Pairing word order with headedness in the history of English: a corpus-based analysis

Javier Pérez-Guerra (jperez@uvigo.es)

Language Variation and Textual Categorisation Research Group
University of Vigo

ISLE 2011, Boston, June 20

Outline

• goal
• background
• description of the corpora: complementation and adjunction strategies
• analysis of the data:
  – Complements-first
  – End-weight
• concluding remarks

Goal

• Goal: check the distributional tendencies affecting the placement of modifiers/adjuncts and complements/arguments in noun phrases, adjectives phrases and verb phrases

• Examples:
  1) Now I will deal [with the construction] [in a way which will lead to odd results].
  1) Now I will deal [in a way which will lead to odd results] [with the construction].
  2) the author [of this book] [from London]
     [N.A. the author is from London]
  2) the author [from London] [of this book]
  3) keen [on music] [to a large extent]
  3) keen [to a large extent] [on music]

Background

• Complements
  – semantically selected or subcategorised
  • Matthews (2007: 187): “unit in a construction either required or specifically taken by an individual member of a lexical category”
  • Matthews (1981: 124-127): impossibility of dropping (if dropped, then latent)
  • exclusion when the pattern is saturated
  – syntactic dependencies:
  • lexical restrictions or formal determination
    (Greenbaum et al. 1996: 76)
    [deal, compliance] + with-PP
    [assume, certain, hypothesis] + that-clause

• Adjuncts
  – loose semantic connection between the adjunct and the head => not required
Background

- Distribution of complements and adjuncts:
  - syntactic explanation:
    * Quirk et al. (1985: 49-50): 'Complements first'
    * Hawkins (2007): 'Arguments precede X'
  - processing explanation:
    * Quirk et al. (1985: 1398): End-weight
    * Hawkins’ (2004) ‘Minimize Domains’ or MiD:
      Given two or more categories A, B, […] related by a grammatical rule R of combination and/or dependency, the human processor prefers to minimize the distance between them within the smallest surface structure domain sufficient for the processing of R. (Hawkins 2000: 234)

Hawkins (2007) hypothesises that MiD is relevant especially to examples of complementation.

The corpora

- Old English: 1.5+ million words (Old English section of the Diachronic Part of the Helsinki Corpus of English Texts, with certain additions, c750–)
- Middle English: 1,155,965 words (Middle English section of the Diachronic Part of the Helsinki Corpus of English Texts, with certain additions and deletions, 1150–1500)
- Early Modern English: 1,737,853 words (the Helsinki directories of the Penn-Helsinki Parsed Corpus of Early Modern English plus two supplements; 1500–1710)
- (Late) Modern English: 948,895 words (1700–1914)

Examples:

(1) Now I will deal [with the construction] [in a way which will lead to odd results].

(1) is claimed to be a better performance solution, on processing grounds [MiD, end-weight], than (1) because of the amount of structure which has to be processed (between the head category and the second constituent in the [local phrase]).
but, if you approve of this, if you please to lett me know y=r= pleasure, I will tell it M=r= Isaac. (ANHATTON-E3-H,2,214.41)

The data

Examples of queries

- **Verb-governed contexts: V + adjunct + complement**
  verb group immediately precedes an adjunct, and the adjunct immediately precedes an object

  **OE:**
  node: IP*
  query: (VB* iprecedes W*|NUMP*|QP*|PP|ADJ*|ADV*|INTJP|XP|NP-ADV*|NP-LOC*|NP-VOC*)

  **ME, EModE and ModE:**
  node: IP*
  query: (VB* iprecedes W*|QP|PP|RRC|ADJ*|ADV*|CP-*) (NP-OB1|NP-OB2|NP-LOC*|NP-VOC*)

- **Verb-governed contexts: V + complement + adjunct**
  verb group immediately precedes an object, and the object immediately precedes an adjunct

  **OE:**
  node: IP*
  query: (VB* iprecedes NP-ACC*|NP-DAT*|NP-LOC*) AND (NP-ACC*|NP-DAT*|NP-LOC* iprecedes W*|NUMP*|QP*|PP|ADJ*|ADV*|INTJP|XP|NP-ADV*|NP-LOC*|NP-VOC*)

  **ME, EModE and ModE:**
  node: IP*
  query: (VB* iprecedes NP-OB*) AND (NP-OB* iprecedes W*|QP|PP|RRC|ADJ*|ADV*|CP-*) (NP-OB1|NP-OB2|NP-LOC*|NP-VOC*)
The data

**OBJ + QP**
so this time will trouble $y=r= Losp$ no more $w=th= y=r= most obedient, duttyful daughter, A. Nottingham. (ANHATTON-E3-H,212.269)

**OBJ + PP**
I thought wilde take some spendyng money wyth me (MERRYTAL-E1-H,31.148)

**OBJ + RRC**
If you have provided those Stockings before spoken of I pray send them. (STRYPE-E3-H,183.69)

**OBJ + ADJ**
and cut it not so close to the Body as to hurt it, nor yet so long that it be a Stump, (LANGF-E3-H,122.269)

The data

**OBJ + ADV**
But my Brother understood the matter aright (HOXINDEN-1660-E3-H,280.162)

**OBJ + IP-SUB**
Contrariwise, there be others againe, that will not endure to read a short composition, bee it never so well approoued: (CLOWES-E2-P1,8.69)

**OBJ + NP-ADV**
I 'll ply him that way, (FARQUHAR-E3-H,9.326)

The data

**OBJ + QP**
so this time will trouble $y=r= Losp$ no more $w=th= y=r= most obedient, duttyful daughter, A. Nottingham. (ANHATTON-E3-H,212.269)

**OBJ + PP**
I thought wilde take some spendyng money wyth me (MERRYTAL-E1-H,31.148)

**OBJ + RRC**
If you have provided those Stockings before spoken of I pray send them. (STRYPE-E3-H,183.69)

**OBJ + ADJ**
and cut it not so close to the Body as to hurt it, nor yet so long that it be a Stump, (LANGF-E3-H,122.269)

The data

**OBJ + ADV**
But my Brother understood the matter aright (HOXINDEN-1660-E3-H,280.162)

**OBJ + IP-SUB**
Contrariwise, there be others againe, that will not endure to read a short composition, bee it never so well approoued: (CLOWES-E2-P1,8.69)

**OBJ + NP-ADV**
I 'll ply him that way, (FARQUHAR-E3-H,9.326)
The data

adjunct + CP-THT
And there was a feeling by no means uncommon, and very deadly, that India would be lost for ever, and with it all the glory of England. (TROLLOPE-1882,177.356)
yea I am sorie, with all my harte, that they be giuen no more to riding, then they be: (ASCH-E1-P1,10R.186)

The data

• Adjective-/Noun-governed contexts: A/N + complement + adjunct
  head [adjective/noun] immediately precedes a complement [that- or infinitive clause], and the complement immediately precedes an adjunct

EModE and ModE:
  node: NP*
  query: (NP* iDoms CP-THT)
  AND (NP* iDoms "P")
  AND (CP-THT iprecedes "P")

node: NP*
query: (NP* iDoms IP-INF)
  AND (NP* iDoms "P")
  AND (IP-INF iprecedes "P")

The data

IP-INF + than/that(so...that)

For I hadde leyer: hafe bene in that payne to demysdaye than hafe come to hevene (CMJUÑOR,55.163)
than have come to heaven

& beon geornran +t+at we Godes bebodu    healdan, +tonne we & be more eager that we God's command keep than we urne teorum yeurncom. (coblick,HomB_10_[BlHom_3]:33.123.444)

The data

CP-THT + than

For we are no less certain that there is a great Town called Constantinople, the seat of the Ottoman Empire, than that there is another called London. (BÜRNETROC-E3-P1,79.231)

& beon geornran +t+at we Godes bebodu    healdan, +tonne we & be more eager that we God's command keep than we urne teorum yeurncom. (coblick,HomB_10_[BlHom_3]:33.123.444)

The data

CP-THT/IP-INF + sentential relative clause

There is a wise saying that nine-tenths of the noble work done in the world is drudgery, which is often misused as if it meant that nine-tenths of the drudgery done in the world is noble work.

[King James sent a Person down to him, with] Offers to mitigate his Fine upon Conditions of ready Payment, to which his Lordship reply'd, that if his Majesty pleas'd to allow him a little longer time, he would rather chose to play double or quit with him: (CIBBER-1740,44.134)

[he also expressed an opinion that mulattoes inherited the vices of both races - a maxim which I had heard often enough before,]

[READE-1863,225.616]
Analysis of the data

(i) Syntactic principle ‘Complements-first’

Analysis of the data

Verb-governed contexts:

<table>
<thead>
<tr>
<th></th>
<th>OE</th>
<th>ME</th>
<th>EModE</th>
<th>ModE</th>
</tr>
</thead>
<tbody>
<tr>
<td>compl_first</td>
<td>8270</td>
<td>17805</td>
<td>7859</td>
<td>13084</td>
</tr>
<tr>
<td>compl_last</td>
<td>3290</td>
<td>6790</td>
<td>1995</td>
<td>2579</td>
</tr>
</tbody>
</table>

Statistical significance for full variation: yes (P<.0001)
Statistical significance for variation OE>ME: yes (P=0.0949)
Statistical significance for variation ME>EModE: yes (P<.0001)
Statistical significance for variation EModE>ModE: yes (P<.0001)

Analysis of the data

Adjective-governed contexts:

<table>
<thead>
<tr>
<th></th>
<th>OE</th>
<th>ME</th>
<th>EModE</th>
<th>ModE</th>
</tr>
</thead>
<tbody>
<tr>
<td>compl_first</td>
<td>10</td>
<td>20</td>
<td>19</td>
<td>32</td>
</tr>
<tr>
<td>compl_last</td>
<td>0</td>
<td>31</td>
<td>29</td>
<td>28</td>
</tr>
</tbody>
</table>
Analysis of the data

• Adjective-governed contexts:

- Statistical significance for variation OE-ME: n.a.
- Statistical significance for variation ME-EModE: n.s. (P=0.2191)
- Statistical significance for variation ME-ModE: n.s. (P=0.8625)

Noun-governed contexts:

<table>
<thead>
<tr>
<th></th>
<th>OE</th>
<th>ME</th>
<th>EModE</th>
<th>ModE</th>
</tr>
</thead>
<tbody>
<tr>
<td>compl first</td>
<td>1</td>
<td>20</td>
<td>26</td>
<td>15</td>
</tr>
<tr>
<td>compl last</td>
<td>22</td>
<td>66</td>
<td>105</td>
<td>107</td>
</tr>
</tbody>
</table>

- Statistical significance for variation OE-ME: n.s. (apparently significant)
- Statistical significance for variation ME-EModE: n.s. (P=0.8415)
- Statistical significance for variation EModE-ModE: n.s. (P=0.1444)
- Statistical significance for variation ME-ModE: n.s. (P=0.2311)
Analysis of the data

• **VPs:**
  - most are *complement-first*
  - ME>ModE: (statistically significant) *increase of complement-first* VPs
  - pivotal period: ME>EModE
• **APs:**
  - OE: all the examples are complement-first (few data)
  - after ME: half are *complement-first*
  - pivotal period: ME
• **NPs:**
  - most are *complement-last* (ME>ModE variation is not significant)
  - pivotal period: ME

So...

• Heads and compliance with Complements-first:
  - VPs clearly comply with Complements-first.
  - NPs do not comply with Complements-first.
  - APs occupy an intermediate position since half comply with Complements-first.
  - *connection head status ≠ Complements-first*
• Chronology of variation:
  - VPs: increase of complement-first from ME onwards
  - NPs: complement-last in all periods; quantitative stability from ME onwards
  - APs: intermediate situation from ME onwards

(ii) Processing principle ‘End-weight’

Complement-first VPs: (no. of words, outliers excl.)

- End weight: × in all periods
Analysis of the data
Complement-last VPs: (no. of words, outliers excl.)
- End-weight not evident, more noticeable in ModE

Analysis of the data
Complement-first APs: (no. of words, outliers excl.)
- End-weight is not clear in the periods investigated

Analysis of the data
Complement-last APs: (no. of words, outliers excl.)
- End-weight: ✓, less noticeable across time

Analysis of the data
Complement-first NPs: (no. of words, outliers excl.)
- End weight: ✓ in all periods; more noticeable in ModE
**Analysis of the data**

Complement-last NPs: (no. of words, outliers excl.)

- End weight: ✓ in all periods, more evident from EModE onwards

---

**Conclusion**

- VP s:
  - complement-first: do not observe End-weight
  - complement-last: do not clearly observe End-weight
- AP s:
  - complement-first: do not clearly observe End-weight
  - complement-last: observe End-weight (in general)
- NP s:
  - complement-first: observe End-weight, and this is more evident across time
  - complement-last: observe End-weight, and this is more evident across time

---

**Concluding remarks**

- Competing forces: syntax (Complements-first) and processing (End-weight):
  - connection complement distribution and head category: the more verbal the head is, the more likely its compliance with Complements-first
  - Headedness prototypicality hierarchy: V>A>N:
    - frequency: the number of bare (intransitive) nouns in NPs outweighs the number of bare (intransitive) VPs
    - paradigmatic functionality: verbs take part in a wider range of complementation patterns than nouns
    - ellipsed nouns in NPs are more common than ellipsed verbs in VPs
  - morphological possibilities (Givón 1993: morphology implies syntactic integration)
- APs: intermediate position since adjectives are more verbal than nouns
Concluding remarks

• Competing forces: syntax (Complements-first) and processing (End-weight):
  – connection complement distribution and head category: the more verbal the head is, the less likely its compliance with End-weight
  – when syntax fails, End-weight is crucial and justifies complement-last examples with the three categories (radically with nouns), especially in ModE
  – when syntax is at work (Complements-first), End-weight is not responsible for the distribution of complements and adjuncts since End-weight is not necessary

Concluding remarks

• So...
  “the biggest single predictor of relative orderings […] is syntactic weight” (Hawkins 2000: 232)
  is not strictly correct since, according to the data, syntax (Complements-first) seems to be a bigger predictor.
  “[in general the light-heavy distribution [end-weight] is no longer a major factor in English word order” (Traugott 1992: 276)

References

References


Pérez-Guerra, Javier and Ana E. Martínez-Insua (2010a) "Diachrony and word order hand in hand: on complementation/adjunction performance solutions in English". Delivered at ICAME 31, Giessen.

Pérez-Guerra, Javier and Ana E. Martínez-Insua (2010b) "Word order and performance optimality get on well in the history of English: on complement+adjunct and adjunct+complement solutions". Delivered at SLE 43, Vilnius.


