Analyzing V2 Triggers in Historical English

Aaron Ecay and Caitlin Light

University of Pennsylvania

June 5, 2011

Introduction

- This talk will present new quantitative data on the behavior of a class of adverbs well known to researchers of historical English, which are known to trigger V2 word order.
- We will compare two analyses of the syntax of these elements, and present evidence in favor of the earlier claim that they are operators in Spec,CP which trigger verb movement to C.
- Based on details of the "loss" of these V2 triggers and data that points towards their continued existence in restricted contexts in Modern English, we will propose a new theory to explain their diachronic behavior.

Table of contents

V2 patterns in Old English

- In Old English, topicalization triggers subject-verb inversion with a full DP subject, but not a pronoun. (cf. van Kemenade, 1987; Pintzuk, 1991)
  - In clauses with a pronominal subject, the subject will appear between a topicalized XP and the finite verb.
  - This has been analyzed as evidence that the verb moves only as high as T in OE matrix clauses, and that a landing site between T and Spec,CP is available to subject pronouns.

V2 patterns in Old English

- In this paper, we reconsider the syntactic properties of a well-known class of adverbs which present an exception.
  - These elements (mainly *þa/þonne* ('then'), *swa* ('so') and *nu* ('now'), trigger subject-verb inversion for *all* subjects.
  - Since Pintzuk (1991), these are generally analyzed as operators in Spec,CP, which trigger V-to-C movement (this aligns them with the general analysis of V2 orders in i.e. German).

An alternate analysis for *þa*

- However, Trips and Fuß (2009, henceforth T&F) proposes that elements like *þa* are not in Spec,CP, but in fact occupy Spec,TP.
  - Under this analysis, subject pronouns are in competition with *þa* to fill the high subject position.
  - Thus, when *þa* fills Spec,TP, the subject is forced to remain lower in the clause.

Þa according to T&F

- T&F's analysis hinges on the claim that OE is "discourse-configurational," unlike modern English.
  - Spec,TP is not filled to satisfy the EPP, but to license discourse-anaphoric and deictic elements.
  - Both personal pronouns and anaphoric adverbs like *þa* are discourse-anaphoric in this sense, and thus both target Spec,TP in the proposed system.

Motivations for T&F's analysis

- T&F outline several reasons why this new analysis would be appealing:
  - These discourse adverbs and personal pronouns form a natural class (which T& F define as the feature *[+anaphoric]*).
  - When these adverbial elements are analyzed as operators in Spec,CP, it becomes difficult to explain why they no longer trigger subject-verb inversion in Modern English, while other operators (like *wh*-operators) still do.
    - Their proposal, in contrast, provides a clean explanation for this change: as T gains an EPP feature in English (and thus no longer serves to license the *[+anaphoric]* feature), *þa* no longer competes with the subject for Spec,TP.

Motivations for T&F's analysis

- T&F outline several reasons why this new analysis would be appealing:
  - Placing *þa*-type adverbs in Spec,CP should predict the absence of V3 word orders where a topicalized XP precedes *þa*:

(1) a. Him **þa** andswarode se biscop.
him then answered the bishop
'Then, the bishop answered him.'
(GD_1_[C]:4.28.5.293, from Trips and Fuß, 2009)

b. For þi **þonne** wacion we
for that then stay-awake we
'because then we stay awake'
(ChrodR_1:14.6.277, from Trips and Fuß, 2009)

Continuing the debate

- We will show, however, that the analysis proposed by Trips and Fuß (2009) does not account for the full range of data.
  - Both diachronic and synchronic studies prove problematic for T&F's account.
  - A number of the predictions entailed by their analysis are contradicted by the data we will present.
- In contrast, the traditional analysis introduced by Pintzuk (1991) handles this new data well.
  - As a result, we will argue that *þa*-type adverbs are operators in Spec,CP, following Pintzuk (and *contra* Trips & Fuss).

Þa and V_4 word orders

- Because T&F claim that *Þa*-type adverbs and and subject pronouns are competing for the same position, they predict the absence of V_4 word orders of the following types:
  1. XP_{top} - $Subj_{pro}$ - *Þa* - V_{fin}
  2. XP_{top} - *Þa* - $Subj_{pro}$ - V_{fin}
- Rather, as long as preverbal *Þa* is present, any (non-topicalized) subject pronoun should occur below the finite verb.

- However, examples with these word order configurations are attested, with a V2-triggering adverb appearing clause-initially and *þa/þonne* following the subject pronoun:

(2) **Nu** he **þonne** costode Godes Sunu þurh idel wuldor
now he then tempted god's son through vain glory

Now, he then tempted God's son through vain glory.
(Blickling Homilies, 29.54.395)

(3) **Swa** he **þa** wæs onhyrigende on þam twam wundrum
so he then was imitating on the two wonders
twegra fædera mægnu
two fathers' mights

Thus he then was imitating two fathers' mights in the two wonders.
(Gregory's Dialogues, Cambridge ms., 7.49.27.565)

Þa and V_4 word orders

- In these examples, we must assume that *nu* and *swa* are in Spec,CP; however, these adverbs are known to generally behave as V2 triggers, just like *þa*. It is thus undesirable to place them in Spec,CP but treat *þa* as a separate case.
- Consider also the following example:

(4) Æfter þissum hi **þa** geweredon
After this they then defended

After this, they then defended ...
(Bede's History of the English Church, 12.52.19.480)

- Here, again, discourse-anaphoric *þa* and the subject pronoun *hi* co-occur above the finite-verb, with a topicalized PP preceding both.

Embedded contexts in T&F

- T&F note that embedded clauses seem to challenge their analysis; subject pronouns tend to precede *þa*:

  (5) Þa hi **þa** hine geornlice beheoldon …
  when they then them carefully beheld …
  'when they then carefully beheld him …
  (van Kemenade and Los, 2006, eust, `LS_8_[Eust]:270.286`; from)

- They suggest that the temporal properties of temporal adverbs like *þa* may be computed from the properties of the matrix clause, such that *þa* need not occupy SpecTP.
- Thus, for T&F, subordinate clauses are a challenge, but can be accounted for.

Þa in embedded contexts

- What T&F *cannot* explain is the absence of clauses with the order *þa* - V_{fin} - $Subj_{pro}$.
  - Although at least some subordinate clauses may have different tense properties than matrix clauses, we expect that *þa* will occupy Spec,TP in some embedded clauses.
  - Alternatively, such an order may be explained much like that in (5), with a verb-first subordinate clause in which *þa* has adjoined high.
- This absence is, however, predicted by the alternate analysis, because if *þa* is an operator in Spec,CP, V2 orders with *þa* should be blocked when C is occupied.

A telling gap in the data

- We considered "strict" V2 orders with only a *þa*-type adverb preceding the finite verb, in subordinate clauses with an overt complementizer.
  - There was only one example with this order in the Old English corpus, and only three in Middle English.
  - All potential examples involved a copular clause which was clearly specificational.

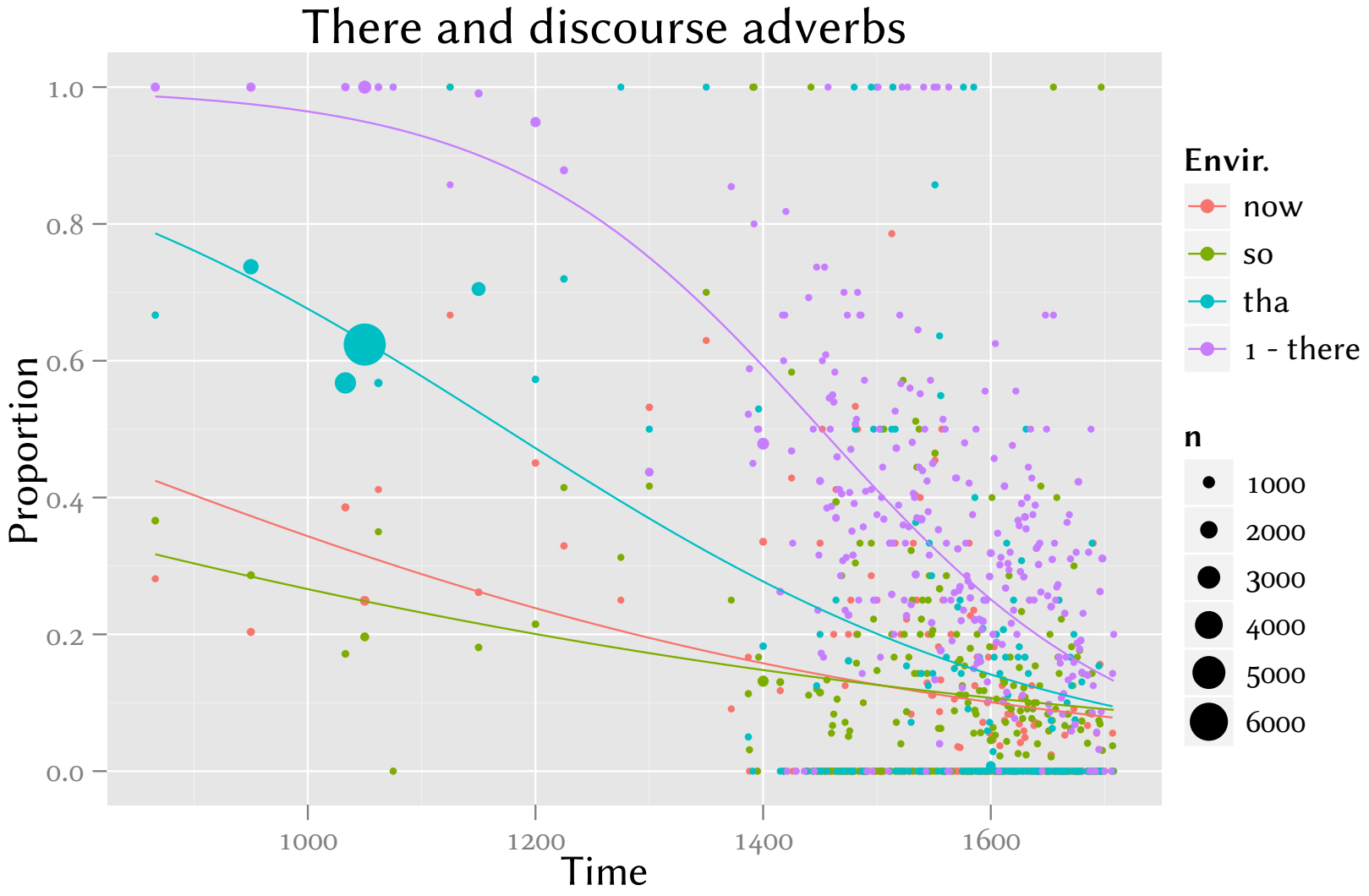
    (6) ...þinking þat **now** was tyme for to entir
    'Thinking that it was now time to enter'
    (Capgrave's Chronicle, 211.3787)

  - The order *þa* - V_{fin} - $Subj_{pro}$ is otherwise unattested.
- This gap in the data is inexplicable under T&F's analysis.

Þa's relation to the EPP

- As previously mentioned, an advantage of the analysis advanced by T&F is that it provides a clear explanation for the loss of V2 orders with *Þa*.
  - Discourse-anaphoric adverbs like *Þa* are expected to occupy Spec,TP only so long as English remains "discourse-configurational," and thus once an EPP feature is associated with T, *Þa* will no longer be licensed in Spec,TP.
  - They therefore make the strong claim that the appearance of the EPP feature in English syntax was the direct cause of the loss of V2 with *Þa*.
- This hypothesis is easily tested, by comparing the declining rate of V2 with *Þa* to the rise of expletive *there* in English (a direct consequence of the appearance of an EPP feature on T).

Þa and expletive *there*



Modern English

- In Modern English, V2 word order is required with *so* in certain contexts:

(7) John kicked the ball...

a. ✓ ...and so did Bill.

b. * ...and Bill so did.

c. * ...and Bill did so.
(on intended reading)

(8) Fluoridated water is a Soviet plot to harm American children.

a. ✓ Yeah, right. So says *John*.

b. * Yeah, right. *John* so says.

c. ?? Yeah, right. *John* says so.

d. ? Yeah, right. So *John* says.

Modern English

- In even more limited circumstances, V2 constructions with *then* appear:

  (9) The engine comes first...

  a. ✓ ...then come the train cars.

  b. * ...then the train cars come.

- Note: this examples appear to involve direct V $\rightarrow$ C, by contrast to other attested English inversion constructions. (cf. Swedish matrix clauses, but this pattern may not be general across English verbs)

Modern English

- In even more limited circumstances, V2 constructions with *then* appear:

(10) The engine comes first...

a. ✓ ...then come the train cars.

b. * ...then the train cars come.

- Note: this examples appear to involve direct V $\rightarrow$ C, by contrast to other attested English inversion constructions. (cf. Swedish matrix clauses, but this pattern may not be general across English verbs)
- ...so what does it mean to say that V2 with *so* and *then* was lost in Old English?

Diachronic data

- Using data from the following corpora, it is possible to plot the evolution of V2 with these lexical items:
  - York-Toronto-Helsinki Corpus of Old English (Taylor et al., 2003)
  - Penn-Helsinki Parsed Corpus of Middle English (Kroch and Taylor, 2001)
  - Parsed Corpus of Early English Correspondence (Taylor et al., 2006)
  - Penn-Helsinki Parsed Corpus of Early Modern English (Kroch, Santorini, and Delfs, 2005)

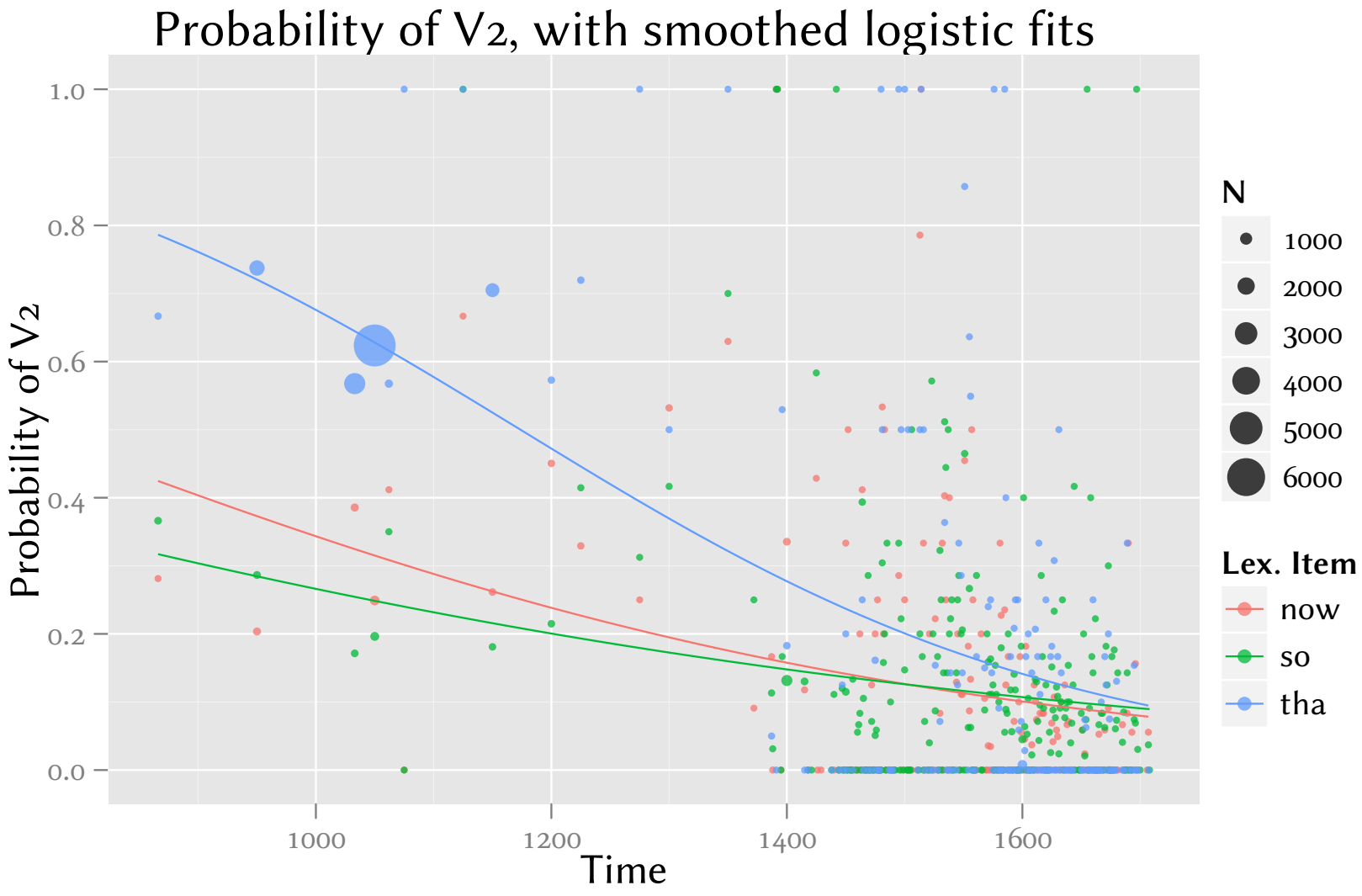
Coding convention

- Problem: how to define the envelope of variation?
- T&F attempt to differentiate between two positions for *þa*, but:
  - As we have seen, other aspects of their account are problematic.
  - Many tokens of *þa* will be perfectly ambiguous between the two putative readings.
  - It is not clear how to extend the theory to *now* and *so*.
- For coding purposes, we used the following definition of V2:

  (11) Potential instances of V2 for a given adverb are sentences where that adverb precedes the tensed verb. Actual instances of V2 are potential instances where the tensed verb additionally precedes the subject.

- It may have problems with "low" *þa* in Infl-final clauses, but it must be approaching correctness, as early attestations of V2 with *þa* are over 70%.

Diachronic data



Diachronic examples

(12) but if men be kind vnto them, and be in their habit; then are they conquered with kindnesse, (JOTAYLOR-E2-H,1,135.C2.217; 1630)

(13) then came the King of Sweth with 38 thousand, and the Duke of Saxe with 12 thousand, whoe led the vanguard and gave the onsett upon Tilly; (CORNWAL,240.148.2057; 1631)

(14) and now have I on a payr of english winter stockins, at which Gondomar iested, (HOLLES,III,489.133.3768; 1636)

(15) Soe is that of Livie expounded by learned men, when in the passage of Annibal over the Alpes hee sayth Rupem muniendam curavit, that is, hee opened a passage thorough the Rock, (BROWNE,307.058.1118; 1658)

- Many (but not all) of the latest examples in our corpus have an auxiliary verb moving from T $\rightarrow$ C, indicating perhaps that we are seeing the tail end of the truly productive use of these adverbs as V2 triggers in a grammar which is in the process of losing V $\rightarrow$ T.

Regression

- We can fit a logistic regression to the data, to check if the individual lexical items matter:

Intercept	2.573×10^{00}
	p: 5.580×10^{-129}
Adverb: *so*/*now*	-1.743×10^{00}
	p: 1.202×10^{-27}
Adverb: *tha*/*now*	2.395×10^{00}
	p: 1.220×10^{-76}
Time	-2.882×10^{-03}
	p: 4.600×10^{-276}
Adverb: *so*/*now* × Time	1.038×10^{-03}
	p: 7.335×10^{-19}
adverb: *tha*/*now* × Time	-1.351×10^{-03}
	p: 6.501×10^{-38}

Regression

- We can fit a logistic regression to the data, to check if the individual lexical items matter:

Intercept	2.573×10^{00}
	p: 5.580×10^{-129}
Adverb: *so*/*now*	-1.743×10^{00}
	p: 1.202×10^{-27}
Adverb: *tha*/*now*	2.395×10^{00}
	p: 1.220×10^{-76}
Time	-2.882×10^{-03}
	p: 4.600×10^{-276}
Adverb: *so*/*now* × Time	1.038×10^{-03}
	p: 7.335×10^{-19}
adverb: *tha*/*now* × Time	-1.351×10^{-03}
	p: 6.501×10^{-38}

Regression

- We can fit a logistic regression to the data, to check if the individual lexical items matter:

Intercept	2.573×10^{00}
	p: 5.580×10^{-129}
Adverb: *so*/*now*	-1.743×10^{00}
	p: 1.202×10^{-27}
Adverb: *tha*/*now*	2.395×10^{00}
	p: 1.220×10^{-76}
Time	-2.882×10^{-03}
	p: 4.600×10^{-276}
Adverb: *so*/*now* × Time	1.038×10^{-03}
	p: 7.335×10^{-19}
adverb: *tha*/*now* × Time	-1.351×10^{-03}
	p: 6.501×10^{-38}

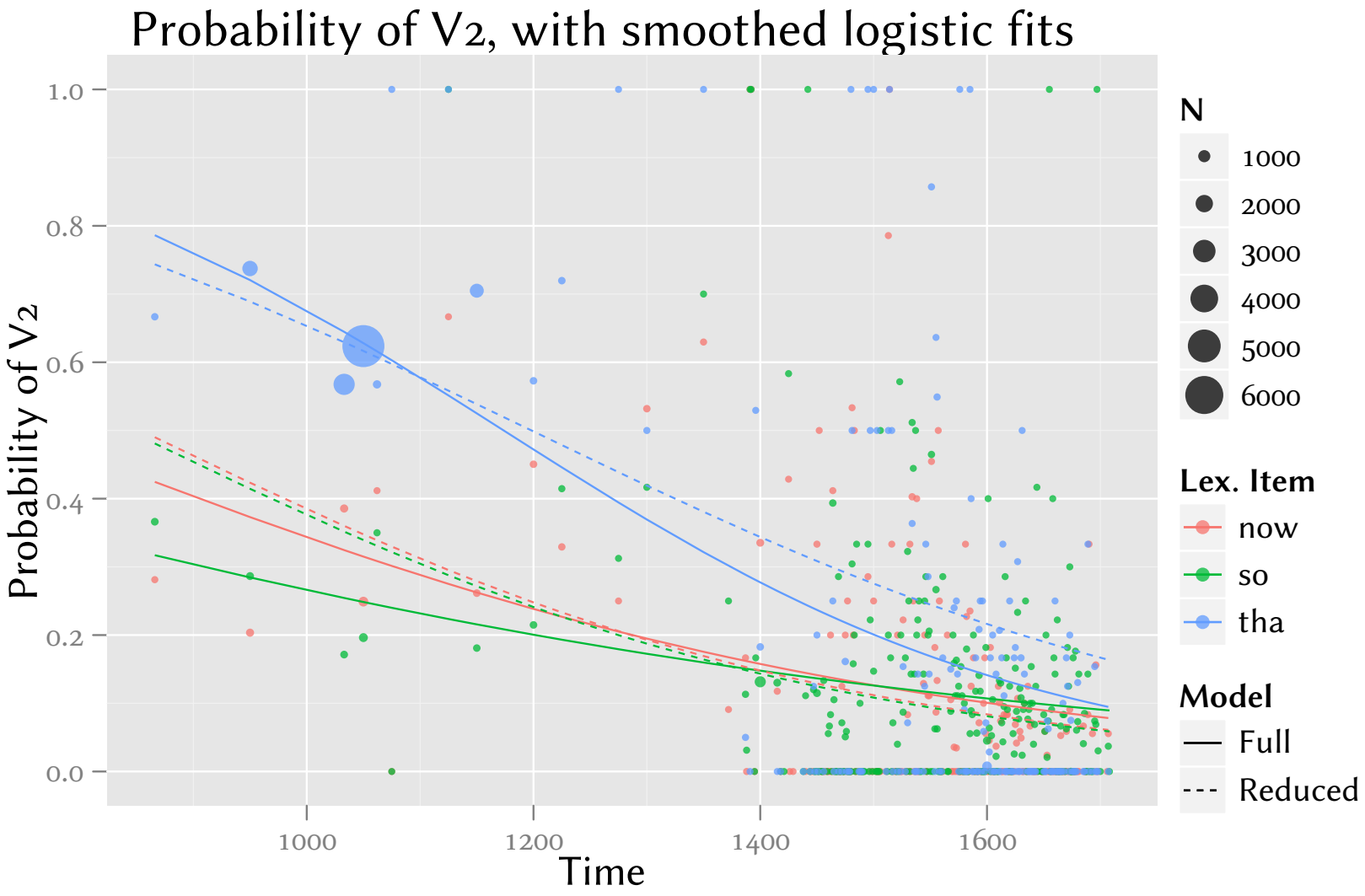
A simpler model

- Using an ANOVA procedure, we can compare this model to one which fits a common slope to all grammatical contexts:

Resid. Df	Resid. Dev	Df	Deviance	P(>|χ|)
26144	28076.42			
26142	27884.51	2	191.91	0.0000

- As can be seen from the small p-value, the more complicated model (with different slopes for each lexical item) offers a substantial reduction in deviance.

Plotting the models



A new proposal

- T&F correctly raise the puzzle of why *then*, *so*, and *now* lose their status as V2 triggers in the history of English, while other triggers (e.g. *wh*-questions) do not.
- There is evidence that strongly suggests that their solution to the problem is not correct.
- Given the diachronic data and the survival of V2 with *so* in certain contexts to the present day, we propose a new approach to this change.

A new proposal

- Specifically, we propose that the change in status of *then*, *now*, and *so* is not a single syntactic change.
- Rather, it is composed of three independent lexical changes.
  - This explains the logistic regression results.
  - It also explains the "splintering" of *so* and *then*, which have survived robustly as a V2 trigger in certain narrowly circumscribed environments.
- This theory does not posit a temporal relationship between the three lexical changes, nor explain why they occur when other V2 constructions in Old English are being lost.
  - Perhaps this is an instance of a syntactic change (loss of IP-V2) influencing a set of lexical changes?

Thanks

- Thanks are due to the following individuals and organizations:
  - The authors of the corpora used in this study
  - Tony Kroch
  - The University of Pennsylvania

Questions?

Bibliography I

- Van Kemenade, A. and B. Los (2006). “Discourse adverbs and clausal syntax in Old and Middle English”. In:
- Van Kemenade, Ans (1987). *Syntactic case and morphological case in the history of English*. Holland: Forris Publications.
- Kroch, A., B. Santorini, and L. Delfs (2005). *Penn-Helsinki parsed corpus of Early Modern English*. University of Pennsylvania. http://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-1/.

Bibliography II

- Kroch, Anthony and Ann Taylor (2001). *The Penn-Helsinki parsed corpus of Middle English*. CorpusSearch; National Science Foundation (US); University of Pennsylvania Research Foundation. http://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-3/index.html.
- Pintzuk, Susan (1991). "Phrase structures in competition: Variation and change in Old English word order". PhD thesis. University of Pennsylvania.
- Taylor, A. et al. (2006). *Parsed Corpus of Early English Correspondence, parsed version.* Compiled by the CEEC Project Team. York: University of York and Helsinki: University of Helsinki. Distributed through the Oxford Text Archive. http://www-users.york.ac.uk/~lang22/PCEEC-manual/index.htm.

Bibliography III

- Taylor, Ann et al. (2003). *The York-Toronto-Helsinki Parsed Corpus of Old English Prose*. http://www-users.york.ac.uk/~lang22/YcoeHome1.htm.
- Trips, C. and E. Fuß (2009). "The syntax and semantics of the temporal anaphor then in Old and Middle English". In: *Advances in comparative Germanic syntax*, p. 171.