

Pushing the boundaries of word classes

David Denison
CBDA, Troyes
7 July 2015

Plan of talk

Part 1 Positives

- Why we need word classes.
 - Some conventional characterisations

Part 2 Problems

- Why word classes don't (always) work.
 - Some case studies

Part 3 Resolution

2

1 Positives

Everyone needs word classes

- Traditional Greek- and Latin-based grammars define them.
- School-children learn them.
- Dictionaries display them.
- Second language teachers usually make use of them.
- Structural and dependency grammars include them.
- Most formal grammatical theories require them.
- Language users (both SP/W and AD/H) couldn't manage without them.

4

An essential generalisation

- For linguists, how else can we capture such facts as
 - Every clause needs a verb.
 - Large sets of words share much the same distributions, e.g.
 - *dog, knife, commitment, rhythm* and thousands of others
 - *bright, heavy, red, exuberant* and many more
 - *above, in, of, with* and dozens more
- For users, language without word classes would place impossible burdens on storage and processing.

5

Conventional view

(crudely simplified):

- Word class is axiomatic.
- Word classes are universal.
- Every word in every grammatical sentence belongs to one and only one word class.
 - If ambiguous, then each reading is a different sentence.
- [Sometimes] All languages possess the same classes.
- [Sometimes] Phrases get their phrasal category from the word class of the head (X-bar Theory).

6

How are word classes defined?

- **Morphologically**
- The classical approach, well suited to highly inflected languages like Greek and Latin.
- For English, inflectional morphology helps potentially to distinguish major word classes N, V, Adj, Adv) from each other, though much variation is potential only, given the reduction of inflectional marking.
- Derivational morphology more haphazard. Some suffixes characteristic of particular major classes, though with many gaps and some exceptions.

Rauh (2010)

7

How are word classes defined?

- **Syntactically**
- Approach favoured within American and British structural linguistics.
- Items satisfy a cluster of distributional tests. For example, functions of adjective are to
 - premodify a noun (attributive Adj, *fat man*)
 - serve as (head of) the complement of *be* and similar verbs (predicative Adj, *He is fat*)
 - postmodify the head of an NP (postpositive Adj, as in *somebody new, a refugee old and lame*)

Rauh (2010)

8

How are word classes defined?

- Huddleston & Pullum add other tests:
 - a modifying adverb is the characteristic dependent of Adj (*very important, so sorry, highly ambitious*)
 - gradability (*unhappier, most dignified*)
This can be morphological or syntactic .
- Syntax ± morphology is privileged over semantics.

Huddleston & Pullum (2002: 527ff.).

9

How are word classes defined?

- **Semantically**
- Traditionally by means of notional definitions such as
 - N = name of a person, place or thing
 - Adj = a describing word
 - V = a doing word
 - etc.
- Often derided as unworkable, but more formal versions defended by some serious linguists.
- Widely used for cross-linguistic comparison.

Langacker (2008), Croft (2001, 2005a, b)

10

Criteria may not coincide

- Best known is semantics~syntax mismatch:
 - *lightning* more of an event than a thing, but always N in English; cf. Hopi, where V
 - *fist* not part of body like *head, hand*, but still N in English
 - *conference dinner* can be construed as a thing (food items, including *summer salad*), therefore N-like
 - or as an event (takes place tonight, with duration), so ?more V-like, but *dinner, salad* always N in English
 - *conference, summer* are describing words, but not Adj

11

Not necessary or sufficient

- Even with formal morphosyntactic criteria, rare for any one criterion to be shared by all members of class.
- Some of this variation handled by subclasses (e.g. mass nouns, non-gradable adjectives, modal verbs).

12

2 Problems

Everyone needs word classes, don't they?

- Traditional ≠ justified.
- Word classes are generalisations. Generalisations can be useful heuristics without being wholly correct.
- Some non-IE languages such as Chinese show weaker support for idea of word classes.
- No overt justification for certain word classes in some languages (e.g. Det in OE).

14

Grammaticality

How apply tests for word class?

- Is a particular string or sentence possible?
- Introspection now disfavoured.
- Test grammaticality judgements with informants.
- Usual assumption that test either passed or failed, i.e. sentence or fragment grammatical or ungrammatical.
- Tests often achieve an uncontroversial consensus:
 - Adj: *dustier, so dusty*, but **very dust, *dustest*
 - N: *the dust's source*, but **the dusty's source*

16

Grammaticality judgements

- Judgements may be influenced *inter alia* by
 - context
 - priming
 - dialect
 - idiolect
 - register
 - prescriptive ideas
 - Observer's Paradox
- c.5% of exceptions to rules because of language change, linguistic creativity, intended ungrammaticality, etc.

Mindt (2002) cited by Mukherjee (2005: 100-1), Sampson & Babarczy (2014)

17

Black & white or colour?

- Generative tradition assumes underlying grammar is discrete and clear-cut.
 - Any graded judgements, whether from individual speaker or from averaging reports of many speakers, 'merely' performance or E-language.
 - Avoid some uncertainty by crafting questionnaires well. But often that means you know the answer you want.
- Usage-based tradition suggests that grammatical knowledge is inherently probabilistic.

Schütze (1996, 2004); Gahl & Garnsey (2006), Gahl & Yu (2006), Bresnan (2007)

18

Underspecified class

Expert fully specified

- Adj
A friend turned his ideas into layouts, although not very expert ones.
so expert a surfer that his fellow chiefs grew jealous
- N
An expert's decision is usually final and binding
Forensic experts were yesterday examining the wreckage.
- So there must be two entries for *expert* in lexicon.
 - (Purely synchronic argumentation here for PDE.)

20

Expert underspecified

- Adj and/or N
You could do it yourself or get expert help.
- AD/H cannot know whether *expert* is noun or adjective here.
- (Does SP/W need to decide?)
- Arbitrary for linguist to privilege Adj or N.
- Vagueness, not ambiguity: makes no difference to interpretation or to constituent structure.

21

Stepwise change of class

Normal, instant change of class

- Addition of affix:
 - *fun* N > *funny* Adj
 - *editor* N > *editorial* Adj (> N by ellipsis) > *editorialise* V
- No segmental morpheme added (zero-derivation):
 - *hammer* N > *hammer* V
 - *import* V > *import* N
- Subtraction of apparent affix (back-formation):
 - *editor* N > *edit* V
 - *pedlar* N > *peddle* V

23

N vs. Adj N > Adj

- Overlaps in distribution:
 - Adj and N can both fill premodifier slot.
 - Adj and some N can be head of predicative complement.
- Suitable N may develop new use as Adj.
- Such change is stepwise, not instantaneous.
- Examples include *ace*, *amateur*, *apricot*, *bandaid*, *cardboard*, *champion*, *core*, *corker*, *cowboy*, *designer*, *dinosaur*, *draft*, *freak*, *(fun)* *genius*, *(key)* *killer*, *landmark*, *luxury*, *niche*, *pants*, *powerhouse*, *rubbish*, *surprise*, *toy*, *Velcro*.

24

N > Adj barely started

- N (probably)
he reeled through four savage rounds before he got the killer punch (1982, COHA)
 - cf. The punch was a killer.

25

N > Adj underway

- Coordination with genuine Adj:
so that the hobby, which often proved fatal, would be safe as well as fun. (1966, Time)
- but cf.
She was a lecturer, very angry and on strike.
- Precedes genuine Adj in NP:
The powerhouse new bestseller from ELIZABETH GEORGE (1996 advertisement, *The Guardian*)
- but cf.
 - *emergency premature delivery; deathbed final words*
N Adj N Adj

26

N > Adj underway

- Modification by intensifier:
It's extremely draft (I think Tom wanted me to post it as an article). (2006)
WOOOOOOOW sexy mistress posing in very killer stainless steel custom made 9inch high heels! (2013)
- but test can be satisfied by proper name:
This is so Woody Allen. Neurotic as ever and so on top of his game. I could hear his voice every time Branagh opened his mouth. (2000)
It's very silly, it's very odd, it's very Woody Allen. Love it.

27

N > Adj all the way

- Postmodifier:
Each track has something killer on offer (2013)
A really lovely tea towel for your husband, wife, girlfriend, boyfriend, best friend, or anyone ace in your world! (2015)
Adler believes in filling your surroundings with all things fun and Joyful (2005, COCA)
- Stem for derivation:
The concept of his art is inherently hard to put into words. But most commonly (and amateurly put), Turrell's Skyspaces can be described as [...] (2013)
Trying to explain the ferry system very draftly (2015)

28

N > Adj all the way

- Gradability:
So therefore that was more key to you than [...]? That was more important to you? (1995, COCA)
It's the most key part of this business. If you don't make the right decisions, you can lose a considerable amount of money [...] (1998, COCA)
I think my key point is going to be this: girls are not wired to do that kind of stuff ... And an even keyer point is the definition of 'stuff' (2005)
Mirror, mirror on the wall, who's the keyest of them all (2001)

29

Varying acceptability of *fun* Adj

- Generally accepted:
Doing something fun like redecorating your room..is really interesting biz for a teen who loves being busy. (1951, OED)
- Acceptance inversely correlated with age:
And they are so fun to eat! (1979, COHA)
- Generally rejected:
Walking and looking is boring. Touching is funner. (1990, COCA)
- Variation both by language user and by context.

30

The march of *long*

Long Adv > Adj?

- Adv
Tyrannes raygne not long. (1562, *OED*)
My lord will be long a commyng. (1539, *OED*)
It won't last long.
- Adv or Adj?
And many oþere londes þer ben , þat it were to long to telle or to nombren (CMMANDEV,96.2317)
and many other lands there are, that it would-be too long to relate/count or enumerate
- Does *long* modify verbs *tell/number* or subject *it*?

Denison (in prep.)

32

Long Adv > Adj?

- More like Adj:
and yn mony oþer myscheves þat he suffurd , þat wern to long to tell (CMMIRK,70.1895)
and in many other misfortunes that he suffered, that would-be too long to relate/count
- Here *long* seems to be predicated of (*myscheves*) *þat*.
- Penn parsed corpora of ME, eModE and IModE always tag *long* as Adj, sometimes parsed as NP-MSR.
 - PPCME2 has 709 Adj but also 11 tagged as adverb (*longe* 8x, *lange* 1x, *lengre/lengyr* 2x)

33

Long Adv > Adj?

- Consider
I won't be long.
I won't be a long time.
x = 'I am not myself a long period of time'
✓ = 'I will not be at the task/away for a long time.'
- Rough semantic equivalence of *long* and *a long time* suggests similar grammatical function, but doesn't imply same word class (cf. *He's miserable/a misery*).
- Both *long* and *a long time* would be adverbials in clause structure, hence *long* (by itself) probably Adv.

34

Long as N?

- *Long* labelled N in *OED*² (but probably not in *OED*³) after Prep:
You shal know before long. (1610)
'Well, Lord, it mayn't be for long', replied Dolly. (1803)
- But Prep can govern any XP, not just NP:

<i>before the game</i>	NP
<i>for real; as new</i>	AdjP
<i>before now; until very recently</i>	AdvP
<i>from beyond the grave</i>	PP
<i>by trying harder</i>	VP

Denison (in prep.)

35

Long as N?

- *Long* labelled N in *OED*² (but probably not in *OED*³) after normally transitive verb:
þe member..nedeþ longe or it be souded.(?a1425)
the member ... needs long before it is healed
We should quickly find, that the largest Stock of Humane friendship would be too little for us to spend long upon. (1694)
How long will it take to be full in this case? (1763)
My son..hastened us to our toilets. Mine did not take long. (1783)

Denison (in prep.)

36

Long as N?

- An even earlier example of *take long*:
These tricks take not long (1656, EEBO)
- Later allows human subject:
He doesn't have/need/spend/take long.
(How long) should we give it?
You don't have/haven't got long.
- But cf.
It should only take until next Thursday. intr. *take*
It shouldn't last long. intr. *last*

37

Long '(for) a long time'

- OED consultancy request in 2014.
- My conclusion: if single label insisted on for the controversial cases of *long*, Adv comes closest.
- Borderline between Adv and Adj use quite blurred, as also e.g. for word after *look* 'have the appearance'
Things had, by that time, begun to look badly for all concerned. (1891)
I was trying to tell her what to do if things look bad. (1948)
Whatsoever looks ill, and is offensive to the Sight (1712)
He looks well (1782)

38

No class

Infinitive marker *to*

- Possible word class:
 - preposition
 - auxiliary verb
 - VP subordinator
 - none

40

Infinitive-marking *to* as Prep

- Prep is historical source.
- Comparison has been made with *of* (for standard *have*) in
He must of done it.
 - Actually more of an adverbial particle indicating irrealis, especially in examples like
If I'd of known
- But behaviour of infinitival *to* too different from ordinary prepositions (and even non-standard *of*) to justify classification as Prep.

Coates (1989), Kayne (1997)

41

Infinitive-marking *to* as Aux

- Main evidence is behaviour before ellipsis site:
I saw the film yesterday, and now Jim will ___ as well.
I saw the film yesterday, and now Jim plans to ___ as well.
- Like other base-form auxiliary verbs – *be*, perfect *have*, substitute *do* – must be unstressed when stranded:
Oh, you couldn't have / *h^{ave} (BNC)
He wants to break away. He means to / *t^o.
- Ingenious analysis, makes grammar economical.
- Would be only aux that is **never** tensed.
 - Is classification psychologically plausible?

Pullum (1982), Levine (2012)

42

Infinitive-marking *to* as unique

- VP-subordinator is choice of *Cambridge Grammar*.
- If word class is set of words that behave similarly, then class with membership of 1 more or less equivalent to claim that *to* is syncategoremic, without category.

Huddleston & Pullum (2002: 1183ff., 1526ff)

43

More than one class

Near as both Prep and Adj

- A few words can satisfy normally non-overlapping classes at same time without ambiguity:
Are we near the end?
How near the end are we?
I thought we were nearer the end.
- Other candidates are now closer to Prep (*like*) or to Adj (*worth*, earlier also (*un*)*worthy*, (*un*)*becoming*, *next*).

45

Uncertain extent of item

Prefabs: multi-word Prep

- Includes e.g. *in front of*, *in terms of*, *for the sake of*.
- Bybee and associates see constituency as emergent and subject to gradual change:

[D]espite holistic processing and chunk-like storage, prefabs can still be related (to varying degrees) to the words and constructions of which they are constituted.

We take the view that it is altogether common even for an individual speaker to have nondiscrete syntactic representations for the same word sequence.

Hoffmann (2005); Bybee & Torres Cacoullos (2009), Beckner & Bybee (2009: 28)

47

Prefabs: *nasty piece of work*

- Famous line in *Hamlet* indirectly responsible for modern application of phrase *piece of work* to human referents:
What a piece of worke is a man! (1623, *Hamlet First Folio*)
- *OED*: 'colloq. (freq. *derogatory*). A person, esp. one notable for having a strong (usually unpleasant) character. Usu. with modifying word; cf. *NASTY adj.* 2c.'
- 12/190 examples of string *piece of work* in BNC have human referent. Shakes. quotation in modern spelling, and all but 1 of others preceded by adjective *nasty*.

48

Prefabs: *nasty piece of work*

'You're some piece of work, Mrs Sutherland, you know that?'

You'd best steer clear of him, Manderley, he's a nasty piece of work.

He was also a member of the Mafia, and he was up to his eyeballs in drugs. Altogether a very nasty piece of work.

It was a nasty piece of work, done with thoroughly malicious intent.

- (Only *nasty piece of work* with non-human referent.)

49

Prefabs: *nasty piece of work*

- Clearly a prefab, somewhat lexically and semantically restricted. Precise extent of the prefab is vague, e.g.

piece of work

a piece of work

a ([optional intensifier]) [pejorative adjective] *piece of work*

a (...) *nasty piece of work*

- Any decision arbitrary. Vagueness as to boundaries and fixity of idiom problematic for purely algorithmic theories of language use (but not for language users).

Denison (submitted)

50

No syntax

Discourse particles

- Morphosyntactic tests are based on sentence syntax.
- Many aspects of language use can only be described at the level of the utterance, or the turn, or discourse.
- The lexical items and idioms called discourse particles operate primarily at discourse level:
Well, I'd like you to consider ...
This is quite a good idea, I think. Sort of, anyway.
- Most of them are lexical items. How do we (should we?) determine their word class?

52

3 Resolution

Vorrei e non vorrei

- Word classes are helpful in some circumstances.
- Word classes are problematic in others.
- They are convenient generalisations made by linguists.
- Perhaps also (to what extent?) by SP/Ws and AD/Hs.
- Most constituency grammars only allow branching downwards, and dependency grammars take lines from one word to another.

54

Vorrei e non vorrei

- Many of the problematic cases seem to show multiple inheritance, e.g. stepwise class change, the weird *take long*, perhaps infinitival *to*, dual-class *near*, prefabs.
- Some theories allow multiple inheritance, e.g. Word Grammar or most flavours of Construction Grammar.
- More psychologically plausible anyway. ☺
- Word classes involve form and meaning. They exhibit prototype rather than Aristotelian membership.
- **Word classes are constructions.**
They can return as a meta-phenomenon.

Hudson (2006); Hoffmann & Trousdale (2013)

55

Presentation available

- Slides online at

<http://tinyurl.com/DD-download>

- Comments very welcome, but please don't quote without permission.

Merci !

56

References

- Beckner, Clay & Joan Bybee. 2009. A usage-based account of constituency and reanalysis. *Language Learning* 59.4, 27-46.
- Bresnan, Joan. 2007. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: Linguistics in search of its evidential base* (Studies in Generative Grammar 96), 75-96. Berlin and New York: Mouton de Gruyter.
- Bybee, Joan & Rena Torres Cacoulios. 2009. The role of prefabs in grammaticization: How the particular and the general interact in language change. In Roberta Corrigan, Edith A. Moravcsik, Hamid Ouali & Kathleen M. Wheatley (eds.), *Formulaic language*, 2 vols, vol. 1, *Distribution and historical change* (Typological Studies in Language 82), 187-218. Amsterdam: John Benjamins.
- Coates, Richard. 1989. A solution to the 'must of' problem. *York Papers in Linguistics* 14, 159-67.
- Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, William. 2005a. Logical and typological arguments for Radical Construction Grammar. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction grammars: Cognitive grounding and theoretical extensions* (Constructional Approaches to Language 3), 273-314. Amsterdam and Philadelphia: John Benjamins.
- Croft, William. 2005b. Word classes, parts of speech, and syntactic argumentation. *Linguistic Typology* 9.3, 431-41.
- Denison, David. 2013. Parts of speech: Solid citizens or slippery customers? *Journal of the British Academy* 1, 151-85.
- Denison, David. in prep. Why would anyone *take long*? In Kristel Van Goethem, Muriel Norde, Evie Coussé & Gudrun Vanderbauwhede (eds.), *Category change from a constructional perspective* (Constructional Approaches to Language). Amsterdam: John Benjamins.
- Denison, David. submitted. Ambiguity and vagueness in historical change. In Marianne Hundt, Sandra Mollin & Simone Pfenninger (eds.), *The changing English language: Psycholinguistic perspectives* (Studies in English Language). Cambridge: Cambridge University Press.
- Gahl, Susanne & Susan Garnsey. 2006. Knowledge of grammar includes knowledge of syntactic probabilities. *Language* 82, 405-10.
- Gahl, Susanne & Alan C. L. Yu. 2006. Introduction to the special issue on exemplar-based models in linguistics. *The Linguistic Review* 23, 213-6.
- Hoffmann, Sebastian. 2005. *Grammaticalization and English complex prepositions: A corpus-based study* (Routledge Advances in Corpus Linguistics 7). London and New York: Routledge.
- Hoffmann, Thomas & Graeme Trousdale (eds.). 2013. *The Oxford handbook of construction grammar* (Oxford Handbooks). Oxford: Oxford University Press.
- Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Hudson, Richard. 2006. *Language networks: The new Word Grammar*. Oxford and New York: Oxford University Press.
- Kayne, Richard. 1997. The English complementizer *of*. *The Journal of Comparative Germanic Linguistics* 1, 43-54.
- Langacker, Ronald W. 2008. Grammatical classes. *Cognitive Grammar: A basic introduction*. Oxford University Press.
- Levine, Robert D. 2012. Auxiliaries: *To*'s company. *Journal of Linguistics* 48.1, 187-203.
- Mindt, Dieter. 2002. What is a grammatical rule? In Leiv Egil Breivik & Angela Hasselgren (eds.), *From the COLT's mouth ... and others: Language corpora studies - in honour of Anna-Brita Stenström* (Language and Computers - Studies in Practical Linguistics 40), 197-212. Amsterdam and New York: Rodopi.
- Mukherjee, Joybrato. 2005. *English ditransitive verbs: Aspects of theory, description and a usage-based model*. Amsterdam: Rodopi.
- Pullum, Geoffrey K. 1982. Syncategorematicity and English infinitival *to*. *Glossa* 16, 181-215.
- Rauh, Gisa. 2010. *Syntactic categories: Their identification and description in linguistic theories* (Oxford Surveys in Syntax and Morphology 7). Oxford and New York: Oxford University Press.
- Sampson, Geoffrey & Anna Babarczy. 2014. *Grammar without grammaticality: Growth and limits of grammatical precision* (Topics in English Linguistics (TiEL) 254). Berlin and Boston: De Gruyter Mouton.
- Schütze, Carson T. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago and London: Chicago University Press.
- Schütze, Carson T. 2004. The nature of graded judgments. In Bas Aarts, David Denison, Evelien Keizer & Gergana Popova (eds.), *Fuzzy grammar: A reader*, 431-46. Oxford: Oxford University Press. Repr. of Section 3.3 of Schütze (1996).