A fresh look at infinitive variation after *help* in 18th–21st century American and British English

Richard Zimmermann

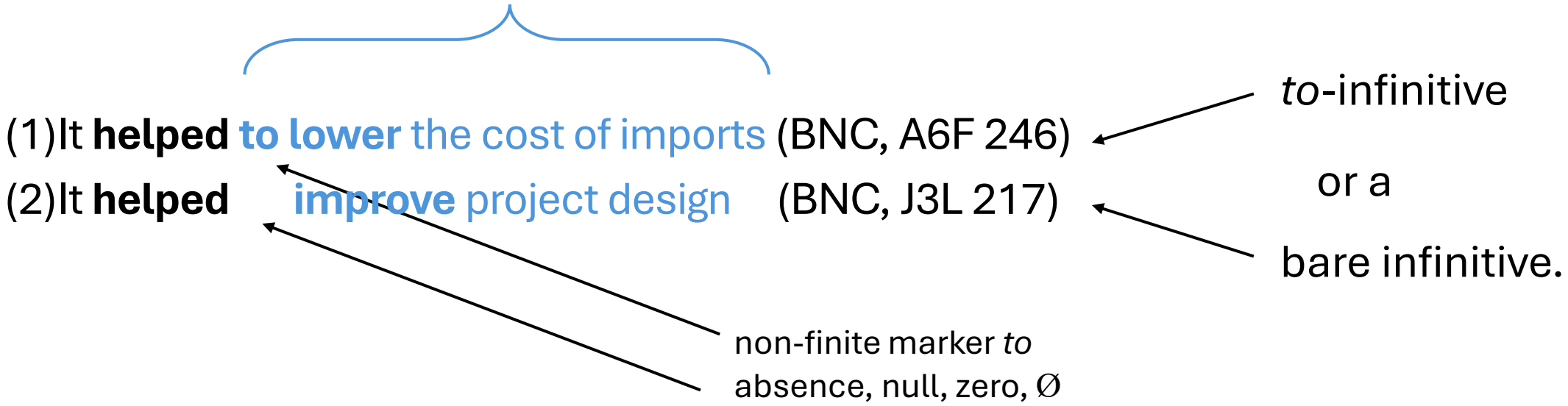
THURSDAY, 28 May 2026, 4.20pm

Location: E114 | Chair: Hanna Schmück

Introduction

Variation in the form of the head of non-finite clauses complementing *help*

- Non-finite complement clauses after *help* can be headed by a



- Biber et al. (1999: 735-7), Kjellmer (1985), Leech et al. (2010: 187-193), Levshina (2018, 2022), Lohmann (2011), Lind (1983), Mair (2002, 2006: 136-140), McEnery & Xiao (2005)

A linguistic replacement change

- The variation is indicative of, i.e., has led to or is the result of, a linguistic change.
- 19th century:

(3) a. a little piece of alum does no harm, and may **help** [to fix the color].
(Child, Maria (1829) *The American Frugal Housewife,* AmE)

historical loss, reduction of *to*, erosion, weakening
NOT: *to*-insertion

- 21st century:

b. a useful tool that, if successful, would **help** [guide millions of young people through their first big life decision].
(O'Neil, Cathy (2016) *Weapons of Math Destruction*, AmE)

Appealing features of the variation

- Simple, minimal, binary: Absence / presence of non-finite marker *to*.
- Help is a comparatively common verb. Allows relatively large datasets.
- Leads into theoretical questions:
  - Grammaticalisation:
    - Phonological attrition, erosion, *to* > Ø, in line with diachronic generalisations
    - "help is undergoing semantic bleaching and starting to take on grammatical properties." (Leech et al 2010: 190)
    - Increase in frequency (Mair 2002, 2006)
  - Realisation of arguments, selection, valency (Vosberg 2006)
  - Role of sociolinguistics, archaism, "formality"
- Raises methodological issues:
  - Test case to showcase multifactorial nature of variation, statistics (Lohmann 2011, Levshina 2022)

Goals for this talk:

- Trace trajectory in American and British English
- Scrutinise the early stages of the change
  - 18th century
- General development
  - 18th to 21st centuries

Background

help do as an Americanism

- Omission of *to* after *help* historically more prevalent in American than British English (Algeo 1988: 22, Kjellmer 1985)
- “AmE has an especially strong preference for the pattern verb + bare infinitive.” (Biber et al. 1999: 735)
- “the bare infinitives being the preferred option in AmE (cf., e.g., Trudgill and Hannah 2002: 67).” (cited from Leech et al. 2010: 188)

References to *help* + infinitive

- Priestley, Joseph (1768) *The Rudiments of English Grammar: with Notes and Observations, for the Use of Those Who Have Made Some Proficiency in the Language*. London: Rivington. (British)

"The verb *dare* is sometimes used without the preposition *to* after it, as if it was an auxiliary verb. *Who durst defy* [...]. This construction, however, does not seem natural, except in such familiar expressions as *I dare say, I dare go,* and the like. It must, I suppose, be according to the Scotch idiom, that Mrs. Macaulay omits it after the verb *help*. *[Archbishop] Laud was promoted as an useful instrument, to help carry on the new measures of the court;* History, vol. 4. p. 150." (p. 132-3)

- Lindley Murray (1795) [1968] *English Grammar*. Menston: Scolar Press (British)

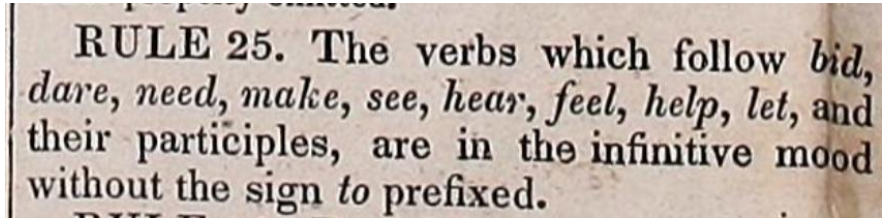
"there are some verbs which have commonly other verbs following them in the infinitive mood, without the sign *to*; as, "Bid, dare, need, make, see, hear, feel;" as also "let," and sometimes "have," not used as auxiliaries; and perhaps a few others" (p. 111)

- Cobbett, William (1818) [2014] *A Grammar of the English Language, in a Series of Letters*. Cambridge: Cambridge University Press (British)

"The *to* is, in some few cases, omitted, when the infinitive is the objective case: "as, *I dare write*." But, I "dare *to* write," is just as neat, and *more* proper. The *to* is omitted by the use of the *Ellipsis:* as, "I like *to* shoot, hunt, and course." But, care must be taken not to leave out the *to*, if you thereby make the *meaning doubtful*. Repetition is, sometimes, disagreeable, and tends to enfeeble language; but, it is always preferable to obscurity." (Letter 19, p. 144)

- Kirkham, Samuel (1823) *English Grammar in Familiar Lectures*. Rochester, N.Y.: Marshall & Dean. (American)

"The verbs which follow *bid*, *dare*, *need*, *make*, *see*, *hear*, *feel*, *help*, *let*, and their participles, are in the infinitive mood without the sign *to* prefixed; as, "He bids me come;" "I dare engage;" "Let me go;" "Help me do it;" i. e. to come, to go, to do it, &c." (Rule XXV, p. 188)


RULE 25. The verbs which follow *bid*, *dare*, *need*, *make*, *see*, *hear*, *feel*, *help*, *let*, and their participles, are in the infinitive mood without the sign *to* prefixed.

- Noah Webster (1828) *A Philosophical and Practical Grammar of the English Language*. New York: Converse. (American)

"The verbs, *bid*, *make*, *see*, *hear*, *feel*, *let*, with the auxiliaries, *may*, *can*, *must*, *shall* and *will*, and *dare* and *need*, when used as auxiliaries, are followed by the infinitive without the prefix to; as, "he bids me come;"" (Rule XXXII)

- Goold Brown (1851) *The Grammar of English Grammars*. New York: Samuel & Wood. (American)

"After *have*, *help*, and *find*, the infinitive sometimes occurs without the preposition *to*, but much oftener with it [...] "Help me do it." [...] But good writers sometimes use the particle *to* after this verb; as, "And Danby's matchless impudence helped to support the knave."—DRYDEN: *Joh. Dict. w. Help*."(pp. 601, Rule XIX, Observation 18)

- Bain Alexander (1875) *A Higher English Grammar*: New and Revised Edition. London: Longman's, Green & Co. (British)

"'To' is not necessary to the infinitive: it is not inserted after the verbs 'have,' 'may,' 'can,' 'shall,' 'will,' 'do,' 'dare,' 'bid,' 'let,' 'make,' 'must,' 'durst,' 'need not,' 'feel,' 'hear,' 'see.' 'To' is also dropped after the preposition 'but:' 'he did nothing but *read*.'" (page 150, Syntax §7)

Mair (2006)'s collection of meta-observations

- "*Help* followed by an infinitive without *to* . . ., once condemned as an Americanism, is now accepted in British English . . ." (Wood 1962: 107). (Mair 2006: 136, fn. 2)
- "Foster (1968: 204): '. . . the constructions accompanying certain verbs quietly change over the years without causing any great outcry. … American idiom, a typical example being seen in the omission of the preposition 'to' after 'help'. Now this phenomenon was not unknown in poetical and somewhat archaic language. … But only in the late nineteen-thirties and early 'forties did the construction really make headway in Britain. Its acceptance into the standard language was very rapid and J. Hubert Jagger … commented on 'the speed with which the American habit of omitting *to* after *help* has invaded Britain' (p. 55). But in spite of the speedy acceptance of the new form the old one is still well entrenched and the two rivals seem destined to battle it out for some time to come.'" (Mair 2006: 136, fn. 3)
- "Of the two constructions with *help*, that with *to* is more common in BrE, and that without *to* is more common in AmE. (Quirk et al. 1985: 1205f)" (Mair 2006: 136)

Methodology

Corpora

- *help* in 18th century corpora
  - American:
    - Evans-TCP (Text Creation Partnership 2012)
  - British:
    - CLMET3.1 (De Smet et al. 2015), Old Bailey corpus (Huber et al. 2016)

Evans-TCP

- Corpus of
  - 4,976 text files
    - Transcribed and digitised as a collaboration between the Text Creation Partnership (TCP) at the University of Michigan, Readex, and the American Antiquarian Society.

  of early American imprints (print dates 1639-1800).
- Catalogued by Evans (1903-1942).
- Public domain text files can be downloaded from Evans-TCP github mirror.
- Other implementations / online search options
  - Evans-TCP (index searches)
  - Readex by NewsBank (paid)
  - Corpus of Founding Era American English (online search interface)

Problems with Evans-TCP

- The corpus mixes
  - text author origin
    - (i) texts authored by first-settler / American-born colonialist (“American”)
    - (ii) British- or European-authored works reprinted in America (“British”)

    e.g., N20265 Philadelphia reprint of London-based print-seller Carington Bowles’ (1724–1793) *Series of Instructional Manuals*
  - a wide range of genres
  - translations with native compositions.
- Potential confounds to variety.

Variety classification

- AI/LLM (Claude 4.6, 4.7) was presented with metadata, prompted to provide a label for variety, genre, translation dimensions:

FileNu	LongT	Incipi	PubYe	AuthorNar	Autho	Pub	Regis	Litera	Fictiona	Genre	TextSegr	Keywords	WordCou	EVANSF	Variety	GenreLLM	Translati	AuthorNotesLLM	EVANSFileID
515	An ora	At a m	1786	Austin, Jor	1748-	Bos	High	Prose	Non-fic	Patriotic or	half title	Fourth of Jul	4114	N15351	American-born	speech-patriotic	native	Jonathan Loring Austin (1748–1826), Boston. Boston-bc	N15351
516	Travel	ROBIN	1786	Defoe, Dar	1661-	Wol	Low	Prose	Fiction	Adventure	illustrati	Robinsonad	3631	N15415	British	novel-juvenile, voya	native	Daniel Defoe (1660–1731), London. English writer; autho	N15415
517	The pc	CONT	1786	Freneau, P	1752-	Phil	High	Poetr	Fiction	War poem	canto	United State	83906	N15445	American-born	poem	native	Philip Morin Freneau (1752–1832), born in New York Cit	N15445

- The result is not perfect, but useful.
  - Hallucinations and other errors cannot be ruled out.
  - I checked many results. They have basically always been correct / reasonable.
  - There are many difficult cases, e.g., multiple authors? AI often flagging, e.g. **`AuthorLifetime 1693–1758:`** `Jonathan Edwards's correct dates are 1703–1758; not 1693.` Therefore: Classified by main author mentioned.
  - Future work could improve. Manual verification. Dissect text by author, Deal with duplicate materials (e.g., N00716 = N05268 = N19488), etc.

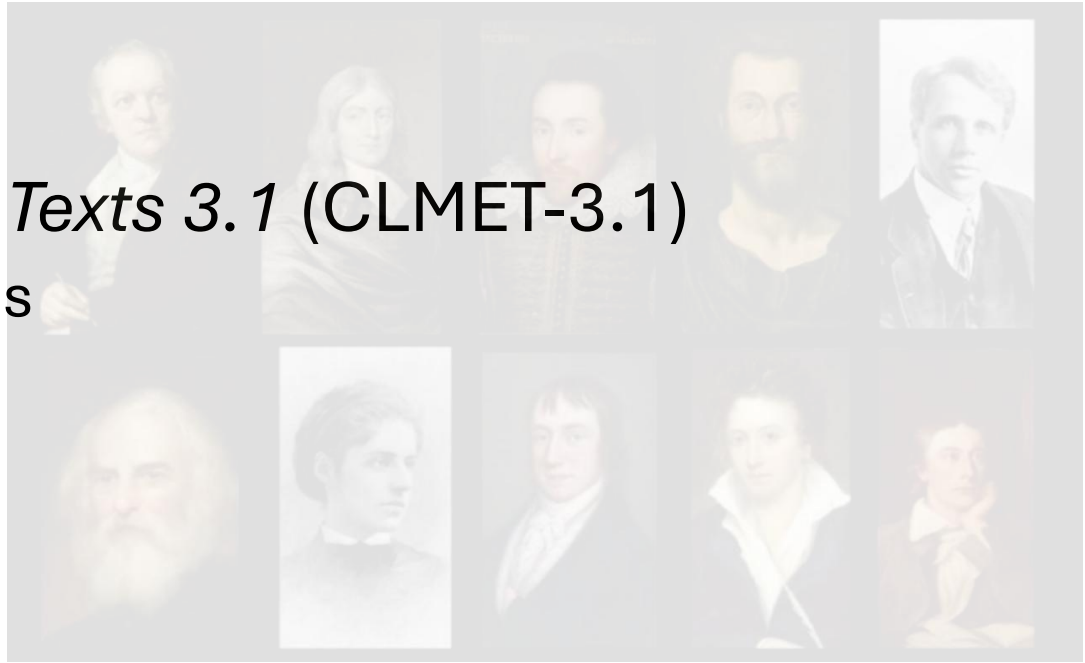
Variety classification

Label	Files	Words
American	3,219	42,158,691
British	1,197	25,483,515
Unknown	560	7,755,660
Total	4,976	75,397,866

Table 1: Distribution of Evans-TCP material by automatically assigned variety category across file and work token count

CLMET-3

- The *Corpus of Late Modern English Texts 3.1* (CLMET-3.1)
  - 333 texts with roughly 34 million words
  - British-authored texts
  - 1710-1920



Duplicate texts

- Substantial overlap between CLMET-3.1 and Evans-TCP
  - Evans-TCP all correctly identified as British
  - Duplicates removed from Evans-TCP, 21 files, loss of c. 800k word tokens

CLMET ID	Evans-TCP ID	Content
9	N22293	Richardson's *Clarissa*
18	N21719	Doddridge's *Life of Col. James Gardiner*
23	N15912	Fielding's *The History of Tom Jones*
25	N08554	Johnson's *Rasselas, Prince of Abyssinia*
35	N31727	Sterne's A *Sentimental Journey*
48	N25990	Smollett' *Adventures of Peregrine Pickle*
58	N09747	Goldsmith's *The Vicar of Wakefield*
64	N10944	Burke's *On Conciliation*
65	N24713	Reeve's *The Old English Baron*
75	N12680	Pratt's *The Pupil of Pleasure*
84	N35515	Chetwood's Captain Robert Boyle
98	N20675	Inchbald's *Every One Has His Fault*
99	N23167	Inchbald's *Nature and Art*
103	N21834	Godwin's *Caleb Williams*
107	N19251	Wollstonecraft's *Rights of Woman*
108	N23927	Wollstonecraft's *Letters*
109	N26724	Wollstonecraft's Maria
115	N22267	Radcliffe's *Mysteries of Udolpho*
129	N26840	Lewis' *The Monk*
130	N25599	Lewis' *The Castle Spectre*
131	N28320	Lewis' translation of The East Indian

Table 2: Identified duplicate material shared between CLMET-3 and Evans-TCP with indication of corpus file IDs and text title

Old Bailey Corpus

- The *Old Bailey Corpus* consists of proceedings of London's Central Criminal Court.
- Cover 1674 to 1913.
- 637 XML-formatted files with 24m word tokens.
  - can be downloaded from Clarin University of Saarland server
- They can be described as "speech-related" texts documents.
- File content of two types:
  - Verbatim files: tend to include many speakers (witnesses, accused, judges, etc.) and much direct speech (utterances inside <u> tags). Corpus provides some social information (gender, historical social class). Written as "shorthand" after c. 1720s.
  - Scribal files: tend to have less or no direct speech; instead, summary of trial with indirect speech. Common especially early in the period.
- Pre-processing. Classify files into verbatim vs. scribal (cut-off: >45% of text contained in <u>). Create separate files for speakers of verbatim files for easier retrieval of social information. 145k files.

Other corpora used

- *help* in 19th-21st century corpora
  - American:
    - COHA (Davies 2010), COCA (Davies 2008)
  - British:
    - Hansard for 19th century, 10IndivCorpus (a custom-made corpus 10 prolific British writers with 10m word tokens), BNC, BNC2014, HUMS19UK

Data collection

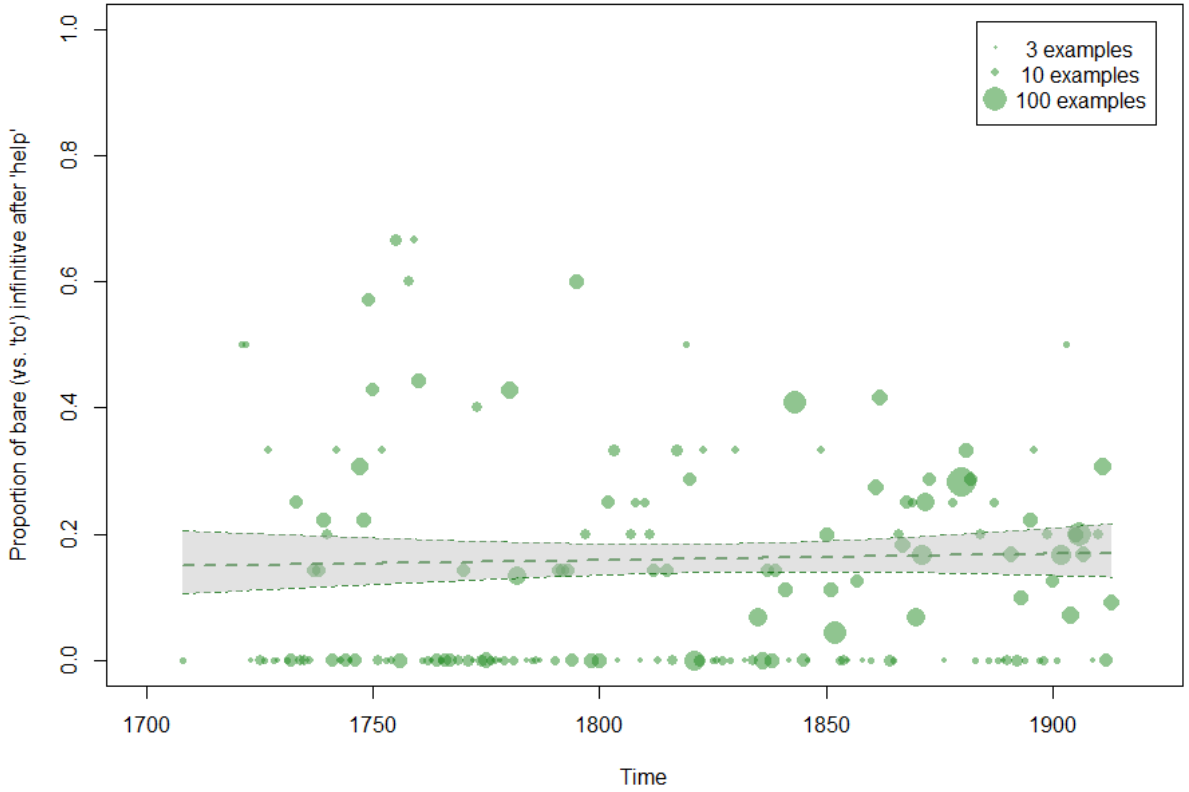
- Data collection:
  - Python script (Python Software Foundation 2026). Every corpus has a custom-made script. Some are similar, but none identical.
  - Pre-processing. Raw text files (strip XML). Lacunas (DISP="...") …
  - Automatic extraction of concordance lines, values for dependent and many independent variables, meta-information.
  - Levels of dependent variable BARE (1) and TO (0) for two.
  - Horror Aequi present (*to help*), absent elsewhere.
  - Minor manual correction, early examples, BARE and TO, but automatic annotation usually correct.

Results

Early phase of the development

- Old Bailey, CLMET-3, Evans-TCP

Distribution in the Old Bailey dataset



Innovative variant attested from the beginning of the sampling period! Surprising.

- at c. ≈16% BARE (vs. TO) in Old Bailey
- N=1030, N_{BARE} =166, N_{TO} = 864
- Stable between c. 1700-1900.
- Origin in Early Modern English (1500-1700)

Figure 1: Bubble plot of the proportion of bare (vs. *to*) infinitives across time (1700-1920), each data point representing a year with point size proportional to the number of examples per year. Visualisation of regression line from generalized additive model (gam) predicting the proportion from a penalised spline of Year, grey bands indicating 95% confidence intervals of the fitted smooth. Data from Old Bailey Corpus.

The earliest example in Old Bailey

(5) William Langly, of Alhallows Barkin,

was indicted for feloniously stealing 28 Pound of Tobacco

[…]

The Prisoner in his Defence said he met a Man who desired him to **help** him **carry** the Tobacco, and that he thereupon lent him his Shirt to put it in; but bringing no body to prove it, the Jury found him guilty

(Old Bailey, 1721, Scribal file, ID 10-17211206)

Early examples from verbatim files

(6) a. the Prisoner desir'd me to help load 'em

(Old Bailey, 1733, verbatim file,
ID 2377-17331205-0251)

b. I saw the prisoner help load the horses

(Old Bailey, 1749, verbatim file,
ID 14870-17490705-0140)

c. if he was an honest man he would come in and help me look about

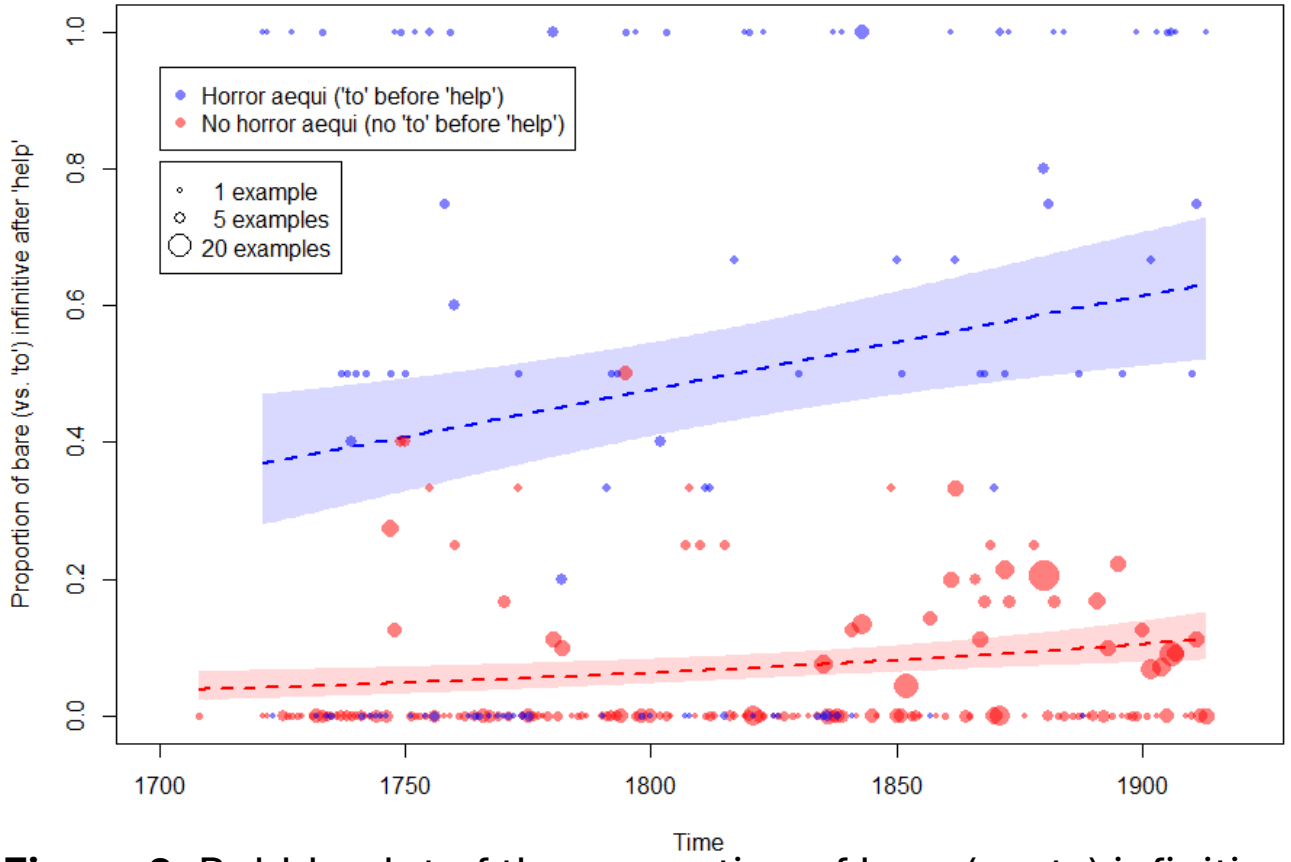
(Old Bailey, 1747, verbatim file,
ID 16512-17510417-0011)

d. he, the Prisoner, helped land them

(Old Bailey, 1747, verbatim file,
ID 13240-17470225-0140)

- 19 examples of BARE in verbatim files (vs. 115 TO) from before 1750
- 13/19 have Horror Aequi (*to help*) (6a)
- 2 ECM/AcI/subject-to-object raising (6b), 2 questions, 2 coordination structure (6c)
- Leaves only 1 "subject – *help* + bare infinitive" order, even here an appositive (6d)
- 9 with objects (helpee), 10 without

Great importance of Horror Aequi



- Horror Aequi (HA) is an extremely strong effect (OR=11.5, 95%CI: 7.8-17.9)

	BARE	TO
to + help	103 (49%)	107
non-*to* +help	63 (8%)	757

Table 3: Contingency table of Infinitive x HA, Old Bailey

- Time trend is positive but weak
- HA x year interaction not significant (when added to this model)

Figure 2: Bubble plot of the proportion of bare (vs. *to*) infinitives across time (1700-1920), each data point representing a year with point size proportional to the number of examples per year, colour-coded for Horror Aequi. Regression lines from a logistic regression model predicting the proportion from year (scaled, mean year 1820) and Horror Aequi, bands indicating 95% Wald confidence intervals for the fitted prediction at each year for each Horror Aequi condition. Data from Old Bailey Corpus.

The earliest examples in CLMET-3

(7) a. I should be glad you was present at the nuptials, to **help** me throw the stocking, and perform other ceremonies peculiar to that occasion.

(CLMET3 file 50, Smollett, Tobias G. (1771) *The expedition of Humphrey Clinker*)

b. i. I went with my Mother to **help** look at some houses in New King Street

ii. A grand thought has struck me as to our Gowns. This 6 weeks mourning makes so great a difference that I shall not go to Miss Hare, till you can come **& help** chuse yourself; unless you particularly wish the contrary. - It may be hardly worth while perhaps to have the Gowns so expensively made up; we may buy a cap or a veil instead; - but we can talk more of this together.

iii. it may enable you to be yet more decided with Fanny, & **help** to settle her faith .

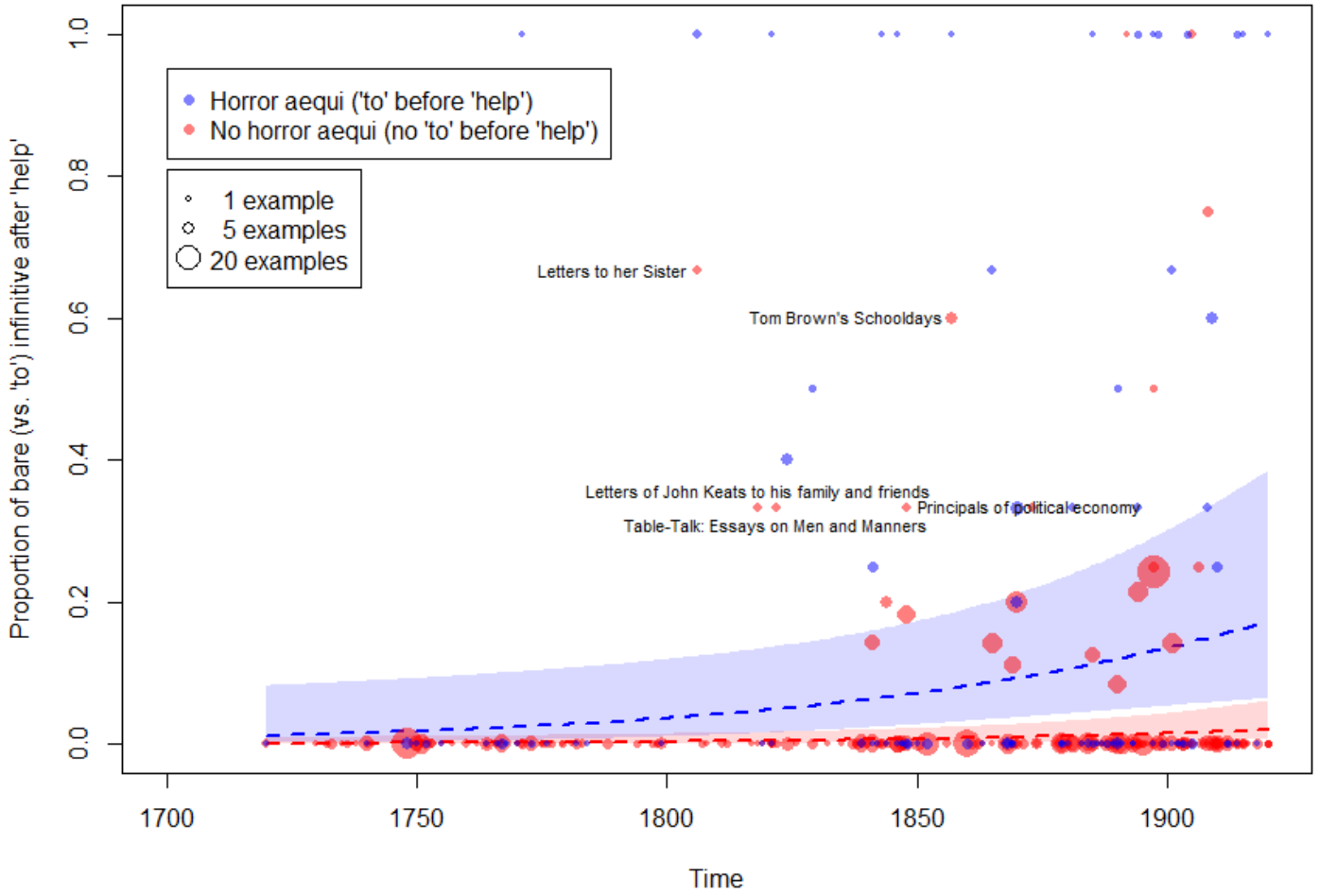
(CLMET3 FileID -135, Austen, Jane (1806) *Letters to her Sister*)

c. and tell me also when you will **help** me waste a sullen day

CLMET3 FileID -154, Keats, John (1818) *Letters of John Keats to his family and friends*)

- No attestations before 1750
- Jane Austen's *Letters to her Sister* 5 BARE vs. 1 TO (*to help*) (7ai-iii)
- first instance with subject – *help* + bare infinitive from 1818
- Early examples from: "colloquial", speech-related contexts: burlesque, personal correspondence

CLMET-3



- Horror Aequi (HA) is an extremely strong effect (OR=5.9, 95%CI: 3.6-9.6)

	BARE	TO
to + help	39 (22%)	136
non-*to* +help	44 (5%)	900

Table 4: Contingency table of Infinitive x HA, CLMET-3

- Bare infinitive after help in British fiction from the middle of the 19th century.
- Time trend significant but weak
- HA x year interaction not significant (when added to this model)
- Individual texts important (Adjusted ICC: 0.598 [0.474, 0.939])

Figure 3: Bubble plot of the proportion of bare (vs. *to*) infinitives across time (1700-1920), each data point representing a file with point size proportional to the number of examples per file, colour-coded for Horror Aequi. Regression lines from a logistic regression model predicting the proportion from year (scaled, mean year 1844) and Horror Aequi, with random file intercepts, bands indicating 95% Wald confidence intervals for the fitted prediction at each year for each Horror Aequi condition. Data from CLMET-3.

Earliest examples in Evans-TCP

(8) a. all such persons as by the advantage of their arts & trades are more enabled to **help** bear the publick charge then common laborours and workm•n

(EVANS-TCP, N00010, 1648, American, *Book of the General Laws and Liberties Massachusetts*)

b. Quest. How should we demean our selves towards Gods Church, and those in Covenant with Him?

A. To Psal: 16 3 & 15 4. Psal 137 6. Col. 1.4. 1 Joh: 3.1• prize, honour, & love such, to entertain & desire fellowship with them, imitate their vertues, **help** bear their burdens,

(EVANS-TCP, N00020, 1657, British, John Fiske' The watering of the olive plant in Christs garden)

Example illustrating variety classifications

(9) The North-West-wind was violent, and the cold such tha• the strongest of us thought We should not out-live that day: having gott through the boggy Marsh and on the Sea-Shoar, our People, Black and White, made all speed, one not staying for another that could not travel so fast; None but I with my Wife and Child, Robert Barrow, my kinsman Benjamin Allen and my Negroe London, whom I kept to **help** carry my Child, keeping together;

(EVANS-TCP, N00716, **1696**, Jonathan Dickinson (1663–1722), born in Port Royal, Jamaica, to English parents)

American examples are more common

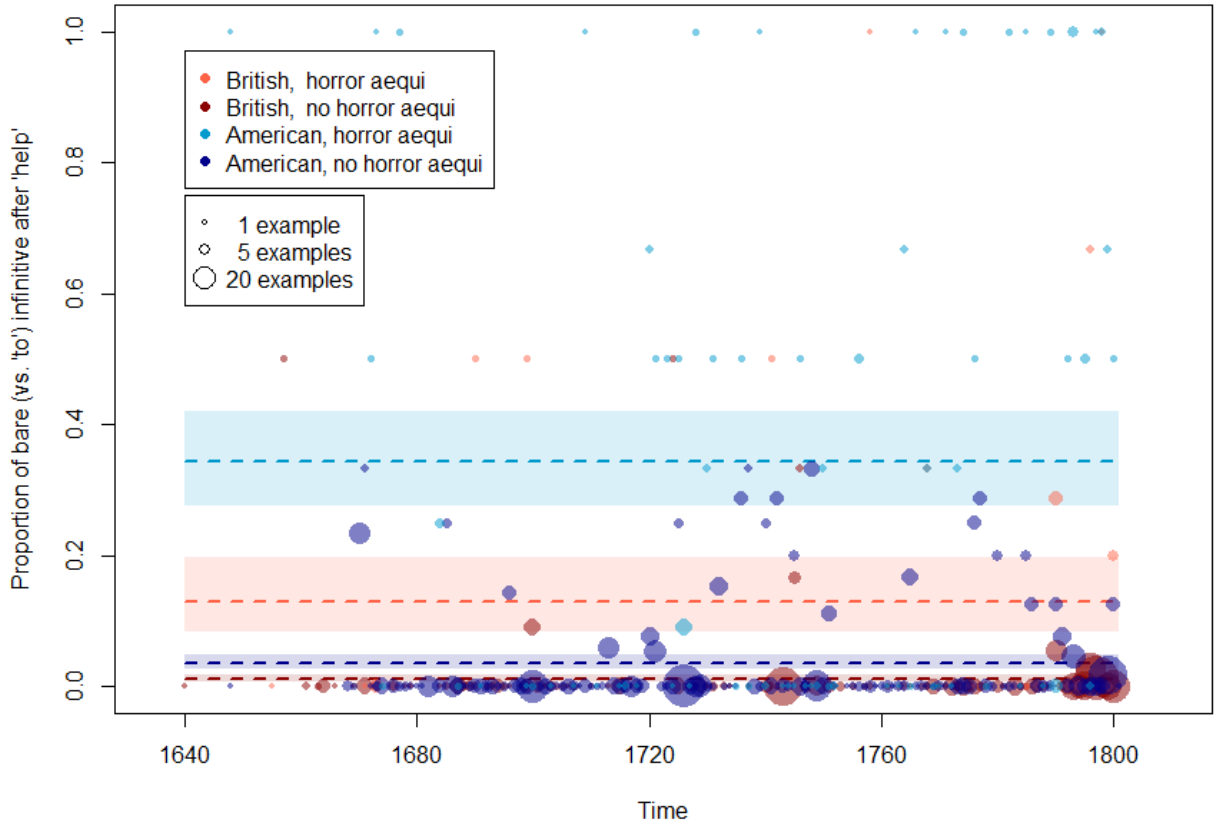
(10) a. they had sent home about twenty men to **help** strengthen the Plantations on Connecticut

(EVANS-TCP, N00179, 1677, Increase Mather (1639–1723), New England Puritan clergyman)

b. the Samaritans, who endeavoured all they could to obstruct the Building of the Temple, were now obliged to pay the King's Tribute to the Iews, to **help** bear the Charge of the Work.

(EVANS-TCP, N02838, **1731**, John Barnard (1681–1770), born in Boston)

Evans-TCP illustration



- Time not significant when added to this model.
- HA not different across varieties.
- Clear preference for bare infinitive after *help* in American English from earliest text records on.

British		
	BARE	TO
to + help	11 (13%)	77
non-*to* +*help*	8 (1%)	716
American		
	BARE	TO
to + help	52 (35%)	98
non-*to* + help	38 (4%)	1034

Table 5: Contingency table of Infinitive x HA x Variety, Evans-TCP

Figure 4: Bubble plot of the proportion of bare (vs. *to*) infinitives across time (1640-1800), each data point representing a year with point size proportional to the number of examples per year, colour-coded for variety and Horror Aequi. Regression lines from a logistic regression model predicting the proportion from variety and Horror Aequi, bands indicating 95% Wald confidence intervals for the fitted prediction at each year for each variety and Horror Aequi condition. Data from Evans-TCP.

Trajectory throughout history

Variety	Corpus	TO	BARE	Time
American	Evans-TCP	2081	120	1642-1800
	COHA	11512	22907	1810-1999
	COCA	10775	71521	2000-2019
British	CLMET3	1037	83	1700-1920
	OldBailey	864	166	1700-1913
	Hansard19thc	49	1454	1802-1899
	HUMS19UK	400	33	1800-1905
	10Indiv	579	611	1890-2000
	BNCWritten	8410	9592	1985-1994
	BNCSpoken	413	534	1990-1994
	BNC2014	115	607	2012-2014

Table 6: Occurrences of TO and BARE in 11 corpora and their time coverages

Envelope of variation

- Exclude cases of “could not help but do”, (11)

(11) you can't **help** but believe
(Evans-TCP, N04873, 1748, American-born)

by ignoring all cases of negation + *help*

- Exclude all cases of *helping*, since those have a higher probability to occur with to (e.g. Levshina 2022)

Trajectory without Horror Aequi

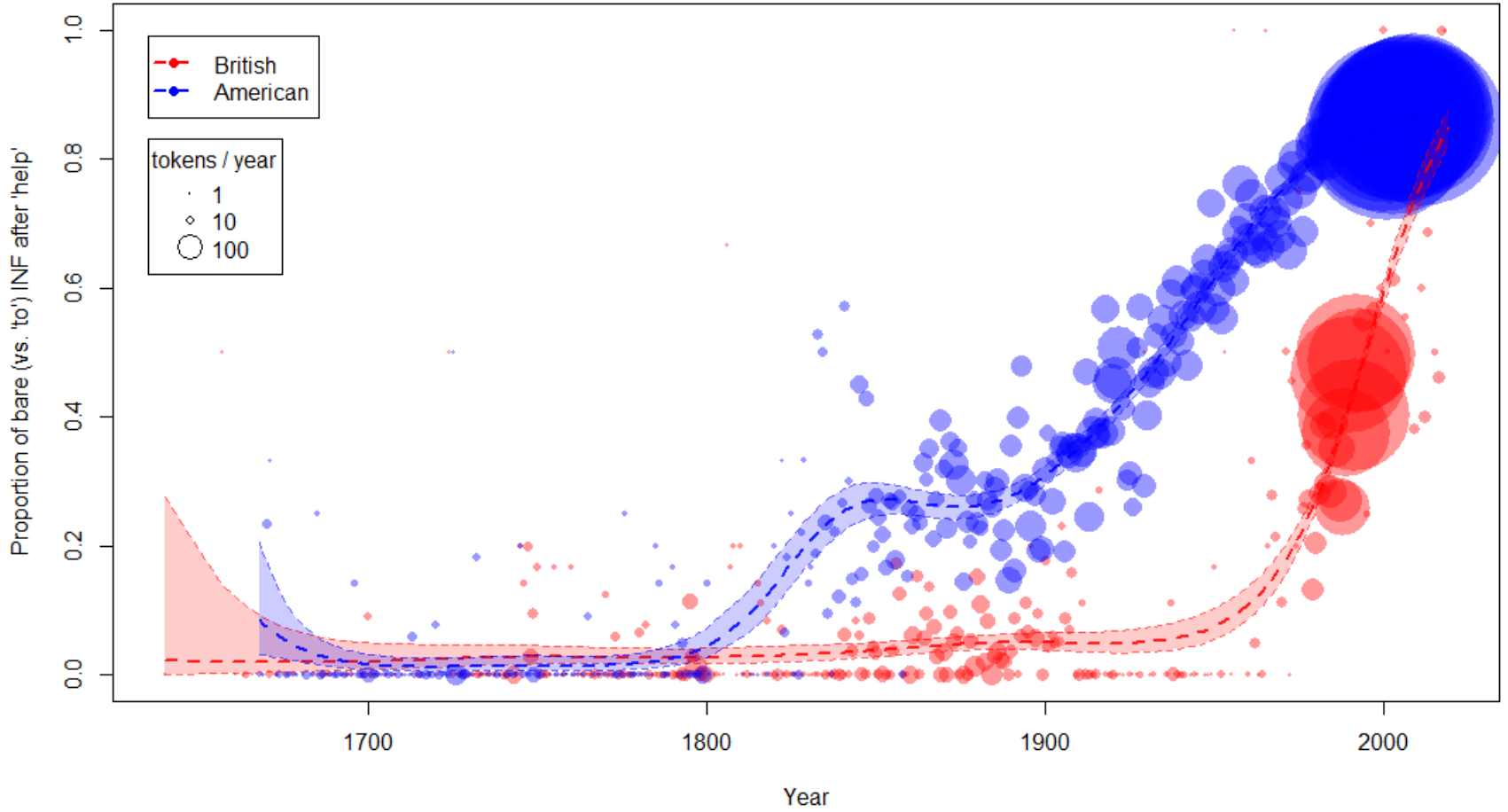


Figure 5: Bubble plot of the proportion of bare (vs. *to*) infinitives across time (1642-2019), each data point representing a year with point size proportional to the number of examples per year, no HA. Visualisation of regression line from 2 generalized additive models (gam) predicting the proportion for AmE, BrE from a penalised spline of Year, bands indicating 95% confidence intervals of the fitted smooth.

Trajectory with Horror Aequi (*to* before *help*)

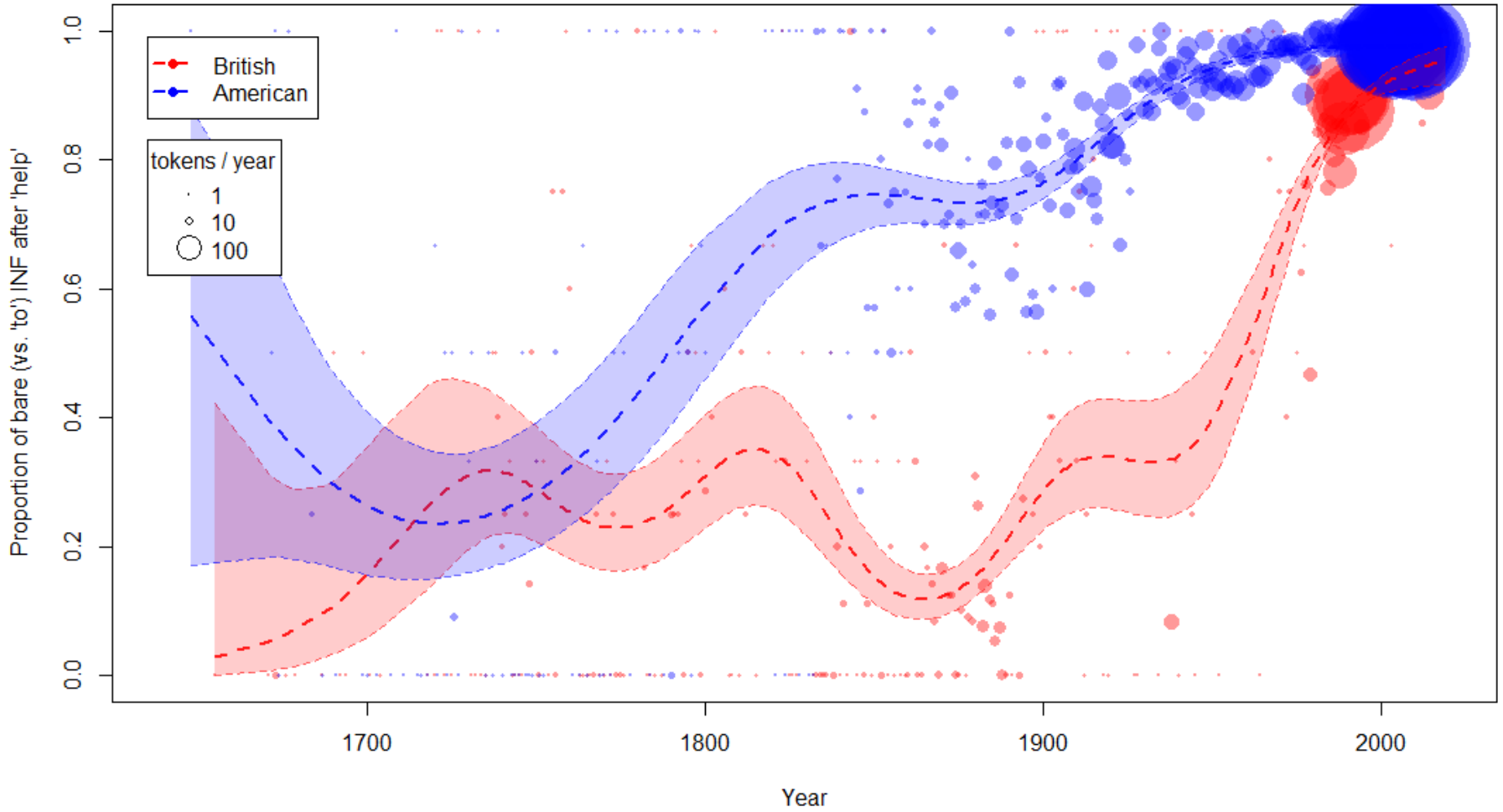


Figure 6: Bubble plot of the proportion of bare (vs. *to*) infinitives across time (1642-2019), each data point representing a year with point size proportional to the number of examples per year, in HA condition. Visualisation of regression line from 2 generalized additive models (gam) predicting the proportion for AmE, BrE from a penalised spline of Year, bands indicating 95% confidence intervals of the fitted smooth.

Core findings

- Trajectory of the rise of bare infinitives after *help* can be traced with substantial precision. Largely confirms previous findings:
  - American and British English "parallel diachronic development", "two parallel developments which, however, differ somewhat in scope, speed and volatility." (Leech et al. 2010: 187, 193)
  - "the bare infinitive is [now] more common than the *to*-infinitive in both dialects." (Biber et al. 1999: 735) (~95% bare-INF in PDE)
- Non-HA condition has a plausible S-shaped curve 1850-today
- Substantial dip in British English in the late 19th century. Artifact?
- In HA condition, no substantial difference before c. 1750.

Discussion

Origin of the change

- Variety is an important predictor for the form of the infinitive after *help* between c. 1700-today.
  - Some evidence for a clear preference for bare infinitive after *help* in American English from earliest text records on (Evans-TCP vs. Old Bailey, CLMET).
  - Earlier rise in American (c.1800-2000) than British (c. late 19th c. – ongoing) English
  - No clear evidence that the change originated or spread from American English
    - Old Bailey attests examples from British English convincingly. Data not clear-cut.
    - Origin and spread in American English remain possible.
- The origin of the change, if there is an activation point, must be explored in Early Modern English (1500-1700).

Horror Aequi

- Horror Aequi is among the most important predictors of the variation (second only to time, variety).
- It may be even more important than previously recognised
  - majority of earliest examples instantiate HA
  - HA may be only the most common of a class of constructions, including AcI (*see someone help do*) and coordination (*go and help do*).
  - HA may be framed as having played a role in the activation of this change in that it provides a rationale for the existence of the new form.

Future research ideas

- Genre control, as a means to control “formality”
- Role of help + –*ing*: in the 19th century *help doing* may have been more common in British than American English.
- Continue search for meta reflections, early grammar books and other remarks on the linguistic development
  - Potentially a dip in bare infinitives in late 19th century British English
  - Contemporary views could help to establish if this is real or an artifact. Overt prestige? Deliberate effort to speak conservatively in Britain? Prescriptive pressure?
- More corpora:
  - Eighteenth Century Collections Online (ECCO-TCP github), (supposedly) 100m words
  - BYU’s *Corpus of Founding Era American English* (COFEA). “the full corpus is only available through an online interface.” Duplicates. Variety mixing.
  - Regional dialects of 17th century English

Final examples

(13) a. It is the Government's aim in the reform of housing finance to concentrate subsidies on people and areas in need. This will **help** ensure a continued high level of building there.
(Hansard, Commons, 27/10/1971, FileID 1399757)

b. I do not see that the hon. Member has explained to the House today how his new Clause would specifically **help** solve this problem of seasonal unemployment.
(Hansard, Commons, 28/07/1972, FileID 1426074)

- Synonyms: "contribute, facilitate, have partial, limited responsibility"
- register like business, government, admin, email speak
- hedging

Conclusion

Conclusion

- Overview over the diachronic trajectory of bare infinitival complementation of *help* from late Modern English to today with special focus on the 18th century, controlling for Horror Aequi.
  - *help do* is attested at least as early as the mid 17th century, recorded, in particular, in speech-related contexts.
  - Noticeable rise in American English from ≈ 1800,
    British English from ≈ 1900 (Figures 5, 6)
- For origin of change, either true activation or lost condition on rare context (lower class speech?), we must look into 17th century.

References

Bibliography

Algeo, John (1988) 'British and American grammatical differences.' *International Journal of Lexicography* 1.1, 1-31. https://doi.org/10.1093/ijl/1.1.1

Biber, Douglas, Stig Johannson, Geoffrey Leech, Susan Conrad & Edward Finegan (1999) *The Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.

Evans, Charles (1903-1942) *American bibliography: a Chronological Dictionary of all Books, Pamphlets and Periodical Publications Printed in the United States of America*. Series I in 13 volumes. Chicago: Blakely Press.

Kjellmer, Göran (1985) 'Help to/help∅ revisited.' English Studies 66.2, 156-161. https://doi.org/10.1080/00138388508598377

Labov, William (1990) 'The intersection of sex and social class in the course of linguistic change.' *Language Variation and Change* **2**, 205-254.

Labov, William (2007) 'Transmission and Diffusion.' *Language* **83.2** 344–387.

Leech, Geoffrey, MarianneHundt, Christian Mair, Nicholas Smith (2010) *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511642210

Levshina, Natalia (2018) 'Probabilistic Grammar and Constructional Predictability: Bayesian Generalized Additive Models of Help + (To) Infinitive in Varieties of Web-Based English.' *Glossa* 3.1.55, 1–22. https://doi.org/10.5334/gjgl.294

Levshina, Natalia (2022) 'Comparing Bayesian and Frequentist Models of Language Variation: The Case of Help + (to-)Infinitive.' In: Schützler, Ole & Julia Schlüter (eds.) Data and Methods in Corpus Linguistics. Cambridge: Cambridge University Press, 224–258. https://doi.org/10.1017/9781108589314.009

Lohmann, Arne (2011) 'Help vs. help to: A multifactorial, mixed effects account of infinitive marker omission.' *English Language and Linguistics* 15.3, 499–521. https://doi.org/10.1017/S1360674311000141

Lind, Åge (1983) 'The variant forms help to/help Ø.' English Studies 64.3, 263-273. https://doi.org/10.1080/00138388308598255

Mair, Christian (2002) 'Three Changing Patterns of Verb Complementation in Late Modern English'. English Language and Linguistics 6.1, 105–31. https://doi.org/10.1017/S1360674302001065

Mair, Christian (2006) *Twentieth-Century English. History, Variation and Standardization*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486951

McEnery, Anthony & Zhonghua Xiao (2005) 'HELP or HELP to: What Do Corpora Have to Say?' English Studies 86.2, 161–187. https://doi.org/10.1080/0013838042000339880

Python Software Foundation. (2026). *Python (Version 3.14)*. https://www.python.org/

Vosberg, Uwe (2006) *Die Große Komplementverschiebung: Außersemantische Einflüsse auf die Entwicklung satzwertiger Ergänzungen im Neuenglischen*. Tübingen: Narr.

Corpora

- Davies, Mark (2012) 'Expanding horizons in historical linguistics with the 400 million word corpus of historical American English.' *Corpora* 7. 121–157.
- Davies, Mark (2010). *The Corpus of Historical American English (COHA)*. Available online at https://www.english-corpora.org/coha/.
- De Smet, Hendrik, Susanne Flach, Jukka Tyrkkö & Hans-Jürgen Diller. 2015. The Corpus of Late Modern English (CLMET), version 3.1: Improved tokenization and linguistic annotation. KU Leuven, FU Berlin, U Tampere, RU Bochum. Available from https://fedora.clarin-d.uni-saarland.de/clmet/clmet.html.
- Text Creation Partnership (2014) *Evans Early American Imprints-Text Creation Partnership (Evans-TCP)*. Github. Accessed 23/5/2026.
- Huber, M., Nissel, M. & Puga, P. (2016). *The Old Bailey Corpus (OBC), version 2.0 (1720–1913)*. Department of English, Justus Liebig University Giessen, Germany. Available at: https:// fedora.clarin-d.uni-saarland.de/oldbailey/
- Love, R., Dembry, C., Hardie, A., Brezina, V. & McEnery, T. (2017). *The Spoken BNC2014: designing and building a spoken corpus of everyday conversations*. International Journal of Corpus Linguistics, 22(3): 319-344. DOI: 10.1075/ ijcl.22.3.02lov

Thank you for your attention

- *The Student Transcribed Corpus of Spoken American English*
  - https://spokencorpus.org/

- *The Parsed Corpus of Middle English Poetry*
  - https://pcmep.net/

Appendices

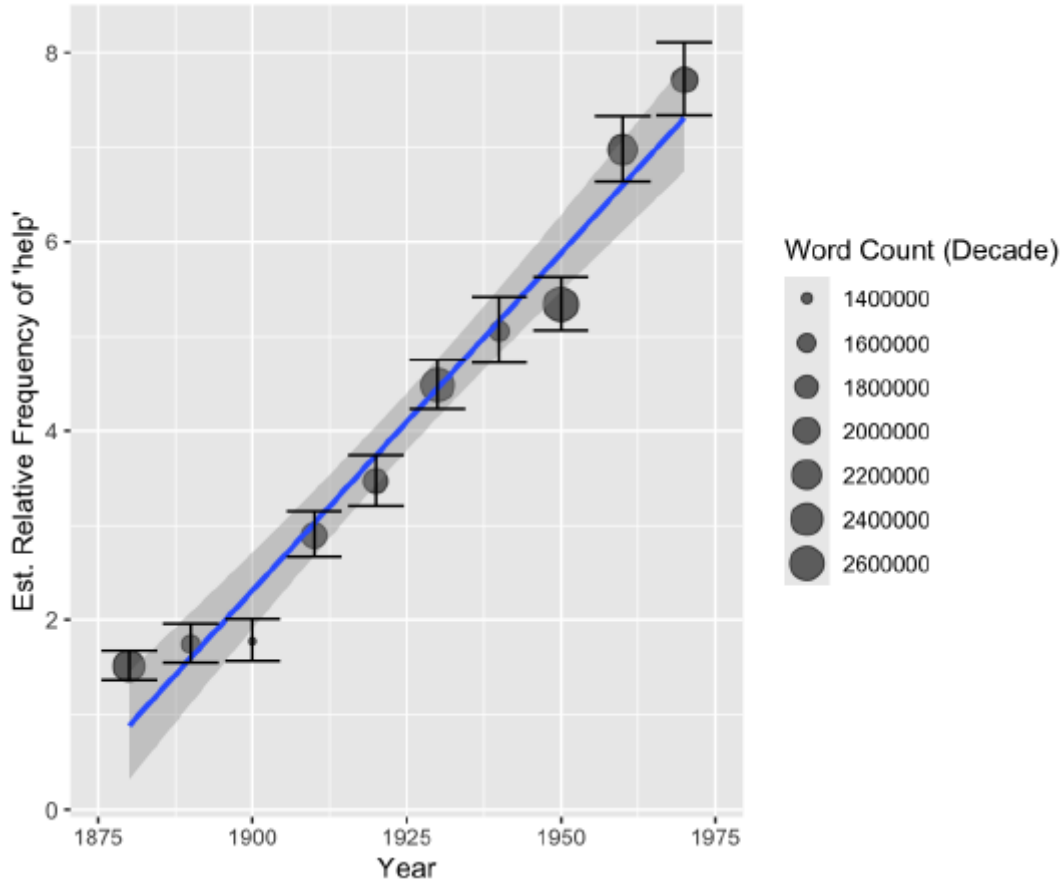


Figure A: Plot of the relative frequency of *help* followed by a non-finite clause (per 10,000 words) across decades in a sample of a time-balanced version of Hansard (1880-1979), 95% Wilson score confidence intervals, trend like from linear regression with 95% Wald confidence intervals of the rate of change.

Figure 9.22
Frequencies of *to* + *help* followed by *to*-clause v. bare infinitive clause, by register

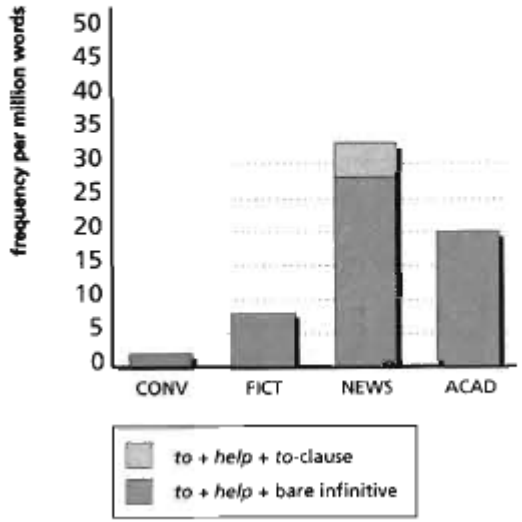


Figure B: From Biber et al. 1999: 737

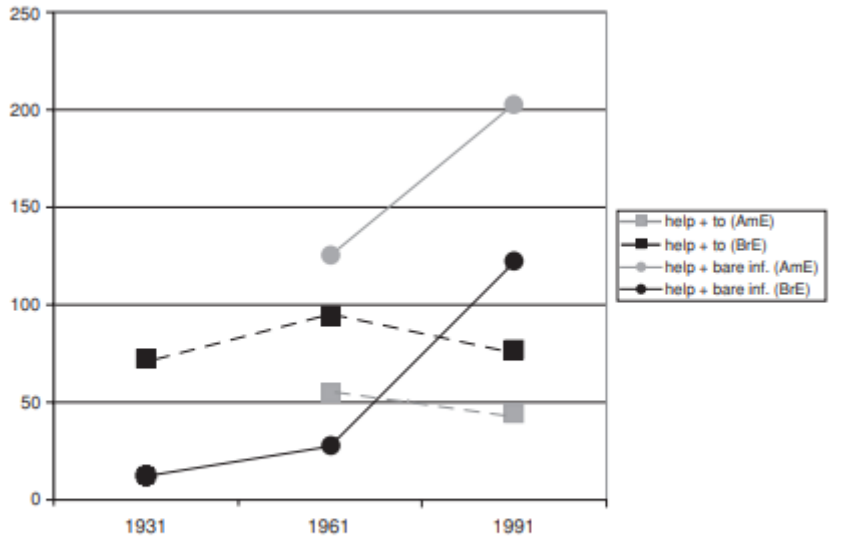


Figure 9.1 *To-* vs bare infinitives with *help* (all construction types) in five corpora – diachronic trends (Broken lines indicate the *to* construction)

Figure C: from Leech et al. 2010: 189

- "going far beyond what would be needed merely to compensate for the decrease of to."
- Supports grammaticalisation

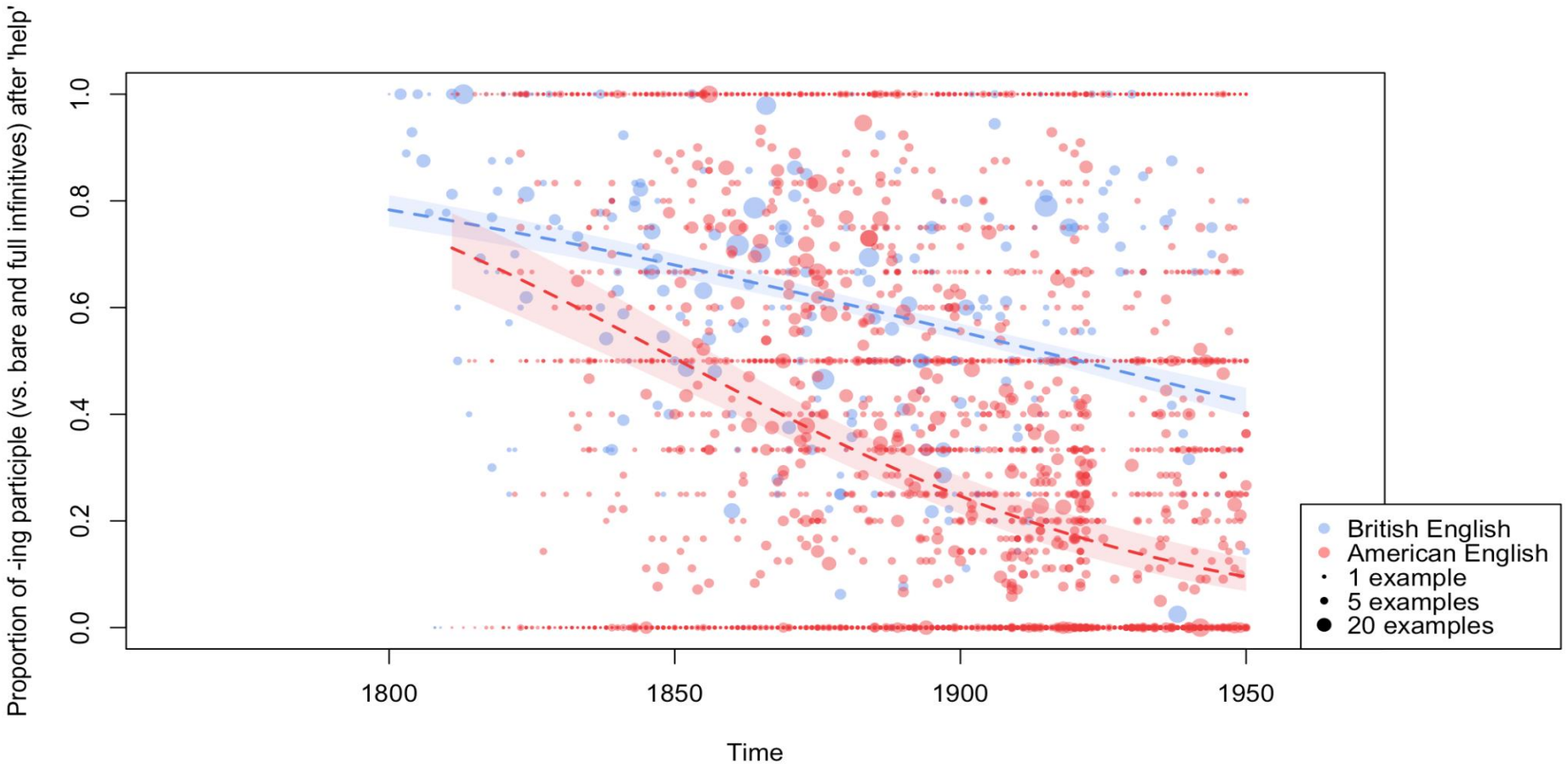


Figure D: Plot of *-ing* participle vs. bare and full infinitives after *help*, colour-coded for Variety, data points represent texts, point size proportional to number of examples per text, Horror Aequi (*to help*) removed, *cannot help but do* patterns removed, regression lines from mixed-effects logistic regression predicting infinitive from year, variety, and their interaction, including random text intercepts.

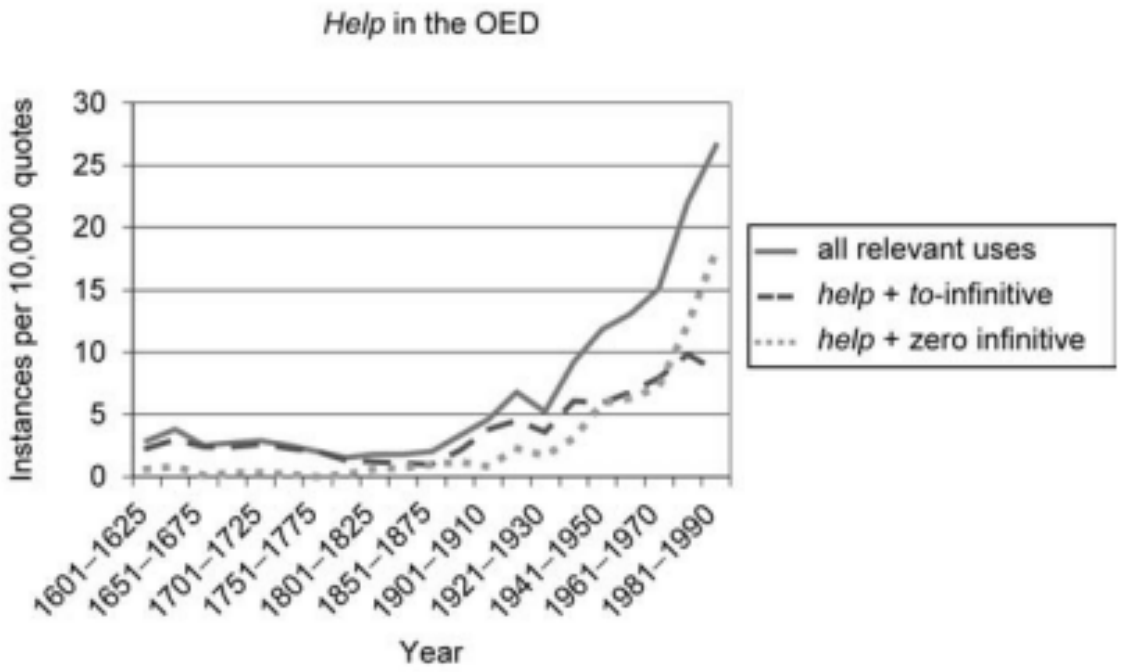


Figure 4.4 *Help* + infinitive 1600–2000 – frequency as n/10,000 citations

Figure E: From: Mair 2006: 138