Agreement syncretisation and the loss of null subjects in Medieval French structural and quantitative models

Alexandra Simonenko, Benoit Crabbé, Sophie Prévost

FWO/UGent, Alpage, LaTTiCe

DiGS, UGent, June 29–July 30 2016
OUTLINE

Introduction

Observations
  Null subjects
  Syncretisation

Model I: Structural

Results I
  Testing the main hypothesis
  Testing other predictions

Model 2: Acquisitional

Conclusions
**AGREEMENT ~ NULL SUBJECTS**

Taraldsen’s (typological) generalisation

- Rich (i.e. non-syncretic) verbal subject agreement implies the possibility of null subjects (Taraldsen 1980)
Debates whether the loss of null subjects was related to the loss of rich agreement in Medieval French (Ewert 1943, Vennemann 1975, Schøsler 2002, Roberts 2014).

Very few quantitative diachronic studies (Duarte 1995 on 1,5 centuries of Brazilian Portuguese), no quantitative data on the loss of rich agreement in French.
Rejection of the Immediate Connection

On the assumption that there is a significant temporal lag between the two changes:

OVERT SUBJECTS $\gg$ SYNCRETISATION

- Schøsler (2002, 196): “…la confusion des flexifs est progressive: elle se produit entre les 12ème et 16ème siècles suivant les personnes, alors que le sujet est exprimé dans moins 50% des cas à partir du 13ème siècle.”

SYNCRETISATION $\gg$ OVERT SUBJECTS

- Roberts (2014): the total loss of the rich agreement (XII c.) precedes by 4 centuries the completion of the loss of null subjects (XVI c.)
OUR CONTRIBUTIONS

▶ First quantitative corpus-based study of syncretisation in Medieval French

▶ Models relating the two changes:
  ▶ “Structural”, relating rich inflection and null subjects as manifestations of the same grammar (building on Kroch 1989)

▶ “Acquisitional”, treating ambiguous inflection as a disadvantage for the null subject-licensing grammar (building on Yang 2010)
Introduction

Observations
   Null subjects
   Syncretisation

Model I: Structural

Results I
   Testing the main hypothesis
   Testing other predictions

Model 2: Acquisitional

Conclusions
OUTLINE

Introduction

Observations
   Null subjects
   Syncretisation

Model I: Structural

Results I
   Testing the main hypothesis
   Testing other predictions

Model 2: Acquisitional

Conclusions
NULL SUBJECTS

(1) de lui ∅ firent segnor et mestre. Puis ∅ of him (they) made sir and lord. Then (they) ont gardé devers senestre have looked towards left...
“They made him their sir and lord. Then they looked to the left...” (Eneas, v. 76-79, XII c.)

➤ Steadily disappearing
**Null Subjects**

MCVF (Martineau et al. 2010) & Penn Supplement to MCVF (Kroch & Santorini 2010)

Finite clauses with either an overt pronominal or null subject (total of 104,485), excluding imperatives, subject relatives, wh-questions targeting subjects, and subject ellipsis under coordination.
SUBJECT AGREEMENT SYNCRETISATION

## Subject Agreement in Old French

- Changes attested in writing:

<table>
<thead>
<tr>
<th></th>
<th>I group</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>present indicative</td>
<td>present subjunctive</td>
</tr>
<tr>
<td>1P</td>
<td>aim (\rightarrow) aime</td>
<td>aim (\rightarrow) aime</td>
</tr>
<tr>
<td>2P</td>
<td>aimes</td>
<td>ains (\rightarrow) aimes</td>
</tr>
<tr>
<td>3P</td>
<td>aimet (\rightarrow) aime</td>
<td>aint (\rightarrow) aime</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>II group</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>present indicative</td>
<td>past indicative</td>
</tr>
<tr>
<td>1P</td>
<td>voi (\rightarrow) vois</td>
<td>vi (\rightarrow) vis</td>
</tr>
<tr>
<td>2P</td>
<td>vois</td>
<td>vis</td>
</tr>
<tr>
<td>3P</td>
<td>voit</td>
<td>vit</td>
</tr>
</tbody>
</table>
### Subject agreement in Modern French

<table>
<thead>
<tr>
<th>Person</th>
<th>Form</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1P</td>
<td>aime</td>
<td>[ɛm]</td>
</tr>
<tr>
<td>2P</td>
<td>aimes</td>
<td>[ɛm]</td>
</tr>
<tr>
<td>3P</td>
<td>aime</td>
<td>[ɛm]</td>
</tr>
<tr>
<td>4P</td>
<td>aimons</td>
<td>[ɛmɔ̃]</td>
</tr>
<tr>
<td>5P</td>
<td>aimez</td>
<td>[ɛme]</td>
</tr>
<tr>
<td>6P</td>
<td>aiment</td>
<td>[ɛm]</td>
</tr>
<tr>
<td>1P</td>
<td>pars</td>
<td>[par]</td>
</tr>
<tr>
<td>2P</td>
<td>pars</td>
<td>[par]</td>
</tr>
<tr>
<td>3P</td>
<td>part</td>
<td>[par]</td>
</tr>
<tr>
<td>4P</td>
<td>partons</td>
<td>[partɔ̃]</td>
</tr>
<tr>
<td>5P</td>
<td>partez</td>
<td>[parte]</td>
</tr>
<tr>
<td>6P</td>
<td>partent</td>
<td>[part]</td>
</tr>
</tbody>
</table>

- No person agreement in singular in spoken French
**Phonology ~ Orthography Problem**

- prior to mid-XIV c., graphemes likely reflect pronunciation very closely (De Jong 2006, 174)

- strict rimes of the first versified texts suggest that the final consonants were pronounced (Bettens 2015)

- grammarians of the XVI c. mention in their work that they still pronounced -s in inflection (Bonin 1992, 56).

Working assumption: at least until the XIV c., the orthographic spread of the “new” endings mirrors the oral syncretisation of the verbal agreement.
FIRST QUANTITATIVE STUDY

- For each text, proportion of the new endings in the relevant environments (subject person + verb group), limited to clauses with overt subjects.

3P I group: \[\frac{#e}{#e+\#t}\]

1P I group: \[\frac{#e}{#e+\#zero}\]

1P II group: \[\frac{#s}{#s+\#zero}\]
Introduction

Observations

Model I: Structural

Results I

Model 2: Acquisitional

Conclusions

Figure: Emergence of new verbal endings
STARTING POINT

- Emergence of overt subjects and syncretic endings – related?
OUTLINE

Introduction

Observations
  Null subjects
  Syncretisation

Model I: Structural

Results I
  Testing the main hypothesis
  Testing other predictions

Model 2: Acquisitional

Conclusions
Null subjects and non-syncretic agreement are related via a **structural property giving rise to both**, e.g. person feature-specified Agr head.
CLASSIC ANALYSIS-BASED MODEL

- Null subjects and non-syncretic agreement are related via a structural property giving rise to both, e.g. person feature-specified Agr head.
  - Non-syncretic endings are spellouts of different person features.
Null subjects and non-syncretic agreement are related via a structural property giving rise to both, e.g. person feature-specified Agr head.

- Non-syncretic endings are spellouts of different person features.

- Null subjects are made possible by person features (they introduce necessary presuppositions about subject’s reference).
**Grammar with Agr (Agr-Gr)**

\[ \emptyset \leftrightarrow [\_, 1P, SG] / V + _ \]
\[ s \leftrightarrow [\_, 2P, SG] / V + _ \]
\[ t \leftrightarrow [\_, 3P, SG] / V + _ \]

```
...                    ...
    AgrP               AgrP
      pro      AgrP       pro      AgrP
        Agr\{P:1\}   TP       Agr\{P:3\}   TP
            \emptyset       t
                T   ...
```

...
Grammar without Agr (TP-Gr)

→ e ←→ [\_, PRES, SG] / V + _

![Diagram](image-url)
Model I main prediction

- Constant Rate Hypothesis: a grammatical change has the same rate of spreading in all grammatical environments (Kroch 1989)

- Emergence of overt subjects and syncretic endings should proceed at the same rate (= underlyingly the loss of Agr).
MODEL I CAVEAT

- A grammar with Agr allows both for null and overt referential subjects (cf. Italian and Spanish, Bates 1976, Otheguy et al. 2007).

- Only expletive subjects must be null.

Restating the prediction:

Emergence of overt expletive subjects and syncretic endings should proceed at the same rate (= underlyingly loss of Agr).
OUTLINE

Introduction

Observations
  Null subjects
  Syncretisation

Model I: Structural

Results I
  Testing the main hypothesis
  Testing other predictions

Model 2: Acquisitional

Conclusions
**Prediction 1: Two changes have the same rate - TRUE**

\[ P(\text{ENDING} = \text{new}|\text{DATE} = d) = \frac{e^{\alpha + \beta \times \text{Date}}}{1 + e^{\alpha + \beta \times \text{Date}}} : \alpha = -5.939 \text{ AND } \beta = 0.0049. \]

\[ P(\text{EXPL SBJ} = \text{overt}|\text{DATE} = d) = \frac{e^{\alpha + \beta \times \text{Date}}}{1 + e^{\alpha + \beta \times \text{Date}}} : \alpha = -6.325 \text{ AND } \beta = 0.0045. \]
INTERPRETATION

- Slope difference is not significant
  - The random slope parameter does not introduce a significant difference between two mixed effects models ($p > 0.37$).

\[
P(Y = \text{new}|\text{Time} = t, \text{Context} = c) = \frac{e^{\alpha + \alpha_c + \beta t}}{1 + e^{\alpha + \alpha_c + \beta t}} \quad (1)
\]

\[
P(Y = \text{new}|\text{Time} = t, \text{Context} = c) = \frac{e^{\alpha + \alpha_c + (\beta + \beta_c) t}}{1 + e^{\alpha + \alpha_c + (\beta + \beta_c) t}} \quad (2)
\]

- Compatible with the hypothesis that the emergence of overt subjects and new endings are part of the same change (on the CRH).
**Prediction 2: Same syncretisation rate across contexts**

- New syncretic endings are expected to emerge at the same rate in different contexts (on the CRH).
PREDICTION 2: SAME SYNCRETISATION RATE ACROSS CONTEXTS - FALSE
Prediction 3: No new ending increase with null subjects

- There should be no increase in new syncretic endings (new TP-Gr) with null subjects (old Agr-Gr)
**Prediction 3:** No new ending increase with null subjects – **FALSE**

![Graph showing probability over time for ending new with and without over overt and null pronouns](image-url)
**Prediction 4: No increase in overt subjects with old endings**

- There should be *no increase* in pronominal subject expression (new TP-Gr) in the context of verbs with old non-syncretic endings (old Agr-Gr)
PREDICTION 4: NO INCREASE IN OVERT SUBJECTS WITH OLD ENDINGS - FALSE
INTERIM SUMMARY

Main prediction borne out:

Emergence of overt subjects and new endings (combined) proceeded at the same rate, as expected for the reflexes of the same change.

Three predictions falsified:

- New -e and -s endings spread at different rates
- New endings raise both with overt and with null subjects
- Overt subjects raise both with new and with old endings
INTERIM CONCLUSIONS

▶ Evidence for a non-accidental relation between the emergence of overt pronominal subjects and syncretic endings.

▶ A model which ties subject expression to a particular agreement paradigm at the clause level fails.

▶ A model is needed which would dissociate subject expression and the choice of ending at the clause level, but would still relate them in language evolution.
Outline

Introduction

Observations
  Null subjects
  Syncretisation

Model I: Structural

Results I
  Testing the main hypothesis
  Testing other predictions

Model 2: Acquisitional

Conclusions
MODEL 2: CHANGE AS DISTURBED ACQUISITIONS

- Syncretisation (independent phonological change) creates a negative bias in acquisition of null subject grammar.
- The progression of syncretism strengthens the bias.
Innate grammatical options: $\mathcal{G} = \{G_i, G_j\}$ with probabilities $P(\mathcal{G} = G_i)$ and $P(\mathcal{G} = G_j)$ of being chosen to analyse a given clause.

- Select a clause $x$ in the data
- Select $G_i$ in proportion to its probability
- Analyse $x$ with $G_i$
  - If $G_i$ succeeds in analyzing $x$ provide $G_i$ a reward: $P(\mathcal{G} = G_i)$ increases.
  - If $G_i$ fails in analyzing $x$ provide $G_i$ a penalty: $P(\mathcal{G} = G_i)$ decreases.

Modelling this process iteratively we can approximate how a grammar can die out when there are data which make it fail.
UPDATING SCHEME

Let $T = t_1 \ldots t_n$ be a sequence of iterations
Let $P(\mathcal{G} = G|E = e, T = t)$ be the probability of grammar $G$ in context $e$ at time step $t$
Then $P(\mathcal{G} = G|E = e, T = t + 1)$ is computed:

$x \sim X$
$G_i \sim P(\mathcal{G} = G|E = e, T = t) \ (x \in e)$

▶ if $G_i \rightarrow x$ then

$P(\mathcal{G} = G_i|E = e, T = t + 1) = P(\mathcal{G} = G_i|E = e, T = t) + \gamma (1 - P(\mathcal{G} = G_i|E = e, T = t))$

$P(\mathcal{G} = G_j|E = e, T = t + 1) = (1 - \gamma)P(\mathcal{G} = G_j|E = e, T = t) \ (\forall j: i \neq j)$

▶ if $G_i \not\rightarrow x$ then

$P(\mathcal{G} = G_i|E = e, T = t + 1) = (1 - \gamma)P(\mathcal{G} = G_i|E = e, T = t)$

$P(\mathcal{G} = G_j|E = e, T = t + 1) = \frac{\gamma}{K - 1} + (1 - \gamma)P(\mathcal{G} = G_j|E = e, T = t) \ (\forall j: i \neq j)$

where $\gamma \in [0, 1]$ is a parameter (Linear Reward Penalty scheme, Bush 1958)
ESTIMATING PROBABILITIES OF THE GRAMMARS

Penalty $c_i$ of a grammar $G_i$:

$$c_i = P(G_i \not\rightarrow x | x \in E)$$

$c_i$ is the probability that $G_i$ fails to analyse an example in a dataset $X$, estimated by the relative frequency of failures.

$$\lim_{t \to \infty} P(\mathcal{G} = G_1 | E = e, T = t) = \frac{c_2}{c_1 + c_2}$$

$$\lim_{t \to \infty} P(\mathcal{G} = G_2 | E = e, T = t) = \frac{c_1}{c_1 + c_2}$$

Narendra and Thathachar (1989)
**CONTEXTS OF FAILURE**

- Ambiguous endings (-e, -s, -oe, -sse, -ais, -ait, -ent, -ai) make Agr-Gr fail (they do not identify the semantics of Agr).
- Null subjects make TP-Gr fail.

<table>
<thead>
<tr>
<th>Ending</th>
<th>Subject</th>
<th>Agr-Gr</th>
<th>TP-Gr</th>
</tr>
</thead>
<tbody>
<tr>
<td>V-a</td>
<td>yes</td>
<td>succeeds</td>
<td>succeeds</td>
</tr>
<tr>
<td>V-a</td>
<td>no</td>
<td>succeeds</td>
<td>fails</td>
</tr>
<tr>
<td>V-ai</td>
<td>yes</td>
<td>fails</td>
<td>succeeds</td>
</tr>
<tr>
<td>V-ai</td>
<td>no</td>
<td>fails</td>
<td>fails</td>
</tr>
<tr>
<td>V-ais</td>
<td>yes</td>
<td>fails</td>
<td>succeeds</td>
</tr>
<tr>
<td>V-ais</td>
<td>no</td>
<td>fails</td>
<td>fails</td>
</tr>
<tr>
<td>V-ait</td>
<td>yes</td>
<td>fails</td>
<td>succeeds</td>
</tr>
<tr>
<td>V-ait</td>
<td>no</td>
<td>fails</td>
<td>fails</td>
</tr>
<tr>
<td>V-as</td>
<td>yes</td>
<td>succeeds</td>
<td>succeeds</td>
</tr>
<tr>
<td>V-as</td>
<td>no</td>
<td>succeeds</td>
<td>fails</td>
</tr>
<tr>
<td>V-at</td>
<td>yes</td>
<td>succeeds</td>
<td>succeeds</td>
</tr>
<tr>
<td>V-at</td>
<td>no</td>
<td>succeeds</td>
<td>fails</td>
</tr>
<tr>
<td>V-e</td>
<td>yes</td>
<td>fails</td>
<td>succeeds</td>
</tr>
<tr>
<td>V-e</td>
<td>no</td>
<td>fails</td>
<td>fails</td>
</tr>
<tr>
<td>V-ent</td>
<td>yes</td>
<td>fails</td>
<td>succeeds</td>
</tr>
<tr>
<td>V-ent</td>
<td>no</td>
<td>fails</td>
<td>fails</td>
</tr>
<tr>
<td>V-es</td>
<td>yes</td>
<td>succeeds</td>
<td>succeeds</td>
</tr>
<tr>
<td>V-es</td>
<td>no</td>
<td>succeeds</td>
<td>fails</td>
</tr>
</tbody>
</table>

... etc.
PROBABILITY OF TP-GR BASED ON ESTIMATED PENALTY PROBABILITIES

![Graph showing probability of TP-Gr based on estimated penalty probabilities over different periods. The x-axis represents the periods from X-I to XVI, and the y-axis represents the probability from 0.0 to 1.0. Three lines are plotted: red for P(TP GRAMMAR = YES), green for P(EXPL SBJ = YES), and orange for P(PRON SBJ = YES).]
Outline

Introduction

Observations
  Null subjects
  Syncretisation

Model I: Structural

Results I
  Testing the main hypothesis
  Testing other predictions

Model 2: Acquisitional

Conclusions
CONCLUSIONS

▶ A reinforcement model which treats syncretisation as an independent change which disadvantages the non-null subject grammar does not make incorrect predictions.
CONCLUSIONS

▶ A reinforcement model which treats syncretisation as an independent change which disadvantages the non-null subject grammar does not make incorrect predictions
  ▶ new endings are expected to raise with null subjects as well (phonological change is not sensitive to that)
CONCLUSIONS

- A reinforcement model which treats syncretisation as an independent change which disadvantages the non-null subject grammar does not make incorrect predictions
  - new endings are expected to raise with null subjects as well (phonological change is not sensitive to that)
  - overt subjects are expected to raise with old endings (TP-Gr doesn’t “care” about endings phonology)
CONCLUSIONS

▶ A reinforcement model which treats syncretisation as an independent change which disadvantages the non-null subject grammar does not make incorrect predictions
  ▶ new endings are expected to raise with null subjects as well (phonological change is not sensitive to that)
  ▶ overt subjects are expected to raise with old endings (TP-Gr doesn’t “care” about endings phonology)

▶ The probability of TP-Gr, computed based on the penalty probability of Agr-Gr, increases parallel to the increase in overt expletive subjects.
CONCLUSIONS

▶ A reinforcement model which treats syncretisation as an independent change which disadvantages the non-null subject grammar does not make incorrect predictions
  ▶ new endings are expected to raise with null subjects as well (phonological change is not sensitive to that)
  ▶ overt subjects are expected to raise with old endings (TP-Gr doesn’t “care” about endings phonology)

▶ The probability of TP-Gr, computed based on the penalty probability of Agr-Gr, increases parallel to the increase in overt expletive subjects.

▶ The latter measure tentatively reflects the probability of TP-Gr to be chosen to produce a clause.


