

Speech, writing and boxsets: a messy linguistic change in English¹

KEN LODGE

University of East Anglia

(Received 15 May 2020; revised 14 July 2021)

This article is an attempt to explain an observable change in present-day English in terms of quite disparate influences. Since the change is not yet complete, it is a messy conspiracy of these influences. By studying life-time changes of this sort we may gain insights into how well-understood historical changes work. The change under discussion is most noticeable in the written form, but its trigger has been the phonetic realizations of the forms to be considered. The forms are exemplified by alternations in noun phrases such as *box(ed) sets*, *skim(med) milk*, *arch(ed) corbel table*. The relationship between the very different structures used in speech on the one hand and writing on the other is also relevant in this case. The NPs with *-ed* have a structure Adj_{pp} N, whereas the forms without it are compound nouns. Some of the Adj_{pp} forms found in such noun phrases are actually pseudo-past participles; that is, they are not formed from a verb, but take the *-ed* ending, e.g. *four-wheeled*, *gate-legged*. Whether native speakers learn such forms from the spoken or written language to some extent determines how they are perceived. This is relevant because the phonetic realization of members of both sets may be the same, so the phonetic form [bɒks set] may be perceived as *boxed set* or *box set*. I also consider the stress patterns of the new compounds, the orthography as a reflection of the structural change, and the ‘Germanic’ tendency towards compounding. The resultant picture is a messy one and the change has certainly not yet been completed, but we can see a conspiracy of disparate areas of the linguistic system putting pressure on certain lexical combinations. It should also be noted that ‘English’ is not a consistent linguistic system: we have to be clear about which variety is being discussed. English ‘belongs’ to many different groups of people, including non-native speakers as a lingua franca, so it is subject to many more influences today than the parochial versions of even just a hundred years ago.

Keywords: language change, phonetic realizations, noun phrases, compound nouns, English, orthography

1 Introduction

This article is focused on a particular set of changes in English, which have had a considerable effect on the structure of the language over several centuries and which continue to affect present-day English. They involve phonetic realizations, syntactic

¹ This article is a revised version of a poster presented to the Fourth Symposium on Historical Phonology at the University of Edinburgh, 9–10 December 2019. I am grateful to Andy Spencer, Heinz Giegerich, Peter Trudgill, Jean Boase-Beier and several of those who attended the Symposium for discussion of my ideas and for providing a number of suitable examples. I am further indebted to two anonymous *ELL* reviewers, whose comments and criticisms have helped me produce what I hope is an improved version for publication. I am, of course, fully responsible for this version.

reinterpretation, nominal compounding and the relationship between speech and writing. It is also the case that the change ends up being ‘messy’, or at least presents this way at the moment, since the change has not finished yet: there are still speakers/writers who use the pre-change forms, so we have a period of variation at present. The changes involved are driven by essentially separate trends in English, which conspire to produce the recent examples to be discussed here.

2 The data

The forms and the changes have been manifested in English over several centuries, but the examples I shall focus on seem to have appeared in the written form relatively recently. For reference I will label the unaltered forms as A and the later, more recent ones as B. The forms are exemplified in (1).

(1)	A	B
	boxed sets	box sets
	fine-toothed comb	fine tooth comb
	skimmed milk	skim milk
	arched corbel table	arch corbel table

The examples in List A are of the form: $NP[Adj_{pp}N]$; those in List B are of the form: $NP[NN]$. (Adj_{pp} = a past participle used as an adjective.)

Although this change may have accelerated during the post-World War II period, it should be noted that there is evidence that it has a much longer time-span than this. Trudgill (personal communication) has pointed out that the earliest recorded instance of *skim milk* in the *Oxford English Dictionary (OED)* is 1598 from Shakespeare’s *Henry IV, Pt 1* II. iv. 32: ‘I could deuide my selfe, and go to buffets, for mouing such a dish of skim milke [1623 edition *skim’d Milk*] with so honorable an action.’ However, the form with *skim* is not necessarily a reduced form of *skimmed milk*, as it is analysable as N + N. (I shall return to the issue of standardized writing below.) Be that as it may, *skim milk* appears on the menu on the wall behind the counter of today’s coffee shops.² Other examples, such as *ice cream*, have stabilized during the twentieth century, as can be seen from the spelling, whereas other items still show variable spelling: *iced tea/ice tea* (see the *OED*).

Written evidence of this kind of phenomenon and the uncertainty surrounding it can be seen in examples such as the following: *triple cook(ed) chips* (Tesco website: the website has *-ed*, but the packaging does not); *six-wheel(ed) van* (railway vehicle, e.g. Hatton’s website, or road vehicle, Google search); *different size boards* (Google search: Images for different size boards). The example of the title, *box set*, is ubiquitous on television on-demand sites and TV adverts, e.g. BBC iPlayer.

² Many of the examples I refer to specifically and similar examples can be found on the Internet. In many cases both the ‘full’ and the truncated forms can be found, sometimes on the same site.

3 Phonology: the phonetic realizations

Most dialects of English have realizations involving the simplification of consonantal sequences. The circumstances under which these occur are very restrictive and the process (if we wish to call it that in a synchronic grammar) is optional but very common. I distinguish here between clusters, which belong to one syllable, and sequences, which belong to more than one. These realizations do not affect initial clusters but may affect syllable-final ones, if there is an intervening morpheme boundary, e.g. [lɪf(t)s]; a word boundary will produce the same effect, as in the examples below. (See Lodge 1984/2015 for many examples of this phenomenon from different locations in the UK.) Examples of such realizations are in (2).

- (2) last time [lɑs taɪm]
 send me [sem mi] (with place assimilation)
 ask me [ɑs mi]
 postman [pəʊsmən]
 corned beef [kɔm brɪf] (with place assimilation)
 boxed set [bɒks set]

(The vowel qualities are irrelevant; I have simply chosen Southern British tokens.) The environment in which these realizations occur is as follows: the second stop consonant of a sequence can be suppressed, iff it has the same phonation as the preceding one and there is a morpheme boundary between the second and third consonants.³ This final requirement is necessary to exclude word-initial consonant clusters such as [spl-], [str-] and [skw-].⁴ All the examples in (2) fulfill this condition, whereas the ones in (3) do not and hence are not subject to such realizations.

- (3) sent me
 help me
 thank me
 milkman
 string
 spleen
 squeak

Although the common examples affect single words and two-word sequences, there may be occasions where more than two words are involved. Ladefoged (2006: 199) gives the example of *I should have thought* in which the auxiliary verbs have no vocoid articulation

³ The common descriptive term ‘*t/d*-deletion’ does not do justice to the subtle environmental constraints on the operation of the rule in many varieties of English.

⁴ Although I give a realization of *clasp me* as [klas me] in Lodge (1981: 35), the loss of labial (bilabial and labio-dental) consonants is unusual, although it has clearly happened in the past, e.g. in *raspberry*. Nasals are not subject to loss either, as in *kiln-dried*, for example. In fact, there are few consonant sequences in morpheme-final position in English that involve C + nasal or labial with matching phonation: *bulb*, *alb*, *wasp*, *gasp*, *film*, *elm*, *kiln*, *shelve*.

and no voicing, so the two words plus the initial consonant of the main verb are realized as the sequence [ʃtfθ], which could also be realized as [ʃfθ] under the same conditions as *last time* etc.

It is important to point out that the length of the contoid articulations can vary. Even if the oral stop phase is not articulated, compensatory lengthening can occur extending the preceding articulation into the final coda place, as in (4), where the third consonant is homorganic.

- (4) [semm mi]
[bɒkss set]

Without detailed spectrographic analysis across a number of tokens it is only possible to postulate that such extended realizations occur in those cases where the following consonant matches the preceding one. Simple structures across the first two consonants of the sequence with a word boundary and no following homorganicity, as in *last time*, *ask me* and *postman*, are unlikely to be subject to what I referred to above as compensatory lengthening, though *last seen* and *ask Sue* might well be because of the homorganicity of the third consonant. Such examples could equally well be treated as instances of assimilation. However, the main point is that the sequences under consideration can be simplified in all cases. It may even be the case that the consonant sequence simplification might be extended to *box set* with the realization [bɒk set], but this is not the same process, as it is not the stop that is affected: rather it is an instance of a simplification of homorganic sequences, as in *Prime Minister*, for example, with one [m].

The evidence for the change in the structure of the Adj_{pp} + N noun phrases (as in (5)) comes from recent spelling conventions, which are now very widespread and which we will consider in the next section. We now see *box sets* advertised on television, or even *boxsets* (cf. also *mixtape* from an advertisement on Spotify), and even technical terminology, such as the art-historical terms *arch(ed) corbel table*, *interlace(d) rib* and *arch(ed) buttress*, has been affected (for examples of a technical nature similar to these, see Huber & Rieth (1988) and parallel volumes). So, what were Adj + N structures have become N + N structures.

- (5) NP[Adj_{pp} N] > NP[N N]

It is only the participial adjectives or pseudo-participles that have changed their syntactic category; the simple adjectival forms, such as *last*, or verb forms, such as *send* or *ask*, have not. So, it seems that the change is not spreading through the language in a consistent way, but only in restricted circumstances, based on realizational norms.

Another realizational change that can be perceived is that the stress patterns do not always conform to the head stress rule. We know that *bläckbirds* can have various colourings, *black birds* cannot. This is a consistent pattern throughout much, but not all, of the English lexicon.⁵ Commonly occurring examples, such as *box set*, as

⁵ Giegerich (2009) draws a distinction between associative and ascriptive relationships in noun phrases, whereby he shows that the stress pattern distinction is not as clear-cut as is usually claimed.

advertised on television, often have their main stress on *box*. The Adj + N structures, on the other hand, have the stress on the N, unless the adjective is being contrasted with another one, as in (6).

(6) They're bláck birds, not red ones.

We seem to have more evidence of the change from Adj + N to N + N, i.e. the move to compound nouns. In English compound nouns are usually spelled separately, but the fluctuation in cases such as *boxset* probably reflects uncertainty as to what the structure and even the semantic relationship is.

Some of the stress variation relates to the place of the expression in the sentence. So, we find:

(7) The mortgage rate is fix(ed) térm.
a fix(ed) term mórtgage (rate)

This is in line with sets of similar disyllabic words, such as *thirteen* and *unknown*, as in (8).

(8) Thírteen people came to the party.
We only catered for thirtéén.
the únknown soldier
He was completely unknówn.

This does not seem to apply to *box set*.

4 Writing

Having looked at the specifics of the change under discussion, I want to consider the relationship between speech and writing a little more insofar as it has a bearing on the change under discussion. Writing is conventional and so subject to cultural influences. Complexity in writing reflects the fact that writing is a learned skill – one which demonstrates variation in level of attainment throughout a society. It also reflects the fact that writing can be reread as many times as the reader wishes, unlike speech, which is fleeting. Early on in the learning process there is an attempt to teach the relationship between the sounds of the spoken language and the letters on the page. Alphabetic writing is often presented as a one-to-one relationship between letters and sounds. Evidence that people who haven't necessarily attained a high level of writing skills will sometimes use the spoken word as the basis of their written forms, witness the commonplace confusion of *your/you're*, *there/their* and even examples driven by assimilated forms, e.g. *Belgium waffle*, to be seen on many café menus.

We should also note that, as written forms developed, the language often used as a model for writing European languages was Latin. This means that the structures required in the written form fitted in with Latin norms. This would explain why so many of the participles and pseudo-participles have been used in attributive position because they are adjective-like. Note, however, that some current commentaries on grammar refer to nouns in this position as attributive, too:

Nouns that modify other **nouns** are called adjectival **nouns** or **noun** modifiers.

For our purposes, **they are called** attributive **nouns**. (www.google.com/search?client=safari&rls=en&q=adjectives+before+nouns&ie=UTF-8&oe=UTF-8)

Such comments do not solve the issue of the structure involved and avoid any discussion of compounding.

It is interesting that spelling pronunciations have long been recognized as a (fairly limited) mechanism of change, but ‘phonetic spellings’, which would describe examples like *box set*, have usually been classed as errors and, therefore, disregarded as a mechanism of change. In the light of what we are discussing here, this seems misguided. There are historic examples of the reinterpretation of phonetic output, suggesting that many people acquired these constructions through speech, such as *spitting image* < *spit and image*, where the syllabic [n] has been reinterpreted as *-ing*. Similarly, there is equivocation about the words of the late nineteenth-century singing game: *Nuts in May*. This was written down in the 1880s in that form, but it makes little sense to be gathering nuts in late spring. Other versions of what must have been the sung version are *knots of may* (= the groups of hawthorn flowers) and *nuts and may* (nuts and haws in autumn). A much older example of reinterpretation of realizations is furnished by Middle English *at the last*, which underwent fricative to stop assimilation and was written as *atte last*, which is then subject to unstressed vowel loss and ends up as modern English *at last*.

Writing is inevitably based on an interpretation of speech. The kind of representation that employs segmental alphabetic writing may be based on morphological units, on phonological units or, occasionally, on phonetic output. A morphologically based system is found in much of the Graeco-Latinate vocabulary of English: *sane/sanity*; *photograph/photography/photographical*; *electric/electricity*, where the base form is maintained in the same form irrespective of the pronunciation. Simple English words, typically of Germanic origin, are spelled with simple segmental letters: *cat*, *dog*, *nut*. The final type of representation, based on phonetic output, is less common because native speakers are unaware of the realizational patterns they use in speech and, it is claimed, perceive speech phonologically. However, examples such as *box set* indicate an awareness of realizational detail, as in the case of *Belgium waffle*, too, given in section 4 above. As this kind of written form gathers currency, it has as a consequence a reinterpretation of the structure, as we have seen above. It does not involve a phonological change, but a reinterpretation of the spoken sequence at a lexical or syntactic level.

5 Compounding

We have already noted fluctuation in the spelling of *box set*. This may not be significant in that English does not always spell compound nouns as one word (unlike German and Modern Greek, for instance). In speech it is the stress pattern that indicates compounding in many cases. But we have another orthographic change, which has an effect on the structure of NPs, namely the lack of apostrophes in ‘possessive’

constructions. At one time these would have been perceived as errors in learning the rules, but the lack of apostrophes, especially in plural possessive nouns, is now so widespread that even in official titles and documents they fail to appear.⁶

- (9) Citizens Advice Bureau
 councillors surgeries
 governors briefing

These are comparable to compound nouns with a plural first element, as in (10).

- (10) High Plains drifter
 antiques fair
 sales manager
 arms race

Here we are witnessing another change: in this case from $N_{\text{poss}} + N$ to $N + N$, which suggests the latter is the preferred structure for native speakers. Harris & Campbell (1995: esp. 200–15) discuss the origins of compounds in a variety of languages. One source is Genitive + Noun. This seems to account for even the more recent examples with plural nouns in (9) and (10). However, in line with avoiding describing current structures in terms of an earlier stage or even a different language, we should avoid referring to the modern English examples as involving any kind of ‘genitive’. Whereas this may well be the origin of the compounds under consideration, modern English has no genitive.⁷

There seems to have been a tendency towards more compounding during the twentieth century; a careful investigation of a selection of both nineteenth- and twentieth-century texts would be needed to establish whether this is merely an impression. Nevertheless, it is the case that the less overt syntax there is, the more the hearer/reader has to supply an appropriate (or otherwise) interpretation of the string. Is (11) concerned with the mental health of the Vice-Chancellor or has the taskforce been set up by him?

- (11) The Vice-Chancellor’s mental health and wellbeing taskforce
 (UEA internal memo)

This extension of compounding may be an influence from American English, but be that as it may, the disparate changes in phonetic realization and orthographic conventions that we have discussed in sections 2 and 3 have accelerated the trend. The lack of morphological markers in English to denote word-class enables changes of category to take place easily.

⁶ Andy Spencer (private communication) has pointed out that apostrophe loss has been occurring slowly since the seventeenth century. We no longer use it to pluralize foreign words such as *concerto’s* or *tomato’s*, or even capitalized abbreviations and numbers, as in *PhD’s*, *MP’s*, *1930’s*. He also points out that loss of the apostrophe probably reflects the fact that in many cases it is unclear whether the ‘possessor’ is singular or plural – and it is probably both or either.

⁷ I am trying here to distinguish the terminology of different linguistic levels. ‘Genitive’ refers to morphological markers in a case system; it is not intended to be used as an indication of a semantic relation.

Compounding is always presented as a major component of the lexicon (see, for example, Bybee 1985; Katamba 1993; Harris & Campbell 1995; Haspelmath 2002). It is a major feature of the Germanic languages in particular. It was a productive part of word-formation in Old English, a time when many new concepts were coined. Many of the compounds of this period have been replaced, as in the following examples in (12).

- (12) læcecræft ‘leech-craft’ medicine
 tungolcræft ‘star-craft’ astronomy
 lofsang ‘praise song’ hymn
 banhus ‘bone house’ skeleton

(See, for example, Hughes 2000.)

Several words are no longer recognized as compounds, though they were originally: for example, *husbonða* ‘house dweller’ *husband*; *wifmann* ‘female person’ *woman*.

The assumption is that the elements making up compound words, in particular nouns, can be classified according to syntactic categories. Because English has few morphological markers of word-class, as noted above, it is often the case that a word gets its category from being placed in a particular syntactic slot. *Kick* cannot be assigned to either V or N until it is in a sentence. In *kick-start* is it a verb or a noun? What criteria do we use to answer the question? This makes compounding relatively free, as the category is irrelevant inside the compound. It is true that N+N is taken to be the norm (cf. Hughes 2000: 344, Haspelmath 2002: 86), but we can find V+N (*think tank*) or Adj+N+*ed* (*green-eyed*) as a compound adjective, and other languages have other possibilities.

In terms of the construction, what is crucial is the Head versus Dependent relation, irrespective of word category. Indeed, Di Sciullo & Williams’ (1987) proposal of a Right-hand Head Rule (see also Williams 1981) works extremely well for English. According to this proposal the rightmost element of a construction is its head. (There are exceptions to this, even in English, but we need not bother with the details here.) This specification includes both inflections and derivational suffixes. Since English has few morphological markers, a suffix will provide a morphological identity in cases where nothing else will, e.g. *box*: *boxes* (plural noun or third person singular, general tense): *boxed* (past tense/participle). For this reason many nouns have become verbs in recent years simply by putting them in a different slot and adding the appropriate endings where necessary, e.g. *progress* (with stress change); *appeal* (rather than *appeal against*); *reference* (rather than *refer to*).

In present-day spoken English and even some forms of the written language, e.g. colloquial journalism, we can find whole verb phrases or even sentences in the dependent position, as in (13).

- (13) a couldn’t-care-less attitude
 an I’ve seen it all before feeling

(Whether these forms are hyphenated depends very much on the writer.) We also find multiple co-ordinated dependents, as in (14). These occur in particular on menus and in product advertising (cf. also the example in (11) above).

- (14) ham-hock, grated cheese and festive chutney wrap
 cutting-edge home security range
 a butternut squash and coconut oil and soya alternative to cream cheese, topped with
 Applewood Vegan coconut oil alternative to cheese and smoke flavoured pea and wheat
 protein pieces

Even technical writing may demonstrate such long nominal constructions, as in (15).

- (15) the choir ambulatory four-part rib vault
 root accent contrast preservation
 consonant sequence simplification paradox

In technical documents relating to local government and planning there are even more extreme examples of compounding, which require considerable effort to understand, even by native speakers, as in (16).⁸

- (16) Draft North Lowestoft Heritage Action Zone Design Guide Supplementary Planning
 Document (East Suffolk Council website)

6 A brief comment on acquisition

One aspect of these orthographic representations to consider is how we, as native speakers, acquire these forms. Either we learn them from the spoken language addressed to and spoken around us, or from written texts. This may have an influence on how we perceive their structure. One example given above in section 4, *spitting image*, suggests that most people learned the expression from speech. To take two examples from my own experience, I learnt *fine tooth comb* from speech and perceived it as an Adj + [N + N] construction (with variable main stress on either *tooth* or *comb*). On the other hand, I learnt *arched corbel table* from written technical art-historical texts. This also has the structure Adj + [N + N], though in this case the Adj is a past participle. In the earlier (A) version of the former example the pseudo-participle is *fine-toothed*, so the structure has apparently changed from Adj + simple N to a structure containing a compound N in my own speech and no doubt that of others. There are a number of pseudo-participial words in English based on nouns: *blue-eyed*, *gate-legged*, *low-backed*, *six-wheeled*. There are no verbs from which these participial-looking forms are derived, with or without the first element in front of the hyphen: **to (fine-)*

⁸ Three speakers of Romance languages at the Edinburgh Symposium pointed out that French, Spanish and Italian could not tolerate such constructions; the Spanish speaker added that he had considerable difficulty in processing such English constructions. French, for example, has some compound nouns (though Head-first), e.g. *pause-café*, *essuie-mains*, but compare *corbel vault* with *voûte en tas-de-charge*. For an extended discussion of the semantics of NN-compounds, see Boase-Beier (1987: esp. 66–8).

tooth, **to (blue-)eye*, **to (gate-)leg*, **to (low-)back*, **to (six-)wheel*. (There are homophonous forms to some of these, but they have a different meaning: *to back*, *to wheel*; neither of these means ‘to put X on something’.) However, in some cases a back formation does occur and we find, for instance, *to triple-cook chips* (Great British Cooks website: cf. the example at the end of section 2). On the other hand, I have always analysed [skim milk] as *skimmed milk*, i.e. as subject to the realization discussed in sections 2 and 3, having been unaware until recently of a noun *skim*.

7 Concluding remarks

It appears that three essentially separate changes are conspiring to consolidate the most recent examples that are under discussion here. One is realizational and part of the phonological implementation component, whereas the other two are orthographical and lexical. They all combine to produce more and more compound NPs rather than analytical ones.

The realizational changes have not altered all Adj_{pp}+N structures because of the restricted environments in which they can apply. So we still have *grated cheese*, *malting milk*, *spilt milk*, *sent mail* (with possible assimilation of the /d/ or /t/), but *box set*, *gate-leg table*, *skim milk*. The last of these examples has already been commented on above: it could be seen as a reduced form of *skimmed milk* or as a N+N compound. However, it is not exocentric either (as *pickpocket* and *killjoy* are): *skim milk* is a kind of milk. Furthermore, there is something decidedly odd about (17).

(17) ?* I'd like two coffees, please, one with ordinary milk and one with skim.

In this case *skimmed* would have to be used, but it would be an orthographic trick to claim that *skim* and *skimmed* were environmentally conditioned alternants.

We may not be able to decide definitely whether we are dealing with a change in progress or stable variation, partly because it only occurs in the restricted circumstances we have been discussing, and partly because there are still people alive who use the older forms with Adj_{pp}s.

It is also the case that there are no phonological consequences to the change in terms of underlying phonological units, only realizational ones. This is a common enough phenomenon in a number of well-known historical instances. For example, in Old English, [θ] and [t] alternated in specific contexts. There was a general constraint on fricative + fricative sequences, which has resulted in alternations in modern English, which are no longer recognized, as it is no longer productive, e.g. *width*, *length* but *height*, *weight*.⁹ The results of this alternation can still be seen in some northern varieties of English: the second-person singular pronoun (*thou* in old-fashioned standard English) alternates depending on the preceding consonant, as described

⁹ The *gh*, of course, represents an earlier voiceless velar fricative, so the nominal suffix appears as [t] not [θ]. The constraint no longer applies, so sequences of two fricatives are quite common. Interestingly, *fifth* and *sixth* often appear as late as Tudor times as *fiht* and *sixt*.

above. The forms of the definite article in such varieties underwent similar variation, but in this case the pattern today has changed the environments in which each variant occurs and the variants have been generalized according to a new pattern: [θ] before vowels and [ʔ] elsewhere. (For details, see Lodge 2010.) What will happen in the case of forms like *box set* that are under consideration here is unpredictable, though there is evidence that in some lexical items the environmentally triggered realizations have been extended into environments that do not fit the template that was presented in section 3, e.g. *good size flat* in accordance with the template, but *good size apartment* where the following noun starts with a vowel.

It is also possible that it is only those participial forms that end up homophonous with an already existing noun that are undergoing the change as indicated by the orthography. If we consider *baked beans*, for example, which can be realized as [beɪk bɪnz] in line with our other examples, I have found no evidence that this is represented orthographically as **bake beans*. *Bake* is a verb and V-first compounds are assumed to be unproductive in English (see Haspelmath 2002: 86), so less likely to be new formations.¹⁰ Similarly with *refined sugar*, since *refine* is only a verb.

One may wish to ask what the drivers of this change might be. As a Germanic language, English has a predisposition to produce compounds (and not just nouns), as we have seen already. This gives the linguistic framework within which this change can take place. In addition there is the element of non-native speaker influence on the perception of normal native-speaker realizations. That English is a worldwide language is part of the picture. ‘English’ is not a consistent linguistic system. We have to be clear about which variety we are discussing, and English ‘belongs’ to many different groups of people, including non-native speakers as a lingua franca, so it is subject to many more influences today than parochial versions were a hundred years ago. Many non-native speakers use English as a lingua franca in all kinds of situations, especially commerce. Speakers of languages that do not themselves tolerate consonantal sequences, such as Chinese or Japanese, will reanalyse sequences. This can be transferred into the written form and will then be well on its way to being normalized, if native speakers also take up the reinterpretation, as seems to be the case in *box sets* and similar NPs.

As with all historical change, another important factor is analogy. In the case of *box set* there are similar N + N structures that it can be aligned with, e.g. *pot plant*. Note that this is not a reduced form from a past participle, which would not fit the realization rule template, i.e. *potted plant* with [-tɪd]. Just as a pot plant is a plant in a pot, so a box set is a set in a box.¹¹

¹⁰ *Bake* on its own does not seem to be a noun, although related compound forms are: *tray-bake*, *bake-off*. In the case of *baked beans* there may be commercial pressure to keep the orthographic form unchanged, as it is part of the marketing of Heinz (and other) baked beans. On the suggestion of one of the anonymous referees I googled *like bake beans*, to which the first response was: ‘Did you mean: like **baked** beans.’

¹¹ Note that in the case of *box set* we are not dealing with a box of anything. Although originally it was used in reference to a set of DVDs or CDs in a box, it is now used to refer to a series of programmes all available at once on catch-up TV. To me a *box set* should mean ‘a set of boxes’, not ‘a set of DVDs in a box’.

In American usage compounds seem to have gone even further into the lexicon. The following are all examples from restaurants and hotels in the United States (thanks to Peter Trudgill): *greet stand*, *share plate*, *handicap lift*, *handicap access*, *shave ice*. The first of these is not from an Adj + N original and the last two are examples of the spread of the consonantal loss to environments beyond those given above in section 3; *share plate* does not fit the environment in non-rhotic accents, either.

The overall picture of the areas affected by this change is messy in that it is incomplete, variable across speakers and varieties, and its interpretation relates to three separate linguistic subsystems. The orthographic change is evidence for the underlying change, but I do not intend to imply in claiming this that orthography is part of native-speaker knowledge of any linguistic system.

We can now turn to a final open-ended question: what is the internal construction involved in these complex NPs? A hierarchical structure such as [(Det) [(Adj) [N]]] seems inadequate for examples such as (16). Whereas the Head of the construction is *document*, is the rest of the construction a cumbersome compound noun? Certainly, there is one Adj + N sequence, *supplementary planning document*, but the rest is a simple string of nouns. *Draft* and *heritage* may appear to be in an adjective slot, but they do not behave like normal adjectives. The examples in (18) are not possible.

(18) *The document is draft.

*The zone is heritage.

On the other hand, some nominal structures can be used both attributively and predicatively, as in (19) and (20).

(19) an eighteenth century painting

The painting is eighteenth century.

(20) resistance bands (= piece of exercise equipment)

My bands are medium tension resistance.

But there seems to be no consistency across the structures – yet another aspect of messiness.

There is a separation of at least three different categories of lexical item that can appear in a NP between Det and N: simple adjectives, other items which cannot be used predicatively, and items which can. To assume these are all adjectival phrases is too simplistic (cf. the statement from the Internet referred to in section 4). Consider the following examples:

- | | | |
|-------------------------------|--|--------------|
| (21) The red house | The house is red. | (Adj N) |
| The box set | *The set is box. | (Compound N) |
| The seventeenth century house | The house is seventeenth century. | |
| Wide-size cards | The cards are wide-size. | |
| The good size flat | The flat is a good size. *The flat is good size. | |
| An interlace rib | The rib is interlaced. | |
| Six-wheel truck | The truck is six-wheeled. | |

To fit the traditional Det Adj N template those items that are (pseudo-)past participles would have to end in *-ed* to match the predicative position. This is clearly no longer the perception of speakers/writers in the examples under discussion. Phonology is assumed to represent abstract speaker knowledge and it is assumed that native speakers are aware of only phonological units, not the details of phonetic implementation. This assumption may be questioned in some cases, such as those under discussion, since awareness of phonetic realization seems to be operating, as evidenced by some of the spellings. But the consequence of some of these changes is that speakers are unsure of what the resultant structure is.

The question here is whether such long NPs are, in fact, examples of a simple concatenation of elements. It is possible to propose that in much modern writing simpler elements of structure from speech are being reintroduced. The claims of Everett (2017) regarding different types of grammar may be of relevance in trying to find an answer to this intriguing problem. There seems to be evidence that both hierarchical and non-hierarchical structures can operate in one and the same language variety. This is not a question I can attempt to answer in this article.

Since much work in linguistics, especially syntactic investigations, takes standard written material as its data, one might legitimately ask to what extent that reflects the grammar of the language as spoken every day by native speakers (see also comments by Miller 2011). Standardized written forms of only a fraction of the world's languages have been around a relatively short period of time in comparison with the history of speech and they have had to be taught/learned (and invented). This means that traditions of linguistic analysis and norms have often been set with reference to some other norm. From the sixteenth to the nineteenth century Classical Latin was used as the model for such grammars. The effects of this have been ridiculed by twentieth-century analysts: the 'declension' of English nouns, the ban on split infinitives, possessive adjectives, though some of these are still to be found in pedagogical grammars. However, it might be said that describing colloquial speech, or even colloquial written forms of the kind under discussion here, in terms of a syntactic theory based on standard written forms makes the same erroneous assumptions as describing English in terms of Latin.

Author's address:

School of Politics, Philosophy, Language and Communication Studies
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
UK
k.lodge111@btinternet.com

References

- Boase-Beier, Jean. 1987. *Poetic compounds*. Tübingen: Max Niemeyer.
Bybee, Joan L. 1985. *Morphology*. Amsterdam: John Benjamins.

-
- Di Sciullo, Anna-Maria & Edwin Williams. 1987. *On defining the word*. Cambridge, MA: MIT Press.
- Everett, Daniel. 2017. *How language began*. London: Profile Books.
- Giegerich, Heinz. 2009. The English compound stress myth. *Word Structure* 2(1), 1–17.
- Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold.
- Huber, Rudolf & Renate Rieth. 1988. *Glossarium artis 6: Gewölbe : Voûtes : Vaults*. Munich: Saur.
- Hughes, Geoffrey. 2000. *A history of English words*. Oxford: Blackwell.
- Katamba, Francis. 1993. *Morphology*. Basingstoke: Macmillan.
- Ladefoged, Peter. 2006. *A course in phonetics*, 5th edition. Boston, MA: Thomson Wadsworth.
- Lodge, Ken. 1981. Dependency phonology and English consonants, *Lingua* 54, 19–39.
- Lodge, Ken. 1984/2015. *Studies in the phonology of colloquial English*. London: Croom Helm.
- Lodge, Ken. 2010. Th^h interpretation of t^h definite article in t^h North of England. *English Language and Linguistics* 14(1), 111–27.
- Miller, Jim. 2011. *A critical introduction to syntax*. London: Continuum.
- Williams, Edwin. 1981. On the notions ‘Lexically related’ and ‘Head of a word’. *Linguistic Inquiry* 12, 234–74.

References to online texts

- BBC iPlayer: www.bbc.co.uk/iplayer/categories/drama-and-soaps/featured
- Hatton’s website: www.hattons.co.uk
- Tesco website: www.tesco.com/groceries
- Great British Chefs website: www.greatbritishchefs.com/how-to-cook/how-to-make-triple-cooked-chips
- East Suffolk Council: www.eastsuffolk.gov.uk/2019/north-lowestoft-heritage-action-zone-design-guide/