Three-way competition and the emergence of *do*-support in English

Aaron Ecay

University of Pennsylvania

July 4, 2012

Table of contents

The phenomenon of *do*-support

- ▶ A syntactic change took place after 1400:
    - ▶ V → T raising lost
    - ▶ auxiliary *do* used in "last resort" contexts (which would otherwise demand V → T movement)
- ▶ Well studied quantitatively since Ellegård (1953), though puzzles remain
    - ▶ why does the change not follow an S-shaped curve through its entire trajectory?
    - ▶ what is the relevance of social factors to the grammatical change? (on both these points, see Warner 2005)
    - ▶ what is the relevance of affirmative declarative *do* to *do*-support?

The phenomenon of *do*-support

- A syntactic change took place after 1400:
  - V → T raising lost
  - auxiliary *do* used in "last resort" contexts (which would otherwise demand V → T movement)
- Well studied quantitatively since Ellegård (1953), though puzzles remain
  - why does the change not follow an S-shaped curve through its entire trajectory?
  - what is the relevance of social factors to the grammatical change? (on both these points, see Warner 2005)
  - what is the relevance of affirmative declarative *do* to *do*-support?

The origin of *do*-support

- ▶ Various theories have been adduced regarding the origin of *do*-support
- ▶ Ellegård proposed that *do*-support arose from a Middle English (ME) causative construction
- ▶ Different ME dialect areas used different lexical items for the causative:

(1)   So he ded smyte of his hed

PPCME2, CMCAPCHR-M4,98.2054

(2)   For he makth serche all the contree

PPCME2, CMMANDEV-M3,127.3087

Dialect contact and *do*-support

- ▶ When the causee is not overt, these are susceptible to reanalysis as auxiliary constructions
- ▶ Thus, tokens of eastern *do* were reanalyzed as auxiliaries by western speakers, for whom *do* could not be a causative

Distribution of ME causatives
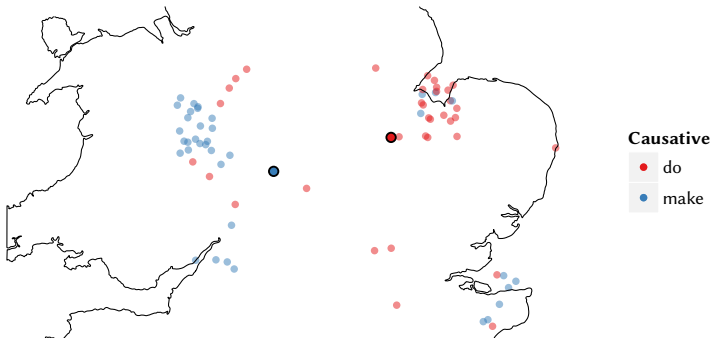
Causative
- ● do
- ● make

Causative origin

- ▶ Ellegård's hypothesis was extended by Denison (1985)
- ▶ "I am proposing four phases" (p. 55)
    1. *do* is one among many causatives
    2. *do* causatives spread at the expense of others
    3. *do* becomes an auxiliary
    4. *do* acquires its modern distribution
- ▶ We will see evidence that this articulation into stages is correct, as well as facts that provide at least circumstantial evidence in favor of the hypothesis that causatives are the origin of *do*-support

Ellegård and the corpora

▶ Ellegård (1953) had a deliberately collected corpus of
  do-support tokens; the parsed corpora of relevant time periods
  offer a sample of comparable (but smaller) size.

Ellegård	
Type	N
Aff. Decl.	7065
Aff. Imp.	77
Aff. Q.	3772
Neg. Decl.	7604
Neg. Imp.	1467
Neg. Q.	753

PPCEME+PCEEC	
Type	N
Aff. Decl.	141454
Aff. Imp.	11560
Aff. Q.	1632
Neg. Decl.	6209
Neg. Imp.	639
Neg. Q.	266

Differences between the two datasets

▶ The two corpora differ in some details, perhaps due to the deliberate collection techniques of Ellegård.
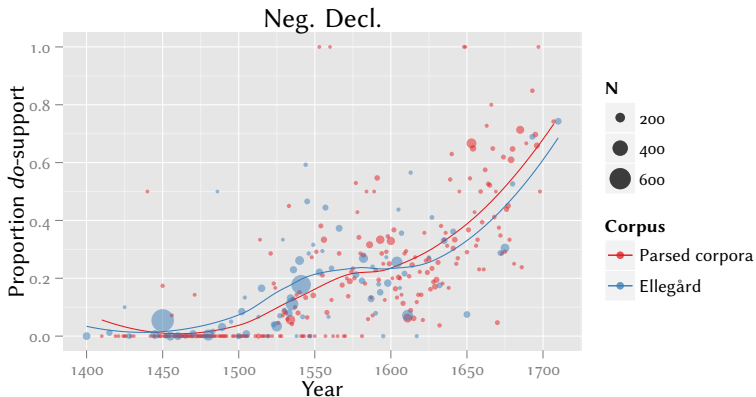
Similarities between the two corpora

- In spite of their differences, the two corpora paint the same general picture of the trajectory of *do*-support.

Similarities between the two corpora

▶ In spite of their differences, the two corpora paint the same general picture of the trajectory of *do*-support.
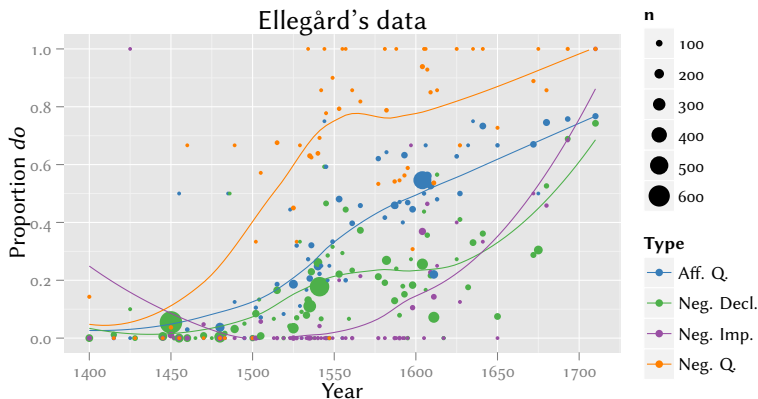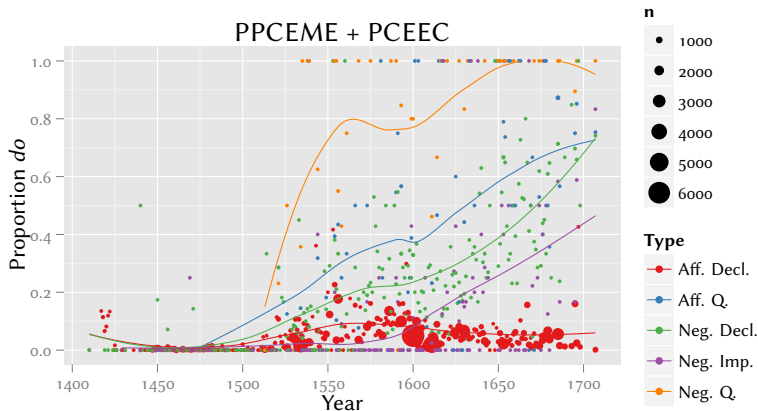
Similarities between the two corpora

▶ In spite of their differences, the two corpora paint the same
  general picture of the trajectory of *do*-support.

Waypoint

Evidence of low *do*
    Co-ocurrence with other auxiliaries
    Adverb position
    Argument structure
    Summary of evidence

Sources of evidence

Three pieces of evidence support the existence of low *do*:

- ▶ *do*'s co-occurrence with other auxiliaries
- ▶ the placement of adverbs relative to *do* and other auxiliaries
- ▶ the behavior of *do* in the absence of an external argument

Coocurrences 1

(3) Examples with duplicated causative:

    a.   He leet the feste of his nativitee
          Don cryen thurghout Sarray his citee,
          'He had the feast of his birthday cried throughout Surrey,
          his city.'
                (Chaucer *Canterbury Tales* "The Squire's Tale" c. 1400)

    b.   gret plentee of wyn þat the cristene men han don let make
          'Great plenty of wine that the Christian men have made.'
                    (PPCME2, CMMANDEV,47.1161 a. 1425)

Coocurrences 2

(4)   Example with duplicated *do*:

    a.   And thus he dide don sleen hem alle three.
            (Chaucer, *Canterbury Tales* "Summoner's Tale" c. 1400)

Coocurrences 2

(4)   Example with duplicated *do*:

    a.   And thus he dide don sleen hem alle three.
           (Chaucer, *Canterbury Tales* "Summoner's Tale" c. 1400)

- ▶ demonstrates that *do* has been bleached of its causative meaning

Coocurrences 3

(5) Example with *have*:

   a. He hes done petuously devour
     the noble Chaucer of makaris flour

     '[Death] has petuously devoured the noble Chaucer, flower
     of makars [=bards]'

              (Wm. Dunbar "Lament for the Makars" c. 1505)

(6) Example with modal:

   a. consequently it wyll do make goode drynke

     'Consequently it [barley] will make good drink'
              (A. Boorde *Introduction of Knowledge* a. 1542)

Coocurrences 3

(5) Example with *have*:

    a. He hes done petuously devour
       the noble Chaucer of makaris flour

       '[Death] has petuously devoured the noble Chaucer, flower
       of makars [=bards]'

                       (Wm. Dunbar "Lament for the Makars" c. 1505)

(6) Example with modal:

    a. consequently it wyll do make goode drynke

       'Consequently it [barley] will make good drink'

                 (A. Boorde *Introduction of Knowledge* a. 1542)

  ▶ demonstrates that *do* is merged lower than T, and lower than
     the head which hosts *have* (= Asp)

Coocurrences 4

- (7) Example in nominalization:

    a. Fro the stok ryell rysing fresche and ying
        But ony spot or macull doing spring
        'From the royal stock rising fresh and young / without any
        spot or blemish springing'
        (Wm. Dunbar *The Thrissill and the Rois* 1503, in Visser
        (1963, §1419))

Coocurrences 4

(7) Example in nominalization:

a. Fro the stok ryell rysing fresche and ying
But ony spot or macull doing spring
'From the royal stock rising fresh and young / without any
spot or blemish springing'
(Wm. Dunbar *The Thrissill and the Rois* 1503, in Visser
(1963, §1419))

▶ demonstrates that *do* is within the domain of nominalization

Adverb position

- A position between T and the subject is available to adverbs in English
- The rate of use of this position is diachronically stable (Kroch 1989)
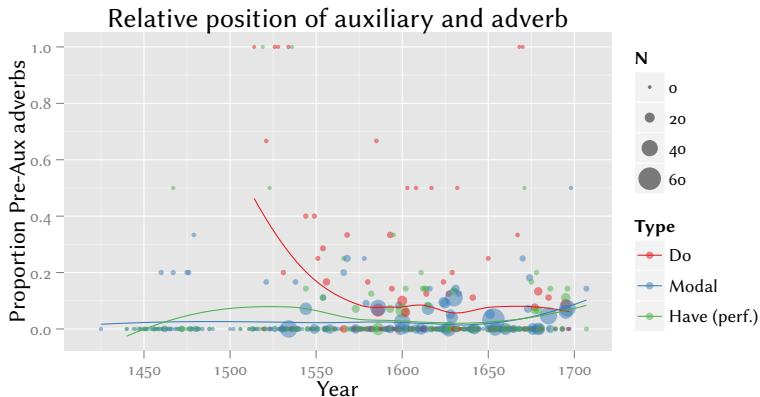
Adverb position

- ▶ A position between T and the subject is available to adverbs in English
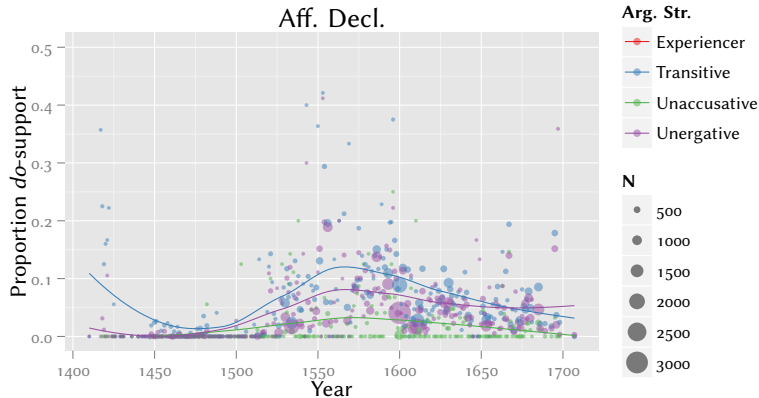- ▶ The rate of use of this position is diachronically stable (Kroch 1989)

Relative position of auxiliary and adverb
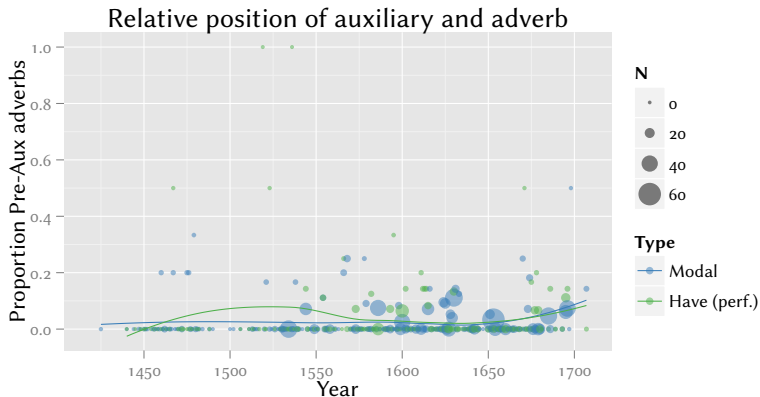
Adverbs and *do*

▶ The behavior of *do* differs from that of other auxiliaries at the beginning of the *do*-support change

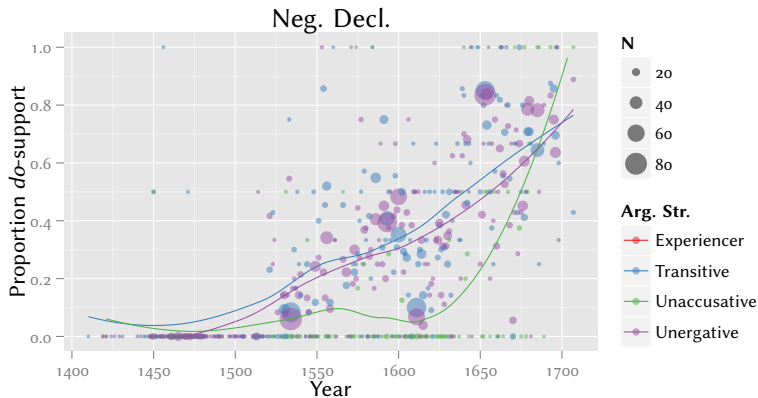Relative position of auxiliary and adverb

Reasons for suspecting argument structure

- ► Ellegard noted an argument structure-affiliated effect: for certain non-agentive verbs (the *know* class), the adoption of *do*-support is delayed.

- ► Evidence that the argument structure of the main verb affects the behavior of *do* will support the idea that the latter is located lower in the functional hierarchy, where it can be in a local relationship with the verb.
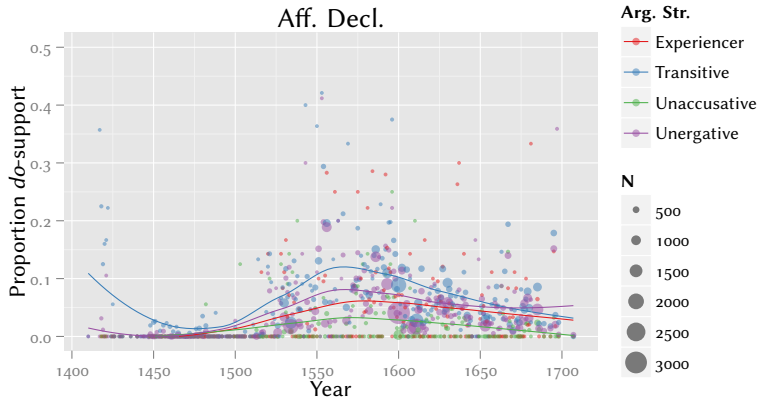
The effect of argument structure

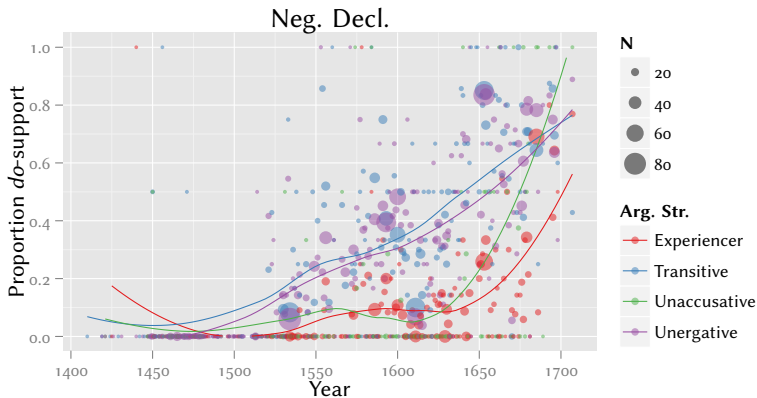# The effect of argument structure

Neg. Decl.

Refining argument structure's effect

Aff. Decl.

Arg. Str.
- Experiencer
- Transitive
- Unaccusative
- Unergative

N
- 500
- 1000
- 1500
- 2000
- 2500
- 3000

Refining argument structure's effect

Solidifying argument structure's effect

▶ To demonstrate that the effect is real, we can turn to a regression model of the data.

▶ Logistic regression model; random effects of author and main verb lemma; fixed effects of year, external argument presence, and sentence type.

	Coef.	Std. Err.	*p*-value
Intercept	-2.30	0.20	$1.539 \cdot 10^{-29}$
Year	1.28	0.21	$7.591 \cdot 10^{-10}$
Aff. Q.	0.85	0.21	$4.540 \cdot 10^{-5}$
Neg. Q.	2.43	0.32	$2.952 \cdot 10^{-14}$
No Ext. Arg.	-1.62	0.49	$8.698 \cdot 10^{-4}$

Further solidifying

- Model comparison statistics also favor the model which includes argument structure

	AIC	BIC
No arg. str. effect	1511	1545
Arg. str. effect	1502	1543

Summary

- In the early stage of the emergence of *do*-support, we have seen:
  - *do* must be an auxiliary verb
  - it must be merged lower than modals or aspectual *have*
  - it must be inside the domain of nominalization
  - It is sensitive to the presence of an external argument
- **Proposal:** *do* is first reanalyzed as an external argument marker, and later as its modern status
- This reanalysis is directly visible in the decline of *do* in affirmative declaratives, and occurs at the same time that *do* use in all contexts suffers a temporary decline

Phrase structure

- Phrase structure of English T domain (omitting irrelevant positions):
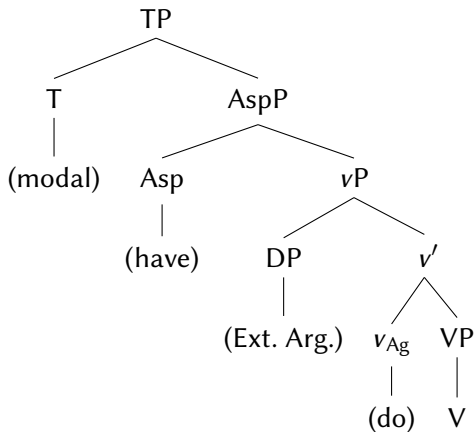
do in Germanic

- ▶ The reanalysis of *do* as an auxiliary or light verb is a common theme in Germanic languages.
  - ▶ It is commonly reported that child acquirers of German and Dutch pass through a stage in which they use *do*(/*doen*/*tun*) periphrasis regularly, perhaps to avoid inflecting "difficult" irregular verbs
  - ▶ Cornips (1998) reports on a (very modest) mixed corpus of L2 and native Dutch: "In all instances [of] the regional doen + infinitive construction, [...] the subjects are construed as agents." (p. 90)
  - ▶ In the southwest of England, there is a dialect which has affirmative declarative *do*
    - ▶ This *do* can occur with unaccusatives and with experiencer verbs (subject to worries about the imprecision of lexical semantics)

do in Germanic 2

- ▶ **So:** the reanalysis of *do* as an auxiliary verb is recapitulated in closely-related dialects, and indeed by language learners in every generation
- ▶ The association between the EME instantiation of auxiliary *do* and agentivity favors (mildly) the account that finds the origin of the construction in the ME causative system

do and grammar competition

- The logistic curve as a model of language change derives from the notion that grammars compete with each other (Kroch 1989)
- The existence of a third grammatical option necessitates a more complex model of grammar competition
- No closed form; can be fit by simulation

A 3-way model of *do*-support
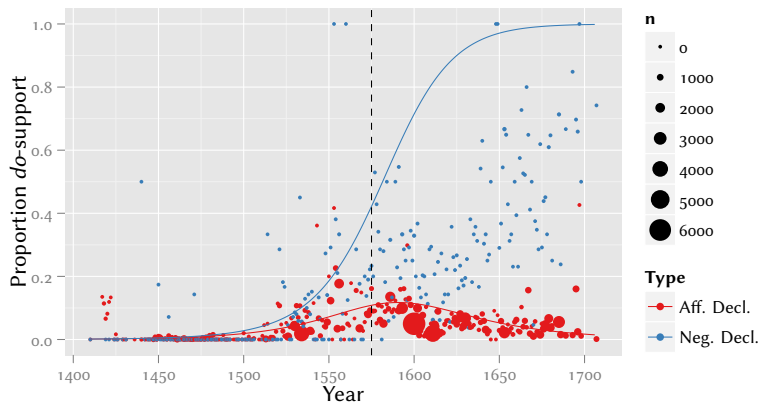
- ► Only data to 1575 is used.
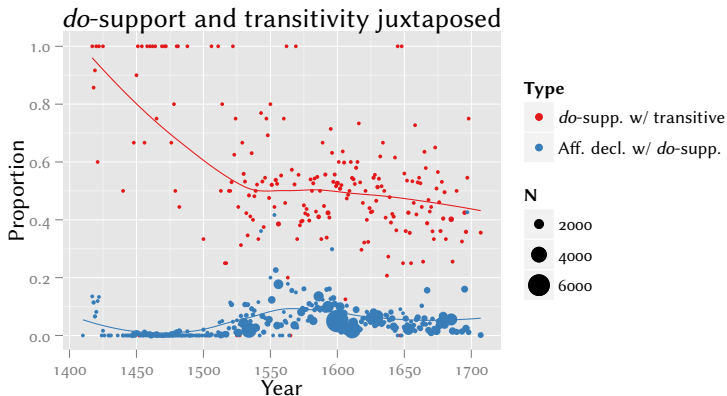- ► Model evaluation is difficult, but fit appears good

Predictions of a competition account

- ▶ Rather than just fitting the parameters of this model, it is possible to use the logical structure of the model to make predictions.
- ▶ **Gold Standard:** to derive *a priori* the model parameters from the distribution of forms in monolingual corpora
  - ▶ Unattainable goal (at present): no explanation, even on a 2-grammar model, why *do*-support wins
  - ▶ More modest predictions are possible, though

Predictions of a competition account

- ▶ Rather than just fitting the parameters of this model, it is possible to use the logical structure of the model to make predictions.
- ▶ **Gold Standard:** to derive *a priori* the model parameters from the distribution of forms in monolingual corpora
  - ▶ Unattainable goal (at present): no explanation, even on a 2-grammar model, why *do*-support wins
  - ▶ More modest predictions are possible, though
- ▶ Specifically, focusing on the intermediate grammar:
  - ▶ plausibly, only transitives (with overt subject and object) count as evidence for/against this grammar
  - ▶ the intermediate grammar can advance only if the proportion of all *do*-support sentences that are transitive is > 50%

Testing the competition model's prediction

do-support and transitivity juxtaposed

Type
- *do*-supp. w/ transitive
- Aff. decl. w/ *do*-supp.

N
- 2000
- 4000
- 6000

▶ This account discards the logical connection between the decline in affirmative and other contexts. More investigation of Warner's evidence that stylistic and social conditions are responsible for the decline is needed.

Summing up

- I have shown that, in Early Modern English, a third grammatical option exists, in addition to the verb-raising grammar of ME and the *do*-support grammar of ModE
    - This grammar has *do* as an external argument-marking auxiliary verb
    - It is related to similar parametric options that appear routinely as near misses in the learning of Germanic dialects
- The presence of this third grammar creates a conceptual challenge for the operational definition of grammar competition, viz. the logistic regression model
    - however, the new data integrate well in the conceptual scheme of grammar competition models, thereby bolstering that hypothesis

Remaining challenges

- ▶ Continue investigating the conditions that allowed *do*-support to take root in English
- ▶ Develop a model relating grammatical and social conditions in the diachrony of *do*-support

Acknowledgments

Thanks are due to the following people and groups:

- ▶ The workers on the PPCEME, PCEEC, and PPCME2 projects
- ▶ Ann Taylor, for digitizing Ellegård's corpus
- ▶ Hilary Prichard, for providing geocoding for ME texts
- ▶ Especially, Anthony Kroch

Questions?

Bibliography I

- ▶ Cornips, Leonie (1998). "Habitual *doen* in Heerlen Dutch:History and present day variation". In. *DO in English, Dutch and German*. Ed. by Ingrid Tieken-Boon van Ostade, Marijke van der Wal, and Atjan van Leuvensteijn, pp. 83–101.
- ▶ Denison, David (1985). "The origins of periphrasic do: Ellegård and Visser reconsidered". In *Papers from the 4th International Conference on English Historical Linguistics*. Ed. by Roger Eaton et al. Vol. 41. Current Issues in Linguistic Theory. John Benjamins, pp.45–60.
- ▶ Ellegård, Alvar (1953). *The auxiliary do: The establishment and regulation of its use in English*. Engelska språket.
- ▶ Kroch, Anthony (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change* **1**.(3), 199–244. ISSN: 0954-3945.

Bibliography II

- ▶ Postma, Gertjan (2010). "The impact of failed changes". In *Continuity and change in grammar*. Ed. by Anne Breitbarth et al. Vol. 159. Linguistik Aktuell. John Benjamins Publishing Company.
- ▶ Visser, F. T. (1963). *An historical syntax of the English language.* E. J. Brill.
- ▶ Warner, Anthony (2005). Why DO dove: Evidence for register variation in Early Modern English negatives. *Language Variation and Change* **17**.(3), 257–280.
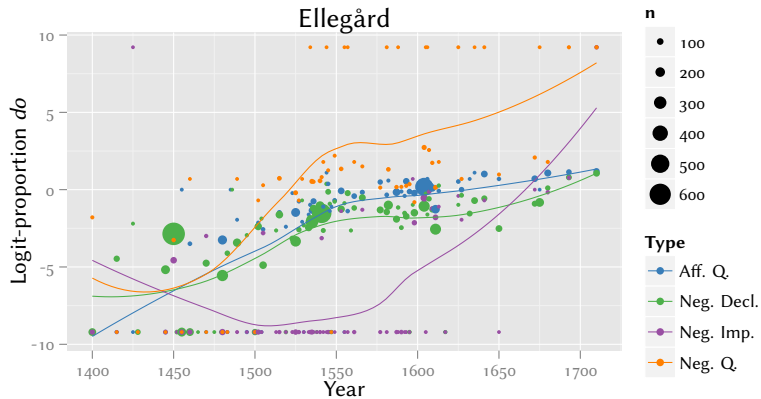
Never and *do*-support

Logit transformed plots

do and never in the parsed corpora
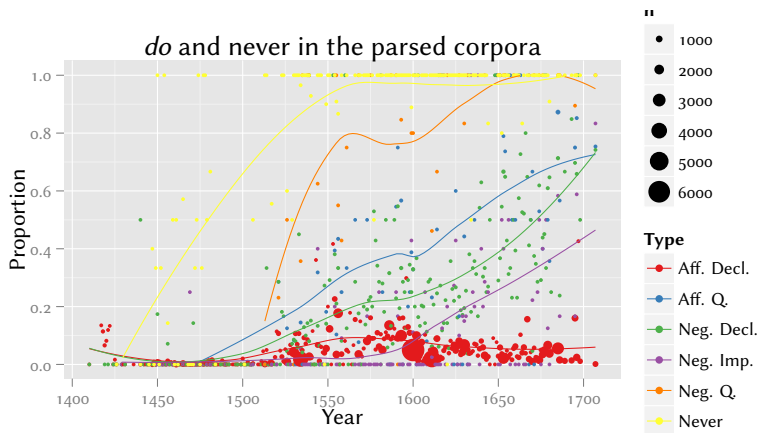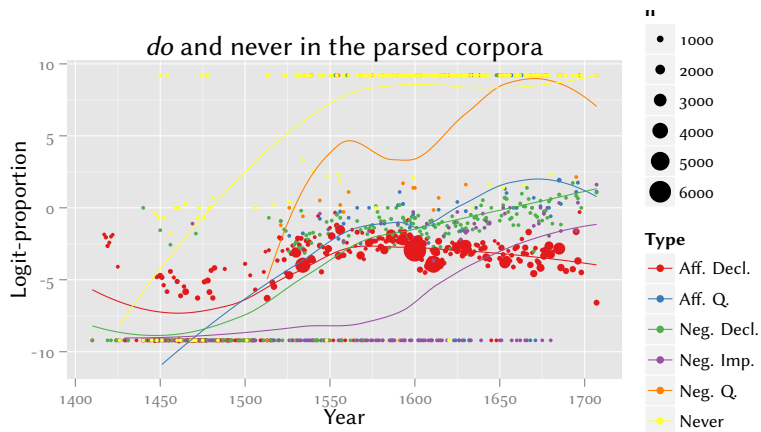
Logit transformed plots

Failed changes?

- ▶ Postma (2010) proposes a model of affirmative declarative do that posits that it is a "failed change"
- ▶ The mathematical relationship between the evolution of a failed change and that of its successful counterpart is that the former is the first derivative of the latter
- ▶ This means that every token of affirmative declarative *do*-support is reinterpreted as a token of modern *do*-support
- ▶ Problems the model faces:
  - ▶ why would speakers (learners) be so grossly misled?
  - ▶ the interpretation of a derivative is scale-dependent