Diachrony as a Laboratory for Syntactic Theory: the two Germanic subject positions and change in English

Joel C. Wallenberg
joel.wallenberg@ncl.ac.uk

January 25, 2012
University of York
Big Questions

- **Is synchronic syntactic theory relevant to diachronic syntax?**
  Definitely.
  (see discussion of the Uniformitarian Principle in Labov, 1994)
- **Is diachronic syntax relevant to synchronic syntactic theory?**
- Language change can be an experimental domain for linguistic theory: we can develop hypotheses about quantitative effects which can only be observed over long periods of time.
  - If surface syntax changed, syntactic theory can suggest quantitative relationships between surface patterns.
  - These patterns are beyond the lifespan of any linguist.

**Experimental Infrastructure:** accurate (i.e. hand-corrected) parsed diachronic corpora.
E.g. YCOE (Taylor et al., 2003), PPCME2 (Kroch and Taylor, 2000a), PPCEME (Kroch et al., 2005), PPCMBE (Kroch et al., 2010), and IcePaHC (Wallenberg et al., 2011).
Outline

1. Two-Subject Parameter
2. Diachronic Experiment
3. Model of Learning and Change
4. Problem: the loss of V2
5. Conclusions
Haeberli (1999) identifies a parameter which divides the Germanic languages: whether a language has two surface positions for definite subjects (e.g. German), or only one (e.g. English).

The possibility of adverbs occurring between C and a definite subject diagnoses the presence of two \( \nu P \)-external subject positions.

(1) daß (später) der Hans dieselbe Uhr kaufen wird
that later the Hans the-same watch buy will
“...that Hans will buy the same watch later”
(Haeberli, 1999, 4)

(2) * I don’t think that later Hans will buy the same watch.
The contrast similarly shows up in V-to-C contexts (which include all matrix clauses in German; it is V2):

(3) Wahrscheinlich wird (später) der Hans (später) Probably will (later) the Hans (later)
dieselbe Uhr kaufen the-same watch buy

“Probably Hans will buy the same watch later”
(Haeberli, 1999, 2)

(4) Will (*later) John (later) buy the same watch?
(Haeberli, 1999, 2)
Dutch patterns like German:

(5) Dit boek heeft (helaas) Jan niet gelezen
This book has (unfortunately) John not read
“John unfortunately didn’t read this book”

(6) dat (ondanks alles) Jan dit boek niet zal lezen
that (despite everything) John this book not will read
(Haeberli, 1999, 16)
West Flemish does not pattern like German and Dutch:

(7) Misschien goa (*loater) Valère (loater) tzelfste
    Probably goes (later) Valère (later) same
    orloge kuopen
    watch buy

(8) *da loater Valère tzelfste orloge goa kuopen
    that later Valère same watch goes buy
    (Haeberli, 1999, 2)
Afrikaans is like West Flemish:

(9) *dat ondanks alles Jan hierdie boek nie sal lees nie
that despite everything Jan this book not will read not

Frisian is like German and Dutch:

(10) dat (ûngelokkichgenôch) myn broer dit boek net lêzen hat
that unfortunately my bother this book not read has

(Haeberli, 1999, 12,19)
Two-Subject Parameter: VO North Germanic

- Danish is like English, West Flemish, Afrikaans:
  
  (11) *Denne bog har desværre John ikke læst
       This book has unfortunately Johan not read
  
  (12) *at trods alt John ikke vil læse disse bøger
       that despite all John not will read this book
       (Haeberli, 1999, 24)

- Icelandic appears to be like Danish:
  
  (13) *Sennilega mun seinna Jón kaupa sama úrið
       probably will later Jón buy same watch-the
Norwegian and Swedish show the two subject positions, though there is a lot of variation in judgments concerning the embedded context.

(14) Den här klockan hade (senare) min gamle far köpt
(Swedish)

(15) Denne klokka hadde (seinere) min gamle far
This watch had (later) my old dad
kjøpt (Norwegian)
bought

(16) %att (??/*tyvärr) Johan inte har läst denna
that (unfortunately) Johan not has read this
boken
book
(Haeberli, 1999, 3,28)
Basic Generalizations

- Assuming that high adverbs occur in roughly the same position in these different languages...
- Some show a high and a low position for definite subjects, while others show only a high one.
- The availability of both subject positions does not pattern according to the North/West Gmc. subgroups.
- The availability of both subject positions is independent of the OV/VO and Infl-final/Infl-medial parameters.
- **Additional Observation:** No language above allows adverbs to intervene between C and a pronominal subject.
  - If a language has the two subject positions, pronoun subjects always occur in the higher one.
  - There is some amount of optionality in subject placement, but not complete optionality.
Two Subject Structure for OV West Germanic

- CP
  - XP
    - C’
      - C
        - AgrSP
          - Pron. Subjects
          - Nom. Subjects
          - AgrSP’
            - TP
              - Nom. Subjects
              - T’
                - vP
                - Tense
Two Subject Structure for OV West Germanic

- **CP**
  - **XP**
  - **C’**
    - **C**
      - **AgrSP**
        - **Pron. Subjects**
        - **Nom. Subjects**
          - **AgrSP’**
            - **TP**
              - **Nom. Subjects**
              - **T’**
                - **vP**
                - **Tense**
Two Subject Structure for OV West Germanic

CP

XP

C’

C

AgrSP

Pron. Subjects
Nom. Subjects

AgrSP’

TP

Nom. Subjects

T’

vP

Tense

AgrS
One Subject Structure for OV West Germanic

```
CP
  /\      \   
XP     C'  
     /\   /\ 
    C  TP
       /\  /\  
  Pron. Subjects  T'
     /\      \   
    Nom. Subjects  Tense
                  \  ...
```

- CP
- XP
- C'
- C
- TP
- Pron. Subjects
- Nom. Subjects
- vP
- Tense
- ...
One Subject Structure for OV West Germanic

CP

XP

C’

C

TP

Pron. Subjects

Nom. Subjects

T’

νP

Tense

...
One Subject Structure for OV West Germanic

```
CP
   /\  
XP   C'
   |
   C  TP

   /\   /\  
Pron. Subjects Nom. Subjects T'
   |
   vP  Tense
   |
   ...
```
Two Subject Structure for (Mainland) North Germanic

CP
  / \  
XP   C'
  /   /
C    AgrSP
   /    /
Pron. Sbj AgrSP'
   /     /
Nom. Sbj AgrS
   /      /
Nom. Sbj TP
   /        /
T'        vP
       /      /
      v ...
Two Subject Structure for (Mainland) North Germanic
Two Subject Structure for (Mainland) North Germanic
One Subject Structure for (Mainland) North Germanic and English
One Subject Structure for (Mainland) North Germanic and English
One Subject Structure for (Mainland) North Germanic and English

Diagram:

```
CP
  XP
    C'
      C
        TP
          T'
            vP
              v
                ...```

- CP: Complementizer Phrase
- XP: Extra-Probabilistic Phrase
- C': Complement
- C: Complement
- TP: Tense Phrase
- T': Tense
- vP: Verb Phrase
- v: Verb
- ...
Problem: Old English

- OE shows regular V2 orders when some constituent is fronted and the subject is nominal.
- OE never shows V2 orders when the subject is a pronoun, aside from well-known V-to-C contexts. (van Kemenade, 1987; Pintzuk, 1991)

(17) Ond eallum þam dagum... he afæste to æfenes
    And all those days he fasted till evening
    “and he fasted all days (except Sundays), till the evening” (Bede’s History of the English Church, 11th c.)

(18) þone wæterscipe beworhte se wisa cyning Salomon
    the conduit built the wise king Salomon
    “The wise King Salomon constructed the conduit”
    (Ælfric’s Catholic Homilies, 10th c.)
Problem: Old English

(19) Ælc yfel he mæg don
Each evil he may do
   “He can do each evil”
(Wulfstan’s Homilies, 11th c.)

(20) & of heom twam is eall mann cynn cumen
And from them two is all mankind come
   “All mankind has come from those two.”
(Wulfstan’s Homilies, 11th c.)
Haeberli’s Solution: (Infl-medial) Old English

```
CP
  fronted-XP
    C'
      C
        AgrSP
          Pron. Sbj
          Nom. Sbj
          AgrSP'
            AgrS
              finite-verb
                Nom. Sbj
                T'
                  Tense
                  vP
                    ...
                    v
```
Haeberli’s Solution: (Infl-medial) Old English

```
CP
  \----\--------
  \      \     
  \      \ fronted-XP
  \      
  \  C’
  \   \--------
  \   \      \ 
  \   \      \ \ 
  \   \      \ C
  \   \        
  \   \--------
  \   \      \ AgrSP
  \   \      
  \   \  Pron. Sbj
  \   \ Nom. Sbj
  \   
  \  AgrS
  \  finite-verb
  \ 
  \ AgrSP’
  \ 
  \ TP
  \ 
  \ Nom. Sbj
  \ T’
  \ 
  \ Tense
  \ vP
  \   ...
  \    v
```
Evidence in favour of Haeberli’s Solution (Speyer, 2010)

- Speyer (2010) shows that a prosodic constraint (avoid accent clash) is responsible for the frequent use of the low subject position when something is topicalized.
- V2 orders prevent a clash between an accented subject and an accented fronted XP, but prosodically weak subjects do not create this problem.
- This is still true in modern English; 90.5% of topicalizations occur with pronominal subjects (based on Gregory Ward’s corpus; see Speyer, 2010)
- For this reason, there is a decline in the overall rate of topicalization in English as these V2 orders were lost.
More evidence in favour of Haeberli’s Solution

- Non-V2 orders with topicalization, a nominal Subject, and a particle verb (Speyer, 2010).
- The finite verb has moved leftward, stranding the particle, and a nominal subject precedes it.

(21) Ac þære ilcan niht... wulfas atugan þa stacan
    And the.DAT same night wolves drew the pin
    up
    up
    “The same night, wolves (= devilish persons) drew up
    the pin.” (Orosius, 10th c.)

(22) æne se geatweard læt in
    this.ACC the gatekeeper let in
    “The gatekeeper let in this one.”
    (Wulfstan’s Gospels, 11th c.)
If OE was not V2, then the loss of V2 was not the loss of V2

- V2 orders disappeared from English during the ME and EME periods, roughly during the years 1200-1700 (cf. Kroch and Taylor, 2000b).
- Haeberli (2002) analyzes this as the loss of the AgrSP projection.
- Once AgrSP is lost, only one subject position remains, above the landing site of the finite verb.
- **Prediction:** If this analysis is correct, then the frequency of pre-subject adverbs should also decline over the same time period.
Finer-Grained Hypotheses

1. C > Adv > Sbj orders should decline with nominal subjects, but should be constant with pronoun subjects (very low freq. or nonexistent).

2. Though the loss of V2 orders is dependent on Infl-medial phrase structure, the loss of AgrSP is not.
   - **Prediction**: If Infl-final clauses are taken separately, C > Adv > Sbj orders should still decline.

3. Old English style C > finite-V > Adv > Sbj orders and Scandinavian-style C > Adv > Sbj > finite-V orders should both show the change.
Finer-Grained Hypotheses: CP-recursion

- Haeberli (1999) reported the following contrast:
  
  (23) * Why will after this John not read her book?
  (24) ...that after this John will not read her book.

- I think this is due to CP-recursion licensing embedded topicalization, thus:
  
  I bought a watch that John likes, ...
  
  (25) and I think that later John will buy the same watch.
  (26) * but I don’t think that later John will buy the same watch.
  (27) * but I doubt that later John will buy the same watch.
Finer-Grained Hypotheses: CP-recursion

If I am correct about the role of CP-recursion, then three predictions follow:

1. It should be possible to see a decline in C > Adv > Sbj orders in subordinate clauses, where C is filled by a complementizer.
2. CP-recursion (embedded root) contexts should show higher frequencies of C > Adv > Sbj orders across the history of English.
3. Even so, the frequency of C > Adv > Sbj in different subordinate clause types should decline at the same rate over time (the Constant Rate Effect; Kroch, 1989, inter alia).
Adverbs preceding nominal subjects, N = 6909 clauses
Adverbs preceding pronominal subjects, N = 13289
Adv preceding nominal subj., $V > Aux$ Clauses, $N = 290$
Sub. Clause Types, Logistic Regression, N = 6909
Sub. Clause Types, Logistic Regression, N = 6909
Sub. Clause Types, Model Comparison

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Deviance</th>
<th>Resid. Df</th>
<th>Resid. Dev</th>
<th>Pr(&gt;Chi)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NULL</td>
<td></td>
<td></td>
<td>6908</td>
<td>6159.4</td>
<td></td>
</tr>
<tr>
<td>Clause</td>
<td>6</td>
<td>41.004</td>
<td>6902</td>
<td>6118.4</td>
<td>2.89e-07</td>
</tr>
<tr>
<td>Year</td>
<td>1</td>
<td>284.589</td>
<td>6901</td>
<td>5833.8</td>
<td>&lt; 2.2e-16</td>
</tr>
<tr>
<td>Clause:Year</td>
<td>6</td>
<td>10.626</td>
<td>6895</td>
<td>5823.2</td>
<td>0.1007</td>
</tr>
</tbody>
</table>

- A model without an interaction between Clause and Year does not fit significantly worse.
- However, the fit is not wonderful, possibly because of *that*-clauses.
Note that the later examples of the declining pattern are the Scandinavian type of AgrSP, without verb movement to AgrS, unlike OE.

(28) **Old English pattern**
for it is a synguler & a specyal ʒyft þat God hath ʒouyn þe, 
- a welle of teerys [ þe whech xal neuyr man take fro þe. ]

(29) **ME/EME pattern**
Sche teld hym [ how sum-tyme þe Fadyr of Hevyn dalyd to 
hir sowle as pleynly and as veryly as o frend spekyth to 
a-noþer be bodyly spech ];
(Book of Margery Kempe, date: 1450)
finite-V > Sbj Clauses, N = 6131
Sbj > finite-V Clauses, N = 6558
A grammar is re-created, reproduced in every generation.

Adults’ I-Language $\rightarrow$ Adults’ E-Language

Children’s I-Language $\rightarrow$ Children’s E-Language
Yang (2000)

- Selectional advantage of a grammar is based on the ability of children to acquire it.
  - Adapts a classic computational model of learning from Bush and Mosteller (1951), Bush and Mosteller (1958) to syntactic acquisition in a state of grammar variation/change.

- Given a mixture of 2 grammars in the input, \( G_1 \) and \( G_2 \), a child is expected to learn both, assign some probability (weight) to each, and then update these weights throughout the learning process.
If an **ambiguous** input is encountered, i.e. either \( G_1 \) or \( G_2 \) can analyze it, then the child will reward whichever grammar he/she happened to be using at the time.

If an **unambiguous** input is encountered, e.g. only \( G_1 \) could have produced the sentence, then either \( G_1 \) will be rewarded, or \( G_2 \) will be punished. Either way, \( G_1 \) ends up with an augmented weight.

Therefore, the grammar which generates more unambiguous sentences of its own type will have its weight augmented more often.

- And over generations as well.
Fitness(G) = proportion of unambiguously “G”-like clauses that G produces, out of all the clauses it produces.

- Advantage(G₁ over G₂) = Fitness(G₁) – Fitness(G₂)
- If Fitness(G₁) > Fitness(G₂), then G₁ must win in the long run (and vice-versa).
  - The outcome of the change is entirely fixed, once it begins.
  - Yang shows that this is true independently of the initial weights of G₁ and G₂.
  - So the initial frequencies of G₁ and G₂ in the population do not matter.
  - This assumes that Fitness is entirely dependent on how well the child can perceive G₁ and G₂ in the input.
Problem: the loss of AgrSP

- **G₁**: Nominal Subjects in Spec(AgrSP) or Spec(TP), pronominal subjects in Spec(AgrSP)
- **G₂**: Nominal subjects and pronominal subjects in Spec(TP)
- G₁ produces surface patterns with both low and high subjects, while G₂ generates only high subjects (or only 1 position).
- G₁ can analyze all of the outputs of G₂, but G₂ can only analyze some of the outputs of G₁.
- According to Yang’s model, Fitness(G₁) > Fitness(G₂).
Problem: the loss of V2

These are analyzable by either $G_1$ or $G_2$, but $G_2$ can produce only these patterns:

(30) Ac þære ilcan niht... wulfas atugan þa
And the.DAT same night wolves drew the stacan up pin up
“The same night, wolves drew up the pin.”
(Orosius, 10th c.)

(31) Ond eallum þam dagum... he afæste to æfenes And all those days he fasted till evening “and he fasted all days (except Sundays), till the evening” (Bede’s History of the English Church, 11th c.)

Note that the modern English glosses provide the modern English word orders.
Problem: the loss of V2

But these are only analyzable by $G_1$:

(32) bone wæterscipe beworhte se wisa cyning Salomon
the conduit built the wise king Salomon
“The wise King Salomon constructed the conduit”
(Ælfric’s Catholic Homilies, 10th c.)

(33) & of heom twam is eall manncynn cumen
And from them two is all mankind come
“All mankind has come from those two.”
(Wulfstan’s Homilies, 11th c.)
Problem: the loss of AgrSP

- Taking verb movement options into account, the OE-style AgrSP generates:
  - C > finite-V > Adv > Sbj
  - C > finite-V > Sbj > Adv
  - C > Sbj > finite-V > Adv

- Scandinavian-style AgrSP generates:
  - C > Adv > Sbj > finite-V
  - C > Sbj > Adv > finite-V
  - C > Sbj > finite-V > Adv

- No AgrSP (G₂) generates (if we are liberal about verb movement):
  - C > Sbj > Adv > finite-V
  - C > Sbj > finite-V > Adv

- At best, G₂ generates fewer patterns than G₁, and at worst, it generates an exact subset.
Possible Solutions to the Problem

- The analysis of OE is wrong.
  - At least, perhaps there is not true optionality in whether definite subjects surface high or low (aside from the case of topicalization), and we have not yet discovered some conditioning factor.

- Yang’s model is missing something.
  - There could be some sort of inherent bias which causes children to privilege one grammar over another, independent of the input patterns.
  - Haeberli proposed that the loss of AgrSP in English is related to the loss of rich subject-verb agreement.
  - “Rich Agreement Hypothesis” (RAH): rich agreement morphology causes verb-movement, and if rich agreement is lost, so is the movement (and maybe the functional projection).
Possible Solutions to the Problem

- Bobaljik (2003) argues against the strong form of RAH, and argues for a weaker version: rich morphology entails verb movement (and the presence of relevant functional projections), but not vice-versa.

- Bobaljik (2003): “morphology may provide a cue to the child about the Split IP Parameter, but the morphology does not cause the syntax”.
  - If the morphology is lost, then there is less evidence to the learner that the relevant functional projections are there.
  - Under a strict Yang, this should not matter.

- **Possible Revision to Yang**: given ambiguous data between $G_1$ and $G_2$, if $G_2$ maps straightforwardly to the morphology and $G_1$ does not, then some of the ambiguous data is counted towards $G_2$. 
The two-subject analysis of Germanic makes a correct diachronic prediction for English: Adv > Sbj orders decline as V2 declines.

- The pre-Sbj adverbs are lost even in Infl/Tense-final clauses, which is a specific prediction of the two-subject analysis, but not predicted by any other theory of the loss of V2.
- Although CP-recursion produces different frequencies of pre-Sbj adverbs in different clause types, these all decline at the same rate. (Though this needs to be investigated further.)
- Adv > Sbj orders decline regardless of where the verb surfaces.

However, the two-subject analysis predicts that the Old English system cannot lose to the modern English system, according to Yang’s acquisition model.
Conclusions and Future Research

- Strongly suggests that learners do not respond to the input in a naive way, but may make systematic mistakes.
  - I have suggested a form of the RAH in which the systematic mistakes are in favour of a one-to-one syntax-morphology mapping.

- The acquisition model needs to be tested further by empirical studies on syntactic change which estimate the evidence to the learner for each variant.
  - The loss of V2 should be formally modeled, including the restatement of the RAH.

- Even if we were not interested in language change, it is an important experimental domain for linguistic theory in general (e.g. syntactic theory, acquisition).
  - This level of quantitative detail is only possible with accurate parsed corpora.
References I


