On the history of English (and) word classes

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
On the history of English (and) word classes

David Denison
University of Manchester
ICEHL17, Zurich, 20 August 2012

1 Introduction
Word classes (categories, parts of speech=PoS) are central to many linguistic theories, descriptive and prescriptive grammar, corpus linguistics. What do they buy us, and what do they cost? Keep in mind three widespread methodological or theoretical assumptions about word classes:

(1) a. Word classes are established by distributional tests.
   b. In a given sentence, each word belongs to one and only one word class.
   c. In a PoS-tagged corpus, consistency of mark-up is of overriding importance.

Each of the assumptions (1) can be held or rejected independently. Will concentrate on case studies from histories of of (over 3m tokens in British National Corpus) and to (nearly 2.6m) – 2nd and 4th place in frequency, discounting 'code', verbal prefix, etc.

For the tag scheme of the Penn Corpora see Santorini (2010). The vast majority of examples have of as a clear preposition in a prepositional phrase, as in (2); only a small proportion of uses of of are coded as adverbial particles, RP, as in (3):

(2) And þa ða he ut eode [þo of, 8am baéde]
    and when he out went from the bath
(3) & sloh þæs sacerdes þeow & [we his eare] ofe, acearf
    and struck the priest’s servant and his ear off cut

These would correspond to off in ModE.

Excursus: Comparison with on

In OE of behaves much like other preposition/adverbs; cf. on, whose statistics are broadly similar. Purely prepositional use overwhelmingly dominant, + small cluster of adverbial particle uses. Note reversal in overall frequency between on and of: In OE, on is 5× as frequent as of, whereas in BNC, of alone is 4× as frequent as on. (On 22.5k pwm OE →7.4k pmw PDE; of 4.7k pmw → 31k pmw.)

Table 1: Tagging of of/off in York Corpus of Old English

<table>
<thead>
<tr>
<th></th>
<th>spelling &lt;of&gt;</th>
<th>spelling &lt;off&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>tag</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>P = preposition</td>
<td>33,183</td>
<td>98.4</td>
</tr>
<tr>
<td>P in þærof</td>
<td>154</td>
<td>0.5</td>
</tr>
<tr>
<td>RP = adverbial particle</td>
<td>358</td>
<td>1.1</td>
</tr>
<tr>
<td>RPX, X = tag for unknown part of speech</td>
<td>17</td>
<td>0.1</td>
</tr>
<tr>
<td>other</td>
<td>10</td>
<td>0.0</td>
</tr>
<tr>
<td>totals</td>
<td>33,722</td>
<td>100.1</td>
</tr>
</tbody>
</table>

Table 2: Tagging of on in York Corpus of Old English

In Middle English and later, of begins to serve as an alternative to inflectional genitive marking, becomes bleached of spatial meaning, and becomes the default relational preposition in many different patterns – hence boost in frequency through the ModE period. Word class status is unaffected, however.

Of uniquely splits into two separate items: of and off. Table 1 suggests that spelling <off> is very rare indeed in OE, and spelling <of> survives till the 17th century or later. The tagging in PPCME2 implies both that the split makes little progress during the ME period, and that the adverbial particle use is slow to take off. Random inspection suggests that the tagging of the Penn corpora is highly reliable, and that P is a preposition rather than (part of) a conjunction.

Table 3: Tagging of of/off in PPCME2

<table>
<thead>
<tr>
<th></th>
<th>spelling &lt;of&gt;</th>
<th>spelling &lt;off&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>tag</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>P = preposition</td>
<td>38,164</td>
<td>98.1</td>
</tr>
<tr>
<td>P in þærof</td>
<td>604</td>
<td>1.6</td>
</tr>
<tr>
<td>RP = adverbial particle</td>
<td>136</td>
<td>0.3</td>
</tr>
<tr>
<td>X = tag for unknown part of speech</td>
<td>3</td>
<td>0.0</td>
</tr>
<tr>
<td>totals discounting ‘code’, verbal prefix, etc.</td>
<td>38,907</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 4: Tagging of of/off in PPEME with two supplements

<table>
<thead>
<tr>
<th></th>
<th>spelling &lt;of&gt;</th>
<th>spelling &lt;off&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>tag</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>P = preposition</td>
<td>59,527</td>
<td>97.3</td>
</tr>
<tr>
<td>P in þærof</td>
<td>1,454</td>
<td>2.4</td>
</tr>
<tr>
<td>RP = adverbial particle</td>
<td>207</td>
<td>0.3</td>
</tr>
<tr>
<td>X = tag for unknown part of speech, ADJ</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td>totals discounting ‘code’, etc.</td>
<td>61,190</td>
<td>100</td>
</tr>
</tbody>
</table>

Figure 1: Proportion of prepositional uses of of/f from OE to eModE

Denison, ‘Word classes’ p.2 of 12
Table 4 suggests that the differentiation of the two words is well under way during eModE. No exact moment when split occurs. By late ModE it is clear that of and off are two wholly distinct words, as shown in Table 5:

<table>
<thead>
<tr>
<th>spelling</th>
<th>of</th>
<th>off</th>
</tr>
</thead>
<tbody>
<tr>
<td>pronunciation</td>
<td>[o]</td>
<td>[ɔf]</td>
</tr>
<tr>
<td>pronunciation</td>
<td>[v]</td>
<td>[f]</td>
</tr>
<tr>
<td>word class</td>
<td>always preposition</td>
<td>preposition or adverbial particle</td>
</tr>
<tr>
<td>semantics</td>
<td>purely relational</td>
<td>spatial (or metaphorical extensions, or Aktionsart etc. values in phrasal verbs)</td>
</tr>
</tbody>
</table>

Table 5: of and off in Present-day English

2.2 Quasi-ellipsis

Unlike most preposition/adverbs, of has always been overwhelmingly transitive (cf. at, with). In late OE/early ME, it could occasionally precede a clause with properties of both adjuncts and complements; see QUASI-ELLIPTIS in Denison (1985: 200, 1993: 141-3). In examples like (4)-(5), never numerous, of is behaving unlike either a traditional preposition or an adverbial particle:

(4) pe deade nis noht of; þah he ligge unburiet (c1230(?a1200) Ancr. 95a.17)
‘the dead(OBL) not-is nothing from though he lie unburied’
(5) ah ful wel he let of hwen ei seîd pet ...
(c1230(?a1200) Ancr. 82b.17)
‘but full well he thinks of when anyone says that …’

Quasi-ellipsis critical to early development of prepositional passive in a family of closely similar combinations (leten of/by, setten of/by, tellen of/by ‘regard, esteem, think of’. By 14/15C, the prepositional passive extended to most prepositions (given appropriate syntactic and semantic relations between V+P and NP complement), and of had reverted to being purely prepositional.

2.3 Multi-word prepositions

In OE there is an incipient multi-word preposition ut of (173 times as a string in YCOE, tagged as RP followed by P). Example (2) shows that the two items can be separated in OE. Gradual loss of separability over time as out of begins to lexicalise, but separate tagging does no harm.

ModE has many so-called compound prepositions ending in of, always purely prepositional: in front of, by means of, etc. Single items, with whole string belonging to word class P, or separate words, where of is P? (Aarts 2007, Denison 2010: 122, Hoffmann 2005, Huddleston & Pullum 2002: 618-23, Quirk et al. 1985: 670-3) As a combination lexicalises, balance of plausibility between two analyses may change. Change is gradual. PPCMBE, a corpus which stops at 1914, is on balance justified in tagging both analyses in PDE (Denison forthcoming,2012). So a historical corpus continued to PDE might justly different leniencies for recent decades as against older periods, violating (1)c. Even stronger case for dint of or in spite of, which have no alternative with ’s. Despite their fossilised and idiomatic nature, Huddleston & Pullum (and others) still argue for a syntactic analysis as three separate words with one of two internal structures (neither involving a compound preposition) (2002: 621). Beckner & Bybee (2009, cited by Reynolds & Pullum in prep.) adduce semantic and corpus evidence for a compound preposition analysis alongside the separate-word analysis:

Yet our position predicts that a word sequence may gradually form a unitary status even while component words are partially activated on each use. […] [Note that even as in spite of strengthens in constituency, it does not instantaneously become fused into an indivisible unit. The sequence continues to maintain some connections to the separate words in, spite, and (most importantly here) of. (2009: 40)]

Much theoretical linguistics has striven to exclude all redundancy from analysis and from mental grammars, but the proof on the lumpers rather than the splitters, as if any string must be considered as made up of independent words until proven 100% lexicalised. This ignores gradual skewing of behaviour as a result of lexicalisation.

2.4 SKT-constructions

I use the term SKT-CONSTRUCTIONS for binominal structures involving N1 = sort, kind or type:

(6) (…) N1 of (…) N2

Among a whole network of related constructions are (examples from BNC):

(7) the Canadians had one sort of sovereign, and the British had another sort (A69 1471) – REFERENTIAL CONSTRUCTION
(8) There was a kind of inevitability about the whole proposal which appalled Alexei. (G17 1172) – QUALIFYING CONSTRUCTION
(9) It kind of built his confidence with each successive flask (A14 937) – ADVERBIAL CONSTRUCTION

See Denison (2002, 2011), Keizer (2007) (I follow her nomenclature), Brems & Davidse (2010). In the referential construction, (7), of is routinely a preposition. The adverbial use of (9) lacks any complement nominal, and of has been incorporated into a unit with a preceding noun’ (Huddleston & Pullum 2002: 621). But at what stage did reanalysis occur? There are several candidates for the best synchronic analysis for (8), but I will confine discussion to two:

(10) a [[kind of\( \text{of}\text{, inevitability}\]]

As with compound prepositions, both analyses, (10) and (11), may remain in play to some extent. In analysis (10), kind remains N and of is P. Apart from maybe ‘explaining’ why of is unstressed, the word classes tell us little. In (11), word class would be assigned if anything only to the unit kind of. Kind of in (8) is not A, or at least very far from prototypical A. Semantically somewhat like D, especially if analysed as a kind of D. When kind of is extended to act as a modifier of various XPs, as in (9), the word class becomes truly problematic: neither D nor A would be appropriate for a VP modifier, while Adv would be odd for an NP modifier. Quirk et al. list it as a COMPROMISER or DOWNTONER

Table 6: Referential vs. qualifying SKT, from Denison (2002) (after Keizer)

<table>
<thead>
<tr>
<th>Referential</th>
<th>Qualifying</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>sort, kind, type</td>
</tr>
<tr>
<td>N2</td>
<td>sort, kind</td>
</tr>
<tr>
<td>number</td>
<td>free</td>
</tr>
<tr>
<td>primary stress</td>
<td>D1 or N1</td>
</tr>
<tr>
<td>N2, omissible</td>
<td>yes</td>
</tr>
<tr>
<td>semantic head</td>
<td>N1</td>
</tr>
<tr>
<td>discourse function</td>
<td>discourse topic or anaphor</td>
</tr>
<tr>
<td>style</td>
<td>neutral</td>
</tr>
</tbody>
</table>

As with compound prepositions, both analyses, (10) and (11), may remain in play to some extent. In analysis (10), kind remains N and of is P. Apart from maybe ‘explaining’ why of is unstressed, the word classes tell us little. In (11), word class would be assigned if anything only to the unit kind of. Kind of in (8) is not A, or at least very far from prototypical A. Semantically somewhat like D, especially if analysed as a kind of D. When kind of is extended to act as a modifier of various XPs, as in (9), the word class becomes truly problematic: neither D nor A would be appropriate for a VP modifier, while Adv would be odd for an NP modifier. Quirk et al. list it as a COMPROMISER or DOWNTONER.
(1885: 598, 602), at least when used as a VP-modifier, a possible function of Adverbial, but the word class is not[?]
discussed, either for the whole string or its two component words.  Marianne Hundt (p.c. 15 Aug 2012) observes
that it is in effect a (compound) pragmatic particle; where the pragmatic function dominates, traditional word
classes are not that helpful.

Kind of can also be used as a style disjunct.  Again, word class is not particularly pertinent.

(12) 'It's a little cartoon, it's just cheap and enj-- like Bart Simpson kind of.' (KPG 4168)
(13) 'Flower is a message?' said the boy.  'Yes.  Kind of.' (CEU 4388-90)

2.5 Have "of"
An entirely new use of of has arisen at some point since the 18C:

(14) the servant to the old Lady I sho'ld'd be of no use:

(15) I should be very happy to of seen m'.  Orford at Leek (?1774 Corpus of 18C Prose)

(16) I would of gone Long before this but I still expected Letters from you as I wrote often to you.  (1796
CORIECOR Chambers)

(17) I never could of thought that force Could turn affection in its course.  (1814 [OED])

(18) I had known of your illness I should not of written in such fiery phrase in my first letter.  (1819 Keats, Letters)

Edith Wharton makes of for have a verbal tic of Zinnie, a spoilt child - merely eye-dialect:  implied speech in example
(19) is actually fully standard; likewise for (20) as far as is concerned (though hadn't ought is not-standard).

(19) "If she really feels like that I'd of thought she'd of sent me a present," she [Zinnie] objected doubtfully.

(1928 Edith Wharton, The Children [Virago, 1985], xxvi.291-2)

(20) "Yes, 'n' Judy hadn't ought to of sneaked away and left us all like that.  -- ought she of?"  Zinnie appealed
indignantly to Mrs. Sellers;  (ibid. xiii.132)

The spelling <of> in examples like (14)-(20) only occurs for infinitive have, never (or rarely) for the general present
form.  Infinitive have is special.  Brian Joseph and Joyce Tang Boyland found that

there are speakers who used the preterite form in modal perfect contexts [e.g. should've took -- DD] and but [sic]
the standard past-participle form in regular perfect contexts.  For these speakers at least, the modal perfect is
losing its connection to the perfect.  (Boyland 1998: 4)

Infinitive have may be grammaticalising separately from rest of have-perfect paradigm, its semantics is becoming
associated with modality/non-assertiveness, hence irreals.  NB. double use would have liked to have gone, frowned
on by prescriptivists but not ungrammatical (Denison 1998: 140).

So <of> in (14)-(20) is more than a mere misspelling, even if made possible by homophony between the unstressed
preposition of and the verb have as [av].  But spelling <have> is relatively unusual in representing only a verb (rarely
as N in a must-have or the haves), while <of> is unusual in spelling a form which is uniquely a pure preposition.
Therefore when writers get it 'wrong', it is almost certainly significant for word class assignment.  There is strong
anecdotal evidence that many educated young Britons do not realise -- let alone assume -- that there is any
connection with the verb have, and the occasional pronunciation [av] is another strong indication.

Back to Zinnie.  The morpheme in her (21) is genuinely non-standard.  That is a pattern with quite a long history,
the element in question variously spelled have/ve/of/ha':a:

(21) "Terry's with his tutor at present," said Blanca ...

"But he said he wouldn't have come down even if he hadn't of been," chimed in Zinnie (1928 Edith
Wharton, The Children, viii.71)

(22) 'If I had ha' seen Nancy, I should ha' given her my mind very different.  ...' (1864-6 Gaskell, Wives &
Daughters [WC] xi.577) [Mrs. Goodenough]

(23) ... but I wish he'd ha' told me (1864-6 Gaskell, Wives & Daughters [WC] lii.646) [Squire Hamley, bluff
old-fashioned gentry]

(24) Little Domby was my friend at old Blimber's, and would have been now, if he'd 'ave lived.  (1848 Dickens,
Dombey xxiii.445.12)

(25) Now if you'd've told us that last Thursday we'd've thought (BNC KCX 3528)

(26) We were offered a flat, which would of mea-- if we'd of accepted it would of meant we would have to of
sold every stick of furniture because the rooms were not large enough (BNC F82 77)

(27) if it hadn't a been for Old Cribb there wouldn't have been none (BNC H09 1152)

When the first verbal morpheme is contracted to 'd, as in (23)-(26), unclear whether its equivalent full form is had or
would.  Possibly the construction licenses a blended contraction where neither full form is widely accepted.

As for the non-standard insertion seen in (21)-(27), the morphosyntax is odd for either
had or
would.  Extra morpheme only occurs in clauses which are non-assertive:  conditional protases, counterfactuals, unrealised
wishes, and so on.  There is no (28) as beside (28)a:

(28) a.  These neighbours had lived there for a number of years (BNC AC7 1413)
b.  *These neighbours had've/'d've lived there for a number of years

What is the word class?  There have been at least two generative attempts to classify it as a preposition because of
the spelling <of>:  Coates (1989) and Kayne (1997); see also (Cheshire & Edwards 1993), (Sampson 2002), (Denison
1998: 140-2, 210-2, 2007a, b).  If anything, it is an invariant grammatical item confined to the verbal group (cf.
egarter n't), so probably Adv.  The development casts doubt on the analysis of what was a clear verb:

(29) I've, eaten it

(30) If I'd, eaten it

(31) If I'd've/ed, eaten it

(32) I could've, eaten it

In (29), 've = have = 1 sg pres of V.  In (30) 'd = had = 1 sg past of V.  In (31), probably 'd = had = 1 sg past of V, while
've = particle (AdvV) and can be regarded as a form of of.  Now 've in (32) resembles both V and Adv uses.  The
former maintains nice generalisations about patterning of English auxiliary system.  For speakers who have (31), we
must revisit the apparently standard string (32) and worry whether (or delight that!) it is now some sort of blend
with dual inheritance.  There is strong but not absolute preference for cliticised forms;  cf. Boyland's data (1998: 3),
where tensed verb + clitic exhibits the NICE properties supposed to apply to tensed operators alone:

(33) N:  "You should'ven't took all the brick cheese!" (child data, Bowerman)

(34) I:  "What would've you done?" (JT B)

(35) C:  "(I should've read the error message.)" "Yeah, so should've I." (JT B)

(36) E:  What could have she been reading?  (Fluent non-native speaker, JTB)

These examples do not involve the 'superfluous' morpheme but rather point to cliticisation (and further
grammaticalisation) of the have-perfect when used with modals.  Some further examples:

(37) a sentiment he would have probably denied (1961 Brown corpus G65 1880)

(38) I should've never went on a stupid blind date.  They never work out." (1992 Armistead Maupin, Maybe the
Moon xx.225)

(39) might've he continued to elude police, had that fourth murder not occurred? (COCA 1999)

Thus (33)-(39) provide evidence of incipient weakening of the neat auxiliary patterning that was the strongest
motivation for assuming that 've in (32) was purely the verb have.
3 To
What are usually regarded as three distinct words in PDE share a common origin, according to the etymological reference works: to as conventional spatial preposition, as infinitive marker, and as intensifier spelled from the 16C with <oo>, too. I am concerned here with the infinitive marker.

3.1 Infinitival to
A syntactic word, usually proclitic to an infinitive (Pullum 1982: 185). Four possible word classes:

(40) a. preposition
b. word without category
c. auxiliary verb
d. VP subordinator

In OE traditionally seen – in origin at least – as normal prepositional to followed by nominal form of verb with dative inflection (OED s.v. to prep., conj., and adv. B). (Los asks whether the verbal ending was inflectional or derivational; she also questions to what extent the infinitive was still syntactically nominal by the historical period of OE (1999).) Some parallel examples where an infinitive represents goal or destination, like an ordinary NP complement of to:

(41) a. Ac hwí ferde ge to seonne þone man mid hnescum reafum gescrydnde (Lk(WScp))
   But why went you to see the man with soft robes clothed
b. hi ferdon to Sinai westene (Exod. B8.1.4.2)
   they went to Sinai desert

Los cites a coordination of to-PP and to-infinitive (1999: 21). Even if to can safely be labelled as P, the infinitive would be a problem for word class: nominal inflectionally and in its relation to to, but verbal syntactically in its relationship with adjuncts and complements. (Cf. the -ing gerund in ModE.) However, preposition-like behaviour weakens from earliest records:

With infinitive in substantial relation. Equivalent to a noun or gerund: to being ultimately reduced to a mere ‘sign’ of the infinitive without any meaning of its own. (OED sense B. III)

When is ultimately? In PDE the P analysis for to is disfavoured (Huddleston 2002: 1184-5). Instead to is generally just called an infinitive marker – in effect a unique member of a minor category, given that at, till and for to are no longer current. Since word classes are for capturing generalisations, belonging to a single-member class is equivalent to belonging to no class at all, what Pullum calls SYNCATEGOREMATICITY (1982), a violation (albeit generally ignored) of assumption (1)b. Pullum prefers to analyse infinitive-marking to as a highly idiosyncratic auxiliary verb. The principal argument in favour is behaviour under post-verbal ellipsis, aka AUXILIARY STRANDING (Huddleston 2002: 1519-26), parallel to auxiliary verbs (‘operators’) (Pullum 1982: 199-200) (my examples from BNC):

(42) He wants to break away. He means to _ (AT7 1835)
(43) But one look at her face told me she had _ (AGE 583)
(44) I would like to go with him if I could _ (EVG 1899)

Clearly, though, if to is a verb, it is an odd one. Pullum argues that it is an obligatorily untensed auxiliary. Like other untensed auxiliaries before an ellipsis site, it cannot carry stress:

(45) a. ‘Oh, you couldn’t have,’ Mrs Yardley protested. (ADD 1084)
   b. *Oh, you couldn’t have
(46) a. He wants to break away. He means to.
   b. He wants to break away. *He means to.

That to can plausibly be categorised as a verb is indeed ‘unexpected’ (Pullum 1982: 184): verbs are usually thought of as predicators, highly inflected, prototypically about events or actions. Infinitival to is none of these. Pullum (1982: 200-1) quotes Zwicky & Levin (1980), who coined the term INFINITOID for a class consisting of the infinitives of be, of perfect have and (in BrE) of supportive do, plus to. These share the property of obligatory stresslessness when just before the deletion site in Verb Phrase Deletion. By analysing to as an untensed verb, Pullum makes INFINITOID a natural class instead of a stipulated disjoint set (and removes the need for a special name).

In a solo chapter of the Cambridge Grammar, Huddleston weighs up this analysis against an alternative, that infinitive-marking to is a VP subordinator that belongs with the clause subordinators that and whether, and comes down for the latter (2002: 1185-7). Levine (2012) reviews their evidence, adds another fact, and swings the pendulum decisively the other way, showing that Pullum’s 1982 analysis of to as an auxiliary verb is far preferable.

Levine pours scorn on classification of to as a subordinator, since it can be preceded by the subordinator whether, whereas ‘subordinator doubling’ is not permitted. But whether can introduce a subordinating conjunction like because (‘preposition’ in H&G’s terms], (55), and even occasionally the clause subordinator that, (56)-(57):

(54) And I didn’t know whether to laugh with pleasure or to cry with pain. (HD7 2455)
(55) Whether because they had ceased to be pleased about her birthday or simply because they were tired, she did not know. (BNC A0L 776)
(56) As long as I do that, at any rate, Pa and Ma will always find something to reproach me with, whether that I don’t obey the rules of decorum or that I have an indecent way of expressing myself or that I’m breaking ties or whether something is of the kind. (WebCorp Vincent van Gogh Letter 186 To Theo van Gogh. Etten, Friday, 18 November 1881. http://vangoghletters.org/letter/let126/letter.html)
(57) (1) Whether creatures always existed.
   (2) Whether that they began to exist is an Article of Faith. (Basic Writings of Saint Thomas Aquinas: God and the order of creation 1447, ed. Anton C. Pegis, c.1945)

Huddleston’s assignment of word class would put it in a class of its own, or possibly a class of three if added to the clausal subordinators that and whether. Pullum’s classification focuses on the verbal component of clause. The strongest evidence for verbal status, ellipsis a.k.a. stranding, is a relatively recent innovation (55B is very early):

(58) I think I could come as often again as I used to. (ARCHER 1766admx.x4a)
but as he had bothered me to, I consented. (PPCMBE HAYDON-1808,1,34.871)

(60) Pufopace Not a farthing, unless I am obliged to, by law. (1812 COHA)

(61) you ever hear the like before? Never, Julia! and hope never to, again. (1822 COHA)

(62) ‘You won’t forget?’ said Newman. ‘I am not very likely to,’ rejoined Nicholas. (1838-9 Dickens, Nickleby xxii.271)

For discussion of early examples, see Visser (1963-73: §1000), Warner (1993: 64), Denison (1998: 201-2). And of course it is not the prototypical V that to is being grouped with, but a set of auxiliary verbs – a minor category that didn’t even exist before the late 16th century, according to Lightfoot, though many have argued that the period of emergence was long-drawn-out and started much earlier. These auxiliary verbs – M, have, be, do – all differ amongst one another, but M, have and be are arguably static, while M and a number of newer accretions to the class are exhibiting less and less inflectional variation. In fact to is assigned to an even smaller subgroup, the untensed auxiliaries. Clearly, whichever of the word class assignments (40)b,c,d is preferred for affixing-marking to, the word class in question has very limited membership, and the number of distributional generalisations involved may be quite small.

4 Open and closed classes

These three polarities in relation to word classes are all slightly different:

In an unpublished paper, Reynolds & Pullum (in prep.) show convincingly, using corpus data mainly from various Google giga- and tera-corpora, that it is impossible to draw a strict line between open and closed word classes: all are open to new members, some more than others. I believe the same goes for the lexical ~ functional polarity, unless of course it is hard-wired into a formal model of grammar as an absolute binary choice. To the extent that the grammatical words of and to belong to word classes, these are classes of small membership, often highly language-specific and indeed period-specific: classes come into existence and change rapidly (e.g. Modal in last few centuries). The so-called open classes are represented throughout the history of English, and although there have been major losses and accretions of membership, the classes themselves seem intuitively to change less, even though they are more open to new members. (The hierarchy of openness tentatively offered by Reynolds & Pullum has noun > verb > adjective > adverb at the top end.) As argued in (Cort, Denison & Spinillo 2006). One [way of generalising from minor to major classes] is to assert that any category is simply the average of its members’ properties, and that when either the membership or individual properties of members change, so must the category. With a small and recent category like M or D, such categorial change is evident. With a larger category like N or V, small changes in membership or distribution make a less perceptible difference to the overall ‘average’, and so such categories look superficially very stable. However, over a long enough period of time, even major categories change their morphosyntax.

5 Intermediate conclusions

Revisit case studies to see what they tell us about word classes and vice versa:

2.1 of ~ of: changing word class of of disguises huge change in semantics, in grammatical significance, in frequency; gradual split with of, likewise arguably involving no change of word class, hides shift in pronunciation.

2.2 Quasi-ellipsis: short-lived change in behaviour which is difficult to model in word class terms, nevertheless implicated in the origins of the prepositional passive.

2.3 Multi-word prepositions: should word class apply to individual components or whole string? And individual analysis becomes decreasingly revealing of behaviour – for any but last element, of ~ the more frozen the compounded preposition becomes.

2.4 SKT: dual analysis possible for qualifying construction, and word class analysis problematic for adverbial usage, which is primarily pragmatic.

2.5 Have ~ of: word class is highly relevant to non-standard usage with extra morpheme, with probable reanalysis of ‘ve from V to particle or Adv. However, this calls into question affinities of infinitive of have in standard.

3.1 Infinitival to: Word class assignment requires use of (mostly) infrequent patterns of fronting and so on and groups to with very few other elements. Best potential grouping is itself rather recent. The older grouping retains some explanatory value. Argues against unique word class.

Elsewhere I have shown that some items are underspecified for word class assignment, where both semantics and structural alignment are unaffected by choice of terminal symbol (e.g. Denison 2010). Situation can obtain when a given word occurs in some contexts clearly as one class, in others clearly as a different word class (but with the same meaning), and in yet other contexts where the word class distinction is neutralised. So in (64), genius is clearly N, in (65) it is A, and in (66) it is unhelpful to choose between N and A (assuming that the writer has both N and A values for genius in their lexicon). In similar fashion, various is A in (67), D in (68), and either or both of A and D in (69):

(64) He’s a genius, unbelievable. (Mo Farah, quoted http://www.telegraph.co.uk/sport/olympics/athletics/9470759/Mo-Farah-move-to-the-US-transformed-me-from-a-weakling-who-can-like-a-girl-into-an-Olympic-champion.html)

(65) This series is arguably the most genius, the most entertaining and the most poignant chapter in the Father Ted chronology. (2001 http://www.amazon.co.uk/product-reviews/B00004ZBXX 8.09.2001 <accessed 3 Nov. 2009> [WebCorp])

(66) everyone else’s genius tactics also came up blank. (http://www.podiumcafe.com/2012/7/28/3198318/no-forever-vinokourov-wins-the-olympic-c-r-from-the-break)

(67) [...] as a way of making his readers see a complex and various world. (BNC ECB 178)

(68) groups who are trying to bring various of the Olympic sponsors to justice (http://files:notgrass.wordpress.com/2012/08/11/olympics-ideal-actual/)

(69) The IAAF said they had been collecting samples from various athletes in the build up to the Olympics (http://www.why.com/olympics/Nike-athletes-banned-for-doping-offences/-/1541752/15701356/-/r12b1r/-/index.html)

In other cases, dual resemblance involves both a word class difference and a structural difference. This is often treated historically as a (catastrophic) reanalysis. Beckner & Bybee (2009) argue that reanalysis can be gradual; cf. have~of in §2.5, both gradual and very much incomplete to date.

An item that belongs to one word class may be coerced in its behaviour by a construction. Thus the verb tweet in its Twitter-related sense (not yet in OED) has developed a consistent pattern of use that resembles various of say, speak, transmit, write:

(70) If you’re at the movies, don’t tweet. (COCO 2012)

(71) [Sarah Palin] tweeted a picture from Alaska. (COCO 2010)

(72) Many Olympic track and field [sic] athletes have tweeted their displeasure with what is known as Rule 40. (http://www.facebook.com/donn.cabrill.run/posts/288260924614171)

(73) David Cameron will face pressure to remove the Tory whip from the Conservative MP Aidan Burley after he tweeted that the Olympics opening ceremony was “multicultural crap”. (http://www.guardian.co.uk/politics/2012/jul/26/olympics-opening-ceremony-multicultural-crappy-tymp)

However, it can be recruited by the way construction (Israel 1996) and used there with a different kind of object:

(74) Celebs tweet their way through the London 2012 Olympic Closing Ceremony (http://www.dailymail.co.uk/football/article-2187842/Celebrities-tweet-way-London-2012-Olympic-Closing-Ceremony.html#ixzz2Yy6dOgH)

This kind of coercion may affect subcategories or even major word classes (e.g. proper name ~ A). Thus there is traffic both ways between countable and mass nouns. Mass nouns behave like count nouns in portion or serving settings (two beers, a coffee), and vice-versa in contexts like showing too much leg.

(75) there are probably around a million tons of pet dog in the US alone (http://what-if.xkcd.com/7/)

(76) The opening ceremony is So Danny Boyle! (http://twitter.com/dansmyrealtor)
6 What are word classes (for)?

I do not believe that word classes are Aristotelian categories with clean boundaries between them. Aarts discusses gradience between classes – what he calls INTERSECTIVE GRADIENCE (e.g. Aarts 2004, 2007) – but attempts to remove most cases by counting up somewhat arbitrarily chosen features and hoping not to get a numerical dead heat. If this approach is designed to save the word class as the basis of algorithmic syntactic analysis, it may be misguided synchronically, and it is quite unhelpful diachronically. If we treat word classes as generalisations (whether by speakers or linguists), then it is obvious that in some cases a given item may share patterning with (= show resemblances to) more than one item simultaneously. A familiar case: near, as P or A. If so, degrees of resemblance may be unequal, but this need not invalidate the lesser resemblance.

Word classes display prototype effects, but whether they are actually best modelled with Prototype Theory is less certain (Taylor 2004). The minor category of Modal verbs certainly shares many morphological, syntactic and semantic properties which distinguish them collectively from all other verbs, but quite apart from recent falls in frequency of core modals (Leech et al. 2009), arguably there is no one core modal which carries all the properties:

- Can/could: retain some normality in the distinction between present and past tense; epistemic meanings limited.
- May/might: obsolete (certainly in AmE), no longer treated as a present ~ past pair, loss of negatives *mayn’t, *mightn’t.
- Must: does not normally have past tense reference, and for most users therefore cannot appear in the apodosis of an unreal conditional.
- Shall: obsolete except in 1st person interrogatives.
- Will/would: lack typical semantics.

Family resemblance or the cluster concept may work better. But without default inheritance and overdies, and especially dual inheritance, word classes are no more than a convenient expository generalisation. You can always go for finer-grained classes, and in the extreme case, every single word may be in a class of its own (Gross 1979).

Many of the mini-histories I have been looking at involve lexicalisation and grammaticalisation. They concern not individual words but words in, and part of, particular patterns – patterns of syntax, semantics, phonology, pragmatics, discourse. It is the patterns that change, the patterns that resemble other patterns. CONSTRUCTIONS. My guess is that my examples are not idiosyncratic, marginal examples but important and even not atypical.

The cost of any linguistic theory is the number of theory-internal assumptions which have to be made in order for the theory to work. The benefits are the predictions, insights and (unexpected) connections revealed by the use of that theory – and not available without it. How do costs and benefits stack up against each other? Even if the benefits are thought to outweigh the costs, could one do even better with a different theory? These are important questions. There is a long western tradition of taking word classes as basic to language description, and in the more highly theorised models, as axiomatic. It makes more sense to me to treat distributional word classes as epiphenomena of constructional behaviour. It is the semantics at least of the major classes that allows us to track the changes in membership and behaviour of word classes over the years in one language, or across languages.

References

Denison, David. 2007a. Playing tag with category boundaries. VARIENG. http://www.helsinki.fi/varieng/journal/volumes/01/denison/
Reynolds, Brett & Geoffrey K. Pullum. in prep. New members of supposedly closed word classes in English: Implications for syntactic theory and grammar education.

Denison, ’Word classes’ p.12 of 12