

What's *that*?

Rob Truswell and Nikolas Gisborne
rob.truswell@ed.ac.uk, n.gisborne@ed.ac.uk

Cambridge, 28/3/17

Two types of competition

- ▶ Part of language **use** is selecting among alternative forms which realize a communicative intention.
- ▶ Part of language **acquisition** involves pairing a given form with grammatical information.
- ▶ Both of these involve **competition**, but in different ways.
 - ▶ Among forms paired with a given function.
 - ▶ Among specifications of the function of a given form.

The functions of functional heads

- ▶ Acquisition of content word meaning has been extensively investigated.
- ▶ But content words are the easy ones.
- ▶ Functional vocabulary is harder in many respects.
 - ▶ Ambiguity is the norm.
 - ▶ Mutual exclusivity not such a strong pressure.
 - ▶ Miscommunications less obvious and/or less serious.
 - ▶ Pairings between category and denotation more fluid.
- ▶ Learners are quick to figure out that *that* is a word.
- ▶ But it is much harder for them to answer a question like 'What is *that*?'.

Functions are not denotations

- ▶ Moreover, a lexical entry *qua* phon/syn/sem triple doesn't directly contain a specification of the functions that the lexical item can realize.
- ▶ Rather, a lexical entry specifies lexical syntactic information (e.g. category) and lexical semantic information (e.g. model-theoretic denotation).
- ▶ The contexts in which a lexical item can be used are a consequence of the compositional interactions among lexical entries, both overt and covert (functional heads, coercion).
- ▶ The distribution of responsibility for these compositional phenomena across component parts is not given.

Change

- ▶ Most well-studied cases of syntactic change involve:
 - ▶ A stable set of forms
 - ▶ A stable set of functions
 - ▶ A dynamically changing set of alignments of forms with functions.
- ▶ A salient fact about the history of English: *gorp*-support never got off the ground.
- ▶ Rather, the grammar of English changed because of a recurring novel answer to the question ‘What is *do*?’.

Innovations do recur

- (1) de fout **wie** hun eigenlijk maken
the mistake who they actually make
'the mistake which they actually make'
(Johan Cruyff, via Boef 2012)
- (2) adnominal adjectives (those **who** are not modifying the noun
predicatively) (Belk 2016: 179)

Grammar competition

- ▶ The major approach to competition in syntactic change comes from Kroch (1989).
- ▶ This presupposes competition among functionally equivalent forms.
- ▶ But the problem is really a two-dimensional one, and the loci of competition less encapsulated.
- ▶ The classic cases of grammar competition are a special case.
- ▶ In the general case, not immediately clear that Kroch's competition-based explanations for
 - ▶ S-curves
 - ▶ Constant Rate Effectare valid.

This talk

- ▶ We will try to understand the brief period in 13th-century English when virtually every relative clause (headed or free) was introduced by *that* (**peak *that***).
 - ▶ This wasn't true in Old or Very Early Middle English.
 - ▶ It hasn't been true since Middle Middle English.
- ▶ Part of this is straightforward.
 - ▶ Demonstrative relatives disappeared as inflected demonstratives disappeared (slowly).
- ▶ Part of it can be understood in Kroch's terms.
 - ▶ Between c.1150–1250, *þe*, which had been the most common complementizer in relative clauses, was replaced by *that*.
- ▶ Part of it (*wh*-relatives) only really makes sense when you consider competition among denotations.

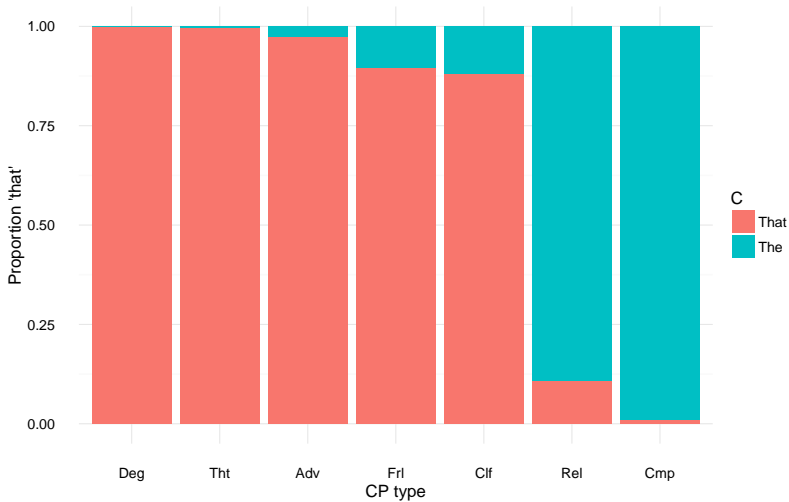
This talk

- ▶ *Wh*-phrases in some contexts must be indefinite descriptions.
- ▶ In other contexts, they must be definite descriptions.
- ▶ The indefinite denotation is old, the definite denotation is newer.
- ▶ The change from indefinite to definite is possible because, within the scope of certain operators, it doesn't make much interpretive difference.
- ▶ (And the denotation of *wh*-forms in interrogatives may be neither of the above).
- ▶ The peak-*that* period corresponds to a lull between the death of indefinite *wh*-phrases and spread of definite *wh*-phrases.

þe and *þæt*

- ▶ OE had two functionally specialized finite complementizers.
- ▶ *That* occurs in complement clauses, adverbial clauses (*if that*), degree clauses (*so much that*), most free relatives, most clefts.
- ▶ *þe* occurs in *the*-comparatives (*the more þe he ate*) and most headed relatives.
- ▶ This specialization is nearly categorical.

p_e and $\beta_{æt}$



Relatives are messier

- ▶ The ‘complementarity’ is more of a strong tendency in relatives (including clefts).
- ▶ *þe* sometimes occurs where *þæt* is expected. I don’t know why.
- ▶ *þæt* sometimes occurs where *þe* is expected.
- ▶ In many cases, relativizer *þæt* is plausibly a demonstrative pronoun (*þæt*: DEM.N.SG.NOM/ACC).
- ▶ But in others, *þæt* displays the hallmark of OE relative complementizers (Allen 1977): P-stranding (Mitchell 1985).

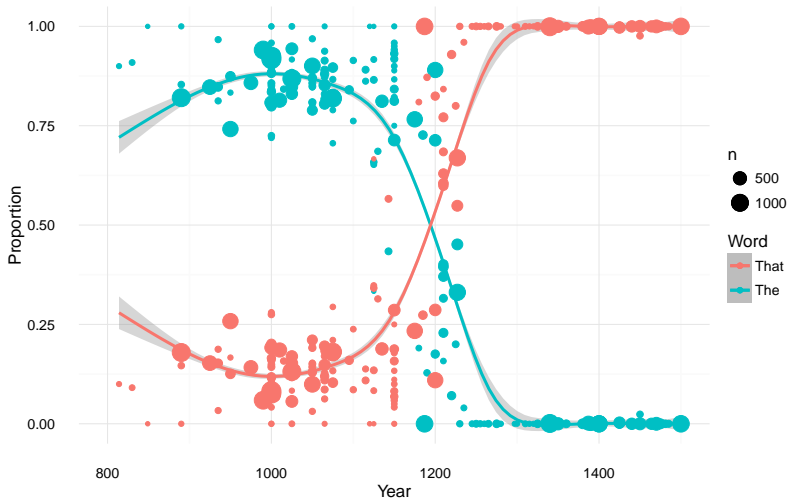
(3) *þa* ... *næfde* *he* *scyld* *æt* *honda*, *þæt* *he* *þone*
When NEG.had he shield at hand that he the
cuning [mid ___] *scyldan* *meahte*
king with shield might
‘When . . . , he did not have a shield to hand with which
he could shield the king.’

(cobede, Bede_2:8.122.19.1160)

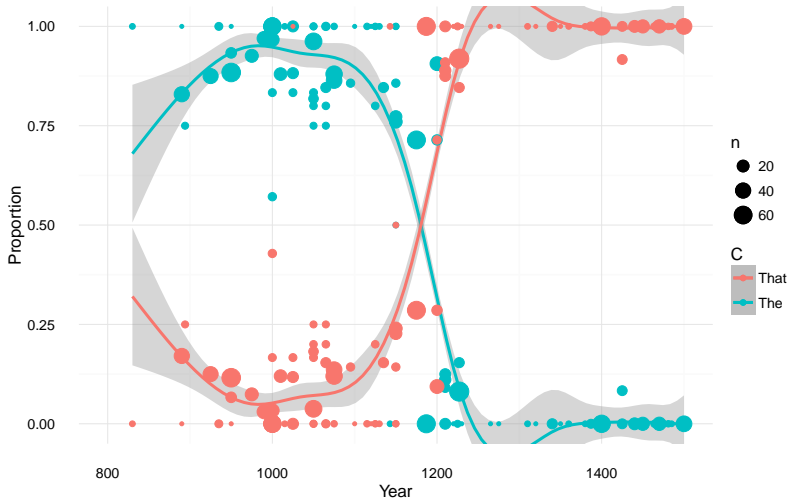
From the back of an envelope

- ▶ *Pe* is undoubtedly a complementizer.
- ▶ It occurs in 16,846 OE relatives, of which 826 involve P-stranding (4.9%).
- ▶ There are 76 occurrences of relative *þæt* with P-stranding (out of 2,715 relative *þæt*).
- ▶ Although we can't know the incidence of relative-complementizer *þæt* as opposed to relative-specifier *þæt*, this suggests an estimate of $76/0.049 = 1550$ occurrences in OE.

pe vs. *βæt*, all tokens



Pe vs. $\beta_{\text{æt}}$, P-stranding only



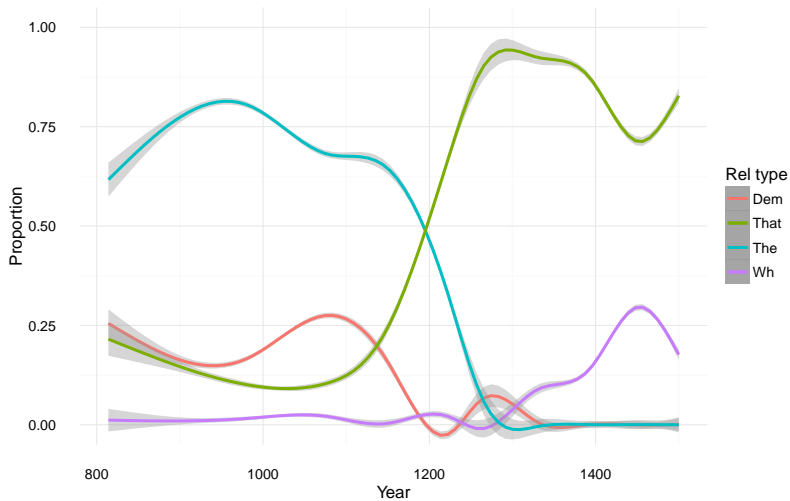
be vs. *þæt*: Summary

- ▶ *þæt* was always a complementizer.
- ▶ It was even always a relative complementizer, occurring as such with low frequency.
- ▶ Other demonstrative forms barely behaved like this (c.10x more common with *þæt* than other demonstratives).
- ▶ *þæt* killed *be* within a couple of generations, following a trajectory compatible with an S-curve (but so abrupt that the middle part of the trajectory is unclear).
- ▶ Compatible with classical grammar competition, though even here, distinctive transient grammars (see McIntosh 1948 on animacy effects in *Peterborough Chronicle*).

A four-way fight

- ▶ The competition between *þe* and *þæt* overlapped with the loss of demonstrative relative specifiers, and subsequent introduction of interrogative relative specifiers.
- ▶ Neither the specifier nor the complementizer has to be present.
- ▶ So the bigger picture could be construed as:
 - ▶ a 4-way fight (DEM, *þe*, *þæt*, WH);
 - ▶ two simultaneous 3-way fights (DEM, WH, $\emptyset \times \textit{þe}, \textit{þæt}, \emptyset$);
 - ▶ a 9-way fight (crossing the two 3-way fights).
- ▶ For the sake of our sanity, we'll stick with a 4-way fight and hope we're not losing much.

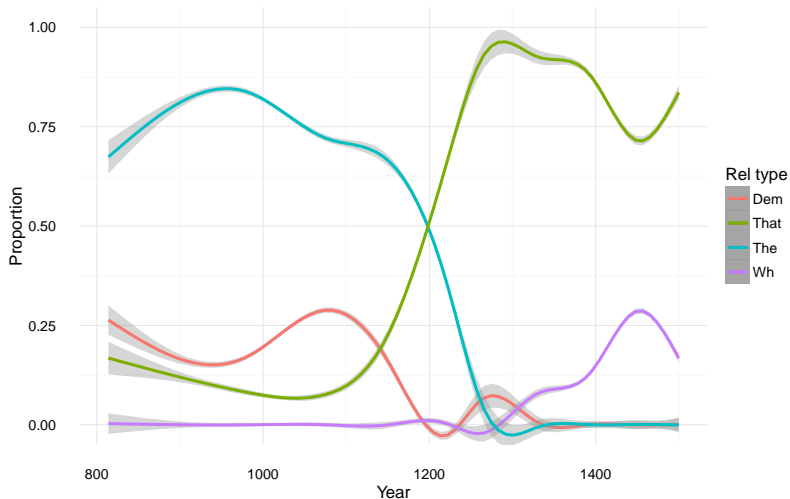
The four-way fight over time



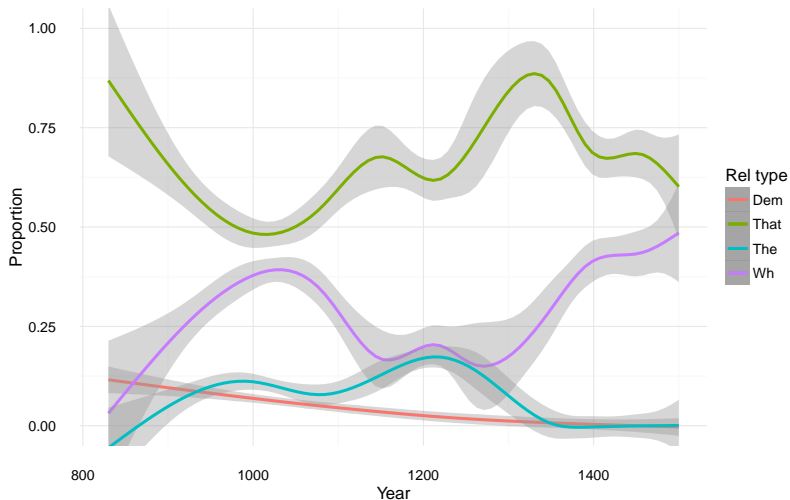
Whatever

- ▶ This shows us the peak-*that* window: from c.1250–1400, around 90% of relatives were formed with *that*.
- ▶ This is much higher than before or since.
- ▶ But it doesn't look very interesting.
- ▶ *Pe* was the dominant strategy, then *that* became the dominant strategy, and it was even more dominant because the demonstratives (secondary strategy) died.
- ▶ The low-level noise c.1300 is uninterpretable, probably caused by sparse data.
- ▶ However, a more interesting pattern is revealed when we factor out headed vs. free relatives.

The four-way fight in headed relatives



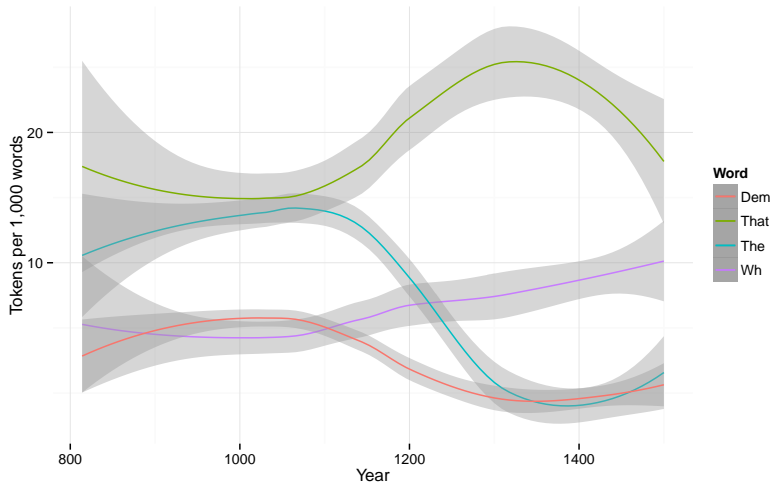
The four-way fight in free relatives



The N-shaped trajectory of free *wh*-relatives

- ▶ Free *hw*-relatives were well-established in OE.
- ▶ They slump appreciably in early ME.
- ▶ They rise again in late ME, and are now the only form of free relative.
- ▶ In its pomp, *that* not only killed *þe*, and replaced demonstratives in [Spec,CP], but took a chunk out of *wh* forms too.
- ▶ This is surprising, because *wh*-forms were in no danger of disappearing.

Word frequencies over time



Wh-form semantics

- ▶ We identify three different denotations for *wh*-forms.

1. Indefinite

- (4) and gif hwa hyt bletsað, þonne ablinð seo dydrung.
and if who it blesses then ceases the illusion
'and if anyone blesses it, then the illusion is dispelled.'
(coaelhom,+AHom_30:4.4082)

2. Definite

- (5) Gemyne, [hwæt Sanctus Paulus cwæð]
remember what Saint Paul said
'Remember what Saint Paul said.'
(cogregdC,GDPref_and_3_[C]:15.207.28.2739)

3. Interrogative

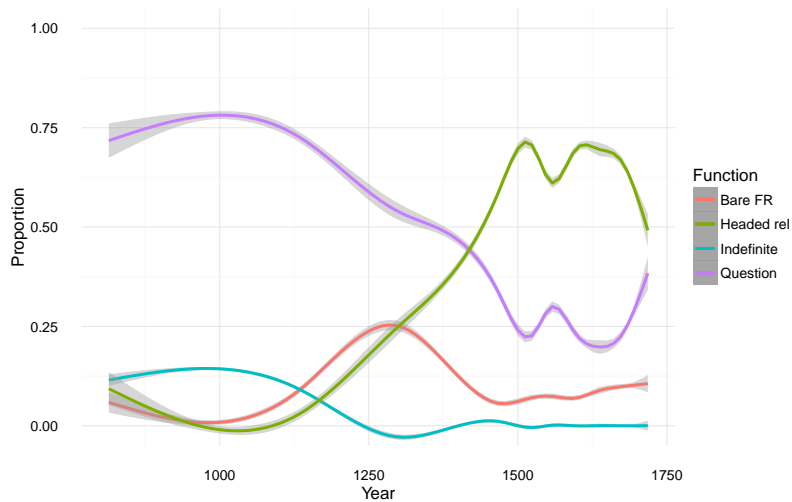
- (6) Hwær lede ge hine?
where lead you him
'Where are you leading him?' (coaelhom,+AHom_6:77.915)

- ▶ We can track the diachronies of these denotations, by tracking the frequency of constructions which require one of them.

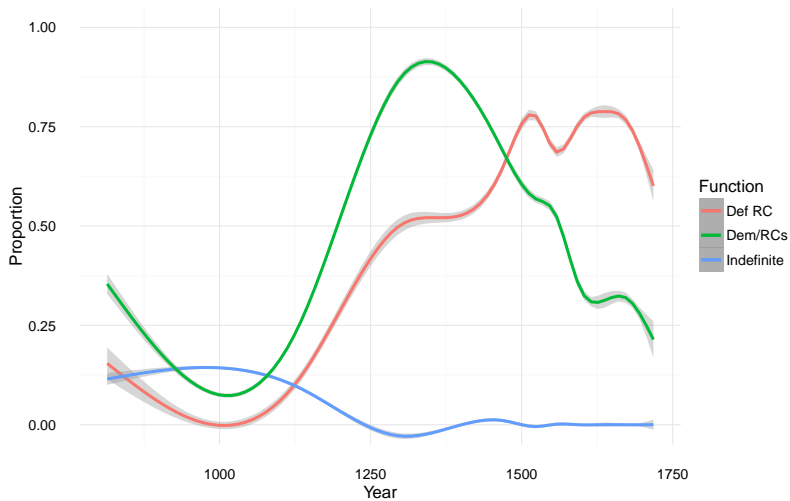
Denotations and constructions

- ▶ *Wh*-indefinites are indicative of an indefinite denotation.
- ▶ Nonrestrictive headed relatives are indicative of a definite denotation.
- ▶ Bare free relatives may be too, though this is less clear (e.g. in scope of subjunctive?)
- ▶ In other constructions, the distinction is unclear or makes less sense (*swa*-free relatives, restrictive relatives, interrogatives).
- ▶ If it's less clear to us, it's less clear to the learner, and so ripe ground for change.
- ▶ Implication: relatively minor *wh*-constructions may be disproportionately important to a learner trying to figure out lexical meaning.

Denotations over time



Alignment with peak *that*



Summary

- ▶ As we approach peak *that*, three things happen:
 1. Inflected demonstratives mainly disappear.
 2. *þe* mainly disappears.
 3. *Wh*-forms lose their indefinite denotation.
- ▶ The spread of definite *wh*-denotations to headed relatives brings an end to peak *that*.
 - ▶ Bare (definite) free *wh*-relatives increase in frequency in the 13th century. This clear trend doesn't much dent peak *that* because free relatives are relatively rare (headed relatives are 10x more common).
 - ▶ Headed *wh*-relatives follow 100 years later (looks impressionistically like a CRE). We have argued elsewhere that early headed *wh*-relatives are nonrestrictive, with the *wh*-form to be analysed as a (definite) discourse anaphor.

How is this possible?

- ▶ Definite and indefinite are typically construed as being in opposition.
- ▶ They're also bound up in separate grammaticalization pathways.
 - ▶ Demonstrative > definite
 - ▶ Numeral > indefinite
- ▶ So a switch from indefinite to definite sounds outlandish.
- ▶ It's only possible because OE/EME *wh*-forms almost always occur in the scope of a quantifier over situations or propositions.

The distribution of *wh*-indefinites

- ▶ *Wh*-indefinites in matrix clauses are rare.
- ▶ Most common context: probably left-peripheral *if*-clauses.

(7) and gif hwa hyt blelsað, þonne ablinð seo dydrung.
and if who it blesses then ceases the illusion
'and if anyone blesses it, then the illusion is dispelled.'
(coaelhom,+AHom_30:4.4082)

- ▶ *If* expresses universal quantification over situations. The indefinite introduces a discourse referent within the scope of that universal quantification (e.g. Kamp 1981).

(8) $\forall s.[\exists x.(bless'(x, y, s) \wedge min(s))][\exists s' \supseteq s.(cease(\text{the illusion}, s'))]$

Two types of free *wh*-relative

- ▶ We have already met bare, postnominal free *wh*-relatives.
- ▶ The better-known OE free *wh*-relatives have *swa* either side of the *wh*-phrase.
- ▶ *Swa* ... *swa* can be paraphrased by *-ever*.

(9) Soðlice [swa hwar swa Israhela bearn wæron], þar wæs
Truly so where so Israel's children were, there was
leoht.
light
'all the children of Israel had light in their dwellings.'

(cootest, Exod:10.23.2788)

- ▶ These relatives are classically called 'indefinite', though the term is rarely justified.

Indefinite *wh*-relatives often look definite ...

- ▶ Truswell & Gisborne (2015) argued that OE *swa* ... *swa* FRs were definite descriptions relative to possible worlds, following recent research on PDE *-ever*-FRs (Jacobson 1995, Dayal 1996, von Stechow 2000).

(10) There's a lot of garlic in whatever Arlo is cooking.

(11) Do whatever the babysitter says.

- ▶ This analysis is clearly appropriate for some clause-final OE examples.

(12) *Gap to Iosepe & doþ [swa hwæt swa he eow secge].*
Go to Joseph and do so what so he you.DAT say.SBJ
'Go unto Joseph; what he saith to you, do.'

(cootest, Gen:41.55.1711)

... but some of them may be indefinite

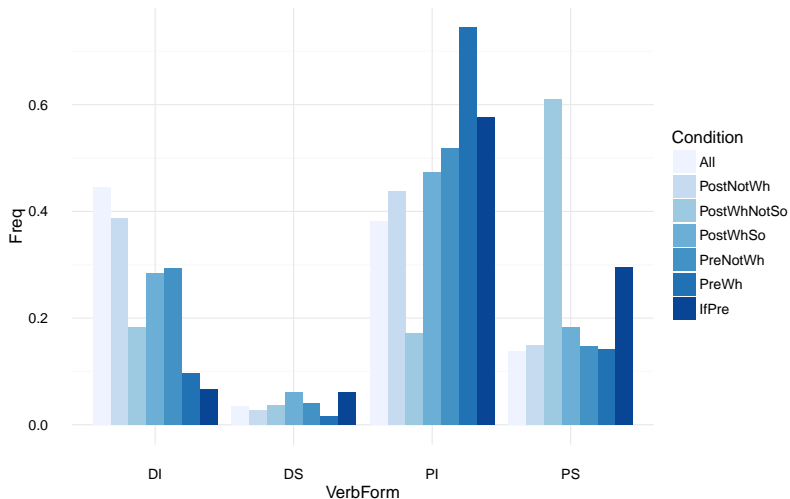
- ▶ However, to our knowledge, all left-adjoined free *wh*-relatives are compatible with conditional + indefinite paraphrases.
- ▶ The left-adjoined examples are demonstrably different to other free relatives, and demonstrably similar to conditionals, in favouring present indicative.

(13) Swa hwær swa þæt hold bið, þider gadriað þa earnas.
so where so that body is thither gather the eagles
'Wherever that body is, the eagles gather there'
(coaelhom,+AHom_19:38.2692)

(14) Drihten, gyf he slæpð, he byð gehealden.
Lord if he sleeps he is healed
'Lord, if he sleep, he shall do well.'(coaelhom,+AHom_6:32.877)

- ▶ These examples are *compatible* with a conditional + indefinite analysis, but also with a modalized definite analysis.
- ▶ That kind of ambiguity suggests a reanalysis context.

Verb forms in free relatives and conditionals



Unfinished business

- ▶ We don't yet know the fine details of the diachrony of left adjoined *swa* . . . *swa* free relatives.
- ▶ They don't disappear when *wh*-indefinites disappear.
- ▶ But we wouldn't necessarily expect them to.
- ▶ We *would* expect them to behave less like conditional + indefinite and more like modalized definite.
- ▶ Early ME also sees the development of new R-pronoun uses of *where* (*wherein* etc., unattested in OE).
- ▶ These are, to our knowledge, never used as indefinites or in indefinite-like free relatives.
- ▶ Tracking forms like *wherein* may be informative about the emergence of definite *wh*-forms.
- ▶ To be continued . . .

Summary

- ▶ Most of the time, *wh*-phrases occur in interrogatives.
- ▶ But in two early constructions, they plausibly have an indefinite denotation.
 - ▶ *Wh*-indefinites
 - ▶ Left-adjoined free *wh*-relatives
- ▶ The earliness of these constructions should not surprise us: the following pathway appears to hold:
 - ▶ Asyndetic conditionals + *wh*-indefinites \rightsquigarrow
 - ▶ *Wh*-correlatives (Belyaev & Haug 2014) \rightsquigarrow
 - ▶ Free *wh*-relatives \rightsquigarrow
 - ▶ Headed *wh*-relatives
- ▶ In constructions from early Middle English on, they have a definite denotation.
 - ▶ Bare free *wh*-relatives
 - ▶ Nonrestrictive headed *wh*-relatives

Summary

- ▶ The shift from indefinite to definite is only possible because interactions with other operators can lead to contexts (especially with free relatives) where definiteness has no clear truth-conditional consequences.
- ▶ The denotation of the *wh*-form is in a symbiotic relationship with the operators which have scope over the *wh*-form.
- ▶ The change in the denotation of *wh*-forms was not obviously triggered by any classical grammar competition.
- ▶ But it contributed to the 'peak *that*' period:
 - ▶ Indefinite free *wh*-relatives declined.
 - ▶ Subsequently, definite relatives emerged.
 - ▶ In between, *that* was all there was.

Implications

- ▶ Grammar change is change in **associations**.
- ▶ Learning the associations between forms and denotations is hard.
- ▶ This suggests a second approach to modelling change, which is orthogonal to classical grammar competition.

References

- Allen, C. (1977). *Topics in Diachronic English Syntax*. PhD thesis, University of Massachusetts, Amherst, MA.
- Belk, Z. (2016). *Attributes of Attribution*. PhD thesis, University College London.
- Belyaev, O. & Haug, D. (2014). The genesis of *wh*-based correlatives: From indefiniteness to relativization. Paper presented at Sinn und Bedeutung 19, Göttingen.
- Boef, E. (2012). *Doubling in Relative Clauses: Aspects of Morphosyntactic Microvariation in Dutch*. PhD thesis, Universiteit Utrecht.
- Dayal, V. (1996). *Locality in Wh-quantification: Questions and Relative Clauses in Hindi*. Dordrecht: Kluwer.
- Jacobson, P. (1995). On the quantificational force of English free relatives. In E. Bach, E. Jelinek, A. Kratzer, & B. Partee (Eds.), *Quantification in Natural Languages* (pp. 451–486). Dordrecht: Kluwer.
- Kamp, H. (1981). Some remarks on the logic of change, part I. In C. Rohrer (Ed.), *Time, Tense, and Quantifiers: Proceedings of the Stuttgart conference on the logic of tense and quantification* (pp. 135–179). Tübingen: Max Niemeyer.
- Kroch, A. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change*, 1, 199–244.
- McIntosh, A. (1948). The relative pronouns *þe* and *þat* in Early Middle English. *English and Germanic Studies*, 1, 73–87.
- Mitchell, B. (1985). *Old English Syntax*. Oxford: Clarendon Press.
- Truswell, R. & Gisborne, N. (2015). Quantificational variability and the genesis of English headed *wh*-relatives. In Csipak, E. & Zeijlstra, H. (Eds.), *Proceedings of Sinn und Bedeutung 19*.
- von Stechow, K. (2000). *Whatever*. In Jackson, B. & Matthews, T. (Eds.), *SALT X*, (pp. 27–39)., Ithaca, NY: Cornell University.