

# Parataxis and hypotaxis in the history of English

<http://walkden.space/ICEHL21.pdf>

**George Walkden**

International Conference on English Historical Linguistics (ICEHL) 21  
Leiden (online), 11 June 2021

## The context

Today's talk is part of a project to assess the following claim(s):

**parataxis > hypotaxis**

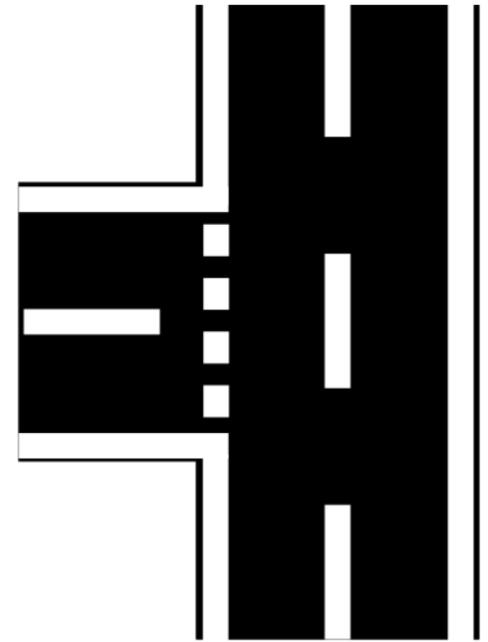
(where “>” is to be read as “precedes”)

# parataxis > hypotaxis

## The idea that parataxis precedes hypotaxis has a long history:

- The term parataxis in its modern sense was introduced by Thiersch (1826) in the context of historical Greek (opposed to syntaxis there; hypotaxis only in later works)
- Very prevalent in historical linguistics before the advent of structuralism (e.g. Gildersleeve 1883; Delbrück 1900: 411; Small 1924: 125)
- Reiterated in more recent works with a functionalist orientation (e.g. Jucker 1991: 203; Deutscher 2001: ch. 11; Dąbrowska 2015: 230)
- Influential in a different form in the generative tradition (O'Neil 1977, Kiparsky 1995)

# Roadmap



## Part 1: establish what the content of the claim(s) is/are

- Spoiler alert: there are several different versions and much unclarity

## Part 2: evaluate parataxis > hypotaxis as structural reanalysis

## Part 3: evaluate parataxis > hypotaxis as change in frequency

- Precondition for this kind of research: parsed diachronic corpora

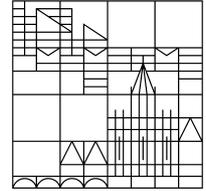
## Content warning



**Ideas don't arise in a vacuum. Some of the ways in which parataxis > hypotaxis is discussed in earlier literature make for uncomfortable reading today.**

- Mitchell (1985) approvingly quotes Small (1924: 125): “It may be laid down as a general principle that in the progress of language parataxis precedes hypotaxis.”
- Small’s following sentence: “The former is associated with the uncultivated mind; the latter, with the cultivated mind of civilized peoples.”
- Andrew (1940: 87): early Old English was characterized by “simply a lack of grammatical subordination such as we find in the language of children and some primitive people”.

**This doesn't mean that (every version of) the hypothesis is wrong, of course. But claims (in science as elsewhere) may persist because of ideology rather than merit.**



# Flavours of parataxis > hypotaxis

## What is parataxis? What is hypotaxis?

Harris & Campbell (1995: 284): “in approaching the question of whether hypotaxis develops out of parataxis we encounter the problem that different linguists have in mind different ideas of parataxis, and that at least some of them are vague”

**Basic intuition: parataxis is a ‘looser’ relation than hypotaxis.**

Unconnected utterances are paratactic. A complement clause is hypotactic.

But there’s a lot in between. What about:

- Coordination?
- Non-finite clauses?
- Embedded clauses other than complementation? (e.g. relativization)
- The non-clausal domain? (e.g. possessor recursion)

Some authors – e.g. Hopper & Traugott (2003: 177–178) – distinguish between hypotaxis and subordination.

**Before a claim about parataxis and hypotaxis can be evaluated, we need to know what these notions mean.**

## A cognitive claim?

In the twentieth and twenty-first centuries, there have been no explicit claims that parataxis > hypotaxis *in the historical record* is a direct result of change in human cognitive capacities.

Working assumption (uniformitarianism; see Walkden 2019):

the human faculty of language (in whatever sense we understand this notion) has not varied or changed during the period in which languages are historically attested.

**The claim that parataxis > hypotaxis is not a cognitive claim.**

- “The cognitive ability to handle finite complementation must have already been a feature of the human brain in the more distant past” (Deutscher 2001: 184–185)

## Variant: Early human languages lacked finite subordinate clauses (Givón 1979; Karlsson 2009; O’Neil 1977: 207, tentatively)

### Givón (1979: 306):

“certain types of languages—those which have only coordination (‘clause chaining’) but no subordination—are found only in preliterate ‘societies of intimates’”

- Givón’s claim is actually stronger: not restricted to finite clauses; uses word “only”.
  - Hard to view the Akkadian or Old Assyrian Empire as a non-literate society of intimates, but Akkadian is one of the languages where the case has been best made for lack of finite subordination (Deutscher 2001)
  - This claim needs to be assessed with reference to present-day languages; obviously, we have no historical records of non-literate societies pre-20<sup>th</sup>-century!
  - The history of English per se can’t tell us much about this version of the claim.

### Working definition of (clausal) subordination:

a CP is subordinated if it is dominated by another CP.

## Variant: Early human languages lacked (finite) subordinate clauses (Givón 1979; Karlsson 2009; O’Neil 1977: 207, tentatively)

### Are there languages without (finite) subordinate clauses?

- Roberts (2007: 174): “the claim that earlier stages of certain languages may have lacked subordination altogether violates the **uniformitarian hypothesis**, the idea that all languages at all times reflect the same basic UG, and so cannot be taken seriously in the approach adopted here.”
  - This of course depends on what we think UG contains.
  - “On the other hand, it is quite plausible that a language may lack **finite clausal subordination** of the familiar type” (2007: 174).
- Delbrück (1900) claims that Proto-Indo-European lacked finite subordination, on the grounds that finite subordinators are not reconstructable.
  - This is bad reasoning; cf. negation, basic vocabulary
  - Languages like Mandarin have subordination but no (overt) subordinator
  - Harris & Campbell (1995: 284): marker/structure fallacy

## Some cautionary notes

For this variant of parataxis > hypotaxis to be correct, there has to be an asymmetry between early and more recently spoken languages in having/lacking (finite) subordination.

- King & Cookson (1890: 204): “We cannot ... suppose that hypotaxis is of recent origin in language; for as far as we can go back in the history of human speech, we find the degradation of sentences to a completely subordinate position fully established.”
- Gildersleeve (1893: xxv): “we have to be on our guard. Hypotaxis is older than our record, and we cannot argue safely as to prehistoric processes”
- Even Karlsson (2009): “Evidence from many language families indicates that **non-finite clausal subordination** and initial stages of **finite clausal subordination** existed already in **preliterate languages**.”

As a categorical claim, it is certainly false that early (attested) human languages lacked (finite) subordinate clauses.

Does it hold statistically? I'm not aware of any studies addressing the question.

## Different claim: Diachronically, embedded structures replace adjoined structures

### Influential presentation: Kiparsky (1995) for Indo-European.

- Proto-Indo-European has only adjoined S; CP is innovated in the history of the subfamilies (e.g. Germanic).
- Wallenberg (2016) presents a supporting quantitative tendency: relative clause “extraposition” has been getting rarer for centuries in English, Icelandic, French and Portuguese. (Cf. also O’Neil 1977, Suárez-Gómez 2006)
- Whether Kiparsky’s version of parataxis > hypotaxis extends beyond Indo-European hasn’t ever really been addressed, to my knowledge.

### Today’s results won’t bear directly on this claim.

## Claims that we can test

The rest of this talk discusses claims that we can actually test with reference to the history of English.

**Claim 1: “Finite clausal complementation emerged from clausal juxtaposition”**  
(in English)

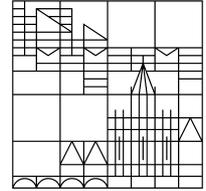
[I think that.] [John is here] > [I think [that John is here]]

This is a categorical claim, about reanalysis events.

**Claim 2: “Subordination gets more frequent over time”**  
(in English)

Early texts have more paratactic structures, later texts have more hypotactic structures.

This is a quantitative claim, about text frequencies.



**Does finite clausal  
complementation emerge  
from juxtaposition?**

## The standard scenario

The English declarative complementizer arises through reanalysis of a cataphoric demonstrative pronoun in a separate clause.

[I think that.] [John is here] > [I think [that John is here]]

(Behaghel 1877, 1928: 130; Paul 1920: 241; Hopper & Traugott 2003: 190–194)

Old English *þæt* “was reanalyzed from a pronoun which was a constituent of the matrix clause to a complementizer that had a whole clause within its scope” (Hopper & Traugott 2003: 192)

## Examples from Old English: cataphora

### Examples from Hopper & Traugott:

- (1) *Ða on morgenne gehierdun bæt þæs cyninges þegnas*  
when/then in morning heard-PL DEM DEM:GEN king's thanes  
*be him beæftan wærun bæt se cyning ofslægen wæs, þa ridon*  
who him behind were COMP the king slain was then rode  
*hie þider.*  
they thither  
'When in the morning the king's thanes who had been left behind heard that he had been killed, then they rode up there.' (ChronA (Plummer) 755.23)

Example "suggest[s] traces of a pronominal origin": the first *þæt* is a cataphoric pronoun.

- (2) *þæt gefremede Diulius hiora consul þæt þæt angin wearð*  
DEM arranged Diulius their consul COMP DEM beginning was  
*tidlice þurhtogen.*  
in-time achieved  
'Their consul Diulius arranged (it) that it was started on time.'  
(c. 880, Orosius 4 6.172.2)

## Examples from Old English: complementizers

### Examples from Hopper & Traugott:

- (3) *And þæs us ne scamað na, ac þæs us*  
And DEM:GEN we:ACC not shames never, but DEM:GEN we:ACC  
*scamað swyþe þæt we bote aginnan swa swa bec tæcan.*  
shames much COMP we atonement begin so as books teach  
'And we are not at all ashamed of that, but we are ashamed of this: of beginning  
atonement in the way that the books teach.' (c. 1010, WHom 20.3 160)

Case (not genitive!) shows that this *þæt* must be a complementizer.

- (4) *... þohte gif he hi ealle ofsloge, þæt se an ne ætburste þe he sohte.*  
... thought if he them all slew-SBJV, that that one not escape-SBJV that he sought  
'... thought that if he slew them all, the one he sought would not escape.'  
(c. 1000, ÆCHoml 5 82.10)

Position of *þæt* and use of subjunctive shows that this *þæt* must be a complementizer.

## Critique and alternative: Axel-Tober (2017)

**Axel-Tober (2017, ‘The development of the declarative complementizer in German’) presents a powerful critique of this standard (ultimately Neogrammarian) diachronic account, and an alternative theory.**

Axel-Tober’s arguments are presented for German, but the crucial change that led to the development of the complementizer could plausibly have been pre-Proto-West-Germanic (2017: e57) or earlier, and her conclusions hold for other West Germanic languages.

The paper has received little attention in the literature on English historical linguistics (cf. also Axel 2009; Axel-Tober 2012: ch. 2), and so I’ll present it in some detail here (with my own additions).

**Axel-Tober’s alternative scenario: the complementizer developed from a relative pronoun in a correlative structure.**

## Obstacles, 1

[I think that.] [John is here]

**Empirical problem: vanishingly little evidence for the input structure to reanalysis.**

Neither Old English nor Old High German robustly exhibits structures that look like the above.

(It could have vanished in the prehistory of all the languages in question, but if so, why?)

**Moreover, this scenario requires an independent clause to be reanalysed as an embedded clause. But matrix and embedded clauses have different word order patterns in all the early West Germanic languages.**

V-in-C is an inheritance from Proto-Germanic (Eythórsson 1995; Walkden 2014), at least in some kinds of matrix clauses. Embedded clauses lack V2.

Old English (van Kemenade 1997), Old High German (Axel 2007), and early Germanic languages in general (Walkden & Booth 2020) all have asymmetries between clause types.

## Obstacles, 2

[I think that.] [John is here]

**Typological problem: the grammaticalization pathway “demonstrative > complementizer” is unattested outside Germanic – no cross-linguistic support.**

Heine & Kuteva (2002: 106–107) adhere to the standard scenario, but their only robust examples are from Germanic, where the change in question could have been as early as Proto-Northwest-Germanic – i.e. there’s only one example.

“Still, more cross-linguistic data are required to establish that the present grammaticalization represents a more general phenomenon. Conceivably, the source of this grammaticalization is not a demonstrative but rather a relative clause marker”

**If the change really took place via reanalysis/grammaticalization of a demonstrative pronoun as a complementizer, this may be unique among the world’s languages.**

## Alternative scenario: correlative > complement

Recall examples (1)–(2), with a correlative structure.

- (2) ***þæt*** *gefremede Diulius hiora consul þæt þæt angin wearð*  
DEM arranged Diulius their consul COMP DEM beginning was  
*tidlice þurhtogen.*  
in-time achieved  
'Their consul Diulius arranged (it) that it was started on time.'  
(c. 880, Orosius 4 6.172.2)

In such examples, the *þæt*-clause is a relative clause headed by the “cataphoric” pronoun in the matrix clause (potentially with extraposition, as in this case).

The cataphoric pronoun may also be absent.

Early Germanic robustly allowed both subjects and objects to be omitted: see Walkden (2014: ch. 5).

Such structures form the input to reanalysis: what is originally a relative pronoun becomes a complementizer introducing complement clauses.

## Advantages of the alternative scenario

The grammaticalization pathway from relative marker to complementizer (introducing a complement clause) is extremely well attested cross-linguistically.

- Slavic *čto*: *wh*-word > relative marker > complementizer (Meyer 2017)
  - Greek *pou*: adverb > relative marker > complementizer (Roberts & Roussou 2003)
  - Chalcatongo Mixtec *xa=*: relative pronoun > complementizer
  - Thai *thîi*: relative marker > complementizer
  - Early Biblical Hebrew *she/asher*: relative pronoun > complementizer
- (Last three examples from Heine & Kuteva 2002: 254; see Lehmann 1995 for discussion)

In X'-theoretic terms, relative pronoun > complementizer is an instance of spec-to-head reanalysis – an extremely common trajectory (Roberts & Roussou 2003; van Gelderen 2004, 2011).

Examples of the input structure to reanalysis are abundantly attested.  
No need to posit reanalysis of a main clause as an embedded clause.

## Input structure: correlatives with unexpressed first element

- (5) **Dæt** *geworht is wæs lif on him sylfum*  
that worked is was life in him self  
'That which was created was life in himself' (coaelhom,+AHom\_1:33.17)
- (6) *for ðamþe ge gebiddað **bæt** ge nyton*  
because you worship that you NEG-know  
'because you worship that which you do not know' (coaelhom,+AHom\_5:44.710)

## Interlude: reanalysis and rebracketing

**The standard scenario involves radical syntactic rebracketing.**

In recent years, this kind of reanalysis has fallen into disfavour – apparent examples turn out to be better analysed in different ways.

Some case studies from the history of English:

- *for ... to* infinitivals (De Smet 2009; Garrett 2012; Whitman 2012)
- *going to* future (Garrett 2012)
- *have*-perfects (Whitman 2012)

**Axel-Tober's alternative scenario still involves a change in category via reanalysis, but no rebracketing (hence no violation of Whitman's 2001 'Conservancy of Structure' principle).**

**This is another point in favour of the alternative scenario.**

## Case study 1: summary

**Does finite clausal complementation emerge from juxtaposition + rebracketing? NO.**

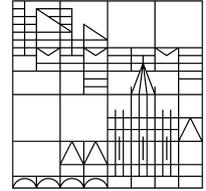
**Axel-Tober's (2009, 2012, 2017) alternative scenario – that complementizers emerge from relative markers – fares better in every respect than the standard scenario.**

- ✓ Examples of the putative input structure to reanalysis are abundantly attested
- ✓ No need to posit a cross-linguistically unique grammaticalization pathway
- ✓ No need to posit a reanalysis that clashes with normal word order patterns
- ✓ Grammaticalization of relative markers > complementizers is cross-linguistically common
- ✓ No need for radical rebracketing

**No parataxis > hypotaxis here: one type of finite embedded clause gets reanalysed as another type of finite embedded clause.**

Cf. also Weiß (2020: 50):

“Contrary to a widespread opinion in historical syntax and grammaticalization research ..., there is no reliable evidence for the parataxis-to-hypotaxis hypothesis nor for the discourse-to-syntax hypothesis. The standard opinion has proven wrong and must be abandoned.”



**Do (finite) subordinate  
clauses become more common  
over time?**

## Claim: Diachronically, hypotactic structures become more common

This is possibly the most frequently-encountered version of parataxis > hypotaxis in the literature. Usually interpreted with respect to (finite) clausal subordination in particular.

Dąbrowska (2015: 230):

- “Further telling evidence can be gleaned from historical data. The earliest written texts in a language are usually highly paratactic ... while later texts typically show more use of subordination. The historical increase in the frequency of subordination is gradual”

Karlsson (2009: 202):

- “It is a well-known fact that, mainly due to Latin influences, German and English were syntactically most complex in the seventeenth century and Swedish in the nineteenth”

This is a quantitative claim. It can only be assessed using quantitative data from historical corpora.

# Causal factors

The causal argumentation for this claim has varying levels of quality.

Three broad groups of possible factors:

- **Latin influence** (e.g. Karlsson 2009)
  - Non-finite clauses leading to an increase in finite subordination?
  - How likely is this to lead to a “real change” in principle?
- **Orality > literacy**
  - Chafe (1982) and Biber (1995) show that finite subordinate clauses are more common in written than in spoken texts
  - But very difficult to disentangle “real change” from genre effects; cf. present day
- **Cultural complexity & communicative needs** (e.g. Givón 1979; Deutscher 2001)
  - Difficult to find a robust proxy; different notions of “complexity”/“needs”
  - Important to avoid discredited notions of “primitiveness” and lower intelligence
  - Causal chain rarely made explicit (though see Deutscher 2001: 166–186)

## Implications of this claim

The claim has no bearing on questions of grammatical architecture. But it is interesting nonetheless for a variety of reasons.

- If PPHH 6 is correct as far as the corpora are concerned, is it a “**real change**” in the sense of differences in knowledge of language between generations?
  - Could in principle be an artefact of the texts available to us from different periods (poetry, literacy)
  - Could in principle be a real, but non-linguistic, change
- If it’s a “real change”, and if the causal argument works (see next slide), it indicates that sociocultural factors have an impact on language change (cf. ethnosyntax, Enfield 2002)
- If it’s not a “real change”, it has important **implications for the variationist approach to syntactic change** (Kroch 1989, Yang 2002, Pintzuk 2003, etc.): how much change in corpus frequency involves change in the weightings of different grammatical options?

**But let’s assess the hypothesis first before speculating further!**

## Methods: investigating parataxis > hypotaxis

Crucially relies on availability of parsed diachronic corpora.

**Hypotaxis level: proportion of all clauses that are subordinate/embedded, including all non-finite clauses.**

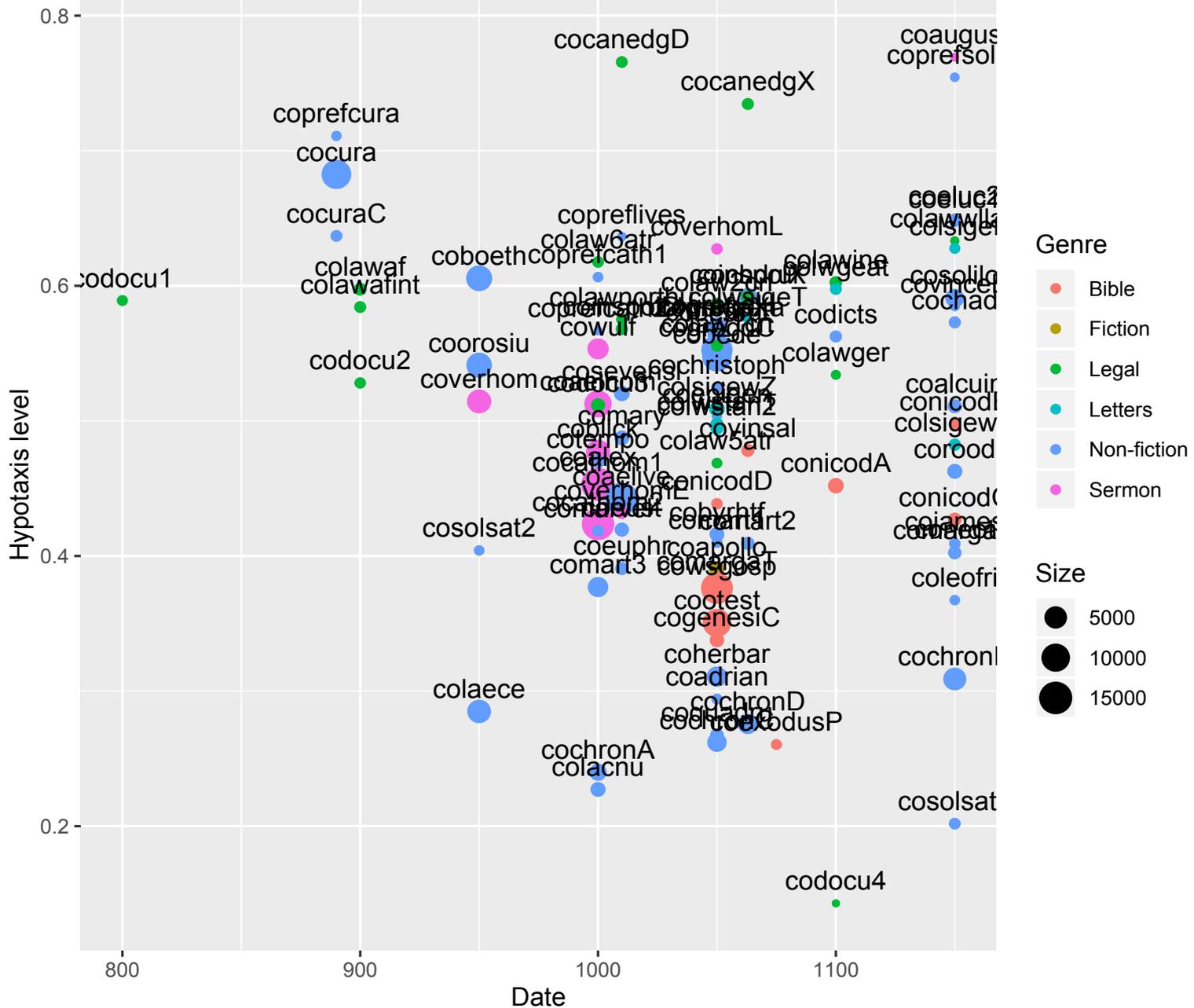
- Finite **unembedded** clauses: IP-MAT\* in Penn-style parsed corpora (includes e.g. imperatives, exclamatives, coordinated clauses)
- Finite **subordinate**/embedded clauses: basically IP-SUB\* (includes e.g. relatives, complement clauses, adverbial clauses)
- **Non-finite** clauses: IP-INF\*

Languages investigated: English, Icelandic, French, Portuguese, Irish, Chinese

**Today I'll be talking mainly about English.**

# Old English

- YCOE (Taylor et al. 2003)
- “Non-fiction” (purple) is something of a dustbin category.
- Legal texts high; bibles and medical texts low







# STAT-1510-E1-P2 – the most hypotactic text

## Acte of Apparell.

Forasmuche as the great and costly arraye and apparell used within this Realme cont=a=rie to good Statut~ thereof made hath ben the occasion of great impov~ysshyng of div~se of the King~ Subgiett~ and p~voked of theym to robbe and to do extorcion and other unlaufall ded~ to mayntayn therby their costely arraye; In eschewyng wherof be it ordeyned by the auctoritie of this p~sent p~liament that ...

And that ...

And that ...

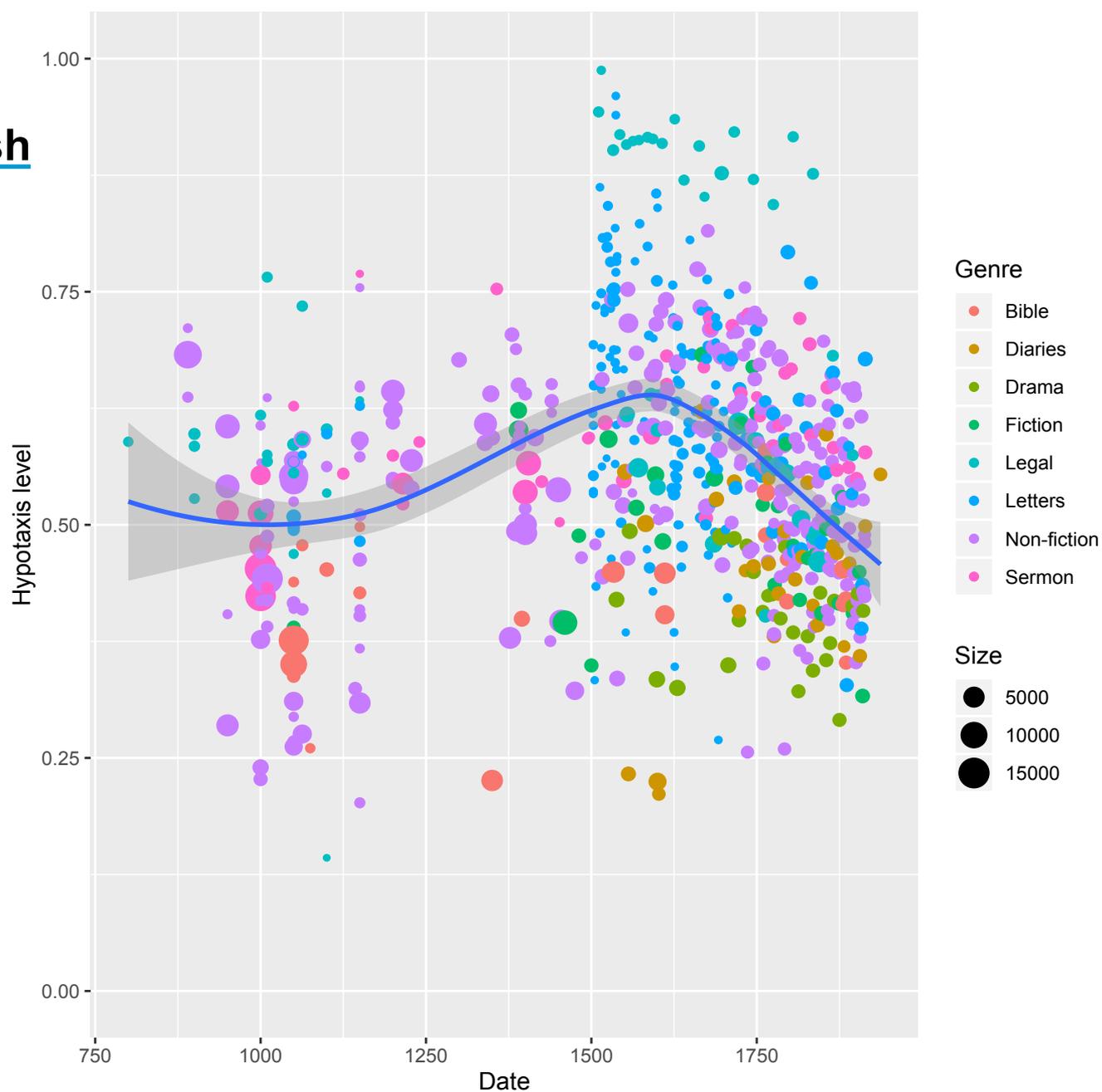
And that ...

And that ...



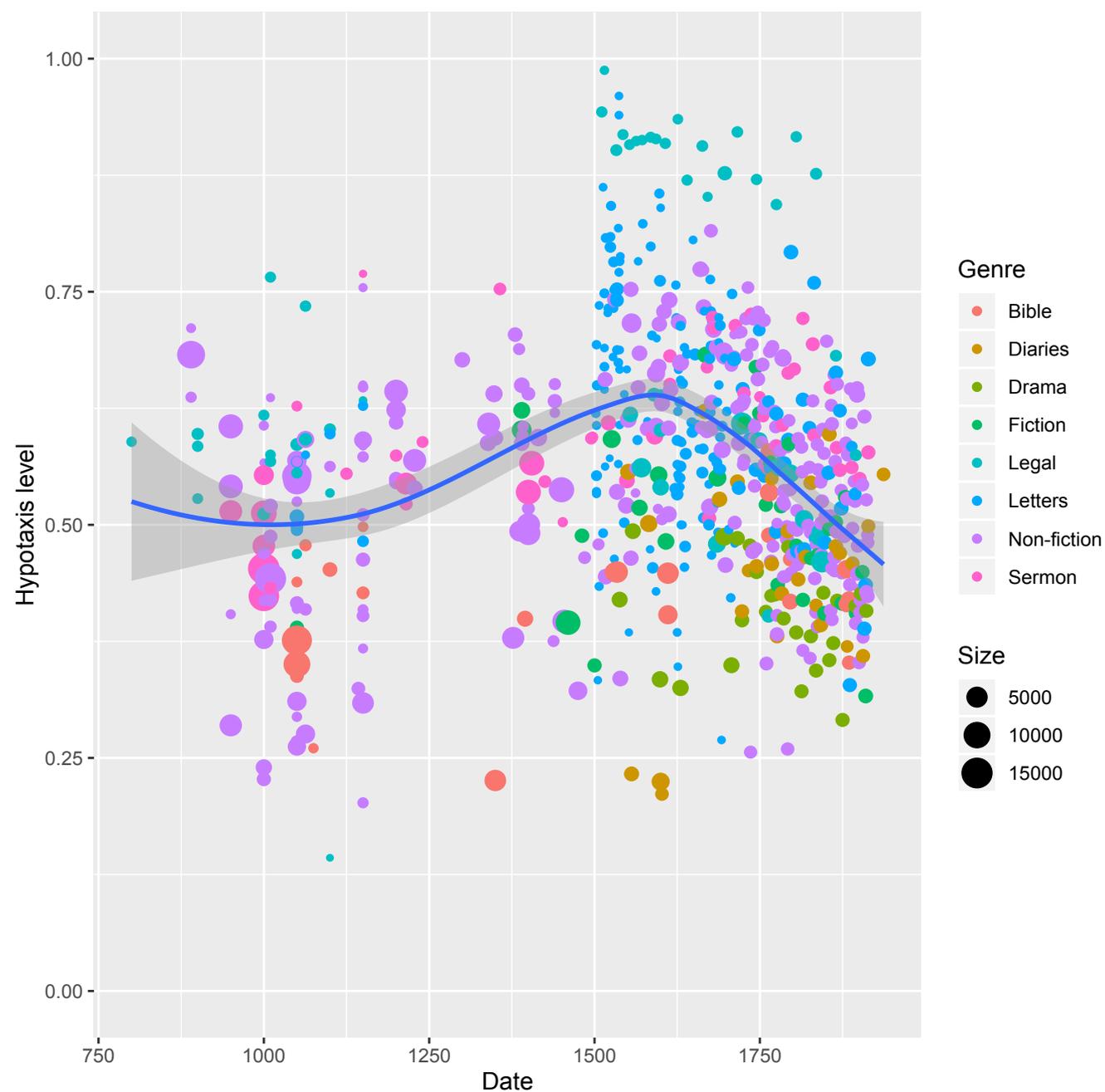
# Zooming out: the history of English

- YCOE  
(Taylor et al. 2003)
- PPCME2  
(Kroch & Taylor 2000)
- PPCEME  
(Kroch et al. 2005)
- PPCMBE2  
(Kroch et al. 2016)
  
- Trend line is locally  
estimated scatterplot  
smoothing (LOESS)  
with confidence  
intervals

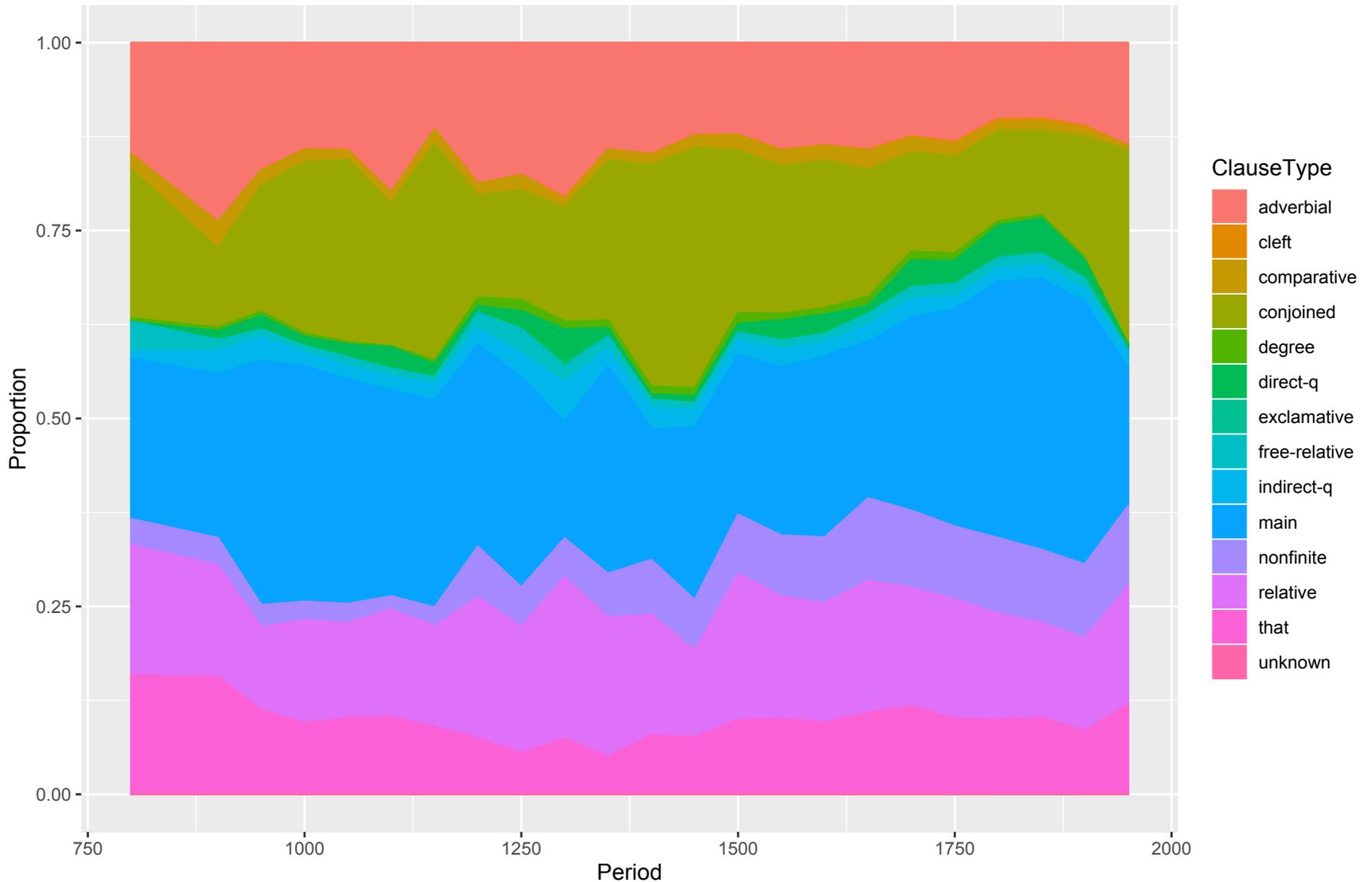


# Generalizations

- On eyeballing, no sign of an overall upward trend
- Pre-1500 not easy to interpret
- High point in 16<sup>th</sup> century (not 17<sup>th</sup>, *pace* Karlsson 2009)
- Apparent decline after 16<sup>th</sup> century
- BUT: genre seems to explain much more of the variation than time does

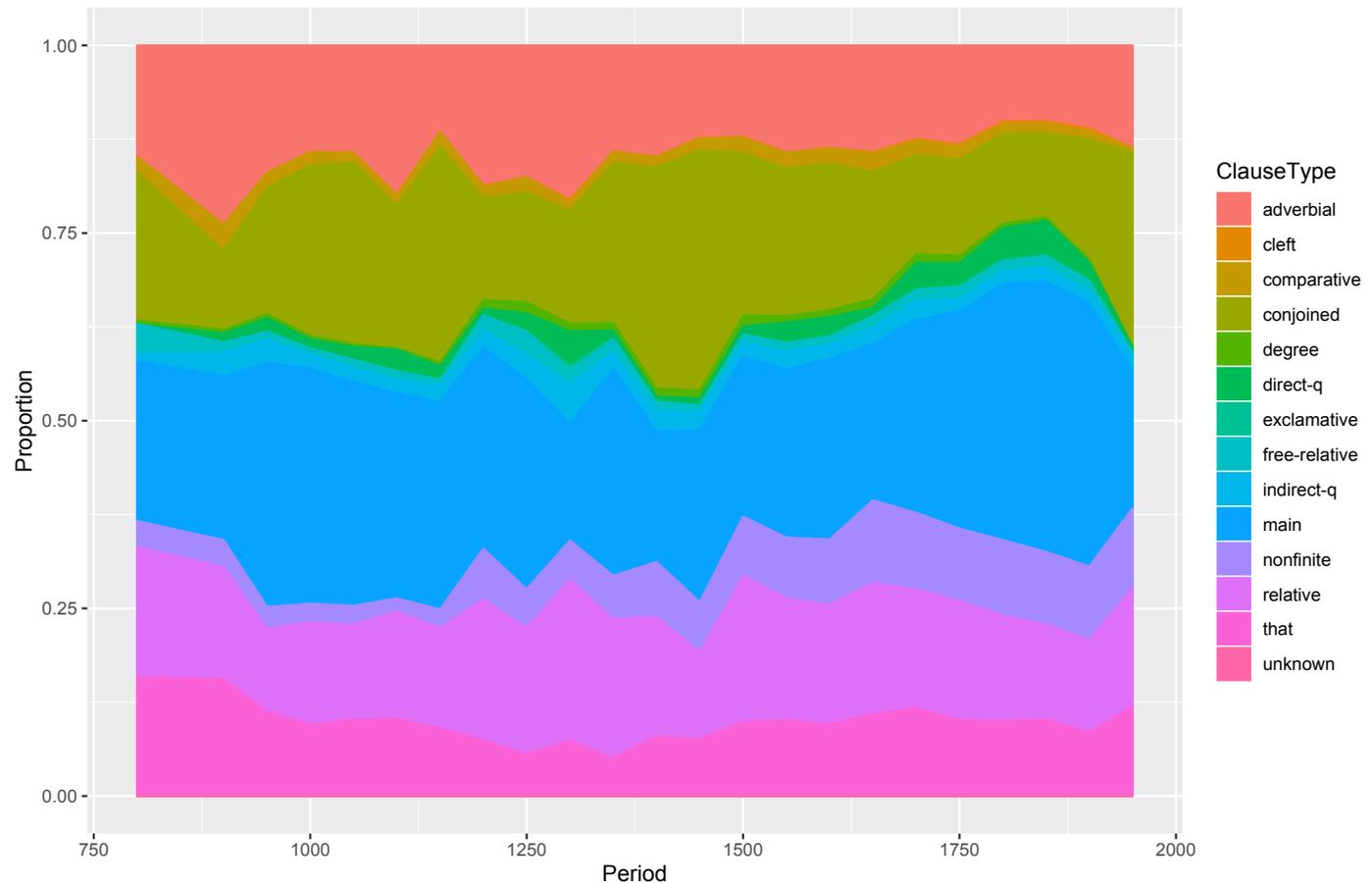


# Distribution across clause types in the history of English



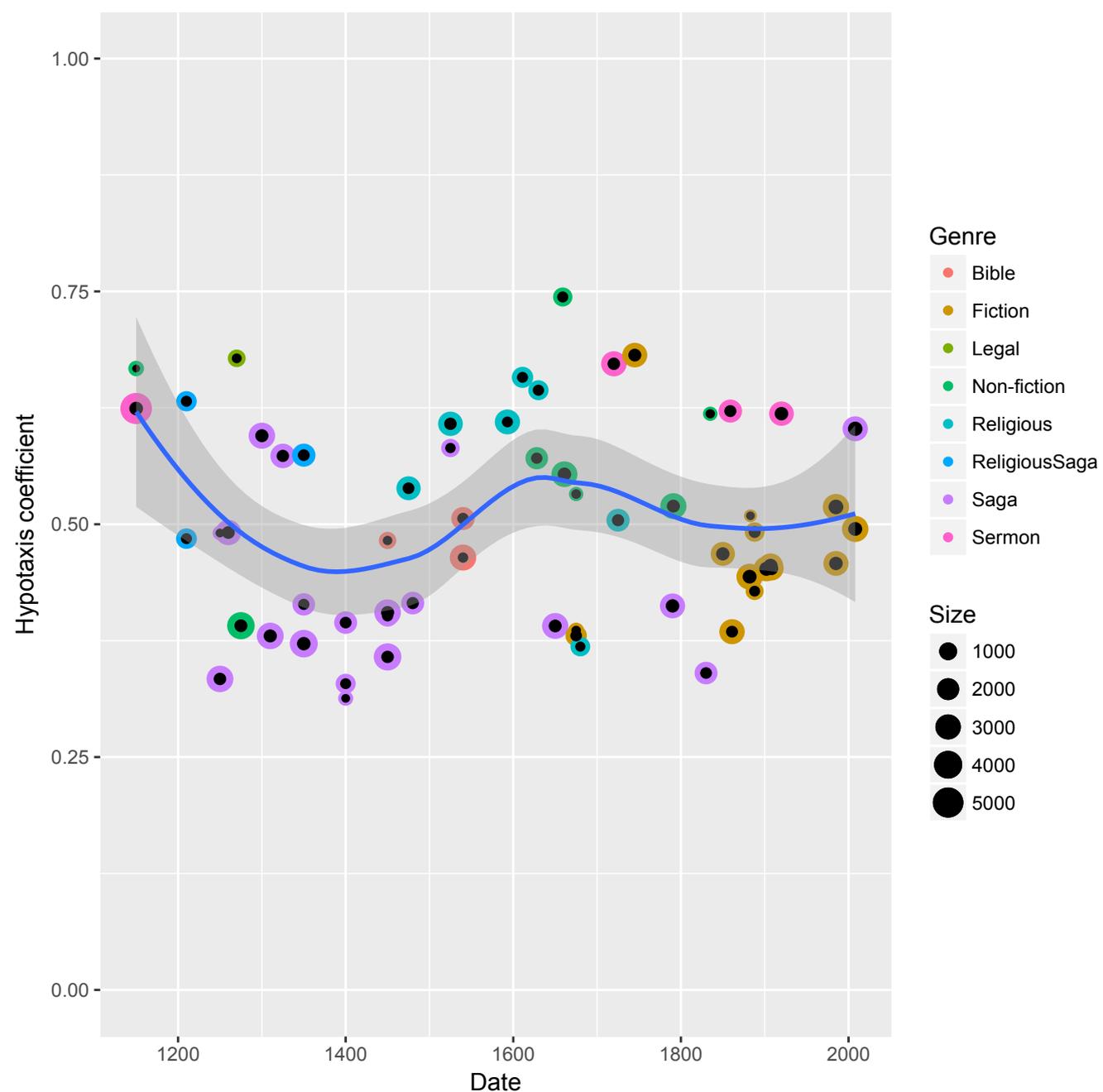
## Clause types: generalizations

- Window: 50 years
- Gentle increase in non-finite clauses between ca. 1150 and 1500
- Change is less striking than continuity here
- Some developments – such as the apparent general decrease in that-clauses, and in adverbial clauses – don't have an obvious explanation
- (The final time point should probably be ignored)



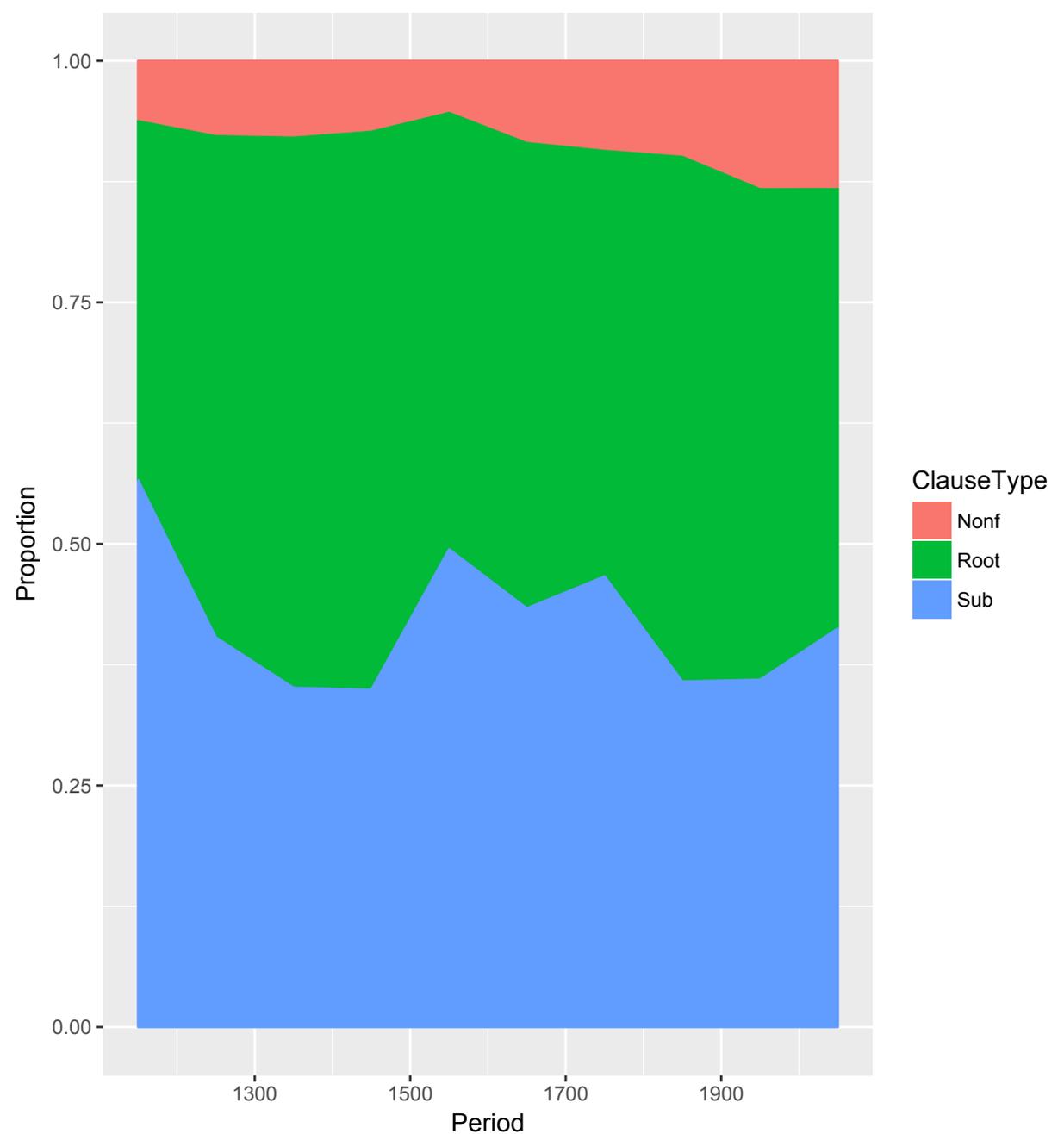
# Icelandic

- IcePaHC (Wallenberg et al. 2011)
- Sagas typically have less than average hypotaxis



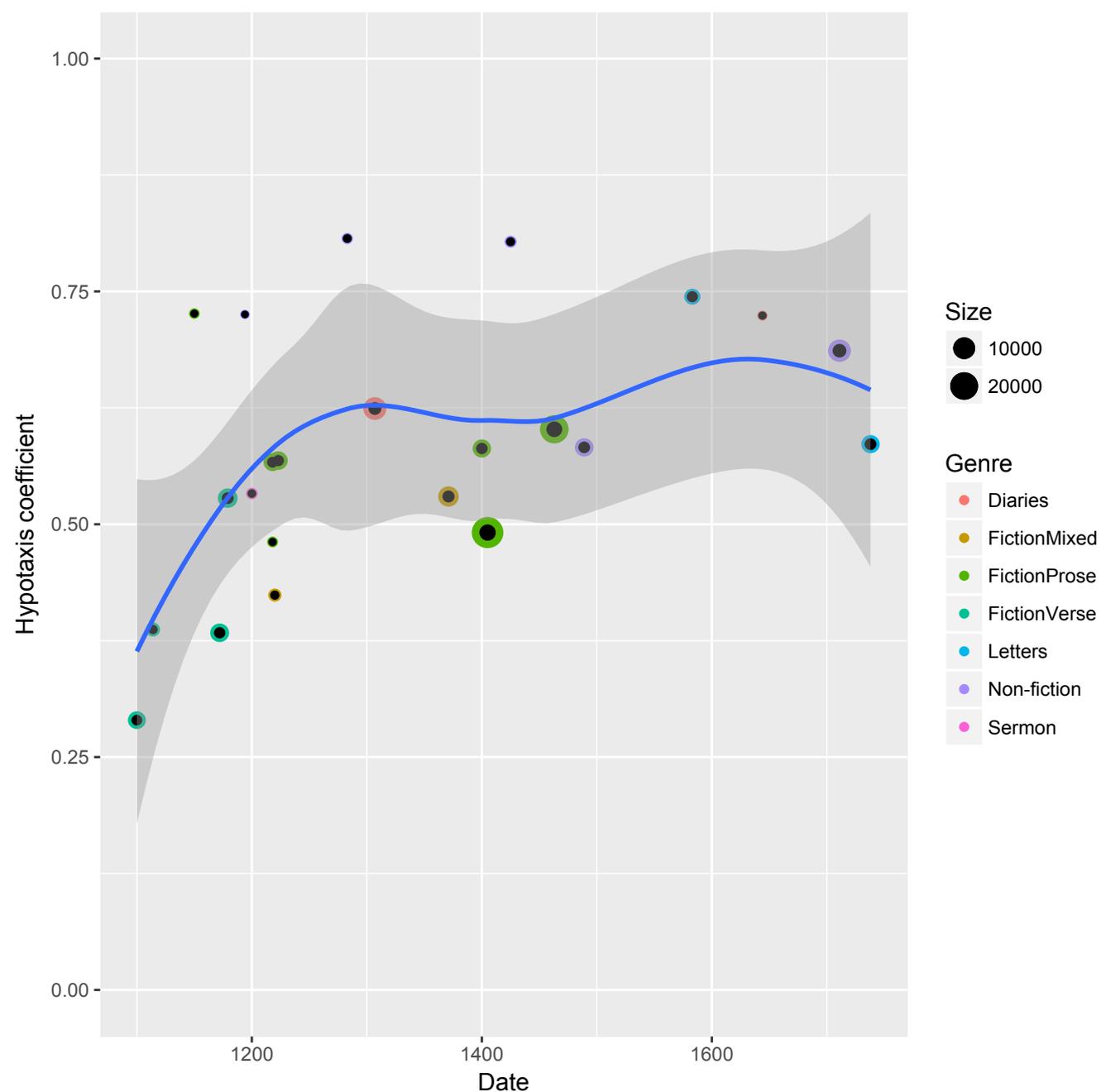
## Icelandic: distribution

- Gentle increase in non-finite clauses between 1500 and 1900.
- Window: 100 years



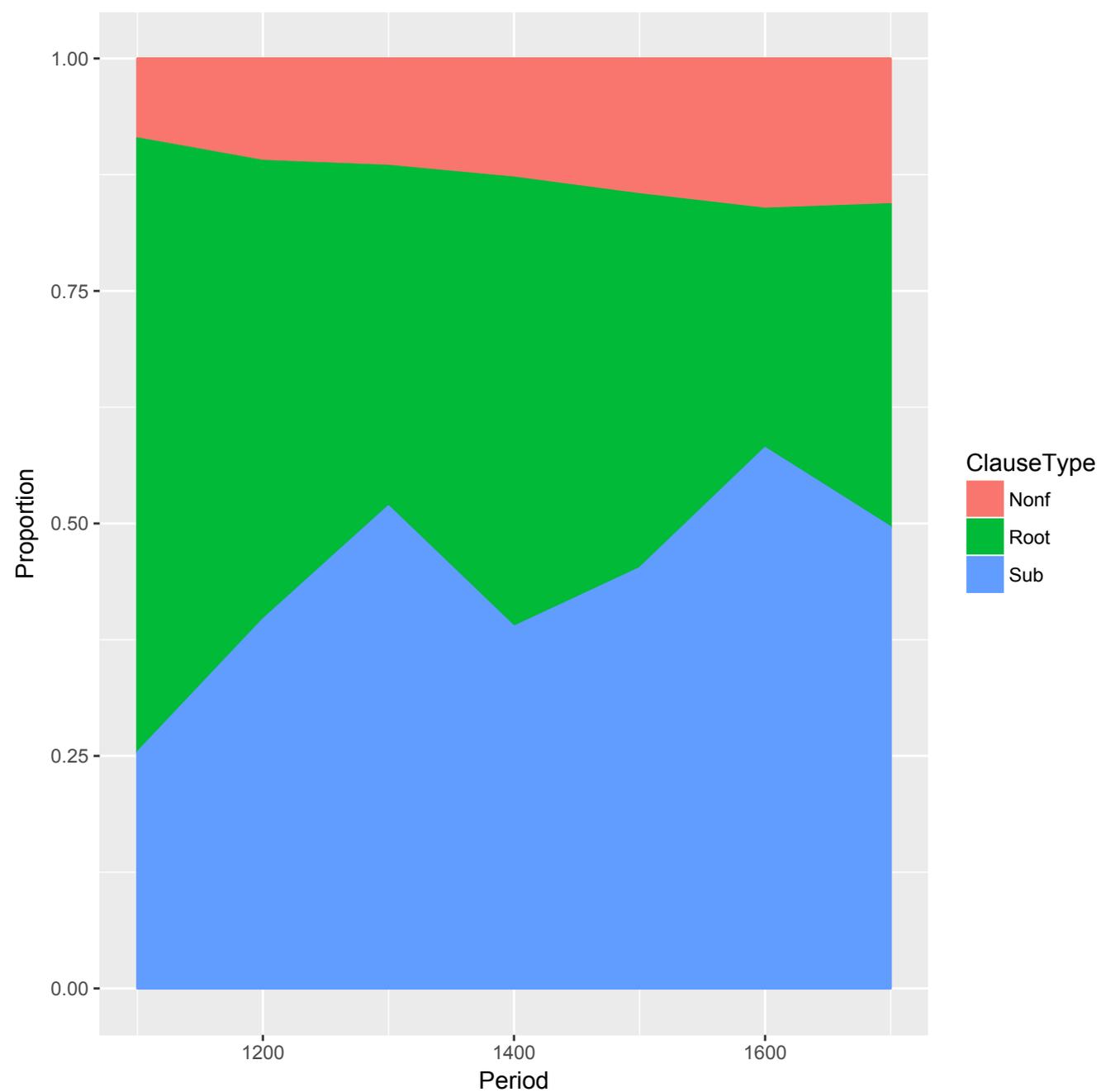
# French

- MCVF (Martineau et al. 2010)
- Apparent early rise is exclusively due to dominance of **verse** texts in this period



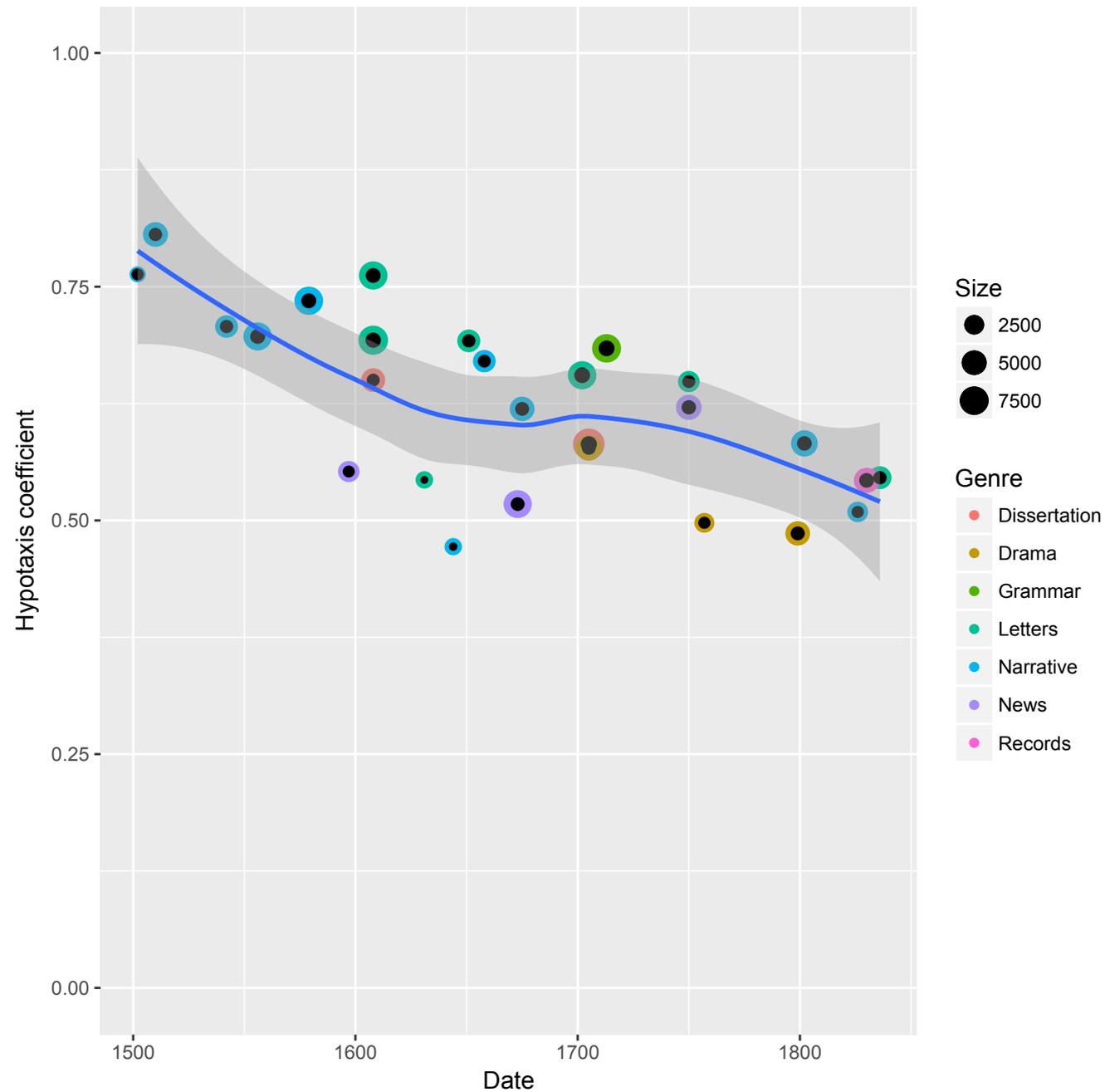
## French: distribution

- Again, gentle rise of non-finite clauses
- Window: 100 years



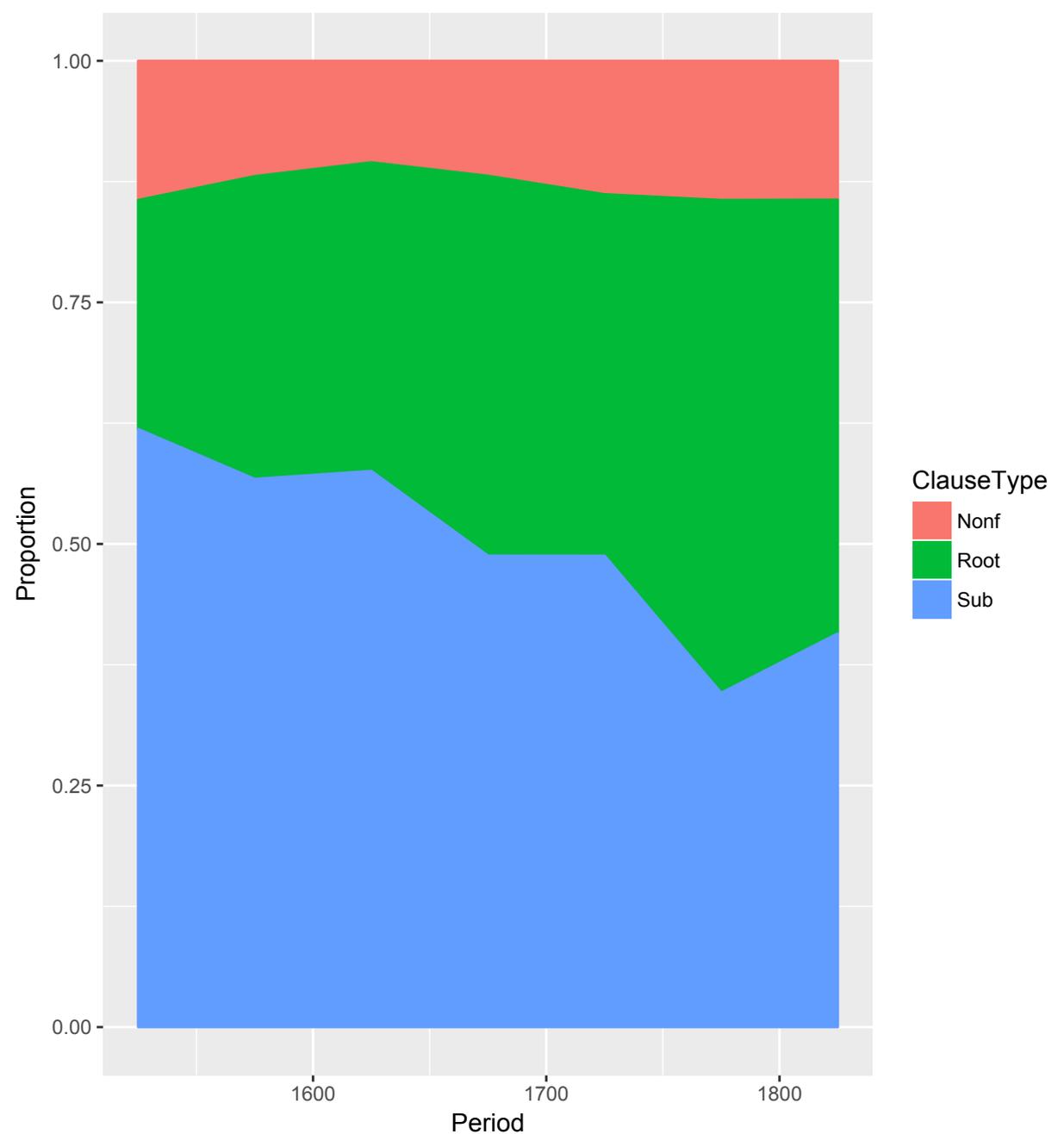
# Portuguese

- Tycho Brahe Corpus (Galves, Andrade & Faria 2017)
- News texts & dramas typically low



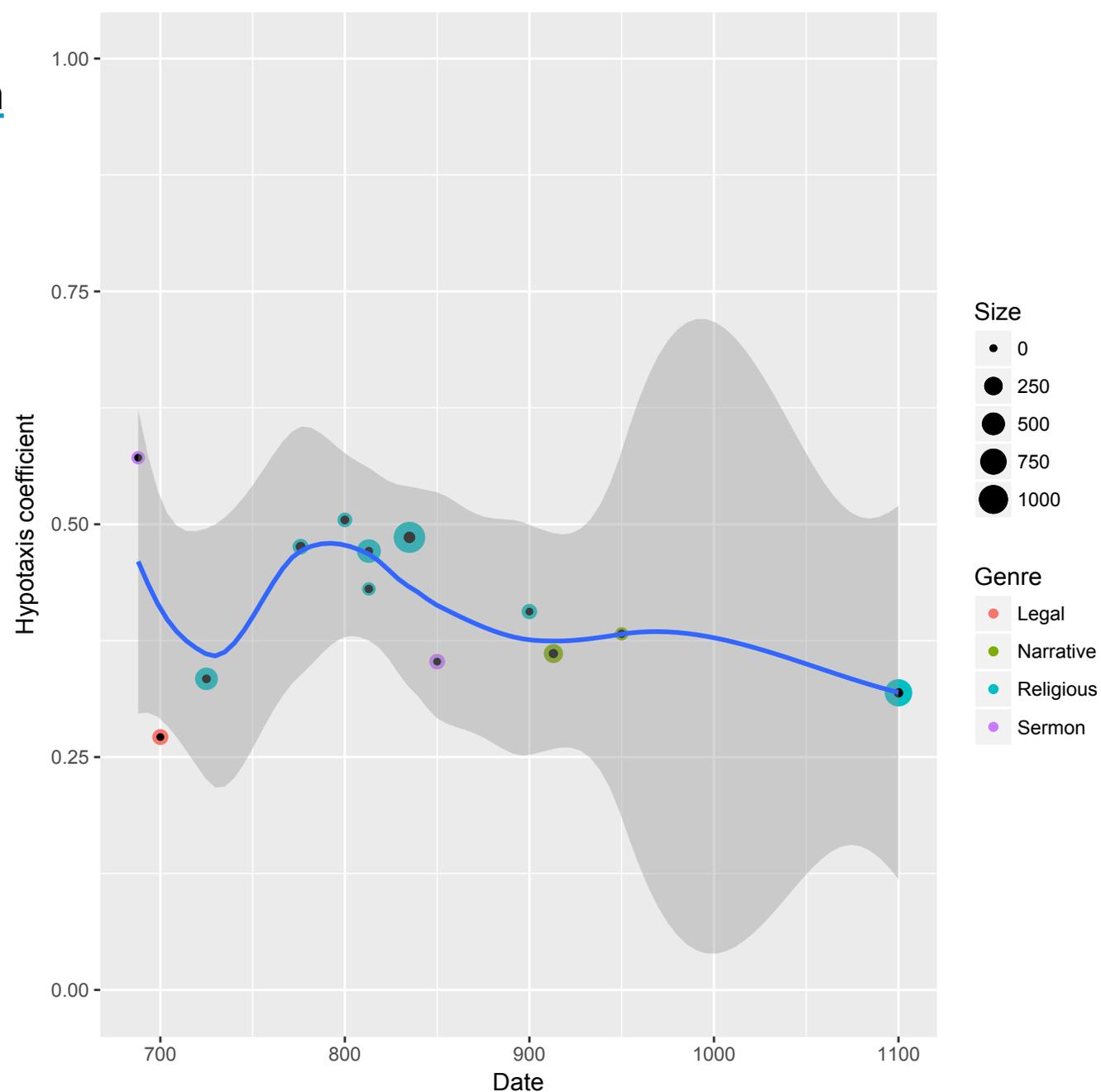
## Portuguese: distribution

- Only clear trend is reduction in finite subordinate clauses
- Window: 50 years



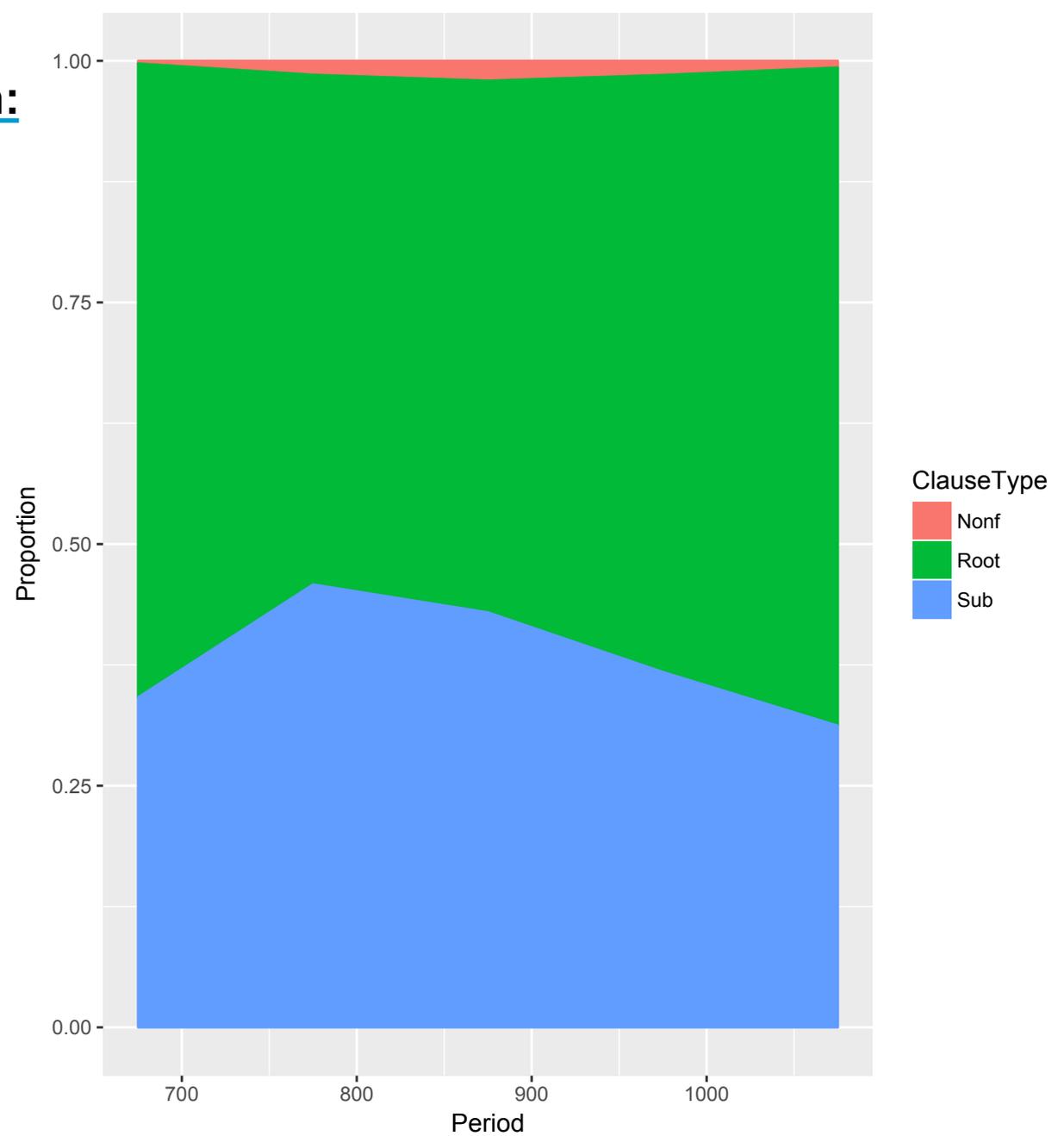
# Old and Middle Irish

- Parsed Corpus of Old and Middle Irish (Lash 2014)
- Hard to generalize about genre



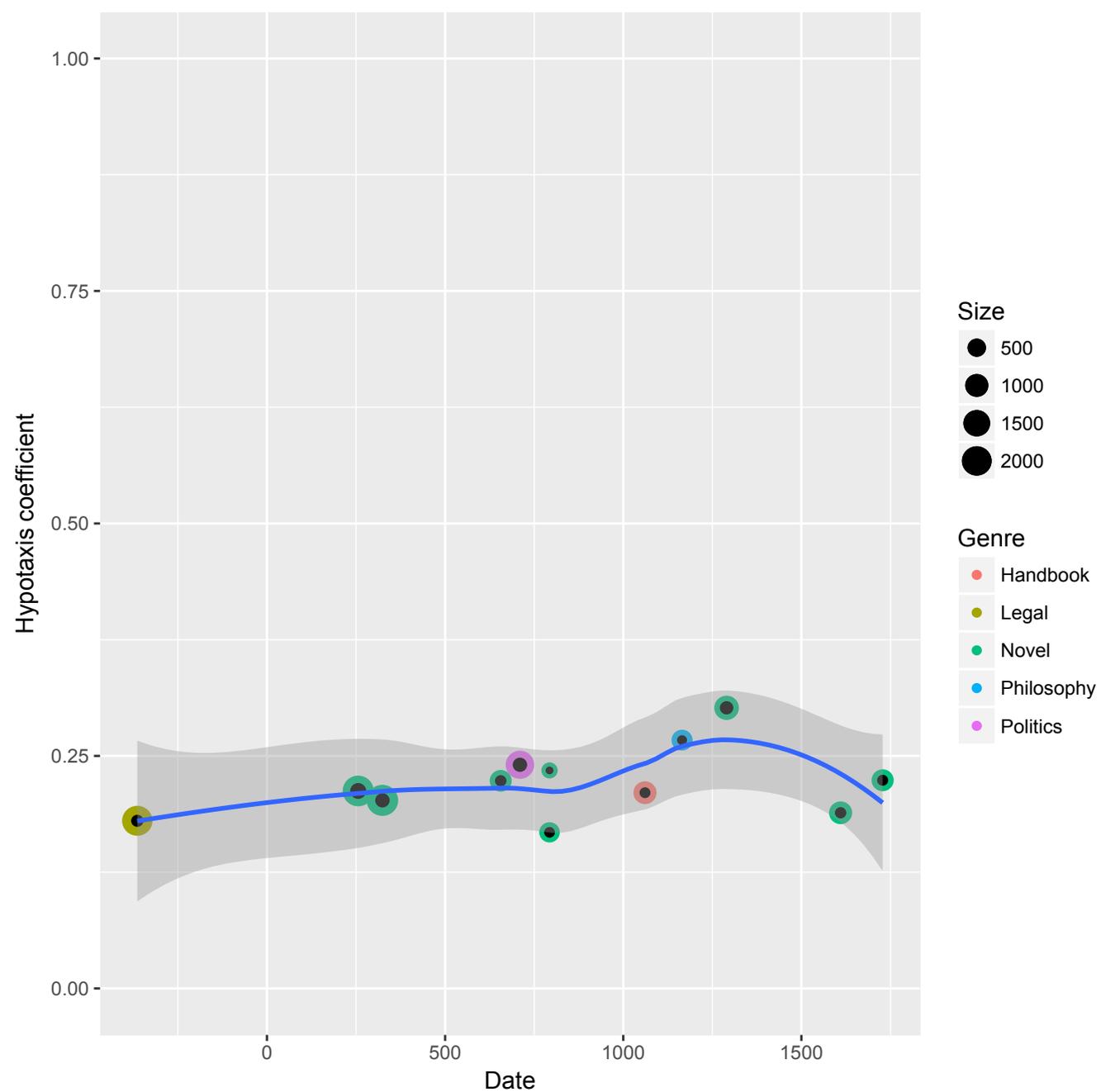
## Old and Middle Irish: distribution

- No clear trends
- Window: 100 years



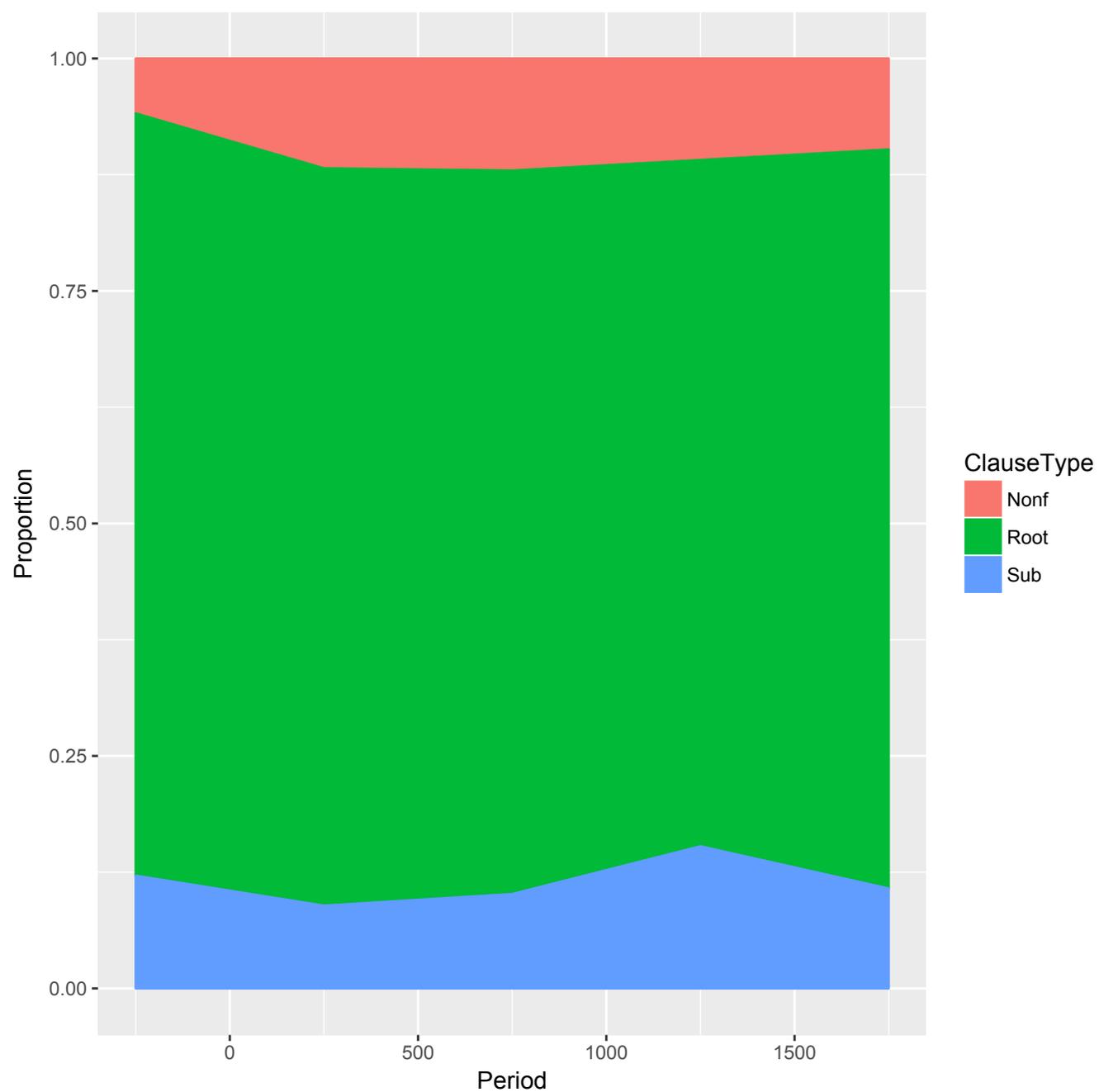
# Chinese

- ChiParHC (Li 2017)
- Again, hard to generalize about genre



# Chinese: distribution

- No clear trends
- Window: 500 years



# Evaluation

## Mixed-effects linear regression using R and lme4 package

- Dependent variable: proportion of unembedded vs. (finite **or** non-finite) subordinate clauses in each text
- Fixed effect: date
- Random intercept: genre

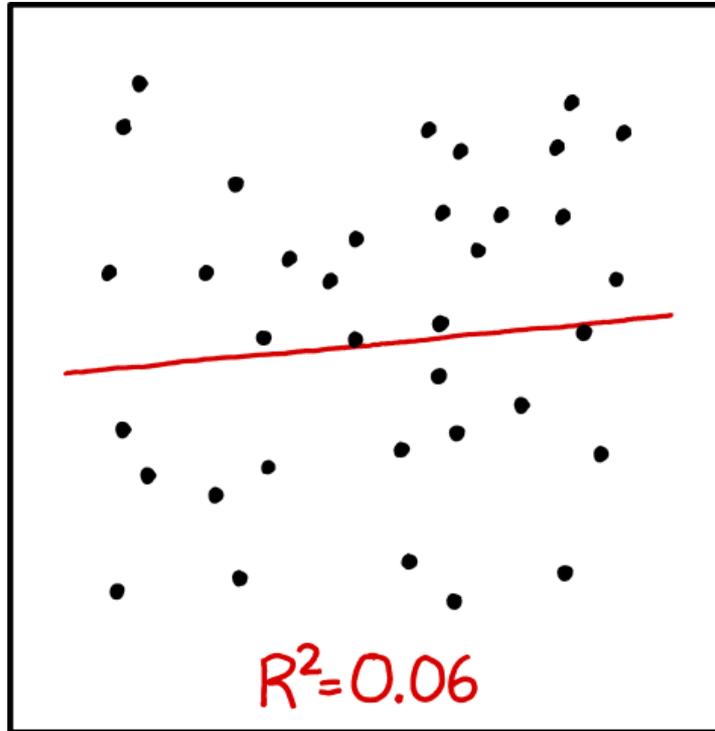
## Positive linear effect of time should at least be detectable if Dąbrowska's (2015) hypothesis is correct!

Nagelkerke  $R^2$ , a measure of goodness of fit, calculated using Nakagawa, Johnson & Schielzeth (2017) method and MuMIn R package. Gives percentage of variance explained by the model.

- Marginal  $R^2$ : only fixed effects (date)
- Conditional  $R^2$ : fixed and random effects (date and genre)

	English	Icelandic	French	Portuguese	Irish	Chinese
Effect of date	-0.00006	0.00001	0.00008	-0.00064	-0.00030	0.00002
Marginal $R^2$	1.2%	0.1%	1.3%	44.4%	11.3%	12.5%
Conditional $R^2$	43.2%	39.9%	54.6%	49.0%	44.6%	12.5%

# Constellations



I DON'T TRUST LINEAR REGRESSIONS WHEN IT'S HARDER TO GUESS THE DIRECTION OF THE CORRELATION FROM THE SCATTER PLOT THAN TO FIND NEW CONSTELLATIONS ON IT.

[xkcd #1725, "Linear Regression"](#) (Randall Munroe, CC-BY-NC 2.5)

## Evaluation (with genre)

	English	Icelandic	French	Portuguese	Irish	Chinese
Effect of date	-0.00006	0.00001	0.00008	-0.00064	-0.00030	0.00002
Marginal $R^2$	1.2%	0.1%	1.3%	44.4%	11.3%	12.5%
Conditional $R^2$	43.2%	39.9%	54.6%	49.0%	44.6%	12.5%
$p$	<0.001	0.829	0.567	<0.001	0.182	0.238

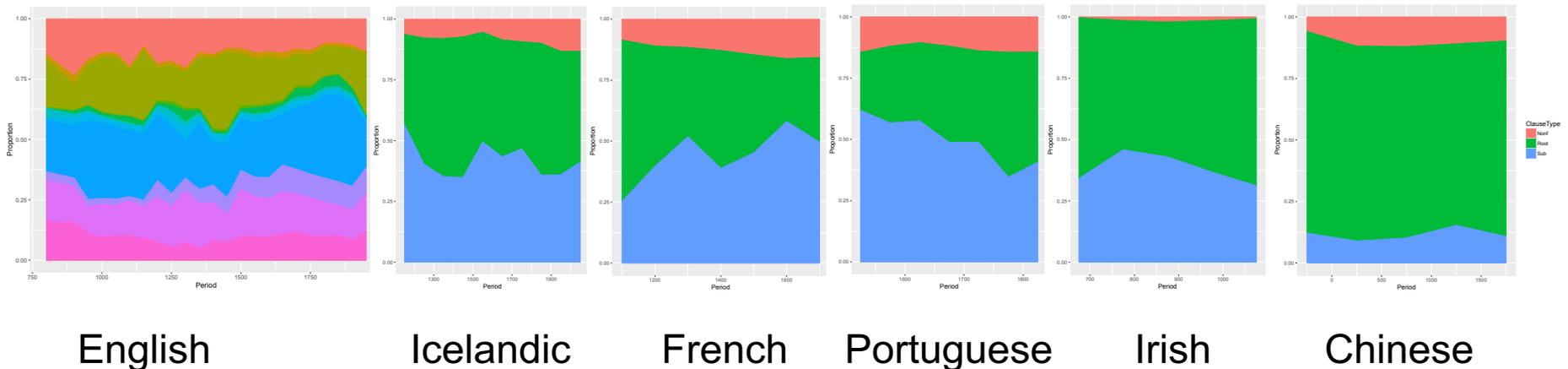
- Marginal  $R^2$ : only fixed effects (date)
- Conditional  $R^2$ : fixed and random effects (date and genre)
- $p$ -value of date effect calculated using package lmerTest

### Effect of date explains little of the data, with the exception of Portuguese.

- Portuguese: effect is in the wrong direction.
- English, Icelandic, French: effect explains almost nothing.
- Irish, Chinese: probably not enough data to be hugely confident.

### Genre explains much, much more of the data, except in Chinese.

# Overview



## No robust support for the quantitative version of parataxis > hypotaxis.

- English, Icelandic, Irish, Chinese: no consistent direction of change.
- French: apparent increase in hypotaxis 1100–1200, but early texts are in verse.
- Portuguese: gentle but steady *decrease* in hypotaxis over the timespan of the corpus.
- Gentle upward trend for non-finite clauses in English, Icelandic and French.

## Does genre play a role? Yes, but irrelevant to the hypothesis as far as we can tell.

- The most hypotactic texts in English are legal texts.
- A consistent role for genre is exactly what we'd predict given Chafe's (1982) and Biber's (1995) results, if performance effects are constant.
- So unless the corpora are unbalanced and genre effects are *counteracting* a real diachronic trend, the result basically stands.

## Conclusion

- It's widely agreed that parataxis > hypotaxis.  
Much less widely agreed what this actually means.
- The idea that finite clausal complementation in English emerged from reanalysis of juxtaposed independent clauses is problematic for several reasons (Axel-Tober 2017). More likely, it emerged from (cor)relative constructions.
- Focusing on the idea that (finite) clausal subordination becomes more prevalent over time, I have found no support for this in parsed diachronic corpora of English, Icelandic, French, Portuguese, Irish, or Chinese.
  - Maybe the corpus annotation is wrong.
  - But insofar as parataxis > hypotaxis is an empirical question, the burden of proof should hopefully be shifting at least somewhat.
- **Much future work suggests itself:**
  - Suggestions welcomed!

**Thank you  
for your attention!**

## References (1)

- **Andrew**, S. O. 1940. *Syntax and style in Old English*. Cambridge: Cambridge University Press.
- **Axel**, Katrin. 2007. *Studies on Old High German syntax: Left sentence periphery, verb movement and verb-second*. Amsterdam: John Benjamins.
- **Axel**, Katrin. 2009. Die Entstehung des dass-Satzes—ein neues Szenario. *Koordination und Subordination im Deutschen* (Special issue of *Linguistische Berichte* 16), ed. by Veronika Ehrich, Christian Fortmann, Ingo Reich, & Marga Reis, 21–41. Hamburg: Buske.
- **Axel-Tober**, Katrin. 2012. *(Nicht-)kanonische Nebensätze im Deutschen: Synchrone und diachrone Aspekte*. Berlin: Mouton de Gruyter.
- **Axel-Tober**, Katrin. 2017. The development of the declarative complementizer in German. *Language: Historical Syntax* 93, e29–e65.
- **Behaghel**, Otto. 1877. *Über die Entstehung der abhängigen Rede und die Ausbildung der Zeitfolge im Altdeutschen*. Paderborn: Schöningh.
- **Behaghel**, Otto. 1928. *Deutsche Syntax: Eine geschichtliche Darstellung. Vol. 3: Die Satzgebilde*. Heidelberg: Winter.
- **Biber**, Douglas. 1995. *Dimensions of register variation: a cross-linguistic comparison*. Cambridge: Cambridge University Press.

## References (2)

- **Chafe**, Wallace. 1982. Integration and involvement in speaking, writing, and oral literature. In Deborah Tannen (ed.), *Spoken and written language: Exploring orality and literacy*, 35–53. Norwood, NJ: Ablex.
- **Dąbrowska**, Ewa. 2015. Language in the mind and in the community. In Jocelyne Daems, Eline Zenner, Kris Heylen, Dirk Speelman & Hubert Cuyckens (eds.), *Change of paradigms – new paradoxes: recontextualizing language and linguistics*, 221–236. Berlin: de Gruyter.
- **Delbrück**, Berthold. 1900. *Vergleichende Syntax der indogermanischen Sprachen*. Vol. 3. Strasbourg: Karl J. Trübner.
- **De Smet**, Hendrik. 2009. Analysing reanalysis. *Lingua* 119, 1728–1755.
- **Deutscher**, Guy. 2001. *Syntactic change in Akkadian: the evolution of sentential complementation*. Oxford: Oxford University Press.
- **Enfield**, Nicholas J. (ed.) 2002. *Ethnosyntax: explorations in grammar and culture*. Oxford: Oxford University Press.
- **Eythórsson**, Thórhallur. 1995. *Verbal syntax in the early Germanic languages*. PhD dissertation, Cornell University.
- **Galves**, Charlotte, Aroldo L. de **Andrade**, & Pablo **Faria**. 2017. *Tycho Brahe Parsed Corpus of Historical Portuguese*.

## References (3)

- **Garrett**, Andrew. 2012. The historical syntax problem: reanalysis and directionality. In Dianne Jonas, John Whitman & Andrew Garrett (eds.), *Grammatical change: origins, nature, outcomes*, 52–72. Oxford: Oxford University Press.
- **Gildersleeve**, Basil L. 1883. On the final sentence in Greek. *The American Journal of Philology* 4, 416–444.
- **Gildersleeve**, Basil L. 1893. Some problems in Greek syntax. *Transactions of the American Philological Association* 24, xxiv–xxvii.
- **Gippert**, Jost, & Manana **Tandashvili**. 2015. Structuring a diachronic corpus: the Georgian National Corpus project. In Jost, Gippert & Ralf Gehrke (eds.), *Historical corpora: challenges and perspectives*, 305–322. Tübingen: Narr.
- **Givón**, Talmy. 1979. From discourse to syntax: Grammar as a processing strategy. In Talmy Givón (ed.), *Syntax and semantics*, vol. 12: *Discourse and syntax*, 81–112. New York, NY: Academic Press.
- **Harris**, Alice C., & Lyle **Campbell**. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

## References (4)

- **Haug**, Dag T. T., & Marius L. **Jøhndal**. 2008. Creating a parallel treebank of the old Indo-European bible translations. In Caroline Sporleder & Kiril Ribarov (eds.), *Proceedings of LaTeCH 2008*, 27–34.
- **Heine**, Bernd, & Tania **Kuteva**. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- **Jucker**, Andreas H. 1991. Between hypotaxis and parataxis: clauses of reason in *Ancrene Wisse*. In Dieter Kastovsky (ed.), *Historical English syntax*, 203–219. Berlin: de Gruyter.
- **Karlsson**, Fred. 2009. Origin and maintenance of clausal embedding complexity. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as an evolving variable*, 192–202. Oxford: Oxford University Press.
- **King**, John E., & Christopher **Cookson**. 1890. *An introduction to the comparative grammar of Greek and Latin*. Oxford: Clarendon.
- **Kiparsky**, Paul. 1995. Indo-European origins of Germanic syntax. In Adrian Battye & Ian Roberts (eds.), *Clause structure and language change*, 140–169. Oxford: Oxford University Press.
- **Kroch**, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation & Change* 1, 199–244.
- **Kroch**, Anthony, & Ann **Taylor**. 2000. *The Penn-Helsinki Parsed Corpus of Middle English (PPCME2)*.
- **Kroch**, Anthony, Beatrice **Santorini**, & Lauren **Delfs**. 2004. *The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME)*.

## References (5)

- **Kroch**, Anthony, Beatrice **Santorini**, & Ariel **Diertani**. 2010. The Penn Parsed Corpus of Modern British English (PPCMBE).
- **Lash**, Elliott. 2014. *The Parsed Old and Middle Irish Corpus (POMIC)*. Version 0.1.
- **Lehmann**, Christian. 1995. Synsemantika. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax: Ein internationales Handbuch*, 1251–1266. Berlin: Mouton de Gruyter.
- **Li**, Man. 2017. *Chinese Parsed Historical Corpus (ChiParHC)*.
- **Martineau**, France, Paul **Hirschbühler**, Anthony **Kroch**, & Yves Charles **Morin**. 2010. *Modéliser le changement: Les voies du français*.
- **Meyer**, Roland. 2017. The C system of relatives and complement clauses in the history of Slavic languages. *Language: Historical Syntax* 93, e97–e113.
- **Mitchell**, Bruce. 1985. *Old English syntax*. 2 vols. Oxford: Clarendon.
- **Nakagawa**, Shinichi, Paul C. D. **Johnson**, & Holger **Schielzeth**. 2017. The coefficient of determination R<sup>2</sup> and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J R Soc Interface* 14, 20170213.
- **O’Neil**, Wayne. 1977. Clause adjunction in Old English. *General Linguistics* 17, 199–211.
- **Paul**, Hermann. *Deutsche Grammatik*. Vol. 5, part 4: *Wortbildungslehre*. Halle (Saale): Niemeyer.

## References (6)

- **Pintzuk**, Susan. 2003. Variationist approaches to syntactic change. In Brian Joseph & Richard Janda (eds.), *The handbook of historical linguistics*, 509–528. Oxford: Blackwell.
- **Roberts**, Ian, & Anna **Roussou**. 2003. *Syntactic change: a Minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.
- **Roberts**, Ian. 2007. *Diachronic syntax*. Oxford: Oxford University Press.
- **Small**, George W. 1924. *The comparison of inequality*. Baltimore: University Press.
- **Taylor**, Ann, Anthony **Warner**; Susan **Pintzuk**, & Frank **Beths**. 2003. The York-Toronto-Helsinki Parsed Corpus of Old English Prose.
- **Thiersch**, Friedrich. 1826. *Griechische Grammatik, vorzüglich des homerischen Dialekts*. 3<sup>rd</sup> edition. Leipzig: Fleischer.
- **van Gelderen**, Elly. 2004. Economy, innovation, and prescriptivism: From spec to head and head to head. *Journal of Comparative Germanic Linguistics* 7: 59–98.
- **van Gelderen**, Elly. 2011. *The linguistic cycle: language change and the language faculty*. Oxford: Oxford University Press.
- **van Kemenade**, Ans. 1997. V2 and embedded topicalization in Old and Middle English. In Ans van Kemenade & Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 326–352. Cambridge: Cambridge University Press.

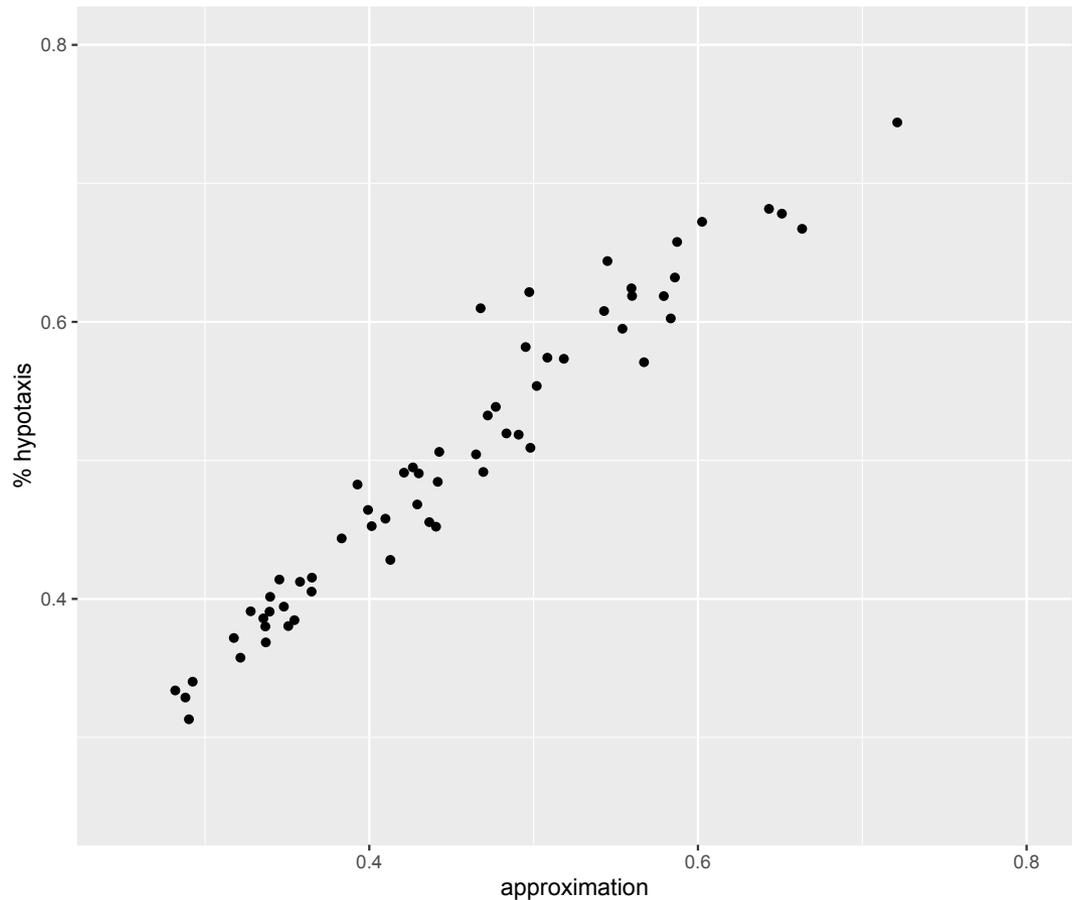
## References (7)

- **Walkden**, George. 2014. *Syntactic reconstruction and Proto-Germanic*. Oxford: Oxford University Press.
- **Walkden**, George. 2019. The many faces of uniformitarianism in linguistics. *Glossa: a journal of general linguistics* 4(1): 52, 1–17.
- **Walkden**, George, & Hannah **Booth**. 2020. Reassessing the historical evidence for embedded V2. In Rebecca Woods & Sam Wolfe (eds.), *Rethinking verb second*, 536–554, Oxford: Oxford University Press.
- **Wallenberg**, Joel C. 2016. Extraposition is disappearing. *Language: Historical Syntax* 92, e237–e256.
- **Wallenberg**, Joel C., Anton K. **Ingason**, Einar F. **Sigurðsson**, & Eiríkur **Rögnvaldsson**. 2011. Icelandic Parsed Historical Corpus (IcePaHC). Version 0.9.
- **Whitman**, John. 2001. Relabelling. In Susan Pintzuk, George Tsoulas & Anthony Warner (eds.), *Diachronic syntax: models and mechanisms*, 220–238. Oxford: OUP.
- **Whitman**, John. 2012. Misparsing and syntactic reanalysis. In Ans van Kemenade & Nynke de Haas (eds.), *Historical Linguistics 2009: Selected Papers from the 19th International Conference on Historical Linguistics*, 69–87. Amsterdam: John Benjamins.
- **Yang**, Charles D. 2002. *Knowledge and learning in natural language*. Oxford: OUP.

## Bonus languages: Latin, Slavic/Russian, Georgian

These corpora don't have constituency parsing.

- Latin: PROIEL
- Slavic/Russian: PROIEL
- Georgian: Georgian National Corpus

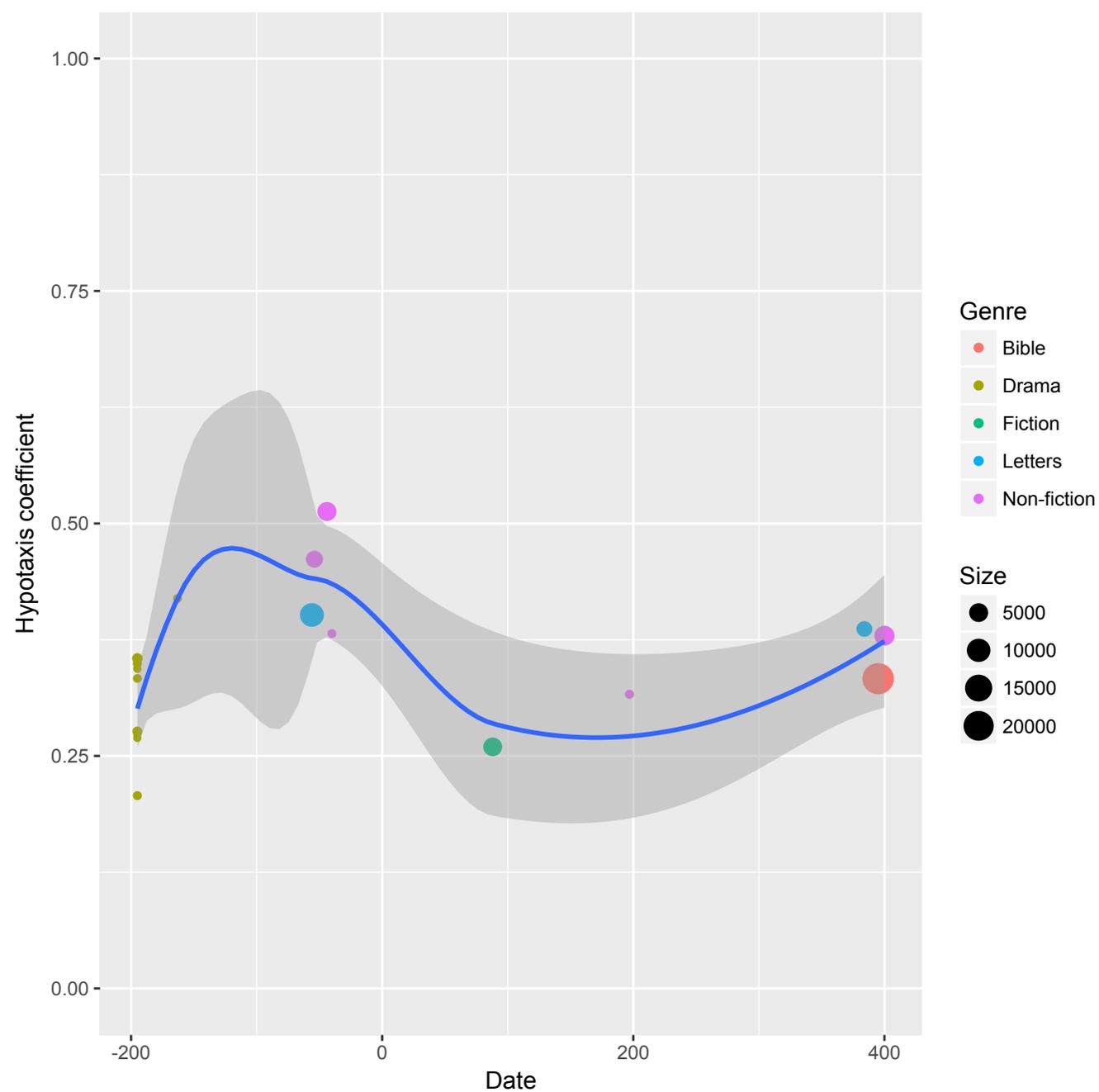


Approximation to the hypotaxis coefficient: number of overt subordinators divided by the number of finite verbs.

This seems to work reasonably well. Correlation for Icelandic shown (incl. nonfinite).

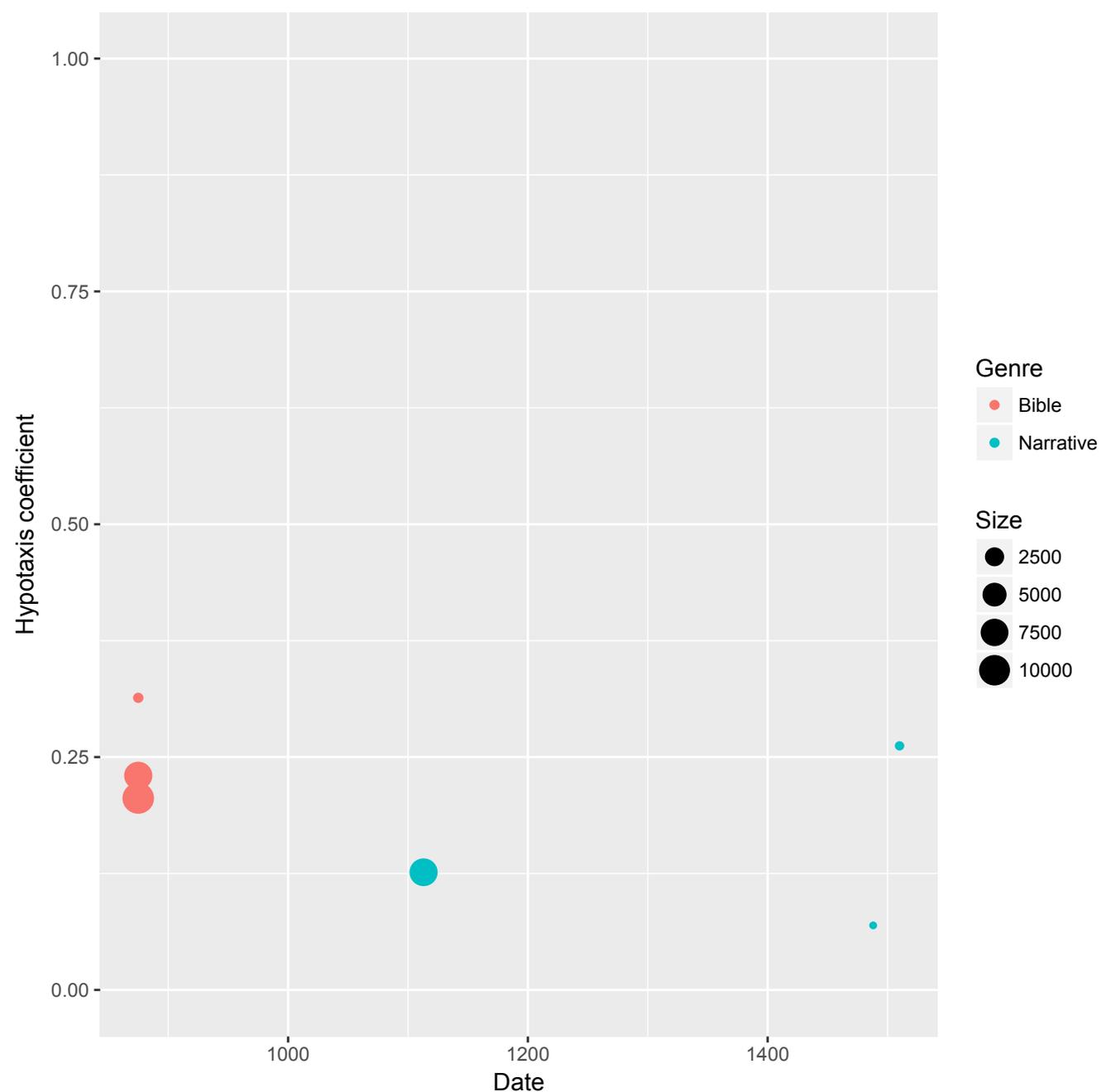
# Latin

- PROIEL (Haug & Jøhndal 2008)
- Again, hard to generalize about genre



# Slavic/Russian

- PROIEL (Haug & Jøhndal 2008)
- Bible texts are Old Church Slavonic; narrative texts are Russian
- Too little here to say anything meaningful at all



# Georgian

- Georgian National Corpus (Gippert & Tandashvili 2015)
- Philosophical and legal texts most hypotactic

