

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 $\sim$ NULL SUBJECTS

Taraldsen's (typological) generalisation

- ▶ Rich (i.e. non-syncretic) verbal subject agreement implies the possibility of null subjects (Taraldsen 1980)
- ▶ Implemented formally: Rizzi 1986, Adams 1987, Alexiadou and Anagnostopoulou 1998, Roberts 2010, Sheehan to appear, and many others

AGREEMENT ~ NULL SUBJECTS IN DIACHRONY

- ▶ 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 > 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 > 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 seignor 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
 (Foulet 1928, Fontaine 1985, Hirschbühler 1992, Schøsler
 2002, Kaiser 2009, Zimmermann 2014, Prévost to appear).

SUBJECT AGREEMENT SYNCRETISATION

- ▶ French went from a non-syncretic to a largely syncretic subject person agreement (Foulet 1935, Dees et al. 1980, Marchello-Nizia 1992, Buridant 2000, De Jong 2006, Bettens 2015).

SUBJECT AGREEMENT IN OLD FRENCH

- Changes attested in writing:

I group		
	present indicative	present subjunctive
1P	aim > aime ^e	aim > aime ^e
2P	aimes	ains > aimes
3P	aimet > aime ^e	aint > aime ^e

II group		
	present indicative	past indicative
1P	voi > vois ^s	vi > vis ^s
2P	vois	vis
3P	voit	vit

SUBJECT AGREEMENT IN MODERN FRENCH

1P	aime [ɛm]	4P	aimons [ɛmɔ̃]
2P	aimes [ɛm]	5P	aimez [ɛme]
3P	aime [ɛm]	6P	aiment [ɛm]

1P	pars [paʀ]	4P	partons [partɔ̃]
2P	pars [paʀ]	5P	partez [parte]
3P	part [paʀ]	6P	partent [part]

- 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.

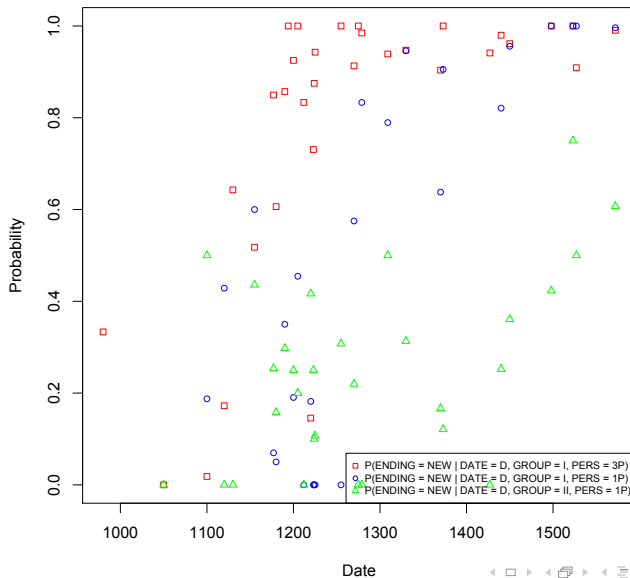
FIRST QUANTITATIVE STUDY

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

$$3P \text{ I group: } \frac{\#e}{\#e + \#t}$$

$$1P \text{ I group: } \frac{\#e}{\#e + \#zero}$$

$$1P \text{ II group: } \frac{\#s}{\#s + \#zero}$$



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

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.

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.

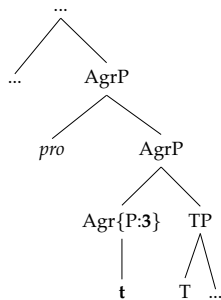
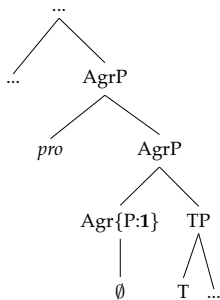
CLASSIC ANALYSIS-BASED MODEL

- ▶ 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 are made possible by person features (they introduce necessary presuppositions about subject's reference).

GRAMMAR WITH AGR (AGR-GR)

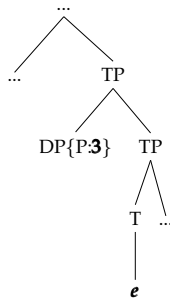
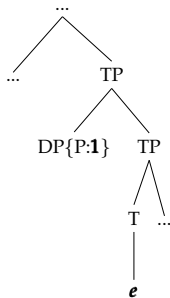
$$\emptyset \longleftrightarrow [_ , 1P,SG] / V + _$$

$$s \longleftrightarrow [_ , 2P,SG] / V + _$$

$$t \longleftrightarrow [_ , 3P,SG] / V + _$$


GRAMMAR WITHOUT AGR (TP-Gr)

► $e \longleftrightarrow [_, \text{PRES, SG}] / V + _$



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

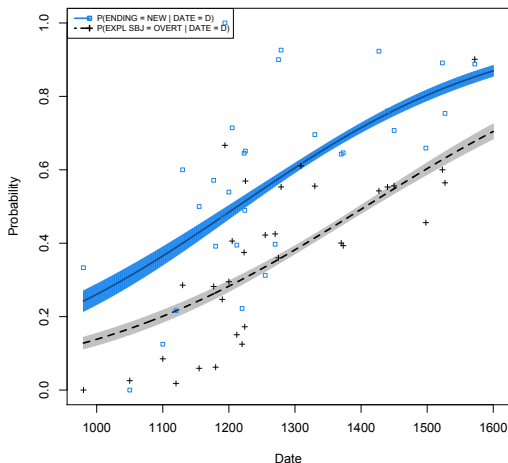
Model 2: Acquisitional

Conclusions

PREDICTION 1: TWO CHANGES HAVE THE SAME RATE - **TRUE**

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

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



INTERPRETATION

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

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

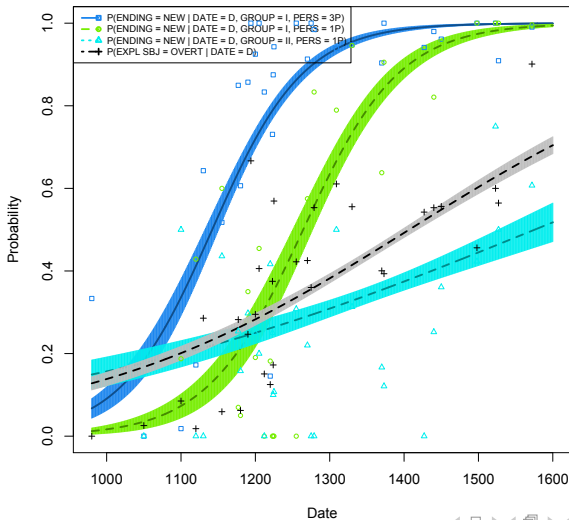
$$P(Y = \text{new} | \text{Time} = t, \text{Context} = c) = \frac{e^{\alpha + \alpha_c + (\beta + \beta_c)t}}{1 + e^{\alpha + \alpha_c + (\beta + \beta_c)t}} \quad (2)$$

- ▶ 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).

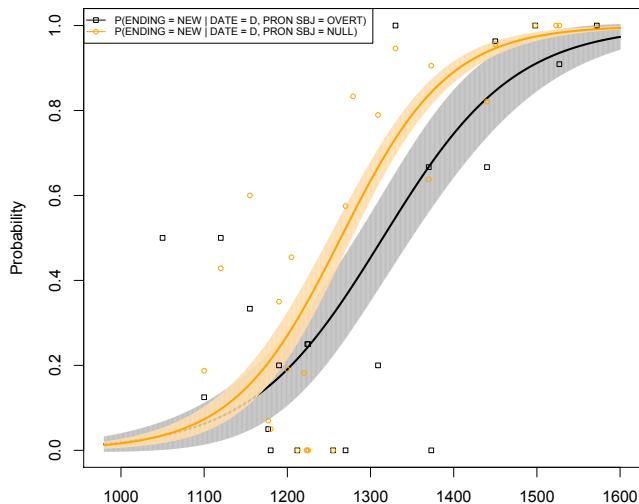
PREDICTION 2: SAME SYNCRETISATION RATE ACROSS CONTEXTS - FALSE



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)

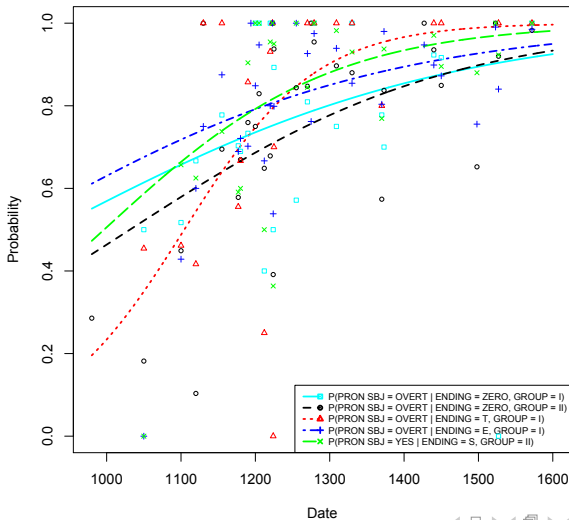
PREDICTION 3: NO NEW ENDING INCREASE WITH NULL SUBJECTS – FALSE



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.

BUILDING ON YANG (2002)

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 $T = t_1 \dots t_n$ 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:

$x \sim X$

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

► if $G_i \rightarrow x$ then

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

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

► if $G_i \not\rightarrow x$ then

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

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

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 :

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

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

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

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

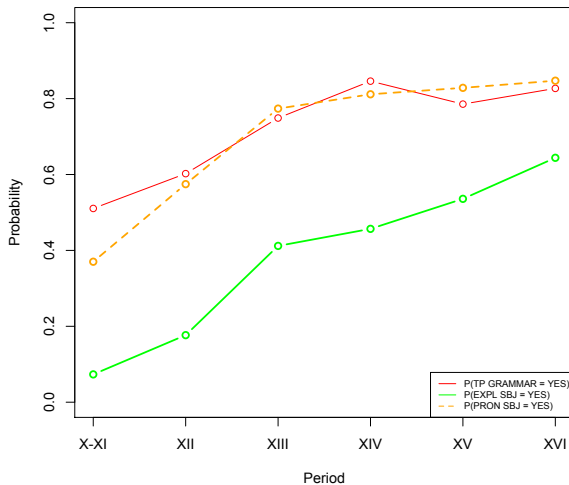
Narendra and Thathachar (1989)

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.

Ending	Subject	Agr-Gr	TP-Gr
V-a	yes	succeeds	succeeds
V-a	no	succeeds	fails
V-ai	yes	fails	succeeds
V-ai	no	fails	fails
V-ais	yes	fails	succeeds
V-ais	no	fails	fails
V-ait	yes	fails	succeeds
V-ait	no	fails	fails
V-as	yes	succeeds	succeeds
V-as	no	succeeds	fails
V-at	yes	succeeds	succeeds
V-at	no	succeeds	fails
V-e	yes	fails	succeeds
V-e	no	fails	fails
V-ent	yes	fails	succeeds
V-ent	no	fails	fails
V-es	yes	succeeds	succeeds
V-es	no	succeeds	fails
... etc.			

PROBABILITY OF TP-GR BASED ON ESTIMATED PENALTY PROBABILITIES



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.

- Adams, Marianne. 1987. From Old French to the theory of pro-drop. *Natural Language & Linguistic Theory* 5:1–32.
- Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural Language & Linguistic Theory* 16:491–539.
- Bates, Elizabeth. 1976. *Language and context: The acquisition of pragmatics*, volume 13. New York: Academic Press.
- Bettens, Olivier. 2015. Chantez-vous français ? remarques curieuses sur le français chanté de moyen Âge à la période baroque. URL <http://virga.org/cvf/>.
- Bonin, Michèle. 1992. Le système verbal français de Luis Meigret, XVIe siècle. MA thesis.
- Buridant, Claude. 2000. *Nouvelle grammaire de l'ancien français*. Paris: Sedes.
- De Jong, Thera. 2006. *La prononciation des consonnes dans le français de Paris aux 13ème et 14ème siècles*. Netherlands Graduate School of Linguistics.
- Dees, Anthonij, S. Meilink, Karin van Reenen-Stein, and Pieter van Reenen. 1980. Un cas d'analogie: l'introduction de -e à la première personne du singulier de l'indicatif présent des verbes en -er en ancien français. *Rapport/Het Franse Boek* 50:105–110.
- Duarte, Maria Eugênia Lamoglia. 1995. A perda do princípio “evite pronome” no português brasileiro. Doctoral Dissertation, Unicamp.
- Ewert, Alfred. 1943. *The French Language*. London: Faber & Faber.
- Fontaine, Carmen. 1985. Application de méthodes quantitatives en diachronie: L'inversion du sujet en français. Université du Québec à Montréal.
- Foulet, Lucien. 1928. *Petite syntaxe de l'ancien français*. Paris: Champion, troisième édition revue. Réédition 1982.
- Foulet, Lucien. 1935. L'extension de la forme oblique du pronom personnel en ancien français. *Romania* 61–62:257–315.
- Hirschbühler, Paul. 1992. L'omission du sujet dans les subordonnées V1: les CNN de Vigneulles et les CNN anonymes. *Travaux de linguistique* 24:25–46.
- Kaiser, Georg A. 2009. Losing the null subject. A contrastive study of (Brazilian) Portuguese and (Medieval) French. In *Proceedings of the Workshop “Null-subjects, expletives, and locatives in Romance”*, 131–156.
- Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language variation and change* 1:199–244.
- Marchello-Nizia, Christiane. 1992. *Histoire de la langue française aux xive et xve siècles*. Paris: Dunod.

- Otheguy, Ricardo, Ana Celia Zentella, and David Livert. 2007. Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language* 770–802.
- Prévost, Sophie. to appear. Emergence and development of personal pronoun subjects: study of a bilingual Latin–Old French corpus.
- Rizzi, Luigi. 1986. Null objects in Italian and the Theory of *pro*. *Linguistic Inquiry* 17:501–557.
- Roberts, Ian. 2010. A deletion analysis of null subjects. In *Parametric variation: Null subjects in minimalist theory*, 58–87. Cambridge University Press.
- Roberts, Ian. 2014. Taraldsen’s Generalization and Language Change: Two Ways to Lose Null Subjects. In *Functional Structure from Top to Toe: The Cartography of Syntactic Structures*, ed. Peter Svenonius, volume 9. Oxford University Press.
- Schøsler, Lene. 2002. La variation linguistique: le cas de l’expression du sujet. In *Interpreting the History of French, A Festschrift for Peter Rickard on the occasion of his eightieth birthday*, ed. Rodney Sampson and Wendy Ayres-Bennett, 187–208. Amsterdam - New York, NY: Rodopi.
- Sheehan, Michelle. to appear. Subjects, null-subjects and expletives in romance. URL <http://ling.auf.net/lingbuzz/002466>, To appear in *Manuals of Romance Linguistics (MRL): Grammatical Interfaces*.
- Taraldsen, Tarald. 1980. *On the NIC, vacuous application and the that-trace filter*. Indiana University Linguistics Club.
- Vennemann, Theo. 1975. An explanation of drift. In *Word Order and Word Order Change*, ed. Charles N. Li, 269–305. Austin: University of Texas Press.
- Yang, Charles. 2002. *Knowledge and learning in natural language*. Oxford University Press.
- Yang, Charles. 2010. Three factors in language variation. *Lingua* 120:1160–1177.
- Zimmermann, Michael. 2014. *Expletive and referential subject pronouns in medieval french*, volume 556. Walter de Gruyter GmbH.