
School + N	198	25%
school or *schools* (not in categories above)	200	25%
Total occurrences	793	100%

For example, in WCWNZE *primary* is used ten times in connection with *health care* or *health services*, ten times in connection with *produce/product/production/producer*, 48 times in connection with education, schools, teachers etc. and only 30 times in all other uses. *Intermediate* is used 12 times with reference to education, and only seven times in any other connection (one of which is an examination!). While we do not have to consult a corpus to tell us that *high*, for example, has a much wider range of uses, there is a sense in which the occurrence of *primary* already predisposes us to expect the word *school*, and the word *school* is insufficiently distinct without the modifier, so that stress on *primary* can be excused, if not explained.

Furthermore, it is interesting to note that *school* itself is used attributively in 25% of its occurrences (or rather more if instances like *primary school teacher*, listed in Table 4 under *primary*, are taken into account). That is, in a quarter of its uses, *school* is actually not defining a class of school, but is being used to define another class. Examples such as *school teacher, School Certificate* might be deemed irrelevant in making the general point about the way in which *school* is premodified contrastively, in which case we might want to claim that in relevant instances, the preponderance of classifying premodification is even higher than is shown in Table 4.

Another survey with a different pair of words provides broadly similar conclusions. In Table 5 on the opposite page, the uses of *society* in the WCWNZE are broken down into various patterns.

Table 5

Uses of society *in* WCWNZE

	Number	% of total	% of relevant meaning
society (no premodifier) = 'people living together'	114	37%	59%
society (premodified) = 'people living together'	78	25%	41%
society (no premodifier) = 'club'	26	8%	23%
society (pre- and/or post-modified) = 'club'	85	28%	77%
Society = 'islands'	4	1%	100%

If we add to this the word *operatic*, which occurs only six times in the corpus, with just one of these modifying *society*, we end up with a similar, although not identical pattern. We might claim to have two lexemes *society*. The one meaning 'club' occurs most frequently with some kind of modifier (*operatic, Royal, building*, etc.). The other *society* occurs most frequently without a modifier; while the modifiers tend to be different, they can overlap: in principle *royal society* or *New Zealand society* could belong to either meaning of *society* until disambiguated by the context. The modifiers of *society* (in both senses, as it happens) are all relatively rare, and thus become particularly important in context. But when they are describing 'people living together' they are not the main focus of the communication, whereas when they are naming the 'club', they are of crucial importance.

An alternative, and perhaps preferable way of looking at this is provided by Ladd (1984). Ladd suggests that heads get de-stressed (and that we therefore get first-element stress) when the modifier is not merely 'descriptive'. This approach seems promising, though we need a more well-defined idea of what it means not to be 'descriptive'. Providing a naming function seems to be important here.

Now let us return to the monosyllabic adjective + noun constructions discussed at the beginning. If the principle of contrast holds for them as it

might be thought to hold for these constructions with longer adjectives, we would expect that a noun like *bird* is relatively frequently premodified (because we find *blackbird* where the modifier gets stress) whereas one like *bear* is usually not premodified (because we find *black bear* where the head noun carries the stress). More accurately, we would expect this to have been the case when the lexical items *blackbird* and *black bear* received their current stress patterns. Note that it is not clear that WCWNZE is a relevant corpus for such a comparison, first because of the period it covers, and secondly because nearly all the bears mentioned are of the stuffed rather than the live variety. Nevertheless, it is disconcerting to find precisely the wrong distribution of modification in Table 6. *Bird* meaning 'young woman' (3 occurrences), *bird* used as a premodifier (12 occurrences) and one instance of *ladybird* are omitted.

Table 6

Modification patterns of bear *and* bird *in* WCWNZE

Word	number	%
bear (unmodified)	6	37.5%
bear (premodified)	10	62.5%
bird (unmodified)	79	65%
bird (premodified)	43	35%

Let us sum up. We have, apparently, two adjective + noun constructions in English, one of which is a single word, the other of which is a phrase. Yet the stress criterion does not match the semantic criterion by which wordhood might be expected to be determined. It turns out, and this is the main observation of this article, that an apparently parallel use of stress is found with a disjunct set of adjectives, and that constructions involving this second set of adjectives have not been traditionally viewed as words at all. For this new set, it seems that there is some sense in which

the first element stress can be correlated with pragmatic contrast. The same does not appear to be true for the original set of adjective + noun 'compounds' (although the data that has been used here is not necessarily as relevant as we could wish).

If, instead of looking at implicit contrast, we consider the non-descriptive de-stressing account provided by Ladd (1984), we seem to be on firmer ground, in that the two types of adjective + noun constructions can be seen as acting rather more in the same way. But then we have the problem that so many apparently relevant constructions end up not being de-stressed at all. While Ladd has further requirements on heads that become de-stressed (for example, that they should be fairly generic, though that is not his terminology) the difference between *blackbird* and *black bear* might be covered, but not, I suspect, the difference between a *'social worker* and a *manual 'worker*.

Yet another possible solution, which has not so far been discussed here, is that first-element stress is simply a matter of lexicalisation. It is hard to know how to measure this, since lexicalisation does not necessarily correlate with absolute frequency in any given corpus. For example, it is one particular text in the WCWNZE corpus that makes *fossil bird*, with 4 occurrences, more frequent than *blackbird* with 2. In other words, a larger corpus might be more revealing. It is clear from the *fossil bird* example that the frequency of individual items might not be significant, but we might nevertheless expect that in general more lexicalised examples would have a greater frequency than non-lexicalised or less lexicalised examples. To test this a number of first-element stressed (single orthographic word) colour-adjective + noun constructions were compared for frequency in the 100-million-word British National Corpus (Burnard 2000) with a number of phrasal-stressed (two orthographic word) equivalent constructions.[5] The results are shown in Table 7 on the next page.

Table 7

Colour adjective + noun constructions: relative frequencies in the BNC

First-element stress	Actual *number*	Second-element stress	Actual number
blackberry	147	black bean	7
blackbird	299	black beetle	8
blackboard	275	black eye	76
blackcock	8	black frost	1
		black tea[6]	14
blueberry	13	blue cheese	24
bluebird	74	blue shift	10
bluefish	1	blue whale	12
bluegrass	14		14
brownstone	26	brown bear	19
		brown coal	17
		brown rice	67
		brown sugar	52
greenfinch	14	green pepper	28
greenfly	51	green tea	9
greyhound	264	grey matter	27
		grey squirrel	16
redworm	7	red carpet[7]	43
		red squirrel	29
whiteboard	9	white gold	16
whitefly	32	white knight[8]	30
whitewood	2	white meat	16
		white tie	27
		white witch	1
TOTAL	1236		549
AVERAGE	77.25		23.87

Table 7 suggests that there may indeed be a function of frequency or lexicalisation which distinguishes the two orthographic conventions in this set of words (particularly when we recall that frequency in one specific domain, such as hunting, may set a stress-pattern and orthography which then becomes general, and that, since orthography and lexicalisation are conservative, the relevant period of high frequency need not be current English).

However, it is not clear that all examples parallel to those in Table 3 can be seen as lexicalised or highly frequent in the same way. While *primary school* (980 occurrences in the BNC) and *secondary school* (609 occurrences)

seem well established, *dramatic society* (30 occurrences) falls far short not only of the clearly lexicalised *building society* (with 1226 occurrences) but even of the descriptive *American society* (93 occurrences). While *social worker* (770 occurrences) can safely be seen as lexicalised in comparison with, for example, *manual worker* (31 occurrences), *floral arrangement* (7 occurrences) seems scarcely different from *financial arrangement* (6 occurrences). A table corresponding to Table 7 is hard to construct here, given the difficulty in finding and in confirming examples of the appropriate types. For example, a search for *feudal system* finds not only *'feudal system,* but also examples of *feudal 'system,* and we also find a wide range of frequencies from the clearer members of the set such as *nervous system* (567 occurrences) and *cardiovascular system* (12 occurrences). It may be that a similar kind of result would emerge on average, but probably not to the same extent. The introduction of a new set of adjective + noun constructions with first-element stress into the discussion of the status of constructions with first-element stress at first looks as if it might be helpful in resolving a problem of some standing. While this new body of data raises a number of interesting questions and suggests some possible solutions, it still seems that first-element stress is doing more than one thing in English. While this does not in itself disprove the notion that there might be two discrete classes of construction involved, it makes it much more difficult to sort out the facts and to provide the kind of description which will be useful to language teachers and lexicographers, such as our honoree. More disturbingly, it raises questions about how lexicographers are supposed to identify lexical items (dictionary words). While stress and orthography have often been taken as contributory criteria, consistent patterns of mismatch between the two, and regular mismatches between either of these and a naming function suggest that the lexicographer needs new strategies for identifying relevant material.

Notes

1. I should like to thank Heinz Giegerich and Winifred Bauer for their helpful comments on an earlier draft of this article. Errors are my own.

2. Lexical items may, of course, be made up of more than one lexeme (as, for example, with idioms). But the prototypical lexical item is a lexeme, and confusion arises about the notion of 'word' in this context. Certainly, it could be claimed, as in Table 1, that *blackbird* is a word, while *black bird* is not; we will see that matters are not always this clear.

3. I assume this list is not complete, since I have had no list of disyllabic adjectives to work from. On the other hand, I considered many more disyllabic adjectives but discovered that they had no relevant examples listed in *Chambers*.

4. In New Zealand a normal school is a school attached to a teacher-training establishment and used as a training-ground for teacher-trainees. It is presumably *normal* in the sense that it provides a norm for new teachers.

5. *Green belt* appears as though it should fit in the second column of Table 7, but is given first-element stress by the *Macmillan English Dictionary for Advanced Learners* (2002). It gets a correspondingly high score with 222 occurrences.

6. Includes both the contrast with *green tea* and the contrast with *tea with milk*.

7. Includes both figurative and literal red carpets.

8. Does not include any mention of chess pieces.

Works cited

Dictionaries

The Chambers Dictionary. Edinburgh: Chambers. 1994.

The Hamlyn Encyclopedic World Dictionary. London: Hamlyn. 1971.

Macmillan English Dictionary for Advanced Learners. Oxford: Macmillan. 2002.

Other works

Bauer, Laurie. 1993. *Manual of Information to Accompany The Wellington Corpus of Written New Zealand English*. Wellington: Victoria University, Department of Linguistics. See also <http://khnt.hit.uib.no/icame/manuals/>

Burnard, Lou. 2000. *The British National Corpus, World Edition*. Oxford: Humanities Computing Centre. <http://www.hcu.ox.ac.uk/BNC>

Kingdon, Roger. 1958. *The Groundwork of English Stress*. London: Longman.

Ladd, D. Robert. 1984. English compound stress. In: Gibbon, Dafydd & Helmut Richter. (eds). *Intonation, Accent and Rhythm*. Berlin and New York: de Gruyter. 253-266.