SYNTACTIC MICROVARIATION IN ROMANCE: BRIDGING SYNCHRONY AND DIACHRONY: THE CASE OF SI

1. Background, Aims and Methods

1.1. Aims of the Paper

• Numerous wide-ranging, data-rich studies of Medieval Romance languages have emerged in the last three decades, drawing on an old intuition that the modern vs. medieval dichotomy is morphosyntactically important (Tobler 1875; Mussafia 1888:145-146; Meyer-Lübke 1889:760-840; Delbrück 1900:375-395; Schoch 1912; Foulet 1928:§389-406; Sorrento 1950).

• Significant puzzles remaining:
  - The extent to which it is possible to identify variation between varieties, akin to the very rich patterns of variation observable in contemporary Romance varieties (i.e. D’Alessandro, Ledgeway and Roberts 2010; Olviéri & Sauzet 2016; Kato & Ordóñez 2016; Panâ Dindelegan, Dragomirescu, Nicula, & Nicolae, 2018).
  - The degree to which we can observe changes within the medieval period (Roberts 1993; Vance 1995; Poletto 1995; Labelle & Hirschbühler 2005, 2012, 2017; Wolfe 2016b), rather than broader brush ‘macro’ distinctions between medieval and modern varieties.
The focus of this paper:

- Discussion of the syntax of a ubiquitous part of Medieval Romance syntax, the particle *si*.

- This seems like a parochial area of interest, but actually intersects a number of the core changes in clausal syntax taking place in this period, namely changes in the null argument system, changes affecting verb second, the syntax of subject positions and the nature of the left periphery.

- **The main thrust of the argument**: the interconnectedness of changes affecting *si* which have not generally been noticed previously and the fact that the diachronic progression of *si* shows several distinct stages.

1.2. Empirical Background on Old French

1.3.1. Verb Movement and the Left Periphery

- Distributional differences between finite verbs and adverbials in Modern English vs. French > finite verb in French but not English moves out of the (v)-VP and into the clause’s inflectional layer (Kayne 1991:648-657; Chomsky 1995; Roberts 2001:121-123, 2010b:258-271; Rowlett 2007:106-107):

  (1) a. John always *kisses* Mary [Modern English]
      b. Jean *embrasse* toujours Marie [Modern French]

  (2) a. [TP John [LAS always [v-VP *kisses* Mary]]]
      b. [TP Jean *embrasse* [LAS toujours [v-VP *embrasse* Marie]]]

- But this ‘classic account’ is a simplification of the empirical facts:

This is an important claim if French is a V2 language:

- Classic account of V2 > Verb and some XP move out of the clausal core targeting C(omplementiser) and SpecCP respectively.


- If ‘V-to-T movement’ is a cover term for a range of verb movement operations within the inflectional layer, where does that leave the V2-related ‘V-to-C’ movement?

  - In a richly articulated let periphery (Rizzi 1997, 2004, 2006, 2010, 2013; Benincà & Poletto 2004; Ledgeway 2010:51; Haegeman 2012) this is non-trivial:

\[
\begin{align*}
\text{Frame } HT, \text{ AdvScene-Setting } & \text{ Force Complementiser}_1, \text{ [Topic CILD, Aboutness Topic } \text{ Focus} \text{ FocusContrastive, QuantifierIndefinite FocusInformation [Fin Complementiser}_2 [\text{TP...}]]}\end{align*}
\]

- This has led to different approaches to Old French V2:

  - V-to-Fin (Salvesen 2011, 2013:134)


  - V-to-Force (Rouveret 2004; Wolfe 2016a) to be the locus of V2 have all been proposed.

- Compare different ways of deriving (4):

\[
\text{(4) messe e matines ad li reis escultet mass and matins have.3SG the king hear.PTCP}\]

‘The King has attended mass and matins’ (Roland 11, 139)

---

1.3.2. Subject Positions

  - Non-pronominal postverbal subjects occur to the right of past participles (Rizzi & Roberts 1989; Hulk & Pollock 2001:5; Sheehan 2006:99, 2010;§3.4):

  (6) J’exige que soit éliminée cette solution
  I=require.1SG that be.3SG.SBJV eliminated.PTCP this solution
  ‘I want this solution to be eliminated’ (Hulk & Pollock 2001:5)

  - Postverbal subjects occur to the left of past participles, infinitives and VP adverbs.3, so unambiguously show that earlier French varieties licensed V-to-C movement:

  (7) a. Par tanttes teres ad sun cors travellent
      over so.many lands have.3SG his body suffer.PTCP
      ‘His body has suffered across so many lands’ (Roland, 540)
  b. Et quant il furent ajusté ensemble, si furent li noir vaincuz
     and when they were joust.PTCP together si be.3PL the blacks defeat.PTCP
     ‘And when they met in the jousting, the blacks were defeated’ (Graal 182, 40)
  c. Et enqui Guillelme de Chanlite brisié le braz d’une pierre
     and there have.3SG Guillelme de Chanlite break.PTCP the arm of=a stone
     ‘and there a stone broke Guillelme de Chanlite’s arm’ (Villehardouin1 167)

- Old French also licenses postverbal subjects akin to those in Modern French:

  (8) a. A cele foiz ne furent mie venu tuit li baron
       at that time NEG be.3PL NEG come.PTCP all the barons
       ‘At that point, not all the lords had arrived’ (Villehardouin1 79)
  b. Si est avenue la semblance de l’évangile...

---

2 For more descriptive discussion of this construction in Old French, see Einhorn (1974:115), Jensen (1990:336) and Muller (2009:246) among others.

3 These non-finite verbal forms demarcate the left edge of the v-VP complex (Cinque 2001, 2006:12; Cardinaletti & Shlonsky 2004:525; Ledgeway in press c).
‘What is said in the Gospel has happened…’ (Graal 166, 17)

  

1.3.3. Null Arguments

- **Standard Modern Metropolitan French > non-Null Subject Language** (Rizzi 1986:400; Roberts 2010a:307, 2010b:§3.5, 2014b:§2-4) vs. **Old French > form of Null Subject Language:**
  
  - System until 1200:
    
    
    - Null subjects occur in a range of embedded clauses (Adams 1987a; Dupuis 1988; Vance 1988; Roberts 1993:136-147) (9b):
      
      (9) a. *Getet* le a terre…
      
      throw.3SG.PST=it to ground
      
      ‘He throws it to the ground…’ (*Roland* 464)
      
      b. *Se veïssum* Rollant einz qu’il fust mort
      
      If see.1PL.IMP.SBJV Roland before that=he be.3SG.PST dead
      
      ‘If we saw Roland before he was dead…’ (*Roland* 1804)
- Change ~1180-1220:
  - Heavy restrictions on embedded null subjects (Adams 1987b:3, 1987c:Ch.1; Roberts 1993:139; Vance 1997:Ch.5; Poletto 2013:160)

1.3.4. Left-Peripheral Constituents and Particles

- Old French preverbal field not specialised for subjects:
  - A point of agreement between traditional, descriptive and formally-oriented approaches:


- Presence of a direct object triggers Verb-Subject ‘inversion’ (4, 8b, 10b):

(10)a. Li reis Marsilie esteit en Sarraguce
    the king Marsile be.3SG.PST in Zaragoza
    ‘King Marsile was in Zaragoza’ (Roland 10)

b. Ce oïrent el palés maint
    that hear.3PL.PST in-the palace many
    ‘Many in the palace heard this’ (Charrette, 80)

c. et a chascun rendra ce qu'il avra deservi
    and to each give.3SG.FUT that which-he have.3SG.FUT deserve.PTCP
    ‘and he’ll give each what he has earned’ (Graal 163, 32-33)

- Particles can appear in the preverbal field:

These often elude satisfying translation, description and formal analysis.

1.4. The Database

- Hand annotated 12th and 13th century corpus, including both poetry and prose:
  - Texts taken from philologically reliable editions in the Base de Français Médiéval.
  - Roland, Thèbes, Eneas and Villehardouin > all si-clauses extracted.
  - Charrette, Clari and Queste > ~ 300 analysed.
  - Annotated for a dozen syntactic and discourse-pragmatic variables.

### Table 1. Textual Database

<table>
<thead>
<tr>
<th>Text</th>
<th>Composition</th>
<th>Poetry/Prose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chanson de Roland</td>
<td>c. 1100</td>
<td>Poetry</td>
</tr>
<tr>
<td>Roman de Thèbes</td>
<td>c. 1150</td>
<td>Poetry</td>
</tr>
<tr>
<td>Eneas</td>
<td>c. 1155</td>
<td>Poetry</td>
</tr>
<tr>
<td>Chevalier de la Charrette ou Lancelot</td>
<td>c.1177-1181</td>
<td>Poetry</td>
</tr>
<tr>
<td>Conquête de Constantinople (Villehardouin)</td>
<td>c. 1199-1213</td>
<td>Prose</td>
</tr>
<tr>
<td>Conquête de Constantinople (Clari)</td>
<td>After 1205</td>
<td>Prose</td>
</tr>
<tr>
<td>Queste del Saint Graal</td>
<td>c. 1225-1230</td>
<td>Prose</td>
</tr>
</tbody>
</table>

2. Previous Approaches to Si

2.1 Si as an Adverb

- Si patterns with other clause-initial adverbs and serves to structure the discourse:
• Advantages of this account:
  - *si* derives from Latin temporal deictic adverb, *sic* ‘thus’ (Salvi 2004:15 n.66; Rosén 2005:228-230; Pinsker 2006:65, 107-108; Ledgeway in press a, in press b), itself derived from a manner adverb ‘like this’:

(11) *et sic plecaremus nos ad montem Dei* and thus arrive.1PL.PST.SBJV we at mountain.ACC God.GEN ‘and we thus approached the mountain of God’ (Peregrinatio 2, 4)

  - *si* triggers inversion like other adverbials (Foulet 1928:301-303; Jensen 1990:472)

(12): *Si luisoient ja les estoiles*  
*si* shine.3PL.PST already the stars

‘The stars were already shining’ (Charrette 31a, 1015)

• Unanswered questions:
  - Sheer abundance of *si*:

    ‣ Foulet (1928:300): in early French ‘there is hardly a page where it doesn’t appear several times’ (Marchello-Nizia 1985:2; Fleischman 1991:261; Buridant 2000:$408).  

    ‣ This corpus: occurs 20 lines in the Roland and every 24 lines in Charrette.


    ‣ Frequency increases significantly in the 13th century

    ‣ Why does an adverbial meaning ‘thus, so’ need to occur after an initial clause 177 times in La Queste?:

(13) *Quant il vint en la valee, si comenza a penser mout durement* when he come.3SG.PST in the valley SI begin.3SG.PST to think.INF very hard

‘When he came to the valley, he began to think very hard’ (Graal, 194d, 8)

---

4 See also Van Reenen & Schøsler (1992:115, 2000:97) who show that the overall frequency shows regional variation, with *si* more frequent in northern texts.
- **Si** co-occurs with adverbials encoding exactly the notions of temporal continuity and succession that **si** is meant to encode! (Fleischman 1991:263; Schøsler & Van Reenen 1996:640-641 and Van Reenen & Schøsler 2000:79-80):

(14) quant a sa proie devoree, donc si *s'endort* gole baee when has.3SG its prey devour.PTCP thus **si** REFL.CL=sleep.3SG mouth open
“When it has devoured its prey, it then goes to sleep with its mouth wide-open” (*EneasI*, 489-490)

2.2. **Si** as a Topic Marker

- A popular analysis where **si** encodes continuity of Topic or more explicitly is a marker of ‘same-subject reference’ (Fleischman 1991, 1992).


- Accounts for a lot of the data like (15):

(15) **Li vasslés entend** bien que **li** empereres li donnoit **boin**
the vassal understand.3SG.PST well that the emperor him.CL=give.3SG.PST good
counsel ; **si** *s'atorna* au plus belement qu'il peut, **si**
counsel **si** REFL.CL=prepare.3SG.PST at most best that=he can.3SG **si**
s'en **vint** avec les messages
**REFL.CL=LOC.CL=come.3SG.PST** with the messengers
‘The servant understood clearly that the emperor had given him good advice. He prepared the best he could and came with the messengers…’(*Clari* 30, 31)


• But this analysis is not without problems:

- Unexpected co-occurrence with a readily ACTIVE or ACCESSIBLE (Lambrecht 1994:165) subject, a construction which is not exceptional (Foulet 1928:301; Marchello-Nizia 1985:7; Fleischman 1991:265-266,271; Ferraresi & Goldbach 2002:12; Salvesen 2013:156; Buridant 2000:509):

  (16) ceste ville **si est** mult riche et mult bien garnie de toz biens
  this town **SI** be.3SG very rich and very well stock.PTCP of all goods
  ‘and this town is very rich and well stocked with provisions’ (Villehardouin 86, 4)

- Unexpected co-occurrence with new information subjects, which is entirely unexpected if **si** is a ‘same-subject marker’ (Fleischman 1992):

  (17) a. E **si i furent e Gerin e Gerers**
  and **SI LOC.CL=be.3PL.PST** and Gerin and Gerers
  ‘And Gerin and Gerer were there’ (Roland 107)

  b. Quant je me fui endormiz **si m’avint une avision**
  when I **REFL.CL=be.1SG.PST asleep** SI **REFL.CL=come.3SG.PST a vision**
  merveilleuse
  marvellous
  ‘When I was asleep a marvellous vision came to me’ (Graal 197a 23-24)

2.3. **Si** as a Focus or Fin Expletive

• Many formal analyses view **si** as a **form of last-resort mechanism to satisfy some part of the V2 constraint**:

  - Merger of a Head **si** is an alternative (Ledgeway 2008:452-465) or complementary (Ferraresi & Goldbach 2002:18-23) strategy to V-to-C movement.
Problems:

- si co-occurs with Germanic inversion, showing that V-to-C movement does obtain in the presence of si. It therefore cannot be a Head (pace Ledgeway 2008:444-445) and must be a phrasal category.:5

(18) Quant eles furent faïtes, si fu la chose teüe que on when they be.3PL.PST do.PTCP si be.3SG.PST the thing keep-quiet.PTCP that one go.3SG.IMPV to Babylon
‘When they [the covenants] were completed, it was kept secret that we would go to Babylon’ (Villehardouin1, 30, 3)

Advantages:


But questions remain:

- What actually is ‘last-resort’?

- How do we reconcile such an analysis with the clear discourse-pragmatic effects si has in certain texts?

---

5 As Ledgeway (2008:439, n.3) highlights, Ferraresi & Goldbach’s (2002) claim that si is merged in Fin in addition to the finite verb makes the incorrect prediction that the finite verb will precede si, following left adjunction to the particle in Fin. As such I do not pursue this account here.

3. Old French sì - Continuity and Variation

- **Table 1** overleaf summarises the main findings of the corpus analysis, showing that the data are far from the homogeneous many treatments of sì would lead us to expect.

- In what follows, I outline some of the main observations that emerge from the corpus analysis, both in terms of points of continuity and variation.
<table>
<thead>
<tr>
<th></th>
<th>Et + SI</th>
<th>SI</th>
<th>DO DP + SI</th>
<th>Subject + SI</th>
<th>Adv + SI</th>
<th>PP + SI</th>
<th>Clause + SI</th>
<th>XP + XP + SI</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Roland, 1100</strong></td>
<td>13</td>
<td>6.47%</td>
<td>138</td>
<td>68.66%</td>
<td>2</td>
<td>1.00%</td>
<td>22</td>
<td>10.95%</td>
<td>8</td>
</tr>
<tr>
<td><strong>Thèbes 1150</strong></td>
<td>28</td>
<td>13.02%</td>
<td>119</td>
<td>55.35%</td>
<td>2</td>
<td>0.93%</td>
<td>9</td>
<td>4.19%</td>
<td>11</td>
</tr>
<tr>
<td><strong>Eneas, 1155</strong></td>
<td>17</td>
<td>13.60%</td>
<td>69</td>
<td>55.20%</td>
<td>1</td>
<td>0.80%</td>
<td>2</td>
<td>1.60%</td>
<td>3</td>
</tr>
<tr>
<td><strong>Charrette, 1177-1181</strong></td>
<td>63</td>
<td>21.43%</td>
<td>135</td>
<td>45.92%</td>
<td>0</td>
<td>0.00%</td>
<td>14</td>
<td>4.76%</td>
<td>19</td>
</tr>
<tr>
<td><strong>Villehardouin 1199 - 1213</strong></td>
<td>32</td>
<td>7.98%</td>
<td>156</td>
<td>38.90%</td>
<td>3</td>
<td>0.75%</td>
<td>32</td>
<td>7.98%</td>
<td>23</td>
</tr>
<tr>
<td><strong>Clari 1205</strong></td>
<td>30</td>
<td>10.53%</td>
<td>113</td>
<td>39.65%</td>
<td>0</td>
<td>0.00%</td>
<td>4</td>
<td>1.40%</td>
<td>26</td>
</tr>
<tr>
<td><strong>Questa 1225 - 1230</strong></td>
<td>15</td>
<td>5.00%</td>
<td>97</td>
<td>32.33%</td>
<td>0</td>
<td>0.00%</td>
<td>6</td>
<td>2.00%</td>
<td>0</td>
</tr>
</tbody>
</table>
3.1. si-initial and Clause + Si configurations

- All texts show orders where:

  1: **si is the only preverbal constituent** (cf. also Marchello-Nizia 1985:Ch.1 and Van Reenen & Schøsler 1992:117-122, 2000:104-105).

  2: **si is the only preverbal constituent and co-occurs with a discourse-OLD or NEW postverbal subject** (on which see also Fleischman 1991:271-274; Van Reenen & Schøsler 1992:121; Vance 1997:204):

(19) Et *si y fu* li abbes de Los en Flandres, qui *estoit de maisons* and *si LOC.CL=be.3SG.PST* the abbot of Los in Flanders *who be.3SG.PST from houses de l’ordre de Chistiax* of the-order of Chistiax

‘And the abbot of Los in Flanders, who is from the house of the order of Chistiax was there’ (Clari 1, 27)


(20) Dolent *estoient et pensif // si descendirent* souz un*if* be.PST and thoughtful *si go-down.3PL.PST* under a *yew* ‘Feeling sorrowful and deep in thought, they went [and sat] under a yew tree’ (Thèbes2 6789-6790)

(21) Quant il *vinrent la, si parlerent au marchis...* when *they come.3PL.PST* there, *si speak.3PL.PST* to-the marquis ‘When they arrived there, they spoke to the marquis...’ (Clari 5, 4)

One of the most striking corpus findings is the concomitant decline in the relative frequency of the si-initial pattern and rise in the Clause + Si configuration (see Figure 1).
This clearly warrants discussion:

- In the paradigm case $SI$ is often described as a competing strategy to the use of a Null Subject (Marchello-Nizia 1985:164-166; Fleischman 1991:258; Buridant 2000:§411).

- We therefore might expect the exact opposite pattern to obtain (i.e. a rise in the $SI$-initial orders) as Null Subjects become near-entirely absent in initial position in 13th century French prose (Simonenko & Hirschbühler 2012:30; Wolfe 2016b).

- The rise in clause-initial configurations also doesn’t follow naturally from any of the existing accounts of $SI$ outlined above.

![Figure 1. SI-Initial vs. Clause + SI Configurations](image-url)
3.2. Subjects

- In line with Fleischman (1991:265) and Marchello-Nizia (1985:169-178) there is a gradual increase in the proportion of SI-clauses which feature an overt postverbal subject in Figure 2, but the diachronic progression isn’t entirely neat (Figure 2).

- In terms of discourse status, Roland, Thèbes and Charrette show a tendency for postverbal subjects to be informationally new, whilst Clari and Queste show more numerous discourse-old postverbal subjects (Table 3).

| Figure 2. Overt Postverbal Subjects |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                  | Roland          | Thèbes          | Eneas           | Charrette       | Villehardouin   | Clari           | Queste          |
| Postverbal Subjects | 19              | 10              | 11              | 18              | 50              | 71              | 38              |
| Total SI Clauses   | 201             | 215             | 125             | 294             | 401             | 285             | 300             |
| % Postverbal Subjects | 9.45%           | 4.65%           | 8.80%           | 6.12%           | 12.47%          | 24.91%          | 12.67%          |

Table 2. Proportion of SI Clauses with Overt Postverbal Subjects
• Preverbal subjects also show some variation:

  - *Roland, Eneas, Clari* and *Queste* > SUBJECT + SI clauses account for between 1 and 1.6% of all attestations.

  - In *Thèbes* this is 4.19%, 4.76% in *Charrette* and a high of 7.98% in *Villehardouin* (Table 1). It does not appear coincidental that these are also the three texts which Table 3 shows have the high proportion of NEW information postverbal subjects.

3.3. PP + SI CONFIGURATIONS

• All the texts show PP-initial SI clauses.

• Adverbial and locative PPs (22, 23) are more uniformly attested than Indirect Object PPs which are only found in *Roland, Eneas* and *Charrette*, three of the 12th century texts:

(22) De cele estoire si *fu chevetaines* Johans de Neele of that fleet SI be.3SG.PST chief Johan de Neele 'Johan de Neele was head of this fleet’ (*Villehardouin* 3, 48)

(23) Sur l’erbe verte si *est caeit envers* on the-grass green SI be.3SG fall.PTCP on-back 'He has fallen down on his back on the green grass’ (*Roland* 2269)

(24) Au roi, si li *comance a dire* to-the king SI him.CL=start.3SG to say.INF 'He starts to say to the king…’ (*Charrette* 27d, 85)

Table 4. PPs

<table>
<thead>
<tr>
<th></th>
<th>PPAdverbial</th>
<th>PPLocative</th>
<th>PPIndirectObject</th>
<th>Total SI Clauses</th>
<th>% PPs</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Roland</em>, 1100</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>201</td>
<td>5.47%</td>
</tr>
<tr>
<td>Thèbes 1150</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>215</td>
<td>1.40%</td>
</tr>
<tr>
<td><em>Eneas</em>, 1155</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>125</td>
<td>3.20%</td>
</tr>
<tr>
<td>Charrette, 1177-1181</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>294</td>
<td>1.02%</td>
</tr>
<tr>
<td>Villehardouin 1199 - 1213</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>401</td>
<td>3.99%</td>
</tr>
<tr>
<td>Clari 1205</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>285</td>
<td>0.35%</td>
</tr>
<tr>
<td>Queste 1225 - 1230</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>300</td>
<td>1.33%</td>
</tr>
</tbody>
</table>
3.4. Adverbial + Si Configurations

- Adverbials are found before si in all texts except La Queste.

- However in the Roland these can be VP-level manner adverbials (25) (Marchello-Nizia 1985:157-160), whereas in all other texts they serve a frame-setting function, anchoring the clause in terms of locative or temporal coordinates (Poletto 2000: 100; Benincà & Poletto 2004: 66; Öhl 2010: 62) (26):

  (25)a. Durement en halt si reclimet sa culpe
      strongly on high profess.3SG his sin
      ‘He loudly professes his sins’ (Roland 2014)
  b. Cunquerrament si finereit li bers.
      conquering ADV end.3SG.COND nobleman
      ‘The lord would end his life as a conquering hero’ (Roland 2867)

  (26)a. puis si se departirent
      then SI REFCL=leave.3PL.PST
      ‘Then they departed’ (Clari 4, 2)
  b. Aprés si s’en alerent en Venice
      after SI REFCL=LOC.CL=go.3PL.PST to Venice
      ‘Afterwards they went to Venice’ (Clari 7, 5)

3.5. Direct Object + Si Configurations

- As predicted by Marchello-Nizia (1985:158) these are shown in Table 1 to be rare.

However, they are found in several texts:

(27) a. Reis Vivien si succuras en Imphe
      King Vivien SI help.2SG.FUT in Imphe
      ‘help King Vivien there in Imphe’ (Roland 3995-3996)
  b. Les forteresces, les citez, si les aiez et departez
      the fortresses the cities SI them=have.2PL.SBJV and leave.2PL.SBJV
      ‘Sack the fortresses and cities and leave’ (Thèbes1 3807-3808)
  c. et tote la terre si tеноit Toldres li Ascre
      and all the land SI hold.3SG.PST Toldre l’Ascre
      ‘And Toldre l’Ascre held all the land’ (Villehardouin2 196, 387)
Intertextual variation:

- The Roland is the only text showing an informationally NEW direct object (27a).
- In all other texts they are discourse-OLD.

3.6 SI-THIRD CLAUSES

- Table 1 > SI-third clauses are never numerous (N=1-17). Nevertheless deserve attention as V3, and V4* orders have become a controversial topic within historical Romance linguistics.

(28) Quant la dame ot piece ploré, // son duel et son complaint when the woman have.3SG a-little cry.PTCP her grief and her complaint mené, // si apela un escuier show.PTCP SI call.3SG.PST a squire 'When the lady had cried a little and shown her grief and sorrow, she called upon a squire' (Eneas1 3385-3388)

- There is a small-scale decline for V3, but Villehardouin does not follow this pattern, with the highest amount of V3. There are only two examples in Clari and one in La Queste.

---

4. SI IN THE OLD FRENCH LEFT PERIPHERY

4.1. THE DATA TO ACCOUNT FOR

• Lots of accounts have viewed SI as a relatively homogeneous entity (Foulet 1928; Moignet 1973; Skårup 1975; Jensen 1990) and more recent treatments (Fleischman 1991; Vance 1995, 1997; Buridant 2000; Ferraresi & Goldbach 2002).

• But any account actually needs to account for the following variation:

  i) Topic continuity contexts, but others where this does not fit the data.
  ii) A link with the V2 property that characterises all texts under discussion
  iii) The fall in SI-initial clauses and rise of CLAUSE + SI configurations.
  iv) The tendency in 13th-century texts for postverbal subjects to be discourse OLD.
  v) The decline of focussed elements before SI.
  vi) The small decline in V3* orders in Clari and La Queste

4.2. SI IN LA CHANSON DE ROLAND


  ▶ Majority of attestations > SI only preverbal constituent > Fits with idea that in a V2 language the first constituent is frequently a ‘default Topic’ (Vikner 1995:41; Westergaard 2009:36; Jouitteau 2010:203; Holmberg 2015:376; Biberauer, Wolfe & Woods forthcoming).
  ▶ Low proportion of overt preverbal (1%) and postverbal (9.45%) subjects > expected under this account.
• But CLAUSE + SI is still found in the *Roland*, where this analysis doesn’t seem the right fit:

(29) Quant l’ot Rollant, si *cumençat* a rire
      when it.CL=hear.3SG Roland SI begin.3G.PST to laugh.INF
‘When Roland heard this, he began to laugh’ (*Roland* 303)

• **Proposal:** SI has been reanalysed from a temporal deictic adverbial in Latin (Ledgeway in press a, b) to a marker of Topic continuity

  - Like other adverbials it can optionally satisfy the V2 constraint, by raising into the left periphery.

• In contexts where there is an initial clause, it may be that certain clauses in Old French do not ‘count’ for V2 (Donaldson 2012)

  - Under standard Relativised Minimality assumptions (Rizzi 1990, et seq.) the adverbial SiC’s high position within the inflectional layer may mean its the closest available constituent to satisfy V2.

(30)[Frame [Force [Topic [Focus [Fin _____ [Fin° V] [TDeixis SI [SubjP… [vP… [VP…]]]]]]]]]

• But where does SI move to within the articulated left periphery assumed for Early Old French (Benincà 2004, Rouveret 2004, Labelle 2007, Wolfe 2016a)?

  - A position within the Topic layer would be a natural fit on pragmatic grounds, but SI can be preceded by topical constituents suggesting Spec-TopP would not be free to host SI.

  - SI can also be preceded by focal elements (new information objects and VP-adverbials) as seen above. If the dedicated position for Information Foci is directly above Fin Cruschina (2006:377-378; 2012:219), SI must be in SpecFinP.

  - This fits with recent accounts of Early Old French V2:

4.3. Si in 12th-Century French - Divergent Paths

- Rise in Clause + Si configurations, which, as noted above, jump from a mere 3.98% in the Roland, to ~20% in Thèbes, Eneas and Charrette > Supports expletive analysis.

- **Hypothesis for 12th-century Si**: Si is first-merged in Spec-FinP when no other constituent can satisfy V2 via internal merge (i.e. movement).

- **Evidence**:
  - Subject + Si clauses which increase in frequency in all three texts, though most notably in Thèbes and Charrette.
  - All these subjects tagged as discourse-old > constituents in the Topic layer base-generated there (Benincà & Poletto 2004:71; Ledgeway 2010:39-40) > Si is merged in Spec-FinP not to encode Topic continuity but to satisfy Fin’s EF.
  - Si can occur with adverbials, but all of them are frame-setting or speaker-oriented adverbials assumed to be base-generated high in the left periphery (Cinque 1999:§1.2; Ernst 2009:§3; Woods 2014:§4-5; Cruschina & Ledgeway 2016:31.2.2.1; Ledgeway in press c:§1.3):

    (31) a. et nequedant si s’an gardot…
    and however Si loc.CL look.3SG.PST
    ‘however, he looks at it...’ (Eneas 9255)
    b. et nepourquant si li est tart
    and nevertheless Si him.CL=be.3sg late
    ‘and nevertheless it is late in coming to him...’ (Thèbes2 8084)
  - In these texts we see a clear syntax-pragmatics mapping: preverbal subjects are consistently discourse-old, new information subjects seem to only occur post-verbally:

    vii) [Frame [Force [Topic SubjectOLD [Focus Fin Si [Fin° V] [SubjP... [VP SubjectNEW... [VP...]]]]]]]]
  - **Why?** Growing incompatibility of Foci with Si:
    - No focussed direct objects in texts other than the Roland and only on focused Indirect Object PP in Charrette and Eneas.
    - No VP adverbs in texts other than the Roland.
Given that French loses preverbal Information focus around 1200 Labelle 2007: 302–305; Labelle & Hirschbühler 2012: 15, 19–21; Mathieu 2012: 341; Wolfe 2016a:480, we could speculated that the incompatibility of foci with SI contributes to its demise.

• **Summary:**

  - Lots of variation between texts, but progressive attestation of grammar where SI loses unambiguous Topic continuity role and increasingly acts as a last-resort mechanism to satisfy V2.

  - Therefore following typical grammaticalisation path where constituent originally moved to a higher functional position come to be base-generated there (Roberts & Roussou 2003:195-202; Van Gelderen 2004:28, 2009:290, 2011:Ch.1; Roberts 2012b: 353).

4.4. SI IN 13TH-CENTURY FRENCH

4.4.1. Villehardouin

• Fits more with the 12th century texts than the other 13th century ones. A FinP expletive analysis fits the data best:

  - No evidence of VP-adverb or Indirect Object focalisation; increase in CLAUSE + SI configurations; increase in overt postverbal subjects.

  - Highest proportion of SUBJECT + SI clauses in the whole corpus at 7.98% (n=32).

  - Shows 4.24% (n=17) of SI-third clauses where Frame-Setter co-occurs with a thematic expression:

(32) En icel termine, li marchis Bonifaces de Monferrat remut de Salenique, si s’en ala a la Serre. "Then at that time, Marquess Boniface of Montferrat who had come from Salenica, moved into the saw formation" (Villehardouin2 270, 456)
4.4.2. Clari and La Queste

• Show major changes in the syntax which appear to suggest the expletive has been reanalysed upwards within the left periphery, in keeping with the Force-V2 hypothesis for 13th century French (Rouveret 2004:220; Wolfe 2016a:478-484):

viii) \[\text{Frame} [\text{Force} \, \text{si} [\text{Force}^\circ \text{V}][\text{Topic} [\text{Focus} [\text{Fin} ...]]]]\]

• Evidence for diachronic change:

  - Only one instance of si-third order in La Queste and two in Clari, but numbers are small across the board.
  - Accounts for rise in \(\text{CLAUSE} + \text{si} > 59\%\) in La Queste, 38.25% in Clari if wh, adverbial and circumstantial clause are typical Frame-Setters (Haiman 1978:585; Chafe 1984; Jacobs 2001:655-658; Sarda, Carter-Thomas, Fagard & Charolles 2014:279; Borreguero Zuloaga 2014:353f)
  - Total absence of indirect and direct objects predicted under the account if fronted objects target positions within the Topic-Focus layer.

5. The Big Picture

• In terms of si’s position within the clausal spine, it seems we can identify multiple stages:

  - A VP manner adverbial stage in Classical Latin
  - A temporal deictic adverbial stage in post-Classical Latin
  - A Topic continuity adverbial stage in the Roland
  - A FinP expletive stage in 12th century verse
  - A ForceP expletive stage in some 13th century prose.
This is also one of many cases of grammaticalisation in the passage from Latin to Romance:

- *Sī* shows the **shift from a lexical to functional element** (Meillet 1912; Heine & Kuteva 2002:2; Hopper & Traugott 2003:4-6; Fischer 2003; Ledgeway 2011)

- *Sī* becomes increasingly ‘**bleached’ of previous pragmatico-semantic content** (Hopper & Traugott 2003:94-98; Roberts 2010:§5; Lehmann 2015:134-136)

- *Sī* shows a **widening of distribution** > ‘defining factor of grammaticalization’ (Lehmann 2015:7).


- We can schematise these reanalyses as follows:

  ix) \( \text{AdvManner > AdvTemporal > AdvTopicContinuity > ExpletiveFin > ExpletiveForce} \)
5. **Summary and Conclusions**

- At first sight si seems a fairly boring, parochial topic with relatively little ‘external significance’. But that is clearly not the case:
  
  - Empirically, it contributes to the growing view that Old French was subject to extensive internal variation in morphosyntax, aside from phonology and morphology where this is generally acknowledged (Price 1971:7-12; Einhorn 1974:Ch.13 and the contributions in Holtus et al. 1995).
  
  - Theoretically, si is an engaging case of upwards grammaticalisation (Roberts & Roussou 2002; Van Gelderen 2008, 2009) and may be unusual in providing a case where a constituent has been successively reanalysed along the entire length of the clausal spine, from a VP-adverbial to an expletive in the very highest reaches of the CP.
  
  - A thorny issue remaining in the historical syntax literature is how to capture the relatedness of certain clusters of change: a topic which was central in Lightfoot 1979, 1989, 1991) amongst others. Si is pertinent here as changes in its syntax are contingent upon changes in the V2 property, but si itself may also have manifestations on other areas of the syntax.
  
  - There are fundamental morphosyntactic splits between Gallo-Romance and (Southern) Italo-Romance in terms of the ease with which constituents are moved into the left periphery and the licensing of left peripheral Focus (see in particular Belletti 2008; De Cat 2009; Rowlett 2007:172-188 Haegeman 2012:Ch.2-3; Cruschina & Ledgeway 2016:§31.2):
    
    ▶ Given si’s ubiquity in Old French texts and growing incompatibility with other moved elements in the left periphery, we may have an answer as to why.
    
    ▶ If this hypothesis is correct, Old Southern Italo-Romance si should not show the same properties. This seems to be correct (Ledgeway 2008; Wolfe 2015).
TEXTS CITED

Ottowa: Laboratoire de Français Ancien.

BIBLIOGRAPHY


