Parametric settings of functional projections in diachrony: Romance clefts and wh-interrogatives

Caterina Bonan[1,*] **and Giuseppe Samo**[2]

[1]University of Cambridge, United Kingdom.
[2]Beijing Language and Culture University, People's Republic of China.
[*]Corresponding author: cb2098@cam.ac.uk

ABSTRACT

This paper is an extract from a study that utilises data from Romance corpora to determine how the functional projections responsible for nominal clefting and for interrogative wh-movement have changed from earlier stages of Romance to present times. The data are assessed through the lens of **?** parameters. Here, we present our preliminary French data.

Introduction

?'s (**?**; see also **?**) study of parameters: functional projections either require overt movement (IM=1) or they don't (IM=0). This, along with parameters as Spell Out (SO), allows a fine understanding of the way languages function, and evolve.

Our ultimate goal is to:

- utilise Rizzi's parameters to evaluate the existing understandings of cleft structure and interrogative wh-movement in Romance;
- redifine those understandings that utilise merely Rizzi's (1997) FocusP and Belletti's (2004) low Foc in their derivations ('focus field', Bonan 2021);
- investigate the parametric settings of the projections involved in these structures, and their evolution over time.

1. ?) parameters

? (**?**: 165): Parameter: "an instruction for the triggering of a syntactic operation, expressed as a morphosyntactic feature associated to a functional head".

Move:

- complex operation (à la **?**);
- involves either a head or a phrase;
- encompasses the establishment of a probe-goal search followed by (internal) merge of the goal.

For Rizzi, a functional head that acts as a trigger of movement may have distinct pairs of features responsible for (1) and (2):

(1) PHRASAL MOVEMENT (**?**: 171 (20))
   a. A search feature at the phrasal level.
   b. The corresponding internal merge feature at the phrasal level (IM) ('EPP feature').

(2) HEAD MOVEMENT (**?**: 171 (21))
   a. A search feature at the lex level ($\text{Search}_{\text{lex}}$ Feature)
   b. The corresponding internal merge feature, again at the lex level (IM_{lex} Feature)

Syntactic operations:

- simple;
- highly learnable;
- restricted to an extremely reduced set for reasons of learnability.

When a functional element enters the syntax and becomes a functional head in the relevant configuration, it triggers one syntactic operation on the structure which is built. The available operations are those in (3):

(3) SYNTACTIC OPERATIONS
  a. Merge
  b. Move
    i. Search: Probe-goal relation at the phrasal level
    ii. IM: Internal merge of phrases; or:
    iii. $\text{Search}_{\text{lex}}$ Probe-goal relation at the head level
    iv. IM_{lex} Internal merge of heads
  c. Spellout

Spell-out ('null subject parameter' etc.):

- deal with "variation in the obligatory, optional or impossible pronunciation of certain heads and of their immediate dependents" (**?**: 175).

2. Explaining language variability

Cartography of syntactic structures (on this, see **??**):

- the functional spine of human language is universal;
- the functional spine comprises of numerous rigidly ordered functional projections;
- **nonetheless**, languages vary to the extent in which they:
  i. activate the functional heads of the spine;
  ii. realise these projections using different strategies.

FocusP (HLP) is not realised/exploited in the same way by all languages. **?**: Focus^0 triggers movement of an XP that bears a relevant focus feature and:

- in languages such as Gungbe this head is phonetically realised (**?**), as in (**??**);

(4) Gungbe (adapted from Aboh 2007: 85(9c))

$[_{\text{FocusP}}$ KÒFÍ$_i$ $[_{\text{Focus}^0}$ wè [ùn yró ___$_i$]]]!
Kofi foc 1sg call
'I called KOFI (as opposed to, for example, Enoch)'

- in languages like Italian, the head is silent (**?** and related), as in (5);

(5) Italian (adapted from **?**: 146 (8))

$[_{\text{FocusP}}$ IL LIBRO$_i$ $[_{\text{Focus}^0}$ ∅ [Gianni ha letto ___$_i$]]]!
the book foc Gianni has read ___
'Gianni read THE BOOK (as opposed to, for example, the article)'

- in V2 languages, the head is activated by moving an already merged head, as in the German example in (6):

(6) German (adapted from **?**: 146 (8))

$[_{\text{SpecFoc}}$ DIESES FRESKO $[_{\text{Focus}^0}$ malte [Giotto]]]
this fresco painted.3sg Giotto
'Giotto painted THIS FRESCO (as opposed to, for example, the one over there)'

The variability of syntactic strategies adopted by different languages thus stems from different combinations of the syntactic operations of Merge, Move and Spell Out:

- Gungbe merges FocusP and spells out Focus°;
- Italian merges FocusP but **does not** spell out Focus^0;

- German requires **both** head movement and phrasal movement.

The parametrisation of the observed phenomena can be viewed as in TABLE I:

	Merge	Spell Out	Search	IM	$\text{Search}_{\text{lex}}$	IM_{lex}
Italian	1	0	1	1	0	0
Gungbe	1	1	1	1	0	0
German	1	0	1	1	1	1

Table 1. Language variability in activating FocusP (**?**: 147 (10)).

The factorial combinations of the boolean operators result in fine cross-linguistic analyses of typological variations.

3. Conceptual challenges

While linguistics is Anglo-centric, cartography tends to be Italo-centric, to the effect that much existing literature still utilises Rizzi's (1997) high FocusP for numerous different type of focalisations (corrective, contrastive, wh-fronting, etc.) and takes this to be merged where it is merged in Italian.
Bonan (2022): the projections related to 'focus' are numerous, and languages vary wrt where they merge these projections.

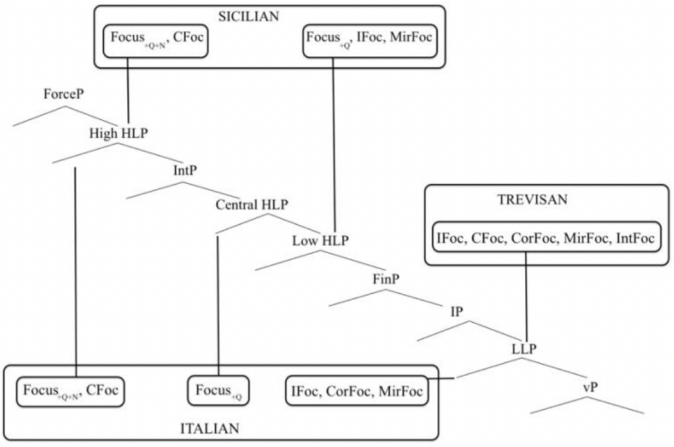

Figure 1. Merging sites for focus projections in 3 Italo-Romance languages (Bonan (to appear)).

The cartography of clefts

?'s (**?**) cartography of clefts derives these structures utilising, depending on the nature of the structure, either **?**'s (**?**) FocusP or **?**'s (**?**) Foc. In both cases, the involved projections require overt movement of the focused element (IM=1), i.e. IM=1.

(7) $[_{\text{ForceP}} \ldots [_{\text{F2b}} [_{\text{FinP}} [_{\text{TP}} \text{COP}_i [_{\text{F1b}} \{\text{focus}\}_j [_{\text{vP}} \_\_j [_{\text{F2}} \{\text{focus}\}_i [_{\text{FinP}} [_{\text{TP}} [_{\text{F1}} [_{\text{vP}} \ldots \_\_i]]]]]]]]]$

(8) Subject cleft (high left-peripheral focus position in embedded clause, F2)

[C'est $[_{\text{F1b}}$ Jean [qui me l'a dit]]
ce=is John who to.me it=has told
'Lit: It is John who told me.'

(9) Non-subject cleft (low peripheral focus position in higher clause, F1b)

[C'est $[_{\text{F1b}}$ à Jean [que je l'ai dit]]
ce=is to John that I it=have told
'Lit: It is to John that I told (this).'

	Merge	Spell Out	Search	IM	$\text{Search}_{\text{lex}}$	IM_{lex}
F2	1	0	1	0	0	0
F1_{b}	1	0	1	0	0	0

Table 2. Parametric settings for the French FocusP: VS.

Based on Table 2, the discussion in the next section, and our results, we shall suggest that the projections involved in Belletti's derivation are indeed focal, but not FocusP and Foc.

FocusP

FocusP has traditionally been considered responsible for the attraction of either wh-elements or contrastive foci. However, while the former are systematically attracted into SpecFocusP in Standard Italian, which suggests a setting as IM=1, the latter can surface either fronted or in situ (**?**), rather suggesting an IM=1/0 setting.

(10) a. Cos'hai mangiato?
what=have_{2PS} eaten

b. * Hai mangiato cosa?
have_{2PS} eaten what
'What did you eat?'

Merge	Spell Out	Search	IM	$\text{Search}_{\text{lex}}$	IM_{lex}
1	0	1	1	0	0

Table 3. Parametric settings for the Italian projection responsible for interrogative wh-movement.

(11) a. UNA MELA ho mangiato, non un gelato!
an apple have_{1PS} eaten not an ice-cream
'Lit: AN APPLE I ate, not an ice-cream!'

b. Ho mangiato UNA MELA, non un gelato!
have_{1PS} eaten an apple not an ice-cream
'I ate AN APPLE, not an ice-cream!'

Merge	Spell Out	Search	IM	$\text{Search}_{\text{lex}}$	IM_{lex}
1	0	0/1	1	0	0

Table 4. Parametric settings for the Italian projection responsible for contrastive focus.

The focus projections defined in Table 3 and 4 cannot be the same.
Clefts are present in the language but require certain context conditions to be met to be licensed (**?**).
Languages like European French, on the other hand, have both shifted and in situ wh-elements (IM=1/0) but no prosodic foci, and productive clefts.

(12) a. Quand (est-ce que) tu l'as vu?
when (est-ce que) you it=have_{2PS} seen

b. Tu l'as vu quand?
you it=have_{2PS} seen when
'When did you see him?'

Merge	Spell Out	Search	IM	$\text{Search}_{\text{lex}}$	IM_{lex}
1	0	0/1	1	0	0

Table 5. Parametric settings for the Italian projection responsible for interrogative wh-movement.

Foc

The existence of Foc was originally posited to account for the existence of VS structures in Standard Italian (**?**) but many pieces of research have suggested that Italian low foci are always unmoved (**???**), thus suggesting an IM=0 setting for the language.

(13) Low focus in Italian
  a. Question: Who's arrived?
  b. Answer A: ?? Gianni è arrivato
     John is arrived
  c. Answer B: È arrivato Gianni
     is arrived John
     'John arrived' (Lit: 'Arrived John')

(14) VS structure in Trevisan[1]
  a. Question: To whom did you give your snack?
  b. ?? Ghe go dato a marenda A GIANI
     3dat have$_{1PS}$ given the snack to John
  c. Ghe go dato A GIANI a marenda
     3dat have$_{1PS}$ given to John the snack
     'Lit: I gave TO JOHN my snack.'

(15) VS structure in Italian (ii)
  a. Question: To whom did you give your snack?
  b. * Ho dato A GIANNI la merendina
     have$_{1PS}$ given to John the snack
  c. Ho dato la merendina A GIANNI
     have$_{1PS}$ given the snack to John
     'I gave the snack TO JOHN'

According to **?**, the different parametric settings for the low Foc in Italian vs Trevisan can be understood as in Table II:

	Merge	Spell Out	Search	IM	Search$_{lex}$	IM$_{lex}$
Italian	1	0	1	0	0	0
Trevisan	1	0	1	1	0	0

Table 6. Language variability in activating FocP.

In French, the low focus position is not exploited in any known structure.

4. Working hypotheses

The working hypotheses behind the present study are as follows:

- The parametrisation of functional projections evolves in the direction of no movement (IM=0 in **?**'s **?** terms, see works on the diachrony of Chinese interrogatives **?** or Japanese, **?**, but also **?**, **?**, a.o.).

	Merge	Spell Out	Search	IM	Search$_{lex}$	IM$_{lex}$
Archaic Chinese (=Trevisan)	1	0	1	0	0	0
Heian Chinese	1	0	1	0/1	0	0
Contemporary Chinese (=Italian)	1	0	1	1	0	0

Table 7. Parametrisations of FocP in the evolution of Chinese.

- Diachronically, one functional projection can display different settings for the same parameter at different stages;
- When what is commonly considered as one single projection displays different parametrisations across structures (e.g., IM=1 in clefts vs IM=0 in interrogatives), the existence of two separate projections ought to be posited instead.

[1] See Bonan(2021) for independent evidence that these orders are not a consequence of rightward movement of what follows the focus.

5. Methodology

Our hypotheses are being tested utilising corpus linguistics techniques. As a preliminary assessment of the theory, we have limited the scope of the present investigation to the study of three standard Romance languages: Italian, French and European Portuguese.

The corpora chosen for this study are:

- Archivio Datini;
- Corpus Epistolare Ottocentesco Digitale;
- Archivio del parlato italiano;
- Banca dati dell'italiano parlato;
- Base de français medieval;
- Groupe d'Observation et de Recherche sur les Documents Epistolaires du Seizième siècle;
- ESLO 1-2;
- 88milSMS;
- Tycho Brahe Parsed Corpus of Historical Portuguese;
- Reference Corpus of Contemporary Portuguese;
- CORP-ORAL: Spontaneous Speech Corpus.

The data, which we classify utilising a parametrisation à la **?**, are starting to shed light on the diachronic evolution of the functional projections under consideration, and especially **?**'s (**?**) FocusP and **?**'s (**?**) Foc. We shall also discuss the consequences of our innovative classification for the received cartography of clefts and the role of FocusP in the theory of interrogatives.

6. The case of French

- interrogative strategies;
- ex situ-in situ alternation;
- regular and inversed clefts (with examples of inversed declaratives from portuguese and trevisan).

7. Preliminary results for French

From predominant ex situ to predominant in situ

	comment		où		quand		quiS		quiO		quoiO	
	EX	IN	EX	IN	EX	IN	EX	IN	EX	IN	EX	IN
1870−1900	60	0	102	0	14	0	94	0	15	0	36	0
1900−1930	40	0	39	0	4	0	33	0	9	0	17	0
1970 (eslo 1)	848	37	233	72	156	38	416	4	42	13	333	198
2014 (eslo 2)	333	156	84	260	32	30	86	6	8	21	24	476

Table 8. Total occurrences of non lexically-restricted wh-elements.

What this means in our framework (the 'n/a' parameters will be discussed in the following two sections):

	Merge	Spell Out	Search	IM	$\text{Search}_{\text{lex}}$	IM_{lex}
late 1800	1	n/a	1	0	n/a	n/a
early 1900	1	n/a	1	0	n/a	n/a
1970s	1	n/a	1	0/1	n/a	n/a
2010s	1	n/a	1	0/1	n/a	n/a

Table 9. Parametric settings for the French FocusP: evolution over time.

Our understanding of French interrogative wh-movement:

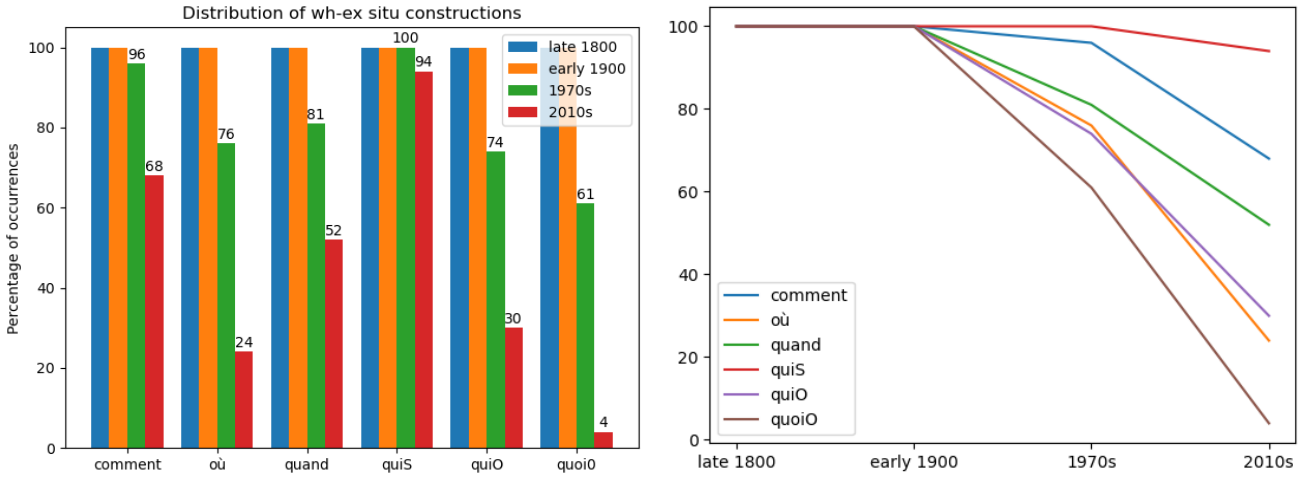

Figure 2. Evolution of the distribution of wh-ex situ structures.

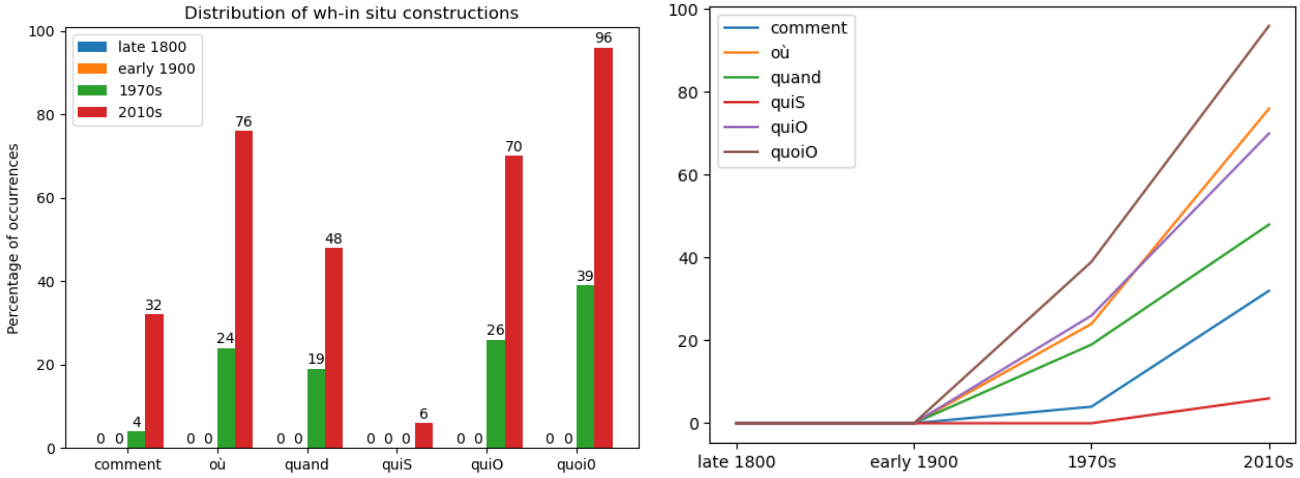

Figure 3. Evolution of the distribution of wh-in situ structures.

- French was a wh-fronting language (IM=0 for FocusP) up until at least the beginning of the 20th century;
- French has been characterized by movement optionality (IM=0/1) since at least the beginning of the 1970s;
- French will evolve to either:
  1. lose fronting altogether (IM=0);
  2. develop a specialised meaning for each strategy (cf. Palasis & Faure 2021).

From verb movement and head activation to predominant non-movement/activation

The case of VS structure vs SV structures.

	comment		où		quand		quiO		quoiO	
	VS	SV	VS	SV	VS	SV	VS	SV	VS	SV
1870–1900	58	2	81	4	11	3	13	1	29	6
1900–1930	35	5	35	3	3	1	8	1	17	0
1970 (eslo 1)	144	435	41	105	62	55	14	23	90	319
2014 (eslo 2)	28	457	12	295	4	45	1	26	2	537

Table 10. Total occurrences of VS and SV (includes ex situ and in situ constructions).

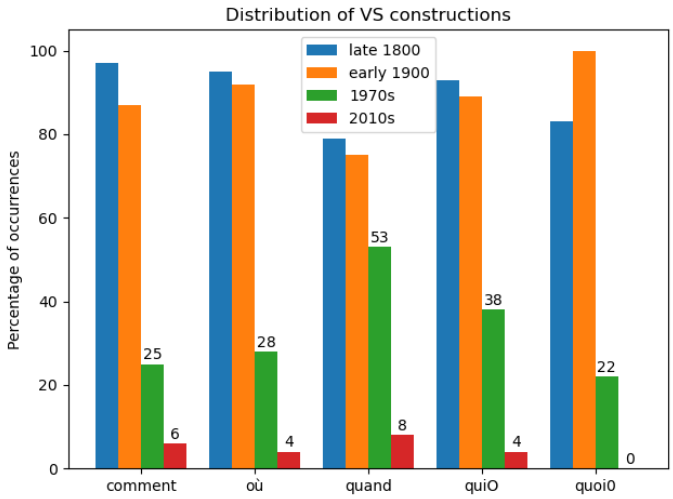

Figure 4. Distribution of VS structures wrt all SV structures.

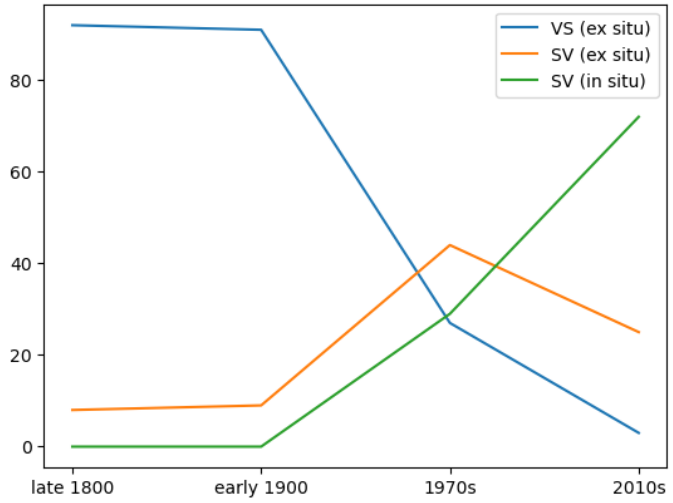

Figure 5. Distribution of VS structures wrt ex situ SV and in situ SV structures.

What this means in our framework:
Two possible parametrisations coexist.
Up until at least the 1930s, French required spelling out the head of Focus (SpellOut=1, see Roberts 2017 for an understanding of French enclitics as an inflectional class) and merging the finite verb (IM_{lex}) almost systematically.

Merge	Spell Out	Search	IM	$Search_{lex}$	IM_{lex}
1	1	1	0	1	1

Table 11. Parametric settings for the French FocusP: VS.

The option of not spelling out $Focus^0$ (SpellOut=0) existed, but was still rare. This systematically correlates to no movement of the finite verb (IM_{lex}=0).

Merge	Spell Out	Search	IM	$Search_{lex}$	IM_{lex}
1	0	1	0	0	0

Table 12. Parametric settings for the French FocusP: SV.

The parametrisation in Table X became dominant before the 1970s.
Forecasted possible evolution: total loss of VS.

From verb movement to predominant non-movement

'Est-ce que/i' structures can be understood as a case of IM_{lex}=1. While head Spell Out cannot exist in the absence of verb movement (VS, see above), verb movement does not require spelling out of the head.

	comment	où	quand	quiS	quiO	quoiO
1870–1900	0	0	0	7	7	0
1900–1930	0	1	0	8	0	0
1970 (eslo 1)	286	157	77	163	13	136
2014 (eslo 2)	44	35	13	17	0	0

Table 13. Total occurrences of VS and SV (includes ex situ and in situ constructions).

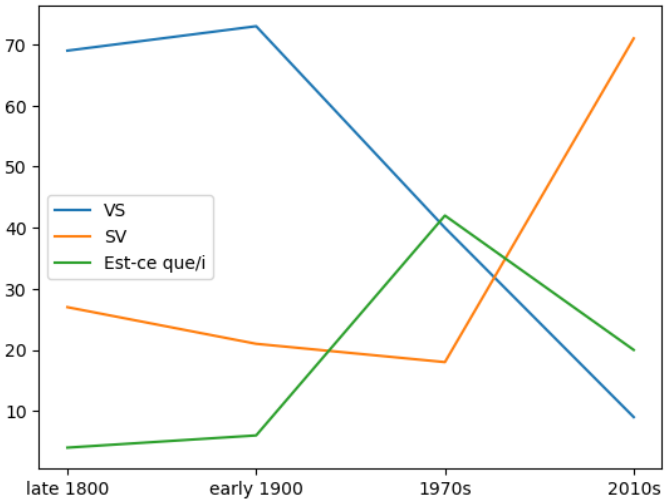

Figure 6. Distribution of 'est-ce que' in constructions with the wh-elements in Table 6.

Est-ce que data in Table 6 seem to suggest that est-ce que was not productive in late 1800, contrary to fact. This is due to the fact that this presentation does not include the data for 'qu(e)'.

Data already collected for 'qu(e)':

- late 1800: 'qu(e)' constitutes 55% of all occurrences:
    1. Que VS: 43%;
    2. Qu' VS: 26%;
    3. Qu'est-ce que: 31%.
- early 1900: 'qu(e)' constitutes 64% of all occurrences:
    1. Que VS: 19%;
    2. Qu' VS: 13%;
    3. Qu'est-ce que: 68%.

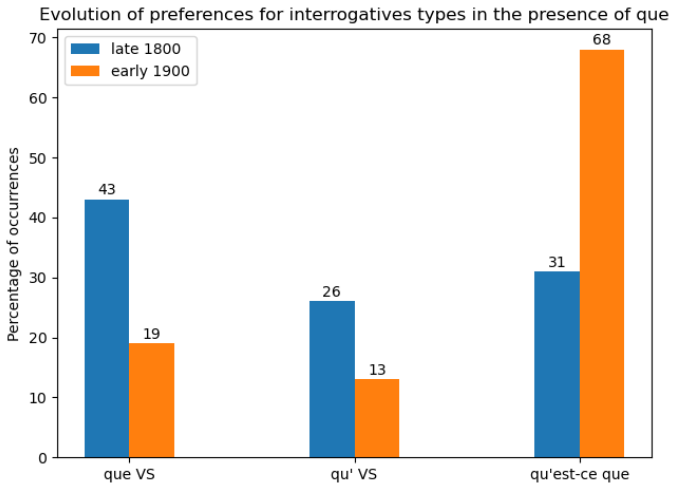

Figure 7. Evolution of preferences for interrogatives types in the presence of que.

What this means in our framework:
There is a third configuration to be added to those in §X. This, as the configuration in Table X (VS), is becoming less common today compared to SV:

Merge	Spell Out	Search	IM	$\text{Search}_{\text{lex}}$	IM_{lex}
1	0	1	0	1	1

Table 14. Parametric settings for the French FocusP: est-ce que.

Possible evolutions:

1. total loss of est-ce que structures;
2. specialisation of est-ce que structures and peaceful coexistence with SV.

8. Concluding remarks

Further steps

The data presented today is only a fraction of the work that we have been doing and plan to do. Certain implementations will include:

- cleaning and analysis of data for 'qu(e)' in eslo1/2 & adjust calculations;
- observation of the evolution from SVS ('Quand Jean a-t-il appelé?') to VS(S) structures ('Quand a-t-il appelé (Jean)?') from 1600 to 2010s;
- observation of the evolution from Wh>Cop interrogative clefts ('Quand c'est qu'il a appelé?') to Cop>Wh ('C'est quand qu'il a appelé?'). For now, we have observed the following:
    1. biclausal interrogative clefts are scarce (DATA);
    2. in late 1800, no clefts were found;
    3. in early 1900, only complex clefts of the 'qu'est-ce que c'est que [...]' type were observed.
    4. in the 1970s, Cop>Wh clefts were already dominant (DATA), although also WH>Cop clefts were present.
    5. in the 2010s, Cop>Wh clefts were not attested.
- repetition of the experience for Italian and Portuguese.

Conclusions

Acknowledgements