Chain shifts in Syntax:
On the replacement of *th-* with *wh*-elements in Middle English

Richard Zimmermann
Université de Genève
ICHL 22

July 30, 2015

1 Introduction

- Old English frequently used the *th*-adverbs *þa* and *þonne* 'then' as subordinators in temporal adverbial clauses

(1) a. ... ðæt hit wære geendodu spæc **ða** se dom   wæs gefylled.
        ... that it were ended   speech then the judgment was fulfilled

     '... that the law suit would be closed when the judgment was carried out'
     (codocu2,Ch_1445_[HarmD_18]:41.62)

   b. Se hyra   flyhð: **þonne** he þone wulf gesihð.
      the hireling flees   then   he the   wolf sees

      'The false shepherd flees when he sees the wolf'
      (cocathom1,ÆCHom_I,_17:314.37.3108)

- In early Middle English, these *th*-based adverbs were increasingly replaced by the *wh*-item *when* in this function (Declerck 1997, 58-63)

- Transitional period with free variation

(2) a. Olibrius þe luðere, **þa**   he þis iherde, changede his chere.
        Olibrius the wicked when he this heard   changed his expression

        'Olibirus the wicked changed his expression when he heard this
        (CMMARGA,58.55, c. 1225 A.D.)

   b. Sothely, þise   wordes, **when** I here thaym or redis þam, stonyes me
      truly,   these words, when I hear them or read them, stupefy me

      'Truly, these words stupefy me when I hear them or read them'
      (CMROLLTR,45.918, c. 1345 A.D.)

- Outline of the talk:

  - Provide quantitative data on the *th-wh* variation
  - Formalize the interpretation of *then* and *when* as main clause adverbs or subordinators
  - Characterize the hypothesis that the rise in *wh*-subordinators is related to word order developments in terms of a chain shift
  - Test this hypothesis empirically

2   A few words on *þo* vs. *þan*

- Middle English had two different words for the temporal *th*-subordinator:
  - *þo* (<OE *þa*)
  - *þan* (<OE *þanne*)

- The differences between these two forms are multi-faceted
  - 1. Dialect differences
    * In Southern and West Midlands texts of thirteenth century, *þan* is extremely uncommon (e.g. *King Horn*, D-Version of *Poema Morale*, Katherine Group, Robert of Gloucester's *Chronicle*) (Kivimaa 1966)
    * In general, however, dialect differences are difficult to distinguish from random text effects
  - 2. Semantic differences
    * Old English:
      "Loosely speaking, we can say that *þonne* and *þa* are synonymous with each other [...]. With more semantic precision, however, it must be asserted that there is a fairly clear distinction in general between these two particles. *þonne* as a conjunction [...] is used when the time of an action or occurrence is indefinite and general or it is to be habitually repeated, and is usually found with a [...] verb in the present tense [...]. On the other hand, *þa* as a conjunction [...] is used when the narrator is going to describe a definite action or occurrence confined to a particular point of time, and is most commonly found with a [...] verb in the past tense" (Yamakawa 1969: 11)
    * This is true in Middle English too:
      all possible forms of temporal *th*-subordinators listed as a set of strings in a definition file:
      `THEN: +ta|+Ta|+da|+Da|+tonne|+Tonne|+donne|+Denne|+denne|+Donne| ...`
      operationalization of Middle English *þo* and *þan* as *o*-based vs. *n*-based words:

```
(P idoms THEN)                    (P idoms THEN)
AND (P idoms *o*)                 AND (P idoms *n*)
AND (P idoms !*n*)
```

Tense	*þo*	*þan*
present	2	109
past	210	54

χ^2=189.05, df = 1, $p<0.001$***

Table 1: Tense in embedded clauses by temporal subordinator form

- Even though dialect and semantic differences are discernable, all instances of Middle English *th*-based temporal subordinators are grouped together in the remainder of the paper,

- Both *þo* and *þan* compete with, and are eventually replaced by, *when*

3 Diachronic Overview

3.1 Methodology

- Initial investigation of Middle English temporal subordinate clauses:

  1. dependent variable: realization of subordinator as *th-* or *wh-*item

  2. only independent variable: time; every text has been assigned a specific year

- Database used: PPCME2 (Kroch and Taylor 2000)

- Subordinators listed as a set of strings in a definition file:
  THEN: +ta|+Ta|+da|+Da|+tonne|+Tonne|+donne|+Denne|+denne|+Donne| ...
  WHEN: Hw*n*|hw*n|W*n*|w*n*|H*n*|h*n*|Q*n*|q*n|Uu*n*|uu*n*|$Hw*n*| ...

- annotation of relevant structures in PPCME2:

```
/~*
So+dfestnesse me ku+d him solue; hwenne he his sunnen unde+d to his proste.
alswa alse he heom haue+d idon
(CMLAMB1,153.423)
*~/
/*
17 PP-SUB:  17 PP-SUB, 18 P, 20 CP-ADV, 19 hwenne
*/

(0  (1 IP-MAT (2 NP-OB1 (3 N So+dfestnesse))
             (5 NP-SBJ (6 MAN me))
             (8 VBP ku+d)
             (10 NP-OB2 (11 PRO him) (13 N solue))
             (15 . :)
             (17 PP (18 P hwenne)
                    (20 CP-ADV (21 C 0)
                               (23 IP-SUB (24 NP-SBJ (25 PRO he))
                                          (27 NP-OB1 (28 PRO$ his) (30 NS sunnen))
                                          (32 VBP unde+d)
                                          (34 PP (35 P to)
                                                 (37 NP (38 PRO$ his) (40 N proste)))
                                          (42 , .)
                                          (44 PP (45 ADV alswa)
                                                 (47 P alse)
                                                 (49 CP-ADV (50 WADVP-1 0)
                                                            (52 C 0)
                                                            (54 IP-SUB (55 ADVP *T*-1)
                                                                       (57 NP-SBJ (58 PRO he))
                                                                       (60 NP-OB1 (61 PRO heom))
                                                                       (63 HVP haue+d)
                                                                       (65 DON idon)))))))))
    (67 ID CMLAMB1,153.423))
```

Figure 1: A temporal adverbial clause in the PPCME2

- Search queries to find relevant temporal adverbial clauses:

```
node:  PP*                          node:  PP*
query:  (PP* idoms P)               query:  (PP* idoms P)
AND (PP* idoms CP-ADV*)             AND (PP* idoms CP-ADV*)
AND (P idoms THEN)                  AND (P idoms WHEN)
```

3.2 Results

- Data set consists of N=4,006 examples (348 *then*, 3,658 *when*)
- Resulting data set for 56 Middle English prose texts from the PPCME2 looks like this:

Text	Year	THEN	WHEN		Text	Year	THEN	WHEN
Elucidarius	1110	5	0		John Trevisia's Polychronicon	1387	0	123
In Festis Sancti Marie	1110	13	1		Purvey's Prologue to Wycliffe	1388	1	129
ASC Cont1	1131	6	0		Julian of Norwich's Revelations	1390	0	8
ASC Cont2	1154	14	0		Chaucer Treatise Astrolabe	1391	0	15
Lambeth Homilies(without OE copies)	1160	18	21		Equatorie of the Planets	1392	0	16
Trinity Homilies(5 sermons removed)	1160	138	8		Chaucer Parson's Tale	1395	1	134
Vices and Virtues	1180	57	3		Chaucer Tale of Melibee	1395	0	69
WooingGroup	1205	3	12		Cloud of Unknowing	1395	0	48
Ancrene Riwle	1215	18	158		Hilton's Eight Chapters Perfection	1396	0	29
Hali Meidhad	1225	5	20		Book of Vices and Virtues	1400	0	41
Saint Juliana	1225	6	7		The Chronicles of England	1400	2	294
Saint Katherine	1225	7	6		Liber de Diversis Medicinis	1410	0	26
Saint Margaret	1225	5	9		Mirk's Festial	1410	1	388
Sawles Warde	1225	1	5		Northern Rule St. Benet	1415	11	66
Kentish Sermons	1250	7	9		Sermons from MS. Royal	1425	0	18
Ayenbite of Inwyt	1340	19	204		Treatise Horses	1425	0	26
Richard Rolle Epistles	1345	1	76		Rievaulx De Institutione(Bodley)	1430	0	25
Richard Rolle Treatises	1345	0	65		Book of Margery Kempe	1435	4	284
Earliest Prose Psalter	1350	0	32		Life of St Edmund	1438	0	13
Dan Jon Gaytryge's Sermon	1357	0	12		Capgrave's Sermon	1452	0	2
Mirror of Edmund(Thornton)	1360	0	74		Capgrave's Chronicle	1460	0	148
Travels of Sir Mandeville	1371	0	216		Gregory's Chronicle	1467	0	33
Mirror of Edmund(Vernon)	1375	0	83		Malory's Morte Darthur	1470	1	222
Rievaulx De Institutione(Vernon)	1375	1	80		Reynes' Commonplace Book	1475	0	16
Chaucer Boethius Translation	1380	0	35		Caxton's Reynard the Fox	1481	1	24
Wycliffe New Testament	1383	0	53		Fitzjames' Sermo die Lune	1495	0	4
Wycliffe Old Testament	1383	0	42		In Die Innocencium	1497	0	8
Wycliffe Sermons	1383	2	189		Siege of Jerusalem	1500	0	29

Figure 2: Variation in temporal subordinators in 56 Middle English prose texts

- Regression analysis:
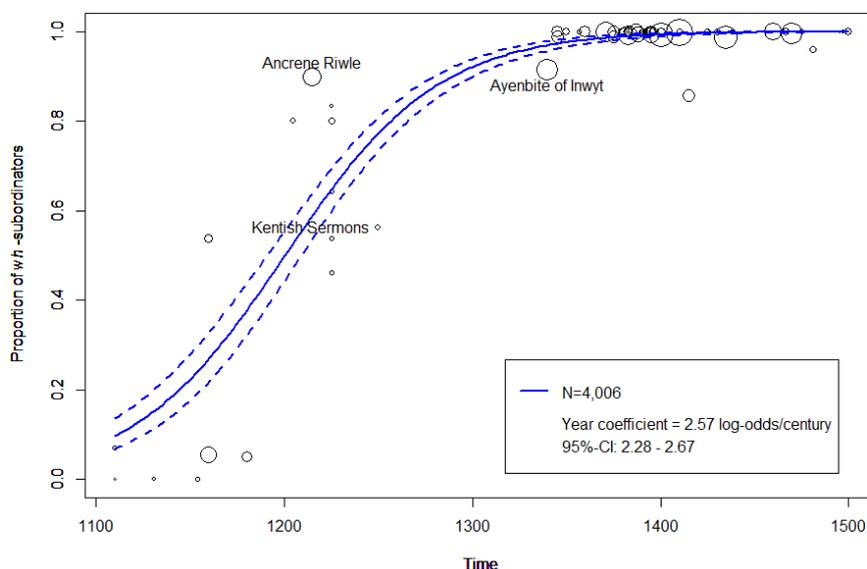
Figure 3: Graph and regression of temporal data

- * There is a coherent increase in *wh*-items during the course of Middle English
- * Main part of the transitional period for this change between 1200 and 1350
  (last examples in very early fifteenth century)
- * The temporal change is quite fast (change in c. 2.5 log-odds per century)
- * Overall, time is an excellent predictor for the realization of the subordinator as a *th* or *wh* item
  (Nagelkerke Pseudo $R^2 \approx 0.8$)

- Unfortunately, precisely the period during which the change takes place is very poorly attested in the PPCME2 (1250-1350, Helsinki M2)

- To fill the prose gap, data from a new corpus will be added to the data set

4 The Parsed Corpus of Middle English Poetry

- The *Parsed Corpus of Middle English Poetry* (PCMEP) is a new corpus of the CorpusSearch family of historical corpora

- Its annotations follow exactly the PPCME2 guidelines
  → compatibility with Penn-Parsed series of corpora
  → researchers do not have to learn a new manual

- It currently includes 37 fully parsed and annotated Middle English verse texts totaling, 105,915 words

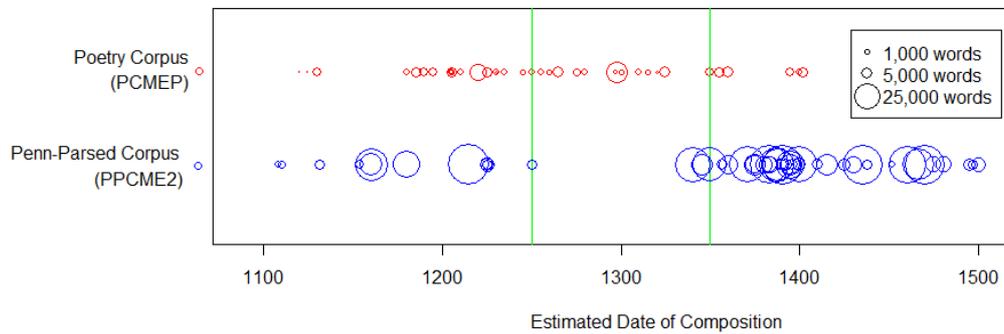- The PCMEP focuses on the period 1250-1350 (Helsinki M2) to close the substantial gap in prose texts at that time

Figure 4: Comparison between PCMEP and PPCME2

- Some well known texts: *Poema Morale, The Owl and the Nightingale, Havelok the Dane*

- Philological information on genre, dialect, manuscript date etc. are available for each text

The corpus can be downloaded for free at: `http://www.pcmep.net`

5   Using the PCMEP data

- Addition of the PCMEP texts (and small number of PPCME2 verse texts) to the previous data set

- Same methodology as before; same search queries

- Resullt:

  – Plot of subordinator type by time in all Middle English texts that are now available:
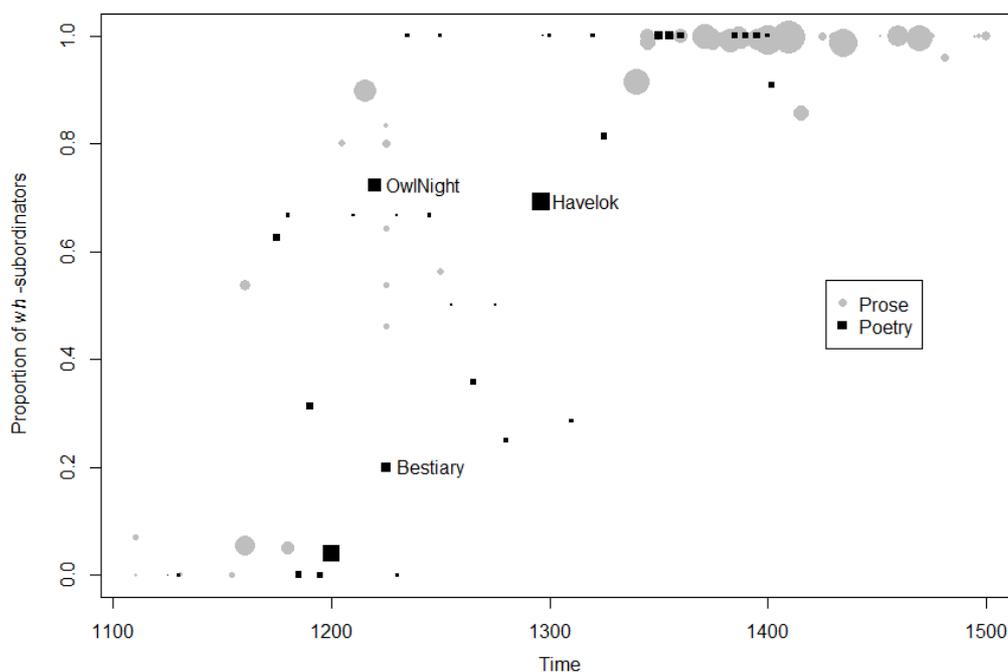
Figure 5: Temporal subordinators in prose and poetry texts

  – Poetry texts enrich the data

    * period 1250-1350 is covered with considerably more detail
    * the text gap is not as pronounced as before
    * variation in subordinator form in the poetic record is roughly where it would be expected
    * hence, poetry can inform the change under investigation and possibly many other syntactic changes as well

  – However, there is a difference between prose and poetry texts; prose texts are significantly more innovative:

```
Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -3.049e+01  1.107e+00 -27.542   <2e-16 ***
Year          2.441e-02  8.832e-04  27.644   <2e-16 ***
GenreProse    1.185e+00  1.333e-01   8.889   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2436.84  on 89  degrees of freedom
Residual deviance:  489.15  on 87  degrees of freedom
AIC: 609.02
```

Figure 6: Predicting *wh*-items from year and genre

- - Poetry and prose change at the same rate of change; Constant Rate Effect (Kroch 1989)
    * the addition of an interaction effect between 'genre' and 'year' to a model with main effects only does not significantly reduce deviance:

```
Analysis of Deviance Table

Model 1: cbind(WHEN, THEN) ~ Year + Genre
Model 2: cbind(WHEN, THEN) ~ Year + Genre + Year:Genre
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1        87     489.15
2        86     488.81  1   0.3409   0.5593
```

Figure 7: Comparison of models with and without genre:time interaction effect

- Some examples of the *th-wh* variation within the same text:

(3) *The Bestiary*

    a. we ben siker dere, So ðis wirm in winter is,
        we are safe there, as this worm in winter is,

        ðan ge   ne   tileð nummore.
        when you not till   nomore

        'We will be safe there, just as this insect is in winter, when you do not till anymore'
        (Bestiary,143.9.290.[Ant_Significance])

    b. **wanne** he is ikindled
        when   he is whelped

        Stille lið ðe leun,
        still   lies the lion

        'When he is first born, the lion lies quietly'
        (Bestiary,9.1.17.[Lion_Nature])

(4) *The Owl and the Nightingale*

    a. þan gode ich fulste to longinge,
        the good I   help   in longing

        vor **þan**   him longeþ, ich him singe:
        for when him longs,   I   him sing

        'I help the good man in longing, for when he feels desire, I sing to him'
        (OwlNight,76.890.503)

    b. an prostes upe   londe singeþ
        a   priest   upon land   sings

        wane þe liȝt   of daie springeþ.
        when the light of day springs

        'Country priests sing when dawn breaks'
        (OwlNight,62.734.423)

(5) *Havelok the Dane*

    a. **þan** he was ded, þere micte men se
        when he was dead, there might men see

        þe meste sorwe   that micte be
        the most sorrow that might be

        'When he was dead, one could see, the greatest sorrow that could ever be'
        (Havelok,8.233.107)

    b. **Hwan** he wore come, sket   was þe erl yare,
        when   he was come, quickly was the earl ready,

        Ageynes denshe men to fare,
        against Danish men to go

        'When he had arrived, the earl was ready right away to advance against the Danish'
        (Havelok,73.2575.1188)

6 Explaining the Change

6.1 List of possible influencing factors

- A number of reasons for the rise of *wh*-items have been proposed (Yamakawa 1969)

    1. Loan syntax from Latin or French sources; subordinators based on *wh*-items in Latin *quando* or Old French *quant*; some Northern Middle English texts spell *wh*-words with *qu* (*quen* for *when*)

    2. Analogy; *wh*-items are rising across the board (nominal relative clauses (*the man the I see* → *the man who I see*), nominal free relatives, locatives (*there / where*), rise of *while* as a subordinator)

    (6)   a.   in the same place **there** the grete batayle was, ys grete tresoure hydde   in the erthe
                in the same place where the great battle   was   is great treasure hidden in the earth
                'In the same place where the great battle was, a great treasure was hidden in the earth'
                (CMMALORY,30.947, c. 1470 A.D.)

                b.   I com   but late oute of the Waste Foreyste **where** I founde the Rede Knyght
                I came but late out   of the waste   forest   where   I found   the red   knight
                'I came late out of the desolate forest where I had found the red knight'
                (CMMALORY,667.4880, c. 1470 A.D.)

- Most importantly, however: loss of disambiguating word order

    3. Word order change; word order was used to distinguish adverbial from subordinating readings of *then*; but the relevant word order patterns begin to disappear

    "... a weak point in functional distinctness that *þanne* and *þo* has when used as a subordinate [...] conjunction. This seems to be an internal motive that caused *þanne* or *þo* to be replaced by *hwanne*." (Yamakawa 1969: 31)

- The third factor is the determinant that will be investigated in this paper

6.2 A model of word order as a disambiguating device

 - The interpretation of *then* as an adverb or subordinator depends at least in part on word order
 - Formalization in LFG (irrelevant detail omitted)

- 1. Verb-second order after *þanne* signaled an adverbial reading

(7)   þa   com   Henri abbot
     then came Henry abbot
     'Then the abbot Henry came' not: # 'when the abbot Henry came'
     (CMPETERB2,54.370)

(8)   a.   *Topicalization Rule*

```
CP_ROOT →          AdvP                    C'_ROOT
               (↑ TOPIC) = ↓              ↑ = ↓
          (↑ TOPIC) ∈ (↑ ADJUNCT)   (↑ CLAUSE-TYPE)= declarative
```

     b.   C'_ROOT →
```
C'_ROOT →   C        IP
           ↑ = ↓    ↑ = ↓
```

     c.   *lexical entry for 'then'*
```
þa    Adv        (↑ PRED)= 'THEN'
                 (↑ ADJ-TYPE)= pres-top
```

     d.   *V-to-C if licensed by clause-initial operator adverb*
```
C →           V
             ↑ = ↓
      (↑ TOPIC ADJ-TYPE)=_c pres-top
```

(9)

```
                              CP_ROOT
                          ╱            ╲
                  AdvP                      C'_ROOT
             (↑ TOPIC)= ↓                  ↑ = ↓
    (↑ TOPIC)∈ (↑ ADJUNCT)      (↑ CLAUSE-TYPE)= declarative
                  │                      ╱        ╲
                 Adv                    C           IP
                  │                     │           │
                 þa                     V           DP
         (↑ PRED)= 'THEN'            ↑ = ↓     (↑ SUBJ)= ↓
      (↑ ADJ-TYPE)= pres-top   (↑ TOPIC ADJ-TYPE)=_c pres-top   Henri abbot
                                      │
                                     com
```

$$\begin{bmatrix} \text{PRED} & \text{'COME}\langle(\text{SUBJ})\rangle\text{'} \\ \text{TOPIC} & \begin{bmatrix} \text{PRED} & \text{'THEN'} \\ \text{ADJ-TYPE} & \text{pres-top} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'HENRY'} \end{bmatrix} \\ \text{ADJUNCT} & \{[]\} \\ \text{CLAUSE-TYPE} & \text{declarative} \end{bmatrix}$$

- 2. In the absence of verb-second order after *then*, Middle English permits an adverbial reading

(10)   And thenne the kynge retornyd ayenne
       And then   the king   returned again

       'And then, the king returned' Context makes it clear that not # 'when the king returned'
       (CMGREGOR,114.448)

(7)   b.''   *Second C' rule without an overt head*
           C'_ROOT →   IP
                        ↑ = ↓

(10)

```
                              CP_ROOT
                          ╱            ╲
                  AdvP                      C'_ROOT
             (↑ TOPIC)= ↓                  ↑ = ↓
    (↑ TOPIC)∈ (↑ ADJUNCT)    (↑ CLAUSE-TYPE)= declarative
                  │                      ╱        ╲
                 Adv                               IP
                  │                           ╱         ╲
                thenne              the kynge retornyd ayenne
         (↑ PRED)= 'THEN'
      (↑ ADJ-TYPE)= pres-top
```

$$\begin{bmatrix} \text{PRED} & \text{'RETURN}\langle(\text{SUBJ})\rangle\text{'} \\ \text{TOPIC} & \begin{bmatrix} \text{PRED} & \text{'THEN'} \\ \text{ADJ-TYPE} & \text{pres-top} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'KING'} \\ \text{DEF} & + \end{bmatrix} \\ \text{ADJUNCT} & \left\{ \begin{bmatrix} \text{PRED} & \text{'AGAIN'} \end{bmatrix} \\ [] \right\} \\ \text{CLAUSE-TYPE} & \text{declarative} \end{bmatrix}$$

- 3.Without verb-second order after *then*, a subordinate clause reading is common

(11) Ac ure drihten eft   of deaþe heo   aræreþ,
      but our Lord   again of death them raises-up

      So he alle men deþ, þonne domes   dai   cumeþ.
      as he all men does, when doomsday comes

      'But our Lord raises them up from death again, as he does all men, when doomsday comes'
      not: # 'He raises them up. And then Domesday comes'
      (BodySoul,185.7.13.FragE)

  – Analysis: temporal subordinate clauses are treated as free relatives (Geis 1970; Grimshaw 1977; Bresnan and Grimshaw 1978).

    * Temporal *when*-clauses have the same distribution as DPs or PPs
    * The subordinator can be construed as modifying embedded complement clauses

  (12) I saw Mary in New York when [$_{IP}$ she claimed [$_{CP}$ that [$_{IP}$ she would leave.]]]
        (i) high construal: at the time that she made the claim
        (ii) low construal: at the time of her presumed departure
        (from: Haegeman 2010, 635, her example (21))

(13) a. *Adjunction of a projection to the main structure; assume form* AdvP

          ...   →   AdvP
                   ↓ ∈ (↑ ADJUNCT)

      b. *The* AdvP *dominates a relative clause with an empty head*

          AdvP →   CP$_{REL}$
                   (↑ PRED) = 'pro'
                   ↓ ∈ (↑ RELMOD)

      c. *The specifier of the relative clause is the topic 'then'*

          CP$_{REL}$ → AdvP   C'$_{REL}$
              (↑ REL-TOPIC) = ↓   ↑ = ↓
              (↑ REL-TOPIC) ∈ (↑ {COMP|XCOMP}* ADJUNCT)(↑ CLAUSE-TYPE)= subordinate

  – Note that the same lexical entry for *then* can be applied in both adverbial and subordinate uses; there are not two *then*s with different featural specifications but only one

(14)

$\ldots$
|
AdvP
$\downarrow \in (\uparrow \text{ADJUNCT})$
|
CP_{REL}
$(\uparrow \text{PRED}) = \text{'pro'}$
$\downarrow \in (\uparrow \text{RELMOD})$

AdvP
$(\uparrow \text{REL-TOPIC}) = \downarrow$
$(\uparrow \text{REL-TOPIC}) \in$
$(\uparrow \{\text{COMP}|\text{XCOMP}\}* \text{ADJUNCT})$
|
Adv
|
þonne
$(\uparrow \text{PRED}) = \text{'THEN'}$
$(\uparrow \text{ADJ-TYPE}) = \text{pres-top}$

C'_{REL}
$\uparrow = \downarrow$
$(\uparrow \text{CLAUSE-TYPE}) = \text{subordinate}$
|
IP
domes dai cume

$$\begin{bmatrix} \ldots \\ \text{ADJUNCT} \left\{ \begin{bmatrix} \text{PRED} & \text{'pro'} \\ \text{RELMOD} & \left\{ \begin{bmatrix} \text{PRED} & \text{'COME}\langle(\text{SUBJ})\rangle\text{'} \\ \text{REL-TOPIC} & \begin{bmatrix} \text{PRED} & \text{'THEN'} \\ \text{ADJ-TYPE} & \text{pres-top} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'DOOMSDAY'} \end{bmatrix} \\ \text{ADJUNCT} & \{[]\} \\ \text{CLAUSE-TYPE} & \text{subordinate} \end{bmatrix} \right\} \end{bmatrix} \right\} \end{bmatrix}$$

- 4. *when* only allows subordinator readings because it is incompatible with declarative clauses

(15) *lexical entry for 'when'*

| *hwan* | Adv | $(\uparrow \text{PRED}) = \text{'WHEN'}$ |
| | | $(\uparrow \text{CLAUSE-TYPE}) \neq \text{DECLARATIVE}$ |

 – This allows straight-forward competition between *when* and subordinator usage of *then* but not between *when* and adverbial usages of *then*

6.3   A syntactic chain shift

- The two readings of *then* were disambiguated by different word order patterns

- Originally, in the Old Engish period, the conditioning factors were extremely rigid

- Simplified summary:

  THEN $\longrightarrow$ $\Bigg\{$ $\longrightarrow$ main clause adverb / __ verb second

  $\longrightarrow$ temporal subordinator / $\sim$ (elsewhere)
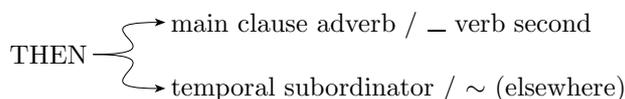
- The conditioning factor weakens; verb-second declines (rule (8d)); the constraint "adverb always if verb second" becomes "adverb usually / sometimes if verb second"; the categorical constraint becomes probabilistic

- Problem: without verb-second order, only context can disambiguate between the readings, but context is not always reliable

- Ambiguity between interpretations

(16)   Tho   the screwe was overcome; Sori   he was and wo.
       then/when the villain was overcome, sorry he was and woe

   (i) 'Then the villain was overcome. He was sorry and miserable.'

   (ii) 'When the villain was overcome, he was sorry and miserable.'
   (Fridesw,43.55)

- In principle, the language system can respond to a loss of conditioning factors in three ways:

  1. Live with the ambiguity. But contrasts without cues are hard to maintain. Subsequent generations may not be able to acquire the contrast

  2. Give up one of the conditions. This is known to happen sometimes but is a rare occurrence. Languages do not just give up important categories like 'temporal subordinate clause'

  3. Replace the conditioning factor. This may be the most common effect given that language systems tend to preserve salient contrasts (e.g. Labov 1994, 119) (Principle of Contrastiveness)

- In the case under investigation, the third option occurred; to preserve the contrast between main clause adverbs and temporal subordinators, *wh*-items were recruited; a kind of repair strategy

- This is a **syntactic chain shift**: the loss of a syntactic conditioning factor promotes another linguistic change to prevent a merger of two linguistic categories that would otherwise have taken place

7 Hypothesis Testing

- The idea that *when* replaced subordinating *then* as a consequence of a syntactic chain shift makes some empirical predictions

- Two specific hypotheses will be tested in this section

- **H1 - Predictive value of verb-second**: The interpretability of *then* should depend on the degree of distinctness of the conditioning word order patterns. Therefore, there should be a correlation between the propensity to invert subject and verb after *then* and the frequency of occurrence of *wh*-based temporal subordinators.

- Procedure:

  1. Compare the proportion of *wh*-based subordinators in temporal clauses against a measurement of verb-second after *then*.

  2. The dependent variable 'type of subordinator' is measured as proportion of *wh*-based subordinators out of all subordinators

  3. The frequency of verb-second after *then* functions as an independent variable:
     - for verb-second: *then* and finite verb must be immediately adjacent; only negation can intervene; the finite verb precedes a non-empty subject anywhere within the same clause; only *then* is contained in the adverb phrase

       (17) **þen maiȝt** þou synge of loue lele.
       then might you sing   of love loyal

       'Then you may sing of faithful love'
       (HowHearMass,134.208.69.[Stanza_18])

     - for absence of verb second: *then* and finite verb (or negated finite verb) must not be immediately adjacent; *then* precedes a non-empty subject and precedes the finite verb in any order; only *then* is contained in the adverb phrase

       (18) Whon he haþ waschen **þen he** walkes
       when he has washed   then he walks

       'When he has done the ritual cleansing, then he walks on'
       (HowHearMass,143.533.221.[Stanza_46])

  4. Also consider time and genre (prose vs. poetry) as additional independent variables

- Caveat:

  - Several confounding factors that were not considered might distort the result to some degree
    * colinearity: since verb-second after *then* is in decline in Middle English, the predictors 'year' and 'verb-second' are probably highly correlated
    * lexical effects: certain words may be more likely to invert after *then* than others, e.g. *be*, motion verbs, quotatives
    * dialect: Northern Middle English has a generalized V2 constraint (Kroch and Taylor 1997); verb-second after *then* as a special case may survive longer in these regions

- Result:
  - 2,037 examples of V2; 2,097 examples of No-V2
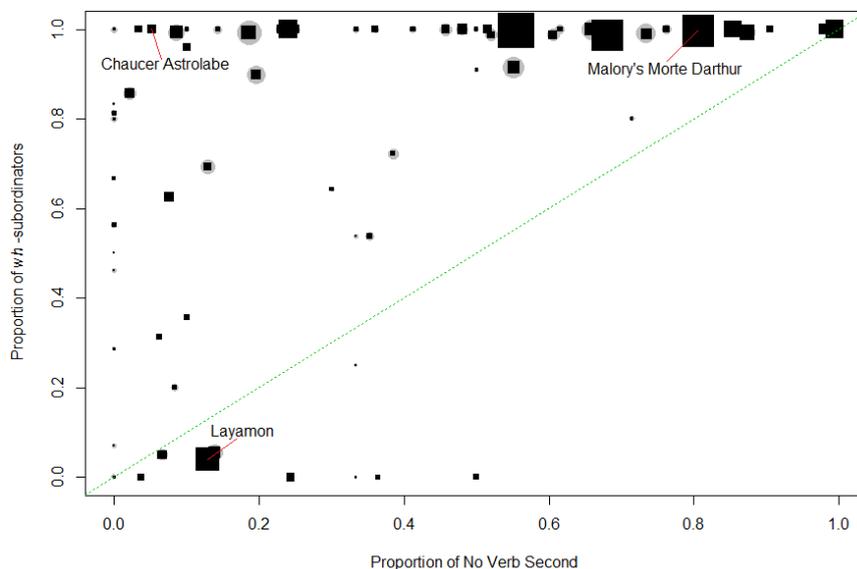 - Relation between subordination and V2 looks as follows:

Figure 8: Correlation between subordination and verb-second

  - * As expected, it is often the case that a low/high proportion of no verb-second correlates with a low/high proportion of *wh*-subordinators
    - *Malory's Morte Darthur*: 80.6% no verb second after adverbial *then*, 99.6% *when*
    - *Layamon*: 12.9% no verb second after adverbial *then*, 3.9% *when*
  - * However, substantial number of texts with a strong, conservative verb second element, but still a large number of innovative *wh*-subordinators
    - *Chaucer's Astrolabe*: 5.2% no verb second after adverbial *then*, 100% *when*
  - * This suggests a one-way conditional relationship: loss of verb-second → rise of *when*
  - * This finding is compatible with the supposed relationship between word order and interpretation of functionally ambiguous items
  - Time course of the two changes:
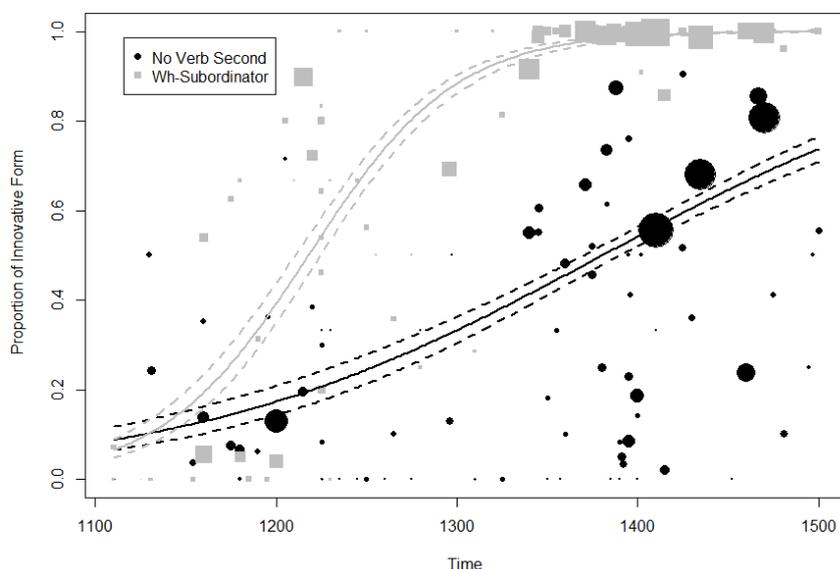
Figure 9: Rise of *wh*-subordination and loss of verb-second by time

14

* Substantial variation in word order after adverbial *then* (black dots)
* Nevertheless, a significant decline of verb-second after adverbial *then* is measurable in Middle English
* The point of origin of the two changes fall roughly together; the loss of verb-second begins just before the rise of *wh*-subordination
* For example, for 1100 A.D., the regression model predicts 8.2% absence of verb second after *then*; but only a 3.6% probability of finding a *wh*-subordinator
* The timing of the changes is compatible with the hypothesis that the loss of verb second promoted the change in subordinator form

– Model comparison shows that the loss of verb-second has explanatory power over and above the 'year' variable

```
Analysis of Deviance Table

Model 1: cbind(WHEN, THEN) ~ Year + Genre
Model 2: cbind(WHEN, THEN) ~ Year + Genre + Log_Prop_NOV2
  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
1        60     420.12
2        59     329.00  1   91.123 < 2.2e-16 ***
```

Figure 10: Comparison of models with and without verb-second predictor

– The variables 'year', 'genre' and 'degreeof verb-second order after *then*' are significant predictors for the realization of temporal subordinator as *then* or *when*

```
Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)   -2.569e+01  1.153e+00 -22.281  < 2e-16 ***
Year           2.193e-02  9.168e-04  23.917  < 2e-16 ***
GenreProse     1.088e+00  1.472e-01   7.390 1.47e-13 ***
Log_Prop_NOV2  1.002e+00  1.024e-01   9.783  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2220.1  on 62  degrees of freedom
Residual deviance:  329.0  on 59  degrees of freedom
AIC: 422.03
```

Figure 11: Predicting *wh*-items from year, genre and verb-second

→ The data supports hypothesis 1

- **H2 - Alternative subordinating strategies**: The interpretation of *then* does not necessarily rely on word order cues. Alternative constructions are available that force a subordinating reading. Therefore, *th*-based subordinators should prevail more robustly when alternative subordinating strategies are used.

- Procedure:

  1. Compare the proportion of *wh*-based subordinators in temporal clauses in contexts with and without alternative subordinating strategies.

  2. The dependent variable 'type of subordinator' is measured as before

  3. The independent variable 'presence of alternative subordinating strategy' is positive in the following three cases:

     · **Overt complementizers**
     · when *then* or *when* co-occur with a complementizers like *that*, only subordinate readings are possible; thus main clause adverbial readings are explicitly ruled out

        (19) *lexical entry for 'that'*

              þat   C   (↑ CLAUSE-TYPE)= subordinate

     (20) a. Anon, **ðo** **ðe** he lokede upen him, he agann to wepen,
              at-once, when that he looked upon him, he began to weep
              'Right away, when he looked at him, he began to weep'
              (CMVICES1,111.1358)

           b. and **when þat** he come þider, þe Porter wende þat it hade bene his owen lorde.
              (CMBRUT3,67.2009)

     · **Correlative þa ... þa constructions**
     · the first *þa* locates the even in discourse, a second resumptive *þa* relates the following event (Kemenade and Los 2006; Links and Kemande 2013); thus, the resumptive adverb can serve as an indication that the first *þa* is used as a subordinator

     (21) a. **þa**   þu þurch delidlich sunne Murðredest godes spuse þi   saule.
              when you through deadly   sin   murdered God's spouse, your soul.

              þa   þu were idemed for tobeon ahonged on berninde wari   treo
              then you were deemed for to-be hanged on burning gallows

              'When you killed God's spouse, your soul, with deadly sin, then you were judged to be hanged on burning gallows'
              (CMANCRIW-1,II.230.3323)

           b. **whan** reynart herd bruyn **tho** wente he Inneward in to his hole
              (CMREYNAR,13.242)

     · **Relativization**
     · *then* can be modified by a relative clause with a temporal relative adverb; therefore "double occurrences" of *then then* or *then when* force a subordinate interpretation

     (22) a. swa hit wæs **þa** **þa**   David sloh Goliam
              so it was then when David slew Goliath

              'so it happened when David killed Goliath
              (ELUCID,141.175)

           b. **Thenne when** the lady was delyverd the kynge commaunded two knyghtes and two ladyes to take the child bound in a cloth of gold
              (CMMALORY,6.149)

  4. Also consider time and genre (prose vs. poetry) as additional independent variables

- Result:
  - 4,168 ordinary temporal subordinate clauses; 411 examples with alternative subordinating strategy
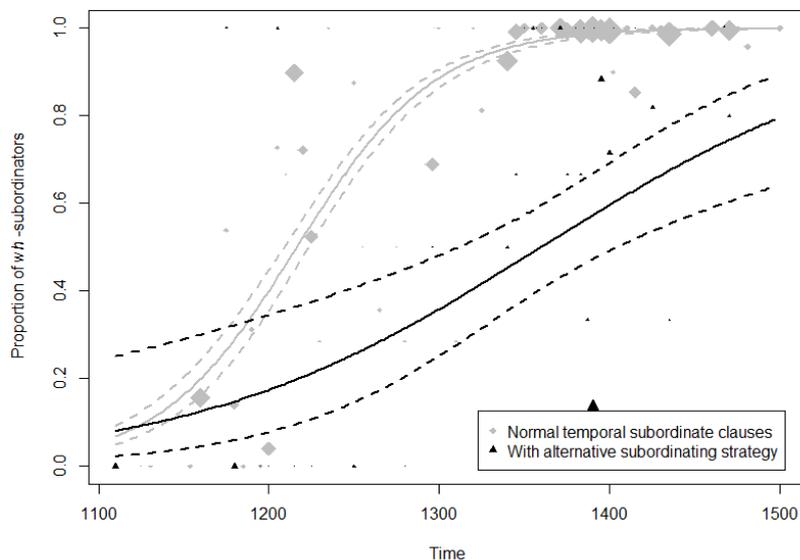 - Regression analysis for the two contexts against time:

Figure 12: Subordinators with alternative subordinating strategies

  - ∗ clear contrast between temporal clauses with and without alternative subordinating strategies
  - ∗ clauses with additional disambiguating devices show a considerably more conservative pattern
  - A model with 'presence of alternative subordinating strategies' predicts the realization of the subordinator as a *th-* or *wh-*item significantly better than a model with 'year' and 'genre' as the only independent variables

```
Analysis of Deviance Table

Model 1: DepVar ~ Year + Genre
Model 2: DepVar ~ Year + Genre + Strategy
  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
1      4576     2164.2
2      4575     1869.8  1   294.43 < 2.2e-16 ***
```

Figure 13: Comparison of models with and without alternative strategy variable

  - The variables 'year', 'genre' and 'alternative subordinating strategy' are all significant

```
Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.839e+01  1.026e+00 -27.660  < 2e-16 ***
Year         2.290e-02  8.227e-04  27.832  < 2e-16 ***
GenrePROSE   9.243e-01  1.273e-01   7.263 3.78e-13 ***
StrategyYES -3.816e+00  2.186e-01 -17.454  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3812.2  on 4578  degrees of freedom
Residual deviance: 1869.8  on 4575  degrees of freedom
AIC: 1877.8
```

Figure 14: Predicting *wh-*items from year, genre and alternative subordinating strategy

→ The data supports hypothesis 2

8 Conclusion

- This paper looked at the rise in temporal *wh*-subordinators in Middle English

- The change is empirically measurable even in poetic records (Parsed Corpus of Middle English Poetry)

- It was suggested that the development of *when* as a subordinator can in part be attributed to the loss of conditioning word order patterns (syntactic chain shift)

- This is not a strictly deterministic conceptualization. Other factors may be relevant too and should be studied in the future:

    - French and Latin influence
    - Analogy with other areas of the language system in which *wh*-items become dominant
    - Dialects (may not be possible with the data currently available)

- Overall, there seems to be good empirical reason to support the idea that the loss of a conditioning factor can contribute to a subsequent change; system-internal motivation for language change

References

Bresnan, J. and Grimshaw, J. (1978), 'The syntax of free relatives in english', *Linguistic Inquiry* **9.3**, 331391.

Declerck, R. H. C., ed. (1997), *When-Clauses and Temporal Structure*, Routledge, London.

Geis, M. L. (1970), *Adverbial subordinate clauses in English'*, PhD Dissertation, MIT Press, Cambridge, MA.

Grimshaw, J. (1977), *English Wh-Constructions and the Theory of Grammar*, PhD Dissertation, University of Massachusetts, Amherst.

Haegeman, L. (2010), 'The internal syntax of adverbial clauses', *Lingua* **120.3**, 628–648.

Kemenade, A. v. and Los, B. (2006), Discourse adverbs and clausal syntax in old and middle english, *in* 'The Handbook of the History of English', Wiley-Blackwell, Oxford, pp. 224–248.

Kivimaa, K. (1966), *þe and þat as clause connectives in Early Middle English with special consideration of the emergence of the pleonastic þat*, Commentationes humanarum litterarum 39, Societas Scientiarum Fennic, Helsinki.

Kroch, A. (1989), 'Reflexes of grammar in patterns of language change', *Journal of Language Variation and Change* **1.3**, 199–244.

Kroch, A. and Taylor, A. (1997), 'Verb movement in old and middle english: Dialect variation and language contact', pp. 45–68.

Kroch, A. and Taylor, A. (2000), *Penn-Helsinki Parsed Corpus of Middle English*, http://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-3 (Accessed 10 April 2013), 2 edn, Department of Linguistics, University of Pennsylvania.

Labov, W. (1994), *Principles of Linguistic Change. Volume 1: Internal Factors*, Wiley-Blackwell, Oxford.

Links, M. and Kemande, A. v. (2013), 'correlative constructions in earlier english: the þa ... þa construction", *Paper delivered at SHES11, Nijmegen, June 2013* .

Yamakawa, K. (1969), 'The development of when as subordinate conjunction or relative adverb', *Hitotsubashi Journal of Arts and Sciences* **10.1**, 8–42.