That-clauses as complements of verbs or nouns

David Denison
LMEC 6, Uppsala
August 2017

Introduction

Innovatory usage?

X <u>articulates that</u>: 'A Standard is that by which we [...]'
(2011, student coursework)

In the first issue of Salvage, Neil Davidson <u>mooted that</u> neoliberalism may be undermining the basis for capital accumulation itself. (2015, China Miéville, *Salvage*)

I <u>propound that</u> an argument could be made for ...
(2015, student coursework)

The group also <u>publicised that</u> they were "re-enacting an Israeli checkpoint outside the university's main library". (2016, BBC News website)

Denison (2009, 2011)   3

Erroneous usage?

- Often relatively unskilled writers, insecure about written expression.
  - Asking them about grammaticality not helpful.

> '... communication verbs controlling *that*-clauses (apart from *say*) are most frequent in academic prose'

- Are writers using a thesaurus? If so,
  - to avoid risk of plagiarism?
  - or for 'elegant variation'?

Biber et al. (1999: 668)   4

Background

- First noticed some years ago and mentioned in presentations. Speculation about
  - who does it (just British students?)
  - since when (very recent?)
  - why (thesaurus?)

- Danger of 'recency illusion'

- Move from anecdotal observation to systematic study.

Zwicky (2005)   5

Plan

- What it is, how it happens
- Corpus linguistic approach
- Problems
- Partial solution
- Why it matters:
  - nature of grammaticality
  - what is a native speaker?
  - computer-mediated language
- Conclusion

6

What

Unexpected V + *that*-clause …

1. X's study (2009), which **advocates that** the subjunctive is in fact replacing the periphrastic [...]
2. this **highlights** once more **that** [...]
3. with X **criticising that** the French influence was sporadic
4. X **defines that** "grammar may be regarded either from a theoretical or practical point of view. [...]"
5. This study has **displayed that** older participants have more stable and confident results than [...]
6. X (1966) [...] **poses that** the informality of *try and* leads to [...]
7. X, (1994) **puts forward that** chimps often just imitate the messages of the trainer
8. which can be reinforced by X, who **utters that**, "In other locations [...]"

✔
?*
*
*
*
*
*
*

where 'shell noun' expected

- **Definition**: Nouns like *fact, claim, idea, worry, perception* that can have a dependent *that*-clause to lexicalise them
- Shell noun + dependents can fill various slots, including object of a matrix verb.
- (not discussed here) Are shell nouns
  - defined by semantics, function, discourse, … ?
  - enumerable or open-ended?
  - best label (cf. 'general nouns', 'signalling nouns', etc.)?

Hunston & Francis (2000), Schmid (2000), Flowerdew & Forest (2015), etc.

How

Analogy?
If so, then why only now?

Shell + *that* ~ V + *that*

If you just **accept the fact that** there's no self [...] (2007, COCA)

You have to **accept that** this could happen again. (2015)

Poland also **espoused the idea that** the COMECON Members should [...] (1990)

by analogy >

⏱The aforementioned authors **espouse that** students from the age of four to eight are aware of racial difference (2011)

Denison (2011)

Minor contribution?

Some verbs disallow *that*-clause in active:

*The grammar expresses [$_{CP}$ that the rule is obligatory]

- even though OK in passive:

[$_{CP}$ That the rule is obligatory] is expressed by the grammar

But rather too formal to be a plausible route here.

Newmeyer (2003)

Where, when, how often

Simple tagged search

Online

Systematic research in corpora

- Need whole of 20th century and up to present.
- No parsed corpora have size or chronological range.
- Only COHA has right coverage.
- Tagged and lemmatised, so online search for V + *that* should distinguish
  - *advance, display* Verb ~ *advance, display* Noun
  - *that* Conj/Relativiser ~ *that* Determiner/Pronoun

Systematic research in corpora

- Need whole of 20th century and up to present.
- No parsed corpora have size or chronological range.
- Only COHA has right coverage.
- Tagged and lemmatised, so online search for V + *that* should distinguish
  - *advance, display* <u>Verb</u> ~ *advance, display* Noun
  - *that* <u>Conj</u>/Relativiser ~ *that* Determiner/Pronoun
- Limit search to *that* immediately after head N/V.

Sample searches in COHA

[display].[v*] that.[cst]
- Accuracy 5/41 = 12%, of which 1/5 somewhat relevant
- (Marginal examples not in tagger's training data?)

[advance].[v*] that.[cst]
- Accuracy 66/137 = 48%, 8/66 relevant

[define].[v*] that.[cst]
- Accuracy 23/42 = 55%, 4/23 relevant

- plus limitations of interface.
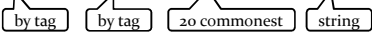- Need to use downloaded COHA.
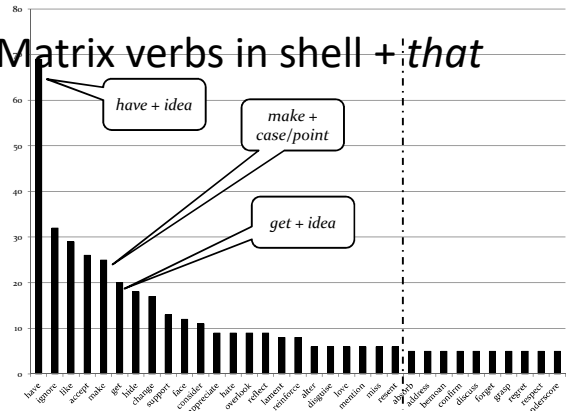
Partly untagged search

Offline, then online

Find shell + *that*

- Python script searches for

  ... V[≠*be*] + article + shell noun + *that* ...

  | by tag | | by tag | | 20 commonest | | string |

- Precision: 76% shell + *that*.
- Filter out 24%, leaving 948 examples from 2000s.
- Sort by matrix verb.
- Choose 20 most common matrix verbs in shell + *that* that could maybe occur in V + *that*.

19

Matrix verbs in shell + *that*

have + idea
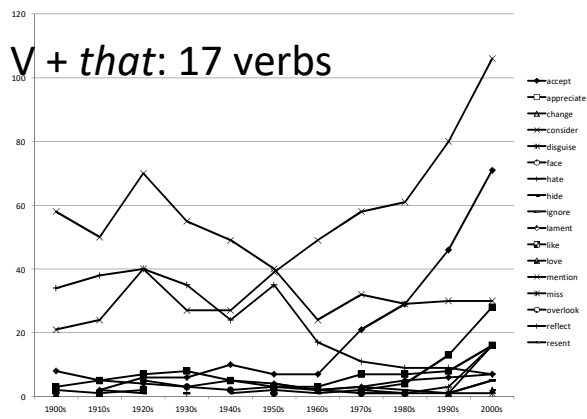
make + case/point

get + idea

20

Find V + *that* with common V

- Script searches for any of

  ... V + *that* + pronoun + V ...
  ... V + *that* + *there* + V ...
  ... V + *that* + N + V ...
  ... V + *that* + N + N + V ...
  ... V + *that* + article/possessive + N + V ...
  ... V + *that* + determiner + N + V ...

- = rough-and-ready shapes of common *that*-clauses
- Greater precision worth loss of recall
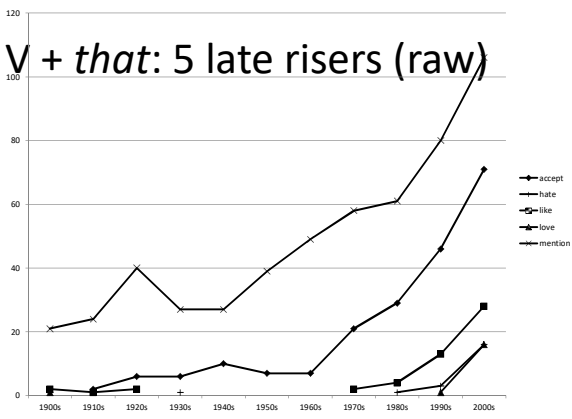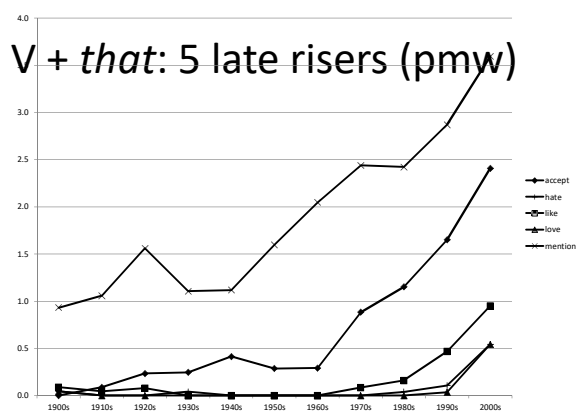- Search all decades from 1900s to 2000s.

21

V + *that*: 17 verbs

22

V + *that*: 5 late risers (raw)

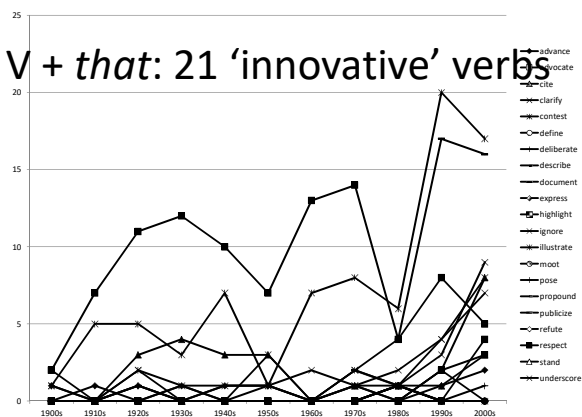

23

V + *that*: 5 late risers (pmw)

24

Find V + *that* with 'innovative' V

- Python search for verbs used in student writing/ journalism collection
- Numbers so low that I went back to online COHA and searched there e.g. for

[**respect**].[**v***] **that**

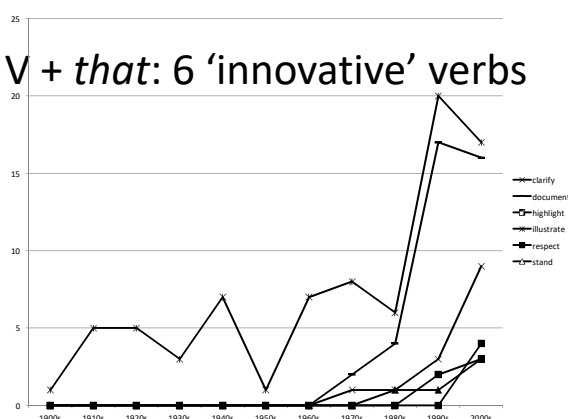- which increased the haul somewhat.

- Are they in fact new usages?

V + *that*: 21 'innovative' verbs

Legend: advance, advocate, cite, clarify, contest, define, deliberate, describe, document, express, highlight, ignore, illustrate, moot, pose, propound, publicize, refute, respect, stand, underscore

V + *that*: 6 'innovative' verbs

Legend: clarify, document, highlight, illustrate, respect, stand

Limitations on searches

- Only search for *that* immediately after shell or V
  - Sample 200 of *case* [… up to 9 words …] *that* found 10/11 shell + *that* were adjacent
  - Sample 200 of *accept* [… up to 9 words …] *that* found 23/25 V + *that* were adjacent
- Assume *that*-clause to left of shell/V too rare to matter.

- But if change 'sneaks in' via non-salient contexts, these rare contexts might be crucial, if they exist.
- Impossible to search for them in unparsed corpus

Results

Inconclusive findings

- Some V + *that* are indeed (fairly) recent, but can't pin blame on online thesauri
- Have no solid explanation for dating of changes – but now less of a problem for claim of analogical change.

- Am guilty of Recency Illusion
- Anglo-American difference, e.g. with *like + that*

- COHA not big enough!

Grammar and grammaticality

- Unreliability of grammaticality judgements, including my own
- Mindt claims that grammatical rule has some 5% of exceptions due to language change, linguistic creativity, intended ungrammaticality, etc.
- Generative tradition assumes individual's grammar is discrete and clear-cut.
- Usage-based tradition suggests that grammatical knowledge is inherently probabilistic.

Mindt (2002), Mukherjee (2005)

32

Whose grammar?

33

Error vs. innovation

- *Native speakers* (mis)using words and patterns *in writing* that would be rare or non-existent in their everyday speech.
- Situation resembles L2 learning or new Englishes(?).
- Studies normally distinguish natural historical change in (L1) language from change due to language contact, and certainly from errors in L2 learning.
- What seems like error in retrospect can be innovation in diachrony. But some of these still seem like errors.
- To what extent is this a natural, internal change?

34

What next?

35

More Qs than As

- And good answers would need more time and resources than questions deserve!
  - psycholinguistic investigation of speakers' knowledge
  - far larger datasets and more sophisticated search algorithms
  - sensitivity to genre and register
    - compare student writing with standard corpora of (mainly) copy-edited material
  - research on influence of technology on language use

36

Acknowledgement

- Comments welcome.
- Bibliography on PDF, on my website.

37

References

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson.

Bresnan, J. W. 1995. Category mismatches. In A. Akinlabi (ed.), *Theoretical approaches to African linguistics*, 19-46. Trenton, NJ: Africa World Press.

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.

Denison, David. 2009. A new class of verbs taking *that*-clause complements. Paper presented at ICLCE3, Symposium on Current Change in the English Verb Phrase, London.

Denison, David. 2011. ISLE highlights? Paper presented at ISLE2, Boston MA.

Flowerdew, John & Richard W. Forest. 2015. *Signalling nouns in English: A corpus-based discourse approach*. 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.

Grimshaw, Jane B. 1982. Subcategorization and grammatical relations. In Annie Zaenen (ed.), *Subjects and other subjects: Proceedings of the Harvard Conference on the Representation of Grammatical Relations*, 53-5. Bloomington IN: Indiana University Linguistics Club.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Hunston, Susan & Gill Francis. 2000. *Pattern grammar: A corpus-driven approach to the lexical grammar of English* (Studies in Corpus Linguistics 4). Amsterdam: John Benjamins.

Jacobson, P. 1994. The lexical entailment theory of Control and the Tough-construction. In I. Sag & A. Szabolcsi (eds.), *Lexical matters*, 269-99. Stanford: CSLI Publications.

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.

Newmeyer, Frederick J. 2003. Theoretical implications of grammatical category-grammatical relation mismatches. In Elaine J. Francis & Laura A. Michaelis (eds.), *Mismatch: Form-function incongruity and the architecture of grammar* (CSLI Lecture Notes 163), 149-78. Stanford: University of Chicago Press for CSLI Publications.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London and New York: Longman.

Ross, John Robert. 1973. Nouniness. In Osamu Fujimura (ed.), *Three dimensions of linguistic theory*, 137-257. Tokyo: TEC for Tokyo Institute for Advanced Studies of Language. Repr. Aarts et al. (eds.) *Fuzzy Grammar* (2004), pp.351-422.

Sampson, Geoffrey & Anna Babarczy. 2014. <u>Grammar without grammaticality:</u> *Growth and limits of grammatical precision* (Topics in English Linguistics (TiEL) 254). Berlin and Boston: De Gruyter Mouton.

Schmid, Hans-Jörg. 2000. *English abstract nouns as conceptual shells: From corpus to cognition* (Topics in English linguistics 34). Berlin and New York: Mouton de Gruyter.

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).

Visser, F. Th. 1963-73. *An historical syntax of the English language*, 4 vols. Leiden: E. J. Brill.

Webelhuth, Gert. 2008. A lexical-constructional approach to "movement mismatches". Paper presented at Fifth international Conference on Construction Grammar, The University of Texas at Austin.

Zwicky, Arnold. 2005. Just between Dr. Language and I. Language Log. http://itre.cis.upenn.edu/~myl/languagelog/archives/002386.html <Accessed 4 May 2006>.