CrossMark

J. Linguistics **52** (2016), 37–69. © Cambridge University Press 2015
doi:10.1017/S0022226714000619 First published online 23 April 2015

Preverbal subjects, information structure, and object clitic position in Old Occitan[1]

BRYAN DONALDSON

University of California, Santa Cruz

(Received 23 July 2013; revised 29 July 2014)

The position of object and adverbial clitics remains problematic in Old Occitan syntax (Wanner 2010). This paper analyzes clitic position specifically in affirmative main declaratives with overt preverbal subjects, in which clitics are either preverbal or postverbal with no apparent semantic distinction. Thus, the phrases *En Constantis* ***s'en*** *anet* and *En Constantis anet* ***s'en*** are semantically equivalent, each meaning 'Sir Constantine left', whether the clitics *s'en* 'himself.from-there' appear before or after the verb *anet* 'went'. Previous analyses have concluded that this variation is random (Mériz 1978) or due to regional or dialectal variation (Hinzelin 2007). Neither approach satisfactorily addresses the underlying grammar or the principles underlying the distribution of the variants. The present analysis draws on claims about the left periphery in medieval Romance (Benincà 2006) and reports empirical data from the troubadour biographies (*vidas* and *razos*) and the *vida* of Saint Douceline. Results from 470 subject–verb declaratives establish that the subject in subject–verb–clitic sequences is left-dislocated, albeit covertly so. This sequence is one of several instantiations of subject left dislocation in Old Occitan and usually signals topic shift.

1. Introduction

The position of object (and adverbial) clitic pronouns remains one of the most problematic areas of Old Occitan (OOc) syntax (Mériz 1978: 145; Wanner 2010). This paper addresses one specific clause type in which the behavior of these clitics is apparently variable: affirmative main declaratives that contain an overt preverbal subject (nominal or pronominal). In this context, henceforth an SV DECLARATIVE, clitics (in bold) can be either preverbal, as in (1a), or postverbal, as in (1b).[2]

(1) (a) E.N Constantis **s'en** ANET.
and.Sir Constantine himself.from-there went
'And Sir Constantine left.'
(*Razo* of 80,20 & 80,32, §8; Boutière 1964: 92)

[1] I thank the audience at LSRL 43 (CUNY, 2013), Barbara Vance, and three anonymous *Journal of Linguistics* referees for their valuable discussion and detailed suggestions. I am also grateful to Amanda Dalola for her assistance. All errors and shortcomings remain my own.

[2] In this and subsequent examples, the matrix finite verb appears in small capitals.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

(b) E.N Guilhem ANET **s'en**.
and.Sir Guillaume went himself.from-there
'And Sir Guillaume left.'

(*Razo* of 208,1, §34; Boutière 1964: 3)

In this paper, I challenge the view that this variation is random (e.g., Mériz 1978: 173). I posit that the subjects in (1a, b) represent distinct syntactic structures and that (1b) is pragmatically marked. I discuss the underlying clausal architecture of both variants, appealing to recent work on the clausal left periphery (e.g., Benincà 2006, Ledgeway 2010, Haegeman 2012, Mathieu 2013). Data from the complete troubadour biographies (Boutière 1964) and the *Vida* of Saint Douceline (Gout 1927) uphold the claim that (1b) represents one of several possible realizations of left-dislocated (LD) subjects in OOc, and that the primary function of LD subjects in OOc is to shift topic.

The development of these claims proceeds as follows: Section 2 presents an overview of clitic position in OOc. Section 3 discusses word-order issues in main declaratives, including verb-second (V2) phenomena, verb-third (V⩾3) orders, the clausal left periphery, the optional saturation of the preverbal position, and the position of clitics. In Section 4, I demonstrate that OOc possessed multiple syntactic positions that could host preverbal subjects. Section 5 discusses information structure and discourse organization. Section 6 presents empirical data, and results appear in Section 7. Section 8 contains discussion, and conclusions are offered in Section 9.

2. Object (and adverbial) clitics in Old Occitan

Like the other medieval Romance languages, OOc developed both tonic (strong, disjoint) and clitic (atonic, weak) object pronouns, reflexes of the Latin demonstrative and personal pronouns. Only the clitic series is of concern here; the most frequent forms, including elided forms and common orthographic variants, appear in Table 1 (Anglade 1921: 245–248; Hamlin, Ricketts & Hathaway 1967: 29–30; Lafont 1967: 105; Smith & Bergin 1984: 107–110; Paden 1998: 302; Romieu & Bianchi 2002a: 43–44; Hinzelin 2007: 147).

In finite clauses, clitics in OOc are adjacent to the verb (Skårup 1997: 45; Romieu & Bianchi 2002a: 127).[3] In most finite contexts, clitic position is regular and governed by a traditional understanding of the Tobler–Mussafia Law (Tobler 1875, Mussafia 1886). Thus, clitics in OOc cannot begin a sentence and must be postverbal if no tonic constituent (XP) precedes the verb (Mériz 1978: 154; Smith & Bergin 1984: 114; Romieu & Bianchi 2002a: 130, 163). For example, in (verb-initial) positive imperatives, total interrogatives, and stand-alone verb-first (V1)

[3] Interpolation of the Old Spanish/Ibero-Romance type (Ramsden 1963, Fontana 1993, Poole 2006), in which preverbal clitics are separated from the verb, is not attested in OOc (Hinzelin 2007: 74, 81; Wanner 2010: 271).

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

Person and clitic type		Clitic forms
1 sg		me, mi, m', ·m
2 sg		te, ti, t', ·t
3 sg	masculine	lo, le, l', ·l – li, l', ·l, ·ll, ·il, ·ill, ·lh, ·ilh
	feminine	la, l', ·l – li, l', ·l, ·ll, ·il, ·ill, ·lh, ·ilh
	reflexive	se, si, s', ·s, ·is
	neuter	lo, o, l', ·l
1 pl		nos, ·ns, ·nz
2 pl		vos, ·us, ·s
3 pl	masculine	los, ·ls, ·lz – lor
	feminine	las, ·ls, ·lz – lor
	reflexive	se, si, s', ·s, ·is
Adverbial clitics		i, y; en, ne, n', ·n

Note: Enclitic forms are preceded by a dot. Where applicable, accusative forms are listed on the left of the dash (–); dative forms are on the right.

Table 1
Object clitics in Old Occitan.

declaratives, clitics are typically postverbal (Mériz 1978: 154; Smith & Bergin 1984: 234; Jensen 1986: 110–111; Kunert 2003: 196; Hinzelin 2007: 205, 207, 224). In most clause types that contain a preverbal XP, clitics are preverbal, in accordance with the Tobler–Mussafia Law. When preverbal negation is present, clitics are preverbal in any clause type (Mériz 1978: 157–158; Jensen 1986: 112; Skårup 1997: 45). In subordinates, clitics are usually preverbal (Ramsden 1963: 73, 92–93; Mériz 1978: 158, 161; Hinzelin 2007: 58; Skårup 1997: 45), although Mériz (1978: 159, 162) and Skårup (1997: 45) cite postverbal cases. In partial interrogatives (with a preverbal *wh*-word), clitics are preverbal (Mériz 1978: 158). On the other hand, the position of clitics in certain other finite contexts presents sufficient variation to preclude facile generalizations. In a subjectless coordinate clause with *e* 'and', for example, clitics are variably preverbal or postverbal (Mériz 1978: 289, 296; Romieu & Bianchi 2002b: 124).

This paper focuses on another finite context in which clitics are variably preverbal or postverbal – namely, SV declaratives, as in (1a, b). Syntactically, (1a, b) form an apparent minimal pair: Both contain an overt preverbal third-person singular nominal subject and the same third-person singular form of *s'en anar* 'to leave'. Only the position of the clitics (reflexive *se* 'himself' and *en* 'from there') differs: (1a) presents preverbal clitics (Subject–Object clitic–Verb, henceforth S–Ocl–V), whereas (1b) presents postverbal clitics (S–V–Ocl).

Previous research has sought to identify the default position of clitics in SV declaratives. Some authors identify the default position as preverbal; as Smith & Bergin (1984: 113) note, 'usually the indirect and direct object and neuter pronouns immediately precede the verb'. Smith & Bergin (1984: 236) note further

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

that 'object pronouns usually come right before the verb except in a number of specific circumstances', although the postverbal circumstances enumerated do not include S–V–Ocl declaratives like (1b). For Romieu & Bianchi (2002a: 130) clitics are preverbal when there is a preverbal subject; examples like (1b) are not presented. Mériz (1978: 170–171) finds the preverbal position to be most common but acknowledges numerous examples of postposition. In contrast, other scholars have argued for postposition as a general rule in OOc: Hamlin et al. (1967: 34) note that 'personal pronouns (even when atonic) and reflexive pronouns frequently follow the verb of which they are a complement' (see also Lafont 1967: 415).

Empirically, S–Ocl–V is more frequent than S–V–Ocl. Mériz (1978: 174) found S–V–Ocl to be rare in extensive corpus data. Hinzelin (2007, also 2009), who examined legal charters from the 11th to the 14th centuries, does not report specifically on SV declaratives, but his data from unmarked affirmative main declaratives (the types subsumed under his type Ix, see Hinzelin 2007: 34) show S–Ocl–V to be the most common. However, Hinzelin finds regional variation: S–V–Ocl accounts for only 18.5% (5/27) of cases in charters from Provence (pages 149, 154) and 24% (3/12) of cases in charters from Languedoc (page 163). However, in charters from Rouergue, S–V–Ocl represents 57% (31/54) of the data (page 158). In the present study, as reported subsequently, S–Ocl–V is more frequent than S–V–Ocl. In the biographies, S–V–Ocl amounts to only 11.3% of SV declaratives; in *Douceline* the proportion rises to 26%. However, as I will argue in Section 8, directly comparing rates of S–Ocl–V to S–V–Ocl is conceptually flawed.

In this paper, I suggest that the position of object and adverbial clitics in SV declaratives, and specifically the distribution of S–V–Ocl, has remained elusive because previous accounts have (i) treated SV declaratives as representing single, unified clausal architecture and (ii) considered the variation in clitic position to be random. For example, Mériz (1978: 173) 'discern[s] no pattern at all governing the choice of one position over the other after the subject of the verb' and characterizes the variation as 'apparently random' (page 174). Hinzelin (2007: 190, 204–205, 217) adopts a similar view, arguing that postposition is optional and that both S–Ocl–V and S–V–Ocl represent unmarked clause types. The underlying structures that Hinzelin assumes for these orders will be discussed in Section 4. Hinzelin (2007: 232) concludes that postposed clitics occur as a result of dialect-internal variation.

In contrast, I will propose that the variation between S–Ocl–V and S–V–Ocl is principled and represents two distinct syntactic structures. In particular, I will claim that S–V–Ocl is pragmatically marked and fulfills certain discourse-pragmatic conditions. As a corollary, I assume that, in the context of SV declaratives, S–Ocl–V is a pragmatically unmarked option.[4] Two observations lend prima

[4] An anonymous *JL* referee asks whether transitivity and verb class affect this claim. In the present data, I do not find any effect of transitivity.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

facie support to these claims. First, S–Ocl–V is globally far more frequent than S–V–Ocl, and the skewed proportions alone belie truly random variation. Second, postverbal clitics are lost in most contexts, including SV declaratives, sometime during or after the 14th century (Posner 1996: 171). In Hinzelin's (2007) data, postposition is in sharp decline in all contexts except affirmative imperatives in the 15th century. Crucially, these changes affected not only OOc, but the Romance languages generally (see Ramsden 1963: 128–129).

3. Main declarative word-order phenomena in Old Occitan and Medieval Romance

To demonstrate that OOc possessed (at least) two distinct syntactic positions capable of hosting preverbal subjects, I first discuss the clausal architecture of OOc and medieval Romance. This section (i) analyzes OOc as a V2 language, (ii) discusses V1 and V⩾3 declaratives, and (iii) discusses Tobler–Mussafia effects and the clausal left periphery.

3.1 *Verb second*

Old Occitan exhibited properties of a V2 grammar, as did the other medieval Romance languages (Roberts 1993: 6; Posner 1996: 249; Salvi 2000: 666; Benincà 2006: 61). Of these languages, Old French provides the most intuitive case of V2, in that the verb in main declaratives is typically the second XP (Herman 1954: 362; Marchello-Nizia 1985: 77; Adams 1987: 5; Vance 1997: 38), as in modern Germanic V2 (Holmberg 2015: 343).[5] Despite the somewhat freer word order that the other medieval Romance languages evince in main declaratives, they nonetheless present characteristic features of V2, and have been analyzed as such: for dialects of Old Italian, see Bocchi (2004: 151), Poletto (2005: 227), or Ledgeway (2008: 437: 438); for Spanish, see Fontana (1993: 52) or Poole (2006: 190). Old Occitan has likewise been analyzed as V2 (Lafont 1967: 442; Smith & Bergin 1984: 237; Vanelli, Renzi & Benincà 1985; Paden 1998: 228; Kunert 2003: 207; Benincà 2006; Vance, Donaldson & Steiner 2010: 313; *pace* Sitaridou 2005: 373). A salient property of a V2 grammar is that, in main declaratives, the subject occurs postverbally when another XP precedes the verb. For Vance (1997: 43), this is 'the distinctive mark of all V2 languages', and this exact feature is a 'fundamental trait of OOc syntax' according to Jensen

[5] See Kaiser (2002) for arguments against Old French V2. Note that the label 'verb second' comes from traditional descriptions of German, a fairly 'literal' V2 language. Descriptively, Old Romance V2 is less literal than modern Germanic, allowing main declaratives that are not descriptively V2 (see Section 3.2 below). Even modern Germanic, however, allows some exceptions to surface V2 orders (Holmberg 2015).

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

(Jensen 1994: 362; also Jensen 1986: 388–390; Romieu & Bianchi 2002a: 164; Kunert 2003: 199). Example (2) is typical:

(2) Mal cosselh DONET Pilat.
bad counsel gave Pilate
'Pilate gave bad counsel.' (Vanelli et al. 1985: 166)

Although the verb often appears (descriptively) as the second XP in a main declarative, in the generative perspective, the V2 property is a formal syntactic or structural requirement (Ledgeway 2008: 446n.; Roberts 2012; Holmberg 2015: 376), rather than a constraint on the surface word order (Roberts 2012: 421). Specifically, in a V2 language, generativists largely agree that the finite verb in main declaratives raises from its base position to a (low) functional head in the C-domain (Rizzi 1997: 328n.; Ledgeway 2008: 446n.; Haegeman 2012: 109; Holmberg 2015). Disagreement remains, however, as to which functional head this is, and the landing site of the finite verb may vary across V2 languages (Roberts 2012: 417).[6] A second structural property is that the specifier position of this head is filled by a raised XP (Vance 1997: 18; Holmberg 2015: 375), but this property is not absolute. As Ledgeway (2008: 446n.) notes:

> V2 should not be understood as a simple linear surface constraint, but, rather, as a syntactic requirement that the finite verb raise to a vacant (lower) head position within the C-domain, IRRESPECTIVE OF WHETHER THIS IS ACCOMPANIED IN TURN BY FRONTING OF ONE OR MORE PRAGMATICALLY-SALIENT CONSTITUENTS [emphasis mine].

That is, although the finite verb raises to a head in C, the specifier position of this head may or may not be filled (see also Holmberg 2015: 353, 376).

Most generative accounts of medieval Romance adopt some version of the analysis just sketched. With some variations, this approach underlies Adams (1987), Roberts (1993), Lemieux & Dupuis (1995), Vance (1997), Poletto (2005), Benincà (2006), and Vance et al. (2010: 303), among others. The clausal architecture for main declaratives given in (3) is based on Adams (1987: 8); this architecture will be refined in Section 3.2 below.

(3) $[_{CP}$ XP $[_{C'}$ V_i [+fin] $[_{TP}$ (subject) $[_{T'}$ $[_{VP}$ … t_i …]]]]]

In light of (3), the V2 declarative in (2) possesses the following clausal architecture:

(4) $[_{CP}$ Mal cosselh $[_{C'}$ DONET_i $[_{TP}$ Pilat $[_{T'}$ $[_{VP}$ … t_i …]]]]]
'Pilate gave bad counsel.'

[6] Roberts (2012: 402) suggests that the verb raises to a head in Force, whereas Haegeman (2012: 109) proposes Fin^0, and Benincà (2006: 67) Focus^0, to name three proposals.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

With Vanelli et al. (1985: 168–169; see also Roberts 1993: 95), I assume that the verb *donet* 'gives' has raised from its base position in VP to C^0 and that the direct object *mal cosselh* 'bad counsel' has raised to SpecCP. The movement of these two XPs explains from a structural perspective the postverbal position of the subject, which is blocked from raising to SpecCP because this position is already filled. Otherwise, however, if no XP raises to SpecCP, the subject is free to do so, yielding surface SV order, as in (5). The V2 grammar of OOc thus licenses both preverbal and postverbal subjects in main declaratives.

(5) [$_{CP}$ Aquisti vertutz$_j$ [$_{C'}$ **l'**ENCLINAVA$_i$ [$_{TP}$ t$_j$ [$_{T'}$ [$_{VP}$ t$_i$]]]]]
'This virtue inclined her . . .' (*Douceline*; Gout 1927: 95)

Because the present paper focuses on SV declaratives, I mention without further elaboration that subjects can be null in OOc, in both main and embedded clauses (Hamlin et al. 1967: 33; Smith & Bergin 1984: 111; Jensen 1986: 86–88; Romieu & Bianchi 2002a: 128; 2002b: 35).

3.2 *Verb-third orders and the clausal left periphery*

The model in (5) above accounts for the large proportion of medieval Romance main declaratives that are descriptively V2. As formulated, however, it is ill equipped to address V⩾3 declaratives, in which the verb appears as the third (or higher) XP. Like the other medieval Romance languages, OOc evinces V⩾3 declaratives, as in (6) and (7).[7]

(6) E [per so qu'illi fos a las autras heissemples], [ill] [si METIA] premiera.
'And, so that she would serve as an example to the others, she put herself first.'
(*Douceline*; Gout 1927: 89)

(7) [Illi] [per amor d'ella] [VOLC] esser dicha paura e vieure pauramens.
'For the love of her (Mary), she wants to be considered poor and to live in poverty.'
(*Douceline*; Gout 1927: 58)

In (6), the verb is preceded by a tensed subordinate (*per so qu'*. . . 'so that. . . ') and the subject *ill* 'she'. In (7), both *illi* 'she' and a prepositional phrase (*per amor d'ella* 'for the love of her') precede the verb. For V⩾3 examples in other medieval Romance languages, see Benincà (2006: 69–70) or Mathieu (2013: 345).

Partially in response to V⩾3 data, Skårup (1975) argued that medieval Romance possessed more than one preverbal position and posited an EXTRAPOSITION ZONE, for XPs that occur to the left of the declarative as traditionally defined (page 179) but that are in some way affiliated with the declarative (page 416). More recent work follows Rizzi (1997) in assuming a multi-layered

[7] In these and similar examples, brackets delimit relevant XPs. In example (6), *si* is the third-person reflexive pronoun.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

CP domain that constitutes the left periphery of the clause (e.g., Benincà 2006, Ledgeway 2010, Haegeman 2012, Mathieu 2013). In this approach, CP contains numerous functional projections, subject to various ordering constraints. The clausal architecture adopted for OOc in the present study is that of Benincà (2006: 55), chosen because it was conceived specifically on the basis of medieval Romance data and because my analysis will draw on Benincà's (2006) insights regarding clitic placement. Benincà's (2006) model appears in its simplest form in (8); for similar approaches to Old French, see Labelle (2007) and Mathieu (2013).

(8) $[_{\text{ForceP}}\ [_{\text{FrameP}}\ [_{\text{TopicP}}\ [_{\text{FocusP}}\ [_{\text{FinP}}]]]]]$[8]

In Benincà's model, the Focus projections comprise the leftward edge of the main declarative as traditionally defined: Focus^0 is the equivalent of Adams' (1987) C^0 or (in philological terms) the second position. Likewise, SpecFocus supersedes SpecCP to denote the traditional first position.[9] For Benincà (2006: 67), the landing site of the finite verb in the medieval Romance V2 system is a C head in Focus (i.e., Focus^0). Whereas the clausal architecture in (3) posits only a single preverbal position, the model in (8) with its expanded left periphery readily accounts for V⩾3 declaratives. If, as generative analyses argue, the fundamental property of the V2 grammar involves verb raising to a functional head in the C domain (e.g., to Focus^0) and (frequently) raising of an XP to the specifier of that head (e.g., SpecFocus), rather than the descriptively V2 surface orders that often result (Benincà 2006: 70), then the V2 property, when coupled with a richly articulated left periphery, should unproblematically generate V⩾3 orders. As a further corollary, XPs to the left of SpecFocus, when present, do not 'count' for V2 purposes (Salvi 2000: 667; Rouveret 2004: 190; Labelle 2007: 303; Ledgeway 2008: 466; Mathieu 2013: 341). As Holmberg (2015: 347) writes in his crosslinguistic discussion of V2:

> [W]hatever category is moved to the left periphery (the C-domain) will count as the first constituent for V2, while categories which are externally merged (first-merged, base-generated) in the left periphery ... do not count as first constituents for V2.

In other words, material in TopicP and FrameP projections is syntactically (if not pragmatically) optional in the sense that it does not satisfy any formal

[8] Following Rizzi (1997), FORCE signals the illocutionary force of the utterance; FRAME and TOPIC host various types of topics, FOCUS houses *wh*- and other (perhaps emphatic) elements, and FIN is roughly equivalent to IP, where the finite verb and its arguments are generated.

[9] Two *JL* referees object to the label Focus (and SpecFocus), observing that in many V2 languages, SpecFocus hosts elements that cannot be pragmatically focused, such as expletives. I fully acknowledge the referees' observation; in Swedish, for example, SpecFocus (SpecCP, *vorfeld*, first position, etc.) hosts both topic and focus information (Bohnacker & Rosén 2008: 513–514). Similar results obtain in Old French: In Steiner's (2014) analysis of 600 matrix declaratives from 13th century prose, only 89 (14.8%) represent focus information. However, as Benincà (2006: 54–55) cautions, labels like TopicP and FocusP need not correspond to the pragmatic properties of the XPs in these positions. Similarly, Roberts (2012: 421) comments on the 'essentially arbitrarily semantically-based labels' for these projections.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

requirement related to the V2 property. V⩾3 declaratives are then V2 declaratives that contain additional XPs base-generated, or externally merged, to the left of SpecFocus.

3.3 *Saturation of SpecFocus*

In the OOc V2 grammar, saturation of SpecFocus is not obligatory (Mériz 1978: 154; Skårup 1997: 46; Kunert 2003: 196; Hinzelin 2007: 172). As a result, V1 declaratives are licit – as in other medieval Romance languages (Benincà 2006: 67–68) – with the structure in (9).[10]

(9) [$_{\text{FocusP}}$ [$_{\text{Focus}'}$ QUERIA [$_{\text{FinP}}$ volentiers luocs solitaris]]] on poguessa orar.
'(She) willingly sought out solitary places where she could pray.'
(*Douceline*; Gout 1927: 44)

In Old French, stand-alone (i.e., non-coordinated) V1 declaratives disappeared after about 1200, when saturation of SpecFocus became obligatory (Foulet 1928: 320; Skårup 1975: 291, 300; Rouveret 2004: 200–201). By contrast, in OOc, SpecFocus could remain unsaturated throughout the period of interest here, and Hinzelin (2007: 172) finds V1 declaratives as late as 1480.[11] Moreover, OOc allowed V1 declaratives more frequently than Old French (Skårup 1997: 46).

3.4 *Tobler–Mussafia effects*

The position of object and adverbial clitics in medieval Romance generally obeys the Tobler–Mussafia Law (Tobler 1875, Mussafia 1886), according to which clitics cannot occur in absolute initial position. Thus, if the verb (to which the clitics are strictly adjacent) is preceded by a XP, clitics appear preverbally. On the other hand, if no XP precedes the verb, then clitics are postverbal.

Tobler–Mussafia effects are illustrated in the following examples, which contrast a V2 declarative (10) with a V1 declarative (11). In the V2 declarative, the XP *aquisti vertutz* 'this virtue' precedes the verb *enclinava* 'inclined', and the clitic *l'* 'her' is preverbal. By contrast, in (11) no XP precedes the verb *fez* 'has', and the clitic *lo* 'him' is postverbal.

(10) [$_{\text{FocusP}}$ Aquisti vertutz [$_{\text{Focus}'}$ **l'**ENCLINAVA [$_{\text{FinP}}$]]] a sentir totas las afliccions que vezia ni auzia.
'This virtue inclined her to take to heart all the sufferings that she saw or heard.'
(*Douceline* ; Gout 1927: 95)

[10] On V1 declaratives, see Skårup (1975: 361, 419–420), Roberts (1993: 96), Rouveret (2004: 231), Benincà (2006: 77), or Ledgeway (2008: 442).

[11] Specifically, Hinzelin's (2007: 172) data show V1 declaratives with postposed object clitics, the type that are of interest here (see Section 3.4 on Tobler–Mussafia effects and the position of object clitics; see also Posner 1996: 171).

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

(11) [$_{\text{FocusP}}$ . . . [$_{\text{Focus}'}$ FEZ **lo** [$_{\text{FinP}}$ lo reis]]] a sa charcer gitar.
'The king has him thrown in his prison.' (Mériz 1978: 154)

In Benincà's (2006: 69) interpretation of the Tobler–Mussafia Law, the position of clitics depends on the saturation or non-saturation of SpecFocus. When SpecFocus contains a XP, clitics – if present – are preverbal. On the other hand, if SpecFocus is empty, clitics are postverbal.[12,13] Linking the position of clitics to the saturation of SpecFocus has the advantage over the traditional formulation of the Tobler–Mussafia Law, in that it accounts for S–V–Ocl examples, like (1b) above, in which clitics are postverbal despite a preverbal XP that is seemingly adjacent to the verb.

As a further illustration, consider the case of *quant* 'when' subordinate clauses that immediately precede main declaratives in Old French. As Donaldson (2012; see also Skårup 1975, Benincà 2006) documents, these initial *quant*-clauses (and most other initial subordinate clauses as well) occupy a left-peripheral Frame position of the main declarative in Old French. As a result, SpecFocus is available to host another XP, frequently the adverb *si*. Terminological differences aside, there is widespread consensus that *si* occupies SpecFocus (Fleischman 1991: 263; Ferraresi & Goldbach 2002: 11; Bocchi 2004: 154; Rouveret 2004: 220; Poletto 2005: 224; Benincà 2006: 75), as in (12):

(12) E [$_{\text{FrameP}}$ quant li rois et li autre qui ou palés estoient voient ce,
[$_{\text{TopicP}}$ [$_{\text{FocusP}}$ si [$_{\text{Focus}'}$ **lor** en POISE]]]] molt.
'And when the king and the others who were in the palace saw this, it worried them greatly.'
(*Queste del saint graal*, §1)

In early Old French (before 1200), however, saturation of SpecFocus was not yet obligatory (Rouveret 2004: 200–201), and configurations like (13) are attested.

(13) Et [$_{\text{FrameP}}$ quant te plest, [$_{\text{TopicP}}$ [$_{\text{FocusP}}$ . . . [$_{\text{Focus}'}$ TOLZ **li** [$_{\text{FinP}}$ la vie]]]]].
'And if/when it pleases you, you take his life.'
(*Roman de Thèbes*, 6092)

In (13), the clitic *li* 'to/from him' is postverbal, revealing that SpecFocus is empty. Despite its surface proximity to the verb *tolz* 'takes', the initial *quant*-clause cannot be underlyingly adjacent to the verb. Rather, it occupies FrameP, as in (12). Unlike in (12), however, SpecFocus is empty in (13), and the clitic is postverbal as a result.

An analogous situation arises in OOc. As Vance et al. (2010: 316) found, initial *cant* 'when' subordinate clauses are overwhelmingly left-peripheral in OOc. Just as in early Old French, examples in which SpecFocus is saturated occur alongside

[12] Similar conclusions, although articulated differently, were reached by Skårup (1975: 397).
[13] I do not espouse a particular theory of cliticization, nor am I concerned here with processes of cliticization.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

examples in which SpecFocus remains empty, and the position of clitics in the declarative changes predictably, witness (14) and (15).

(14) [$_{\text{FrameP}}$ Cant illi retornava de sos autz raubimens, [$_{\text{TopicP}}$ [$_{\text{FocusP}}$ enaissi [$_{\text{Focus}'}$ **si** MOSTRAVA]]]] paura e defallens e humana a totas.
'When she came out of her holy rapture, she appeared poor and feeble and human to all.' (Vance et al. 2010: 314)

(15) [$_{\text{FrameP}}$ Cant li maires fon morta, [$_{\text{TopicP}}$ [$_{\text{FocusP}}$ . . . [$_{\text{Focus}'}$ MUDERON **si**]]]] az Ieras.
'When the mother died, they moved themselves to Hyères.'
(*Douceline*, Gout 1927: 44)

The same approach yields fruitful results when applied to S–Ocl–V and S–V–Ocl.

4. Multiple positions available for preverbal subjects

I propose that preverbal subjects in OOc affirmative SV declaratives occupy one of two distinct positions: In the unmarked case, the subject occupies SpecFocus, syntactically adjacent to the verb in Focus^0. I will refer to this default case as a CONVENTIONAL (preverbal) subject. In the pragmatically marked case, the preverbal subject occupies a TopicP position, to the left of SpecFocus and hence not adjacent to the verb (regardless of the eventual surface order). The spirit of this proposal is not new, as Skårup (1975: 396–397) posited (leftward) extraposed subjects in OOc and Old French (see also Skårup 1997: 46).

The present analysis treats left-peripheral subjects as instances of subject LD, following work on both historical and contemporary Romance (e.g., Lambrecht 1981, Ashby 1988, Anagnostopoulou 1997, Marchello-Nizia 1998, Cardinaletti 2004, Frascarelli 2007, Labelle 2007, Mathieu 2013). From these studies emerge two relevant observations. The first, broad but crucial, is the presence of multiple positions capable of hosting preverbal subjects. The second is a consensus that LD subjects occupy a TopicP position (Rizzi 1997: 324; Benincà 2006: 58; Frascarelli 2007: 693; Labelle 2007: 303) in the languages under discussion.

As a first step, I will show that OOc possesses LD subjects. Importantly, in the following examples, SpecFocus is overtly saturated, and the subject is visibly to the left of SpecFocus. That OOc possesses conventional preverbal subjects (in SpecFocus) is evident from examples (1a) and (10) above, where no XP intervenes between the subject and the verb, and clitics are preverbal.

In the biographies, sequences of subject + adverb *si* + verb (S–*si*–V) are particularly common in declaratives (Jensen 1972: 25). If *si* occupies SpecFocus as claimed, then it follows that a subject to the left of *si* is left-peripheral or

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

left-dislocated. Accepted theory places it in a TopicP projection, as in (16).

(16) [$_{\text{ForceP}}$ [$_{\text{FrameP}}$ [$_{\text{TopicP}}$ Bertrans de Born [$_{\text{FocusP}}$ si [$_{\text{Focus}'}$ FO]]]]]
uns castellans.
'Bertran de Born was a nobleman.'
(*Vida* XIA; Boutière 1964: 65)

Left-dislocated subjects followed by *si* can also be pronominal:[14]

(17) Mas ella si AMAVA N'Ugo de Lasigna.
'But she loved Sir Hugo of Lasigna.'
(*Razo* of 167,52; Boutière 1964: 180)

Left-dislocated subjects can also be reprised by a preverbal subject pronoun (Sp), rather than *si*:

(18) En Richartz, cant saup que tuich aquist l'avion abandonat, el s'en VENC denant Autafort.
'When Sir Richard learned that all those had abandoned him, he came before Hautefort.'
(*Razo* of 80,21; Boutière 1964: 103)

Both the LD subject and the clause-internal resumptive element may be subject pronouns:

(19) Et el, per gran aver del rei Enric, el li DIS ...
'And he, on account of the great wealth of King Henry, he told him ...'
(*Razo* of 80,32; Boutière 1964: 107)

I take these examples as evidence that OOc licenses subjects to the left of SpecFocus, and that both conventional and LD subjects are possible in main declaratives.[15] Regardless of eventual discourse-pragmatic differences, these two types of preverbal subject are structurally distinct, a difference that has not always been understood or acknowledged. For example, Hamlin et al. (1967: 73n.1) and Paden (1998: 228) treat *si* in S–*si*–V as a clitic that co-occurs with the verb, with the subject in SpecFocus.[16] Instead, drawing on Benincà (2006), I propose the

[14] Subject pronouns in medieval Romance are tonic, at least when preverbal (Vanelli et al. 1985: 164); for OOc, see Smith & Bergin (1984: 111) or Romieu & Bianchi (2002b: 43). In these data, all the subject pronouns have an identifiable discourse referent (none are pleonastic).

[15] Like other V⩾3 orders, LD subjects occur in other medieval Romance varieties. For example, S–*si*–V occurs in Old Italian, illustrated in (i), and Old French, illustrated in (ii):

(i) L'altre ami si EST la moiller.
'The other friend is the wife.' (Benincà 2006: 69)

(ii) et la raison si ESTOIT ...
'and the reason was ...' (*Roman de Cassidorus*, §251)

[16] This is not to criticize Hamlin et al. 1967 and Paden (1998); it is commendable that these authors discuss S–*si*–V at all, given its absence from other works (e.g., Anglade 1921; Jensen 1986, 1994).

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

following structures for SV declaratives: Conventional preverbal subjects have the representation in (20a), whose terminology supersedes that of (5), and LD subjects that in (20b):

(20) (a) [ForceP [FrameP [TopicP [FocusP Subject [Focus′ VERB]]]]]
(b) [ForceP [FrameP [TopicP Subject [FocusP (XP) [Focus′ VERB]]]]]

As shown in Section 3.3 above, OOc allows SpecFocus to remain empty. With this in mind, the minimal pair of SV declaratives in (1a, b) can be represented as in (21a, b). In the interests of syntactic accuracy, the elided sequence *E.N* 'and Sir' of the original is (artificially) separated in (21a, b) into its components *E* 'and' and *En* 'Sir'.

(21) (a) E [ForceP [FrameP [TopicP [FocusP En Constantis [Focus′ **s'en** ANET]]]]]
'And Sir Constantine left.'
(*Razo* of 80,20 & 80,32, §8; Boutière 1964: 92)

(b) E [ForceP [FrameP [TopicP En Guilhem [FocusP ... [Focus′ ANET **s'en**]]]]]
'And Sir Guillaume left.'
(*Razo* of 208,1, §34; Boutière 1964: 325)

Adopting Benincá's (2006) claim that the position of clitics depends on the saturation of SpecFocus, the subject *En Constantis* 'Sir Constantine' in (21a) occupies SpecFocus, and the clitics are preverbal; this is a conventional preverbal subject. By contrast, in (21b), the subject occupies a TopicP position in the left periphery, so it is an LD subject. SpecFocus is empty, however, and as a result, the clitics are postverbal. The subject in (21b) is thus COVERTLY left-dislocated: Although superficially next to the verb in the surface structure, it is crucially NOT underlyingly adjacent the verb.

Although Skårup (1975) proposed multiple preverbal positions in OOc, the idea has gained little currency in subsequent work on OOc syntax. Mériz (1978), for example, does not cite Skårup, and neither do recent manuals or reference works (e.g., Smith & Bergin 1984, Paden 1998, Romieu & Bianchi 2002a, b). Vance (2011) posits multiple preverbal positions, but Hinzelin (2007: 118) rejects Skårup's analysis on the grounds that it flouts the conventional (i.e., linear, pre-Benincà) understanding of the Tobler–Mussafia Law, although Hinzelin alludes to the possibility of LD in OOc, as discussed in Section 5.3. Instead, Hinzelin (2007: 204–205, 221) posits that a single position – SpecFocus – hosts subjects in SV declaratives, regardless of whether clitics are preverbal or postverbal. Furthermore, Hinzelin considers both S–Ocl–V and S–V–Ocl to be unmarked and attributes variation between them to regional differences.

However, Hinzelin (2007: 118n.) raises a valid critique of an approach à la Skårup (1975) that posits multiple positions for preverbal subjects: Without a principled justification for choosing one structure over the other, it is circular to claim that the subject occupies the left periphery in S–V–Ocl sequences but SpecFocus in S–Ocl–V cases.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

5. INFORMATION STRUCTURE

5.1 *Topic and focus*

To resolve this difficulty, I turn to information structure, and particularly the notion of TOPIC. Following Reinhart (1981: 54), the topic is what an utterance is ABOUT. Similarly, Lambrecht (1994: 131) writes:

> A referent is interpreted as the topic of a proposition if in a given situation the proposition is construed as being about this referent, i.e., as expressing information which is relevant to and which increases the addressee's knowledge of this referent.

Topics are generally already known and readily accessible to the discourse participants, usually because they have been mentioned previously in the discourse. FOCUS information, in contrast, is new information predicated about the topic; it is the comment made about the topic. Focus is the unpredictable, unanticipated information that moves the discourse forward (Firbas 1964: 270; Lambrecht 1994: 213), and most assertions contain both topic and focus (i.e., topic–comment).[17]

5.1.1 *Sentence topics and discourse topics*

Following van Dijk (1976), Reinhart (1981), Givón (1983), and Lambrecht (1994), I distinguish between two types of topic – SENTENCE TOPIC (ST) and DISCOURSE TOPIC (DT) – on the basis of their prominence in the discourse, that is, how long they remain topical. In previous research, the longevity of topics has been measured in sentences. For Reinhart (1981: 54), STs remain topical only within the confines of a single sentence (matrix clause).[18] In contrast, a DT '[designates] a topic expression whose referent is pragmatically salient beyond the limit of a single sentence' (Lambrecht 1994: 117; see also van Dijk 1976: 50; Reinhart 1981: 54). To return to Reinhart's (1981) notion of ABOUTNESS, a DT is what a stretch of discourse (comprised of multiple sentences) is about. Unlike an ST, whose longevity as a topic is quite restricted, a DT creates topical or thematic continuity, serving as a *leitmotif* of a thematic paragraph (Givón 1983: 8). In addition, the DT is usually the participant 'most crucially involved in the action sequence' of the thematic paragraph (Givón 1983: 8; see also Lambrecht 1994: 117). In the present study, a DT is defined as a referent that remains topical across two or more contiguous main declaratives.

Although a thematic paragraph is globally ABOUT a certain DT, other discourse referents may appear within this thematic paragraph, and these referents sometimes become STs. As van Dijk (1976: 53) emphasizes, however, such STs are 'momentaneous' and are usually related to the larger DT; they serve as comments

[17] TOPIC and FOCUS are discourse-functional terms, separate from the labels TopicP and FocusP.
[18] Main declaratives generally carry the narrative burden, with subordinates providing incidental, collateral information (Fleischman 1990: 188; van Reenen & Schøsler 1992: 106).

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

on the larger DT without changing it. Nevertheless, the presence of topics other than the DT within the thematic paragraph creates potential interference or competition between topics (Givón 1983: 11, 14).[19]

5.2 *Topic marking*

Discourse-pragmatic notions like topic and focus can be marked grammatically to make an utterance easier for a hearer to comprehend (e.g., Lambrecht 1994: 3). Crosslinguistic research has revealed both that the grammatical subject usually represents the topic of the utterance (Reinhart 1981: 62) and that LDs are typically topics (Ashby 1988: 212; Prince 1997: 124; Lambrecht 2001: 1072). It ensues that subjects are the most frequently left-dislocated XPs (Lambrecht 2001: 1054).

Within the general description of topics, more specialized functions can be identified. Here, I follow Frascarelli & Hinterhölzl (2007: 87), who distinguish between aboutness topics, contrastive topics, and familiar topics; in particular I will retain Frascarelli & Hinterhölzl's further opposition between SHIFTING TOPICS and CONTINUING TOPICS. Frascarelli & Hinterhölzl's definition of aboutness topics adheres to Reinhart (1981) and Lambrecht (1994). With respect to discourse organization, Frascarelli & Hinterhölzl (2007: 89) note that many aboutness topics possess the function of shifting topics, in that they encode a referent that is 'newly introduced, newly changed or newly returned to' (Givón 1983: 9). That is, a shifting topic (which is always an aboutness topic) signals a rupture with the previous topic by orienting the discourse toward a new topic. A contrastive topic 'induces alternatives which have no impact on the focus value and creates oppositional pairs with respect to other topics' (Frascarelli & Hinterhölzl 2007: 88). Finally, familiar topics involve referents that are already accessible in the discourse, and familiar topics sometimes represent continuations of the already established DT (pages 88, 95), in which case they are continuing topics. Both shifting topics and continuous topics qualify as aboutness topics (page 111), but whereas shifting topics signal a rupture in topic, continuing topics extend the existing topic.

5.3 *Information structure in Old Occitan*

Previous studies of OOc syntax have not linked discourse structure to clitic position. Ramsden (1963) examined the nature of the preverbal constituent to explain clitic position, but although he looked beyond grammatical categories in an effort to capture 'the whole sense of the utterance' and the strength of the 'union of verb and preceding element' (page 38), his analysis did not extend beyond individual phrases to encompass the surrounding discourse. Similarly,

[19] For similar observations on competition between multiple potential topics in Old Neapolitan, see Ledgeway (2008: 462).

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

both Skårup (1975, 1997) and Mériz (1978) focused on individual phrases devoid of discourse context. Although Combettes (1971) and Kunert (2003) examined OOc from an information-structural perspective, neither examined LD subjects or specifically addressed the position of clitics in SV declaratives, and in fact, Combettes (1971: 298, 299, 301) subsumed S–*si*–V sequences under SV. Hinzelin (2007) invoked information structure (e.g., pages 149, 202–203, 232), but only in passing, and without considering how it interfaces with syntactic structure. For example, Hinzelin's data from early OOc charters provide three occurrences of the sequence NOMINAL DIRECT OBJECT + VERB + CLITIC(S), with postverbal clitics. Although Hinzelin (2007: 149–150) analyzed these examples as topic constructions and instances of (clitic) LD, they are absent from the formal syntactic analysis (pages 217–230), and the syntactic representation of LDs is not made explicit. Hinzelin (2007) analyzed no further examples from his OOc data as LD (see pages 154–172). Importantly, although the same early charters (Rouergue region) evince examples of S–V–Ocl, Hinzelin analyzed these not as topic or LD structures but rather as unmarked SV declaratives (pages 204–205). As I argue in Section 7 below, however, S–V–Ocl marks topic in the present data (and perhaps in the Rouergue charters as well; see Section 8).

Data from visibly left-dislocated subjects in the OOc biographies demonstrate the link between LD subjects and topic marking. This section synopsizes the results of Donaldson (to appear), based on 326 tokens from the biographies in which the presence of one (or more) XPs between the subject and the verb shows visibly that the subject is left-dislocated. In these data, the most common configuration is NP S–*si*–V (n = 296); in addition, there are 20 tokens of Sp–*si*–V, eight tokens of NP S–Sp–V, and two tokens of Sp–Sp–V. With few exceptions, these LD subjects represent shifting topics. As seen in Table 2, 35.0% elevate

	Shifting topic			Continuing topic	
	DT	New referent + DT	ST	DT	Total
S–*si*/Sp–V	114 (35.0%)	148 (45.4%)	45 (13.8%)	19 (5.8%)	326

DT = discourse topic, ST = sentence topic

Table 2
Discourse function: Occurrences and discourse functions of S–*si*–V and S–Sp–V in the biographies.

a known referent to the status of DT, and another 45.4% fulfill a dual function of introducing a new referent to the discourse and simultaneously marking it as the DT.[20] A smaller portion of the LDs (13.8%) represents shifting topics that

[20] This is how LD functions in modern English (Prince 1997). These same functions appear to characterize the Neapolitan double-subject construction discussed by Ledgeway (2010).

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

establish an ST. Finally, a small number (5.8%) represent continuing topics, in that they clarify or confirm the existing DT, often in the presence of other referents potentially competing for topichood.

Thus, 94.2% of these data (307/326) contain shifting topics. Within this subset of the data, LD subjects preferentially mark DTs as opposed to STs: Of the 307 tokens of shifting topics, 85.3% (262/307) establish a DT. The following extract illustrates how S–*si*–V LDs function as shifting topics. (I have introduced internal section divisions, shown in brackets, and the relevant LDs are underlined.)

(22) [1] Ben aves entendut d'En Raimon de Miraval qui fo ni don, en la razo que es escriuta denan las soas chansos; per que no vos vueill dire d'aiso qu'ieu vos ai dig. [2] Mas el si AMAVA una dona de Carcases, que avia nom la Loba de Pueinautier, filha d'En R. de Puegnautier, et era moiller d'un cavalier ric e poderos de Cabares e parers del castel de Cabaretz. [3] La Loba si ERA sobravinens e bela e corteza e ensenhada, e volontoza de pretz e d'onor; [4] e tuit li valen home e·ill valen baro de l'encontrada, e l'estranh que la vezion si entendion en ela: lo coms de Fois, En Oliviers de Saisac, En Peire Rotgiers de Mirapeis, En Aimericx de Monreial e·N Peire Vidal, que fes maintas bonas chansos de leis. [5] En Raimons de Miraval si l'AMAVA sobre totz coralmen.
'[1] You have certainly heard about Sir Raimon de Miraval, who he was and where he was from, in the *razo* that is written in front of his songs, I do not want to tell you what I have already said. [2] But he loved a lady from Carcassès, who was called la Loba de Pueinautier, daughter of Sir R. de Puegnautier, and was the wife of a rich and powerful knight, from the castle of Cabaretz. [3] La Loba was extremely gracious and beautiful and courteous and educated, and desired esteem and honor; [4] and all the worthy men and the worthy barons of the country, and all the strangers who saw her desired her: the count of Foix, Sir Olivier de Saissac, Sir Peire Rotgier de Mirepois, Sir Aimeric de Montréal, and Sir Peire Vidal, who composed many good songs about her. [5] Sir Raimon de Miraval loved her, more than any of the others, from his heart.'

(*Razo* of 406, 4, 27, 38 & 364, 21; Boutière 1964: 384)

After the narrator's introduction in [1], the S–*si*–V at the beginning of [2] establishes Sir Raimon de Miraval as the DT. Sir Raimon remains the topic throughout [2], in which we learn of his love for la Loba ('she-wolf'), a woman from the region of Carcassès whose background is sketched. At the beginning of [3], a S–*si*–V shifts to a new DT, la Loba. All the new information proffered in [3] and [4] is about la Loba: her desire for recognition, her numerous suitors (named as purely collateral information), and that songs have been written about her. By contrast, in [5], the DT shifts back to Sir Raimon, signaled by the S–*si*–V at the beginning of [5].

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

The distribution of visibly left-dislocated subjects (S–*si*–V, S–Sp–V) in the biographies confirms their status as a pragmatically marked word order whose primary function is that of a shifting topic, usually to a DT. The data also confirm Fleischman's (1991: 265) claim that *si* divides an utterance into a topic and comment as well as Kunert's (2003: 207) observation that OOc phrases usually possess a topic–comment structure (also Combettes 1971: 297). Given the tight association between LD and topic marking, the discourse-pragmatic function of S–*si*–V corroborates analyses of medieval Romance that have considered S–*si*–V to be LD on primarily structural (i.e., purely syntactic) grounds, including Marchello-Nizia (1985: 169), Fleischman (1991: 261), Vance (1997: 61), and Ferraresi & Goldbach (2002: 17).

This section gave the rationale for positing multiple preverbal positions available to host subjects in OOc, on the basis of examples in which SpecFocus is overtly saturated and the left-peripheral nature of the preceding subject is visible (i.e., S–*si*–V, S–Sp–V). In the biographies, these sequences fulfill a particular discourse-pragmatic function, that of a shifting topic. In most cases, the sequence establishes a DT, either by using an already known discourse referent or by introducing a new discourse referent that is immediately topical. In Section 7, this analysis will be extended to the distinction between S–Ocl–V and S–V–Ocl main declaratives.

6. Data

The data come from complete analyses of the entire body of OOc troubadour biographies as well as the *Vida* of Saint Douceline. The biographies were composed in the 13th and 14th centuries (Boutière 1964: viii; Jensen 1972: 22), and *Douceline* dates from the early 14th century (Henrichsen 1955: 186; Garay & Jeay 2001: 14). The biographies are in prose, although they regularly cite short extracts of verse works; these verse extracts were excluded. *Douceline* is entirely in prose.[21]

Considerable dialectal variation existed within the OOc zone (Bec 1963: 34–58), and few if any OOc texts are dialectally pure (Lafont 1967: 23; Smith & Bergin 1984: 4). The emergence of a stable written *koiné* sometime after 1100 (Bec 1963: 69; Smith & Bergin 1984: 4) did not preclude inconsistencies in morphology and pronunciation (Lafont 1967: 23–24). Although the language of the biographies reveals Italian influences, especially in morphology and spelling (Boutière 1964: x; see also Henrichsen 1955), it is accepted that the authors were native OOc speakers (Schutz 1938: 232; Boutière 1964: x). Importantly, given the relative paucity of OOc prose (Henrichsen 1955: 12), the biographies constitute the largest extant body of OOc literary prose (Schutz 1938: 225) and form a

[21] When available, prose is preferable to verse for syntactic analysis (Henrichsen 1955: 12) to avoid possible effects of meter and rhyme on word order (Ramsden 1963: 33–34; Mériz 1978: 152; Kunert 2003: 195).

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

literary genre with considerable linguistic and literary consistency (Boutière 1964: viii). Rohlfs (1970: 45) considers them 'splendid examples of Provençal prose', and selections of the biographies appear in most anthologies and manuals on OOc.

The biographies and *Douceline* were searched for all instances of affirmative SV declaratives that contained object or adverbial clitics (preverbal or postverbal). Only affirmative declaratives were retained because clitics are systematically preverbal in the presence of a preverbal negator. All overtly left-dislocated subjects (i.e., those reprised by *si* or an Sp) were also excluded; these data are discussed in Donaldson (to appear). Finally, SV declaratives devoid of clitics were excluded, as the absence of object renders uncertain the position of the subject. All tokens were coded for type of subject and the position of clitics. In addition, all the tokens of S–V–Ocl were coded for discourse function (shifting topic vs. continuing topic; ST vs. DT; focus). As mentioned, only referents that remained topical across at least two contiguous main declaratives qualified as DTs. As a control measure, the first 100 tokens of S–Ocl–V from each text were coded for the discourse function of the subject referent.

7. Results

The biographies yielded 301 tokens of SV declaratives with clitics, and *Douceline* yielded 169; the data concerning clitic position appear in Tables 3 and 4.

	Clitic position		
Subject type	S–Ocl–V	S–V–Ocl	Total
Nominal	145	27	172
Pronominal	122	7	129
Total	267 (88.7%)	34 (11.3%)	301

Table 3
Overall frequency of preverbal and postverbal object clitics in SV declaratives: Biographies.

	Clitic position		
Subject type	S–Ocl–V	S–V–Ocl	Total
Nominal	69	33	102
Pronominal	56	11	67
Total	125 (74.0%)	44 (26.0%)	169

Table 4
Overall frequency of preverbal and postverbal object clitics in SV declaratives: *Douceline*.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

In the biographies and *Douceline* alike, S–Ocl–V is more frequent than S–V–Ocl, but S–V–Ocl appears productive in these texts. Both types of subject (nominal and pronominal) occur with both possible clitic positions (preverbal and postverbal). Nominal subjects are more common, as they are in these data more generally.

7.1 *Discourse functions*

If S–Ocl–V and S–V–Ocl represent distinct clausal architectures and fulfill different discourse-pragmatic functions, as I hypothesize, then their distribution should vary according to discourse function. Specifically, I claim that subjects in S–V–Ocl are dislocated and, like S–*si*–V and S–Sp–V, represent shifting topics. The data support this hypothesis. The biographies contain 34 tokens of S–V–Ocl, and all encode topic (Table 5). Specifically, 94.1% of the S–V–Ocl data are shifting topics, and the remaining 5.9% serve as continuing topics. Predictably, none of the LD subjects can be construed as focus information. Looking only at the LD subject shifting topic data, in 68.8% of the cases (22/32 tokens), the subject is established as a DT, either by selecting an existing discourse referent or by introducing a new one. The remaining 31.3% of the cases (10/32 tokens) of shifting topics create an ST, distinct from the topic of the previous main declarative, but which ultimately does not supersede the existing DT. The dominant function of S–V–Ocl sequences in the biographies is thus as a shifting topic, often to a new DT.

	Shifting topic			Continuing topic		
Sequence	DT	New referent + DT	ST	DT	Focus	Total
S–V–Ocl (dislocated)	20 (58.8%)	2 (5.9%)	10 (29.4%)	2 (5.9%)	0	34
S–Ocl–V (conventional)	5 (5%)	0	43 (43%)	51 (51%)	1 (1%)	100*

DT = discourse topic, ST = sentence topic
* Only the first 100 tokens of conventional subjects were coded for discourse function.

Table 5
Discourse functions of S–V–Ocl and S–Ocl–V: Biographies.

The data from conventional subjects reveal a different picture. In the 100 tokens of S–Ocl–V whose discourse function was coded, 48% represent shifting topics, 51% represent continuing topics, and a single token contains focus information (Table 5). Comparing the results from LD and conventional subjects in the biographies reveals a significant difference in the correspondence between discourse-pragmatic function (shifting vs. continuing topics) and sentence form

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

(LD vs. conventional subjects), χ^2 (2, N = 4) = 20.12, $p < .0001$. However, another major difference exists within the shifting topics: Although both LD and conventional subjects can serve as shifting topics, LD subject shifting topics most frequently establish DTs (68.8%), as seen previously, whereas conventional subject shifting topics rarely establish DTs (10.4%; 5/48 tokens). In fact, when conventional subjects serve as shifting topics, it is overwhelmingly to establish an ST (89.6%; 43/48 tokens). Within the data for shifting topics, the distribution of DTs and STs differs significantly between LD and conventional subjects, χ^2 (2, N = 4) = 26.67, $p < .0001$. Globally, then, whereas LD subjects most often signal a shift of (discourse) topic in the biographies, conventional subjects maintain the existing DT, either by introducing an ST (whose topicality is only momentary) or by restating the existing DT, usually in the presence of other referents that could potentially be misconstrued as topics.

Similar and even more strongly pronounced patterns obtain in *Douceline* (Table 6). Of the 44 tokens of S–V–Ocl, 95.5% represent shifting topics, and only 4.5% represent continuing topics; none of the tokens can be construed as focus. Within the S–V–Ocl shifting topics, 95.2% (40/42 tokens) establish a new DT; only 4.8% (2/42 tokens) establish STs. As in the biographies, the predominant function of S–V–Ocl is to shift to a new (discourse) topic.

The data from conventional subjects in *Douceline* again reveal striking similarities with the biographies. Of the 100 S–Ocl–V tokens analyzed, 27% represent shifting topics, 64% represent continuing topics, and 9% represent focus. As in the biographies, the correspondence between discourse-pragmatic function (shifting vs. continuing topics) and sentence form (LD vs. conventional subjects) is significantly different, χ^2 (2, N = 4) = 48.77, $p < .0001$. Again, within the data from shifting topics, the proportions of DTs to STs differ significantly according to subject type (LD vs. conventional), χ^2 (2, N = 4) = 46.01, $p < .0001$. As in the biographies, S–V–Ocl in *Douceline* usually shifts to a new DT, whereas S–Ocl–V maintains the existing DT.

Sequence	Shifting topic			Continuing topic	Focus	Total
	DT	New referent + DT	ST	DT		
S–V–Ocl (dislocated)	31 (70.5%)	9 (20.5%)	2 (4.5%)	2 (4.5%)	0	44
S–Ocl–V (conventional)	3 (3%)	0	24 (24%)	64 (64%)	9 (9%)	100*

DT = discourse topic, ST = sentence topic
* Only the first 100 tokens of conventional subjects were coded for discourse function.

Table 6
Discourse functions of S–V–Ocl and S–Ocl–V: *Douceline*.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

7.2 *Illustration of shifting topic to establish new discourse topic*

The passage in (23) illustrates how S–V–Ocl functions as a shifting topic to signal a new DT; the passage is an extract from a *razo* that purports to explain a poem by Guillem de Balaun.

(23) [1] Et ela dis que, puois qe a lui tant plaisia, q'ela volia prendre la vendeta e far lo perdon; mas si volia qe la vendeta fos aitals: [2] q'En Guilelms se traises la ongla del det plus long e que la·il aportes. [3] Bernarz d'Andusa s'en tornet a·N Guilelm e dis li la vendeta q'ela volia e·l perdon q'ela li fasia. [4] Et el FO·**n** lo plus legres hom del mon. [5] Et ades se fetz liar lo deit e la ongla traire fora. [6] Et montet a caval com En Bernart d'Andusa e venc s'en a ma dompna Guilelma, a Jaujac.
'[1] And she said that, since he insisted, that she would exact vengeance and then forgive [him]; but she wanted the vengeance to be such: [2] that Sir Guillem would pull off the fingernail of his longest finger and bring it to her. [3] Bernard d'Anduze returned to Sir Guillem and told him the vengeance that she wanted and the forgiveness that she would give. [4] And he was the happiest man in the world on account of this. [5] And right away he had his finger tied and the fingernail pulled off. [6] And he mounted a horse with Sir Bernart d'Anduze and went to Lady Guillelma, at Jaujac.'
(*Razo* of 208,1; Boutière 1964: 327–328)

At the beginning of the extract, Lady Guillelma is the DT. In the previous discourse, she has been wronged and humiliated by Sir Guillem. In [1], Lady Guillelma agrees to pardon Sir Guillem, but at a price. In [2], she stipulates to Bernard d'Andusa – a messenger – that Sir Guillem is to pull out a fingernail and bring it to her. In [3], Bernard d'Andusa communicates this condition to Sir Guillem. *Bernarz d'Andusa*, no more than an ST here, appears as a conventional subject; the larger discourse is not about Bernard. In [4], an LD subject shifts the DT to Sir Guillem: In the sequence *E el fo·n lo plus legres hom del mon*, the clitic *en* 'of it' is postverbal with respect to the verb *fo* 'was', indicating that the subject pronoun *el* 'he' is left-dislocated. In the rest of the passage, the DT remains Sir Guillem: In [5] he removes his fingernail, per Lady Guillelma's conditions, and in [6] he returns to her.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

7.3 *Illustration of new discourse referent introduction*

Like S–*si*–V dislocations, S–V–Ocl can introduce into the discourse a new referent and simultaneously mark a shift of topic to this new referent, as in (24):

(24) E una de sas compainnhieras ERA **li** de pres; e comenset pensar sobre alcun vet de sancta obediencia que·ll sancta maire avia fach.
'And one of her companions was there close to her, and she started to think about a certain prohibition that the holy mother had made, out of holy devotion.'

(*Douceline*; Gout 1927: 190)

In (24), *una de sas compainnhieras* 'one of her companions' is a new referent, although it is inferable from the preceding discourse; this referent remains the topic of the immediately following discourse.

7.4 *Illustration of continuing topic*

Conventional subjects, by contrast, generally extend the existing DT, either by encoding an ST, as in section [3] of (23), or by functioning as a continuing topic, as in (25), extracted from the *razo* of 80,34, about a poem by Bertran de Born.

(25) [1] En Bertrans si·n fo molt alegres. [2] E sabia que·il reis en era fort dolens et iratz d'aisso que ill dizion e del castel de Nontron e d'Azgen que ill avian tout, [3] e fez un son sirventes per far saillir lo rei Richart a la guerra. [4] E cant el ac fait son sirventes, [5] el **lo** MANDET a·N Raimon Jauzeran.
'[1] Sir Bertran was very happy about it. [2] And he knew that the king was afflicted and angry about what they were telling him and about the castle of Nontron and Agen that they had taken, [3] and he [Bertran] composed a poem to incite King Richard into battle. [4] And when he had written this poem, [5] he sent it to Sir Raimon Jauzeran.'

(*Razo* of 80,34; Boutière 1964: 136–137)

In the discourse preceding [1], Sir Bertran has learned of the political intrigue between King Richard and a number of barons in the southwest of France. Bertran, who is disliked by these barons, is happy to learn of this news in [1]. Bertan is the DT of the entire extract, although in [2], we learn – via Bertran's thoughts – of King Richard's losses to the barons. In [3], Bertran writes a poem (*sirventes*) to incite King Richard into war with the barons. In [4], the subject pronoun *el* 'he' of the *cant* 'when' clause refers to Bertran. Likewise, in [5], the conventional subject pronoun *el* 'he' of the SV declarative again refers to Bertran; here, the S–Ocl–V acts as a continuing topic. From a purely information-structural perspective, this subject appears redundant, as OOc often evinces null subjects when the subject has not changed from that of preceding utterances (see Hamlin et al. 1967: 33; Smith & Bergin 1984: 111; Jensen 1986: 86–88; Romieu & Bianchi 2002a: 128; 2002b: 35). That is, if the declarative in [5] contained a

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

null subject, the reader would presumably have no difficulty identifying Bertran as the referent of the subject. This example typifies how a large portion of the conventional subjects are used in the biographies and *Douceline* (Tables 5 and 6 above). Elsewhere, as discussed, conventional subjects mark STs, whose referent is almost always easily recoverable from previous discourse (unlike DTs, whose referents can be new to the discourse).

8. Discussion

The present analysis accounts for the variable position of object and adverbial clitics in OOc SV declaratives by positing that preverbal subjects can occur in (at least) two distinct syntactic positions, depending on discourse-pragmatic conditions. From a purely syntactic point of view, this approach builds on Skårup (1975) and draws on models of a richly articulated clausal left periphery (Rizzi 1997), in particular that of Benincà (2006). I follow Benincà (2006), who claims that the position of clitics depends on whether SpecFocus is overtly saturated. Applied to the present data, this means that preverbal clitics result from the fact that the subject occupies SpecFocus, whereas postverbal clitics arise when two conditions obtain: (i) the subject occupies a TopicP in the left periphery and (ii) SpecFocus remains unsaturated. Thus, despite surface similarities, the subjects in S–Ocl–V (1a) and S–V–Ocl (1b) occupy different positions.

The position of preverbal subjects is not random, nor does it need to be stipulated on a (circular) case-by-case basis. Rather, the choice of a conventional versus LD subject depends on the discourse-pragmatic role of the referent placed in subject position. By relating the syntax of an individual phrase to the larger discourse context, the present study reveals consistent information-structural differences between conventional and LD subjects. The data show that LD subjects most often act as shifting topics and usually establish a new DT, in contradistinction to conventional subjects, which maintain the existing DT, either by shifting to a momentary ST or by acting as a continuing topic.

As a pragmatically marked word order, subject LD was probably an expressive option rather than a strict grammatical requirement in OOc. First, the functions of LD and conventional subjects overlap to some degree in these data. For example, although infrequent, a few tokens of conventional subjects appear to encode changes in DT and introduce new referents into the discourse. Similarly, although STs typically appear as conventional subjects, LD subjects can also encode STs, and both LD and conventional subjects can serve as continuing topics. In addition, LD subjects constitute only one strategy among others to manage information structure; for example, changes in DT can also be accomplished via postverbal subjects, as can new referent introductions (Kunert 2003: 203). Furthermore, the syntactic marking of information structure sometimes depends on text genre and other variables (van Reenen & Schøsler 1992).

Inherent limitations arise when working with historical data. Some definitions of topic, for example, reference the 'mental states of speakers and hearers'

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

(Lambrecht 1994: 1), insofar as a speaker constructs utterances and stretches of discourse based on assumptions regarding what the hearer knows, does not know, and is likely to successfully infer (Reinhart 1981: 59). For historical data, such information is clearly inaccessible. For practical reasons, in the present study, DTs are distinguished from STs on the basis of duration, as measured in main declaratives. Such a criterion may not entirely capture the importance that the medieval author accorded to a given referent, and some instances of subject LD classified as STs in the present analysis may have, in the mind of the author, represented DTs worthy of heightened prominence.

Turning to syntactic structure, analyzing S–V–Ocl as covert subject LD entails that it is a variant of S–*si*–V and S–Sp–V. The fact that overt and covert subject LD accomplish the same discourse-pragmatic functions corroborates the claim. I propose that the same underlying clausal architecture, which places left-peripheral subjects in TopicP, can yield at least the three distinct surface realizations in (26a–c).[22]

(26) (a) S – *si* – (Ocl) V
[$_{\text{ForceP}}$ [$_{\text{FrameP}}$ [$_{\text{TopicP}}$ Subject [$_{\text{FocusP}}$ si [$_{\text{Focus}'}$ VERB]]]]]
(b) S – Sp – (Ocl) V
[$_{\text{ForceP}}$ [$_{\text{FrameP}}$ [$_{\text{TopicP}}$ Subject [$_{\text{FocusP}}$ Sp [$_{\text{Focus}'}$ VERB]]]]]
(c) S – V – Ocl
[$_{\text{ForceP}}$ [$_{\text{FrameP}}$ [$_{\text{TopicP}}$ Subject [$_{\text{FocusP}}$ ... [$_{\text{Focus}'}$ VERB]]]]]

The claim that (26a–c) represent surface variants of a single underlying architecture is supported by at least one instance of manuscript variation in the biographies.[23] The variants in (27) and (28) represent the same passage, and they form a near minimal pair, with an alternation between S–V–Ocl in (27) and S–*si*–Ocl–V in (28). In both cases (underlined), the NPs referring to the king are topicalized.

---

[22] As an anonymous *JL* referee observes, one should perhaps add another possibility to the list in (26), a (26d), where SpecFocus hosts a XP other than *si* or an Sp:

(i) S – XP – (Ocl) V
[$_{\text{ForceP}}$ [$_{\text{FrameP}}$ [$_{\text{TopicP}}$ Subject [$_{\text{FocusP}}$ XP [$_{\text{Focus}'}$ VERB]]]]]

Both example (7) and the following example probably possess this architecture:

(ii) E car [$_{\text{ForceP}}$ [$_{\text{FrameP}}$ [$_{\text{TopicP}}$ illi [$_{\text{FocusP}}$ aquestas doas vertutz
[$_{\text{Focus}'}$ DEZIRAVA]]]]]
'And for she desired these two virtues.' (*Douceline*; Gout 1927: 81)

Given that the present paper examines variation in the placement of object clitics in SV declaratives, I have not analyzed data such as these, where clitics are uniformly preverbal. In (i), however, the subject is compatible with a topic interpretation.

[23] *Douceline* exists in a single manuscript (Garay & Jeay 2001: 14), precluding an analysis of manuscript variation.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

(27) Don lo rey N'Anfos, que entendia en ela, s'aperseup que volia ela gran be ad Arnaut; e·l rey FO **ne** fort gilos e dolens.
'So King Alphonse, who desired her, learned that she had great affection for Arnaut, and the king was very jealous and saddened.'
(*Razo* of 30,19; Boutière 1964: 36–37)

(28) Lo rei nanfos, que sentendia en ella, si **se** PERCEUP quella ...
'King Alphonse, who desired her, learned that she ...'
(*Razo* of 30,19; Boutière 1964: 36–37)

Previous research has found S–V–Ocl to be infrequent compared to S–Ocl–V (Mériz 1978, Hinzelin 2007). However, I argue that this type of direct comparison is flawed, in that it compares two separate clausal architectures and assumes – tacitly at least – that the two architectures occur, or should occur, in equal proportions. Rather, the crucial observation is that the frequency of S–V–Ocl depends both on how frequently a text left-dislocates preverbal subjects and, more importantly, on how LD subjects are realized. In the biographies and *Douceline*, LD subjects are not rare. The biographies contain 316 tokens of S–*si*–V, 10 S–Sp–V, and 34 S–V–Ocl for a total of 360 SV declaratives with LD subjects (Table 7). For comparison, the biographies contain 267 tokens of S–Ocl–V (with a conventional subject in SpecFocus). Even when one factors in SV declaratives that do not contain clitics, LD subjects cannot be considered rare.[24] As a configuration of LD subject, however, S–V–Ocl accounts for only 9.4% of the data from the biographies and is thus a relatively infrequent realization of LD. In *Douceline*, however, a different picture emerges. This text contains 44 tokens of S–V–Ocl but only a single token of S–Sp–V. More surprising yet, S–*si*–V is entirely absent in *Douceline*, whereas this is the preferred realization of subject LD in the biographies.

	LD subject realization			
Text	S–*si*–V	S–Sp–V	S–V–Ocl	Total
Biographies	316 (87.8%)	10 (2.8%)	34 (9.4%)	360
Douceline	0	1 (2.2%)	44 (97.8%)	45

Table 7
Realization of left-dislocated subjects.

I hypothesize that the variable realization of subject LD is partly due to how SpecFocus is treated in the grammar of a given OOc text, and that this behavior

[24] Without data from object clitic position, the position of the subject in SV declaratives is ambiguous.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

of SpecFocus evinces intertextual variation, reflecting either the idiolects of individual authors or, as discussed subsequently in this section, regional syntactic variation. Pursuing this claim, the grammar of *Douceline* differs from that of the biographies by more freely allowing SpecFocus to remain empty, an observation that explains the preference for visible LD in the biographies and for covert LD in *Douceline*. Further support comes from the behavior of subordinate clauses that immediately precede a main declarative. Here, I focus on initial *cant*/*quant* 'when' clauses. In OOc, as in Old French, initial *cant*-clauses that immediately precede a main declarative almost always occupy a position in the left periphery rather than SpecFocus of the main declarative (Vance et al. 2010, Donaldson 2012).[25] The biographies contain 26 tokens of an initial *cant*-clause that immediately precedes a main declarative. In 25 of the 26 tokens, SpecFocus of the main declarative is overtly saturated (as in 14), as detailed in Table 8. In these examples, clitics are uniformly preverbal when present. In the single remaining example, shown in (29), SpecFocus is empty, and the clitic is predictably postverbal; this example mirrors the Old French example in (13) and the OOc example in (15).

	SpecFocus			
	Filled		Empty	
Text	*si*–(Ocl)–V	XP–(Ocl)–V	V–Ocl	Total
Biographies	7 (26.9%)	17 (65.4%)	1 (3.8%)	26*
Douceline	0	35 (66%)	18 (34%)	53*

* Unambiguous cases only: SpecFocus overtly filled, and/or object clitics present on matrix verb.

Table 8
Saturation of SpecFocus after initial *cant* 'when' subordinate clauses.

(29) E [$_{\text{FrameP}}$ qant s'en venia per lo Roine en una barca, [$_{\text{TopicP}}$ [$_{\text{FocusP}}$ ... [$_{\text{Focus}'}$ PREIREN **lo**]]]] li pescador d'En Aimar.
'And when he came via the Rhone in a boat, Sir Aimar's fishermen took him.'

(*Vida* LXXII; Boutière 1964: 485)

In *Douceline*, however, examples like (29) in which SpecFocus is empty are more frequent, accounting for 34% of the data (Table 8). As anticipated, clitics are postverbal in these examples. In the remaining examples, SpecFocus hosts the subject of the declarative or a non-subject XP, as in (14). Recall that

[25] Examples in which the initial *cant*/*quant*-clause unambiguously occupies SpecFocus of the following main declarative are rare; the biographies contain two such tokens, and *Douceline* contains an additional two tokens. For similar findings in Old French, see Donaldson (2012).

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

Douceline contains no examples of subject LD constructed with *si*. Crucially for my argument, none of the initial *cant*-clauses are followed by *si* in *Douceline*. Thus, whereas the grammar of the biographies frequently fills SpecFocus with *si* after an LD subject or an initial *cant*-clause, *Douceline* leaves SpecFocus empty in these same contexts, with the resulting effects on the placement of clitics. These same tendencies are almost certainly responsible for differences in stand-alone V1 declaratives as well (i.e., those not introduced by a coordinating conjunction). Whereas *Douceline* contains 44 tokens of a stand-alone V1 declarative (in which SpecFocus is empty), the entire set of biographies yields only four. Finally, evidence from Old French corroborates the view that the grammar of SpecFocus underlies the different realization of subject LD in the biographies and *Douceline*, as well as the behavior of *quant*/*cant*-clause + main declarative sequences and V1 declaratives. In early Old French, saturation of SpecFocus was not obligatory (Foulet 1928: 320; Skårup 1975: 291, 300; Rouveret 2004: 200–201; Simonenko & Hirschbühler 2012); V1 declaratives (where SpecFocus is empty) were attested, and so were sequences of *quant*-clause–V–Ocl like (13). Between about 1160 and 1200, however, SpecFocus came to be obligatorily saturated. As a result, V1 declaratives were lost (Simonenko & Hirschbühler 2012), as were sequences like (13), as shown in Donaldson (2012). Following this change, SpecFocus frequently hosted *si* and other semantically weak adverbs (Foulet 1928: 302–303; Herman 1954: 85; Marchello-Nizia 1985: 199–200).

I can only speculate at this juncture about what underlies the different treatment of SpecFocus in the biographies and *Douceline*. The close dating of the texts (late 13th for the biographies, very early 14th for *Douceline*) points to factors other than diachronic change. Regional variation likely plays a role. On this topic, an anonymous *JL* referee mentions data from early OOc legal charters from the Rouergue regions (Brunel 1926, 1952), discussed in Hinzelin (2007: 157–162; also Hinzelin 2009: 74). The referee asserts that the Rouergue data 'presumably show real regional variation' that the present analysis cannot account for. Although limitations of space preclude a full investigation of the Rouergue charters, a preliminary information-structural analysis leads me to disagree. The Rouergue data are of interest because Hinzelin (2007: 161) finds postverbal clitics to be more frequent than preverbal clitics in affirmative SV declaratives, unlike in charters from other regions, where preverbal clitics predominate. However, the information-structural properties of most of Hinzelin's (2007: 158–160) examples with postverbal clitics conform to the predictions of my proposal. In particular, in five of the eight tokens, the leftmost XP represents a shifting topic, either to a new discourse referent or to a currently inactive referent; this XP is then followed by matrix V+Ocl. Another two tokens involve matrix V+Ocl immediately following an initial subordinate, mirroring the structure of (14) and (29).[26] In all these

[26] The two examples cited by Hinzelin (2007: 159, examples (20) and (21)) contain initial *se* 'if' clauses rather than initial *cant* 'when' clauses. Donaldson (2012) shows that, in Old French, initial *quant* 'when' and *se* 'if' clauses occur almost exclusively in the left periphery.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

cases, if the initial XP occupies a left-peripheral position (here, a TopicP or a FrameP), as I argue, and SpecFocus is empty (i.e., no other XP intervenes between the left-peripheral XP and the matrix verb), then postverbal clitics are expected. Only a single token (example (22) in Hinzelin 2007: 159) runs counter to these predictions; I leave this question as well as a more complete information-structural analysis of the charter data to future research.

To return to the question of regional variation, I argue that the relevant syntactic property involves the ability or inability to leave SpecFocus unsaturated. Rates of preverbal versus postverbal clitics in matrix clauses will then be a secondary result of the interaction between the presence or absence of left-peripheral XPs (e.g., shifting topics or initial subordinates) and the concomitant saturation or non-saturation of SpecFocus. Thus, in texts whose grammar requires SpecFocus to be overtly saturated in matrix declaratives, clitics will be preverbal whether or not any XPs are present in the left periphery. On the other hand, in texts whose grammar allows SpecFocus to remain unsaturated in matrix declaratives, if the initial XP or XPs occupy left-peripheral positions, and SpecFocus remains unsaturated, clitics will be postverbal when present. In other words, in my view, the correct measure of regional variation lies not in raw percentages of preverbal versus postverbal clitics but rather in how the grammar of a given text treats SpecFocus and, as a result, whether that grammar preferentially realizes left-dislocated shifting topics as (26a, b) or (26c). If this reasoning is correct, the fact that *Douceline* and the Rouergue charters evince relatively robust rates of postverbal clitics in matrix declaratives follows from the fact that SpecFocus can remain empty, and that there are numerous instances of left-dislocated shifting topics (and/or initial subordinate clauses) in matrix declaratives with unsaturated SpecFocus.[27]

I leave open the question as to why the grammars of different texts vary in their treatment of SpecFocus. Henrichsen (1955) speculated that the biographies contained syntactic influences from Italian, perhaps due to an OOc-speaking author's prolonged contact with Italian (Henrichsen 1955: 117) or from later copyists (Boutière 1964: xi; Lafont 1967: 24). Old Italian texts employ *si* quite frequently (see e.g., Bocchi 2004, Poletto 2005, Ledgeway 2008), including in S–*si*–V, but further work is necessary to pinpoint an Italian influence in the present case, especially given the frequency of *si* in Old French, which also enjoyed close contact with OOc. On the other hand, Old Catalan and Old Spanish, culturally and geographically close to much of the OOc zone, evince S–V–Ocl (Ramsden 1963: 79, 163–164), although it remains to be seen if those languages pattern according to the discourse-pragmatic conditions described here. Text genre may be another factor, as van Reenen & Schøsler (1992) documented for Old French, and stylistic factors may also come into play.

[27] In these texts, the non-saturation of SpecFocus is an option rather than a requirement.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

9. Conclusion

This paper analyzed the position of object and adverbial clitics in OOc affirmative SV declaratives. I claim that postverbal clitics arise from the combination of an LD subject and empty SpecFocus, drawing on Benincà's (2006) work on the clausal left periphery in medieval Romance and the relationship between saturation of SpecFocus and the position of clitics. The present proposal assumes multiple preverbal positions capable of hosting a subject in OOc declaratives. I place LD subjects in a left-peripheral TopicP, although it should be noted that the present data do not permit as fine-grained an analysis as in Frascarelli & Hinterhölzl (2007: 97), who distinguish three discrete positions within TopicP (partially on the basis of prosody). Multiple surface realizations of LD subjects are available in OOc, with variation both within and across texts, but these different instantiations share the primary discourse-pragmatic function of shifting topic. The different discourse functions observed for conventional versus LD subjects justifies their analysis as syntactically and pragmatically distinct, and the position of clitics – far from being random – follows from the position of the subject and the saturation (or not) of SpecFocus. The analysis departs from Mériz (1978) and Hinzelin (2007), who treat preverbal subjects as a unified class occupying (in Hinzelin's analysis) a unique position (SpecFocus).

Further research is needed to determine to what extent comparable results obtain in other OOc texts and, eventually, in other medieval Romance varieties that used S–V–Ocl productively. Similarly, an analysis of left-peripheral objects may shed further light on the left periphery in OOc, and further examinations of the interactions between information structure and syntax are warranted for OOc more generally, following Combettes (1971) and Kunert (2003). Finally, the present analysis is purely synchronic, and a diachronic investigation of how syntactic information structure marking is affected by the eventual loss of V2 in Occitan (and along with it the loss of S–V–Ocl) remains a topic for future work.

References

Adams, Marianne. 1987. From Old French to the theory of pro-drop. *Natural Language & Linguistic Theory* 5, 1–33.

Anagnostopoulou, Elena. 1997. Clitic left dislocation and contrastive left dislocation. In Elena Anagnostopoulou, Henk van Riemsdijk & Frans Zwarts (eds.), *Materials on left dislocation*, 151–192. Amsterdam: John Benjamins.

Anglade, Joseph. 1921. *Grammaire de l'ancien provençal*. Paris: Klincksieck.

Ashby, William J. 1988. The syntax, pragmatics, and sociolinguistics of left- and right-dislocations in French. *Lingua* 75, 203–229.

Bec, Pierre. 1963. *La langue occitane*. Paris: Presses Universitaires de France.

Benincà, Paola. 2006. A detailed map of the left periphery of medieval Romance. In Rafaella Zanuttini, Héctor Campos, Elena Herburger & Paul H. Portner (eds.), *Crosslinguistic research in syntax and semantics: Negation, tense, and clausal architecture*, 53–86. Washington, D.C.: Georgetown University Press.

Bocchi, Andrea. 2004. 'Sì' nel 'Livero de l'abbecho'. In Michelangelo Zaccarello & Lorenzo Tomasin (eds.), *Storia della lingua et filologia*, 121–159. Florence: Sismel.

Bohnacker, Ute & Christina Rosén. 2008. The clause-initial position in L2 German declaratives: Transfer of information structure. *Studies in Second Language Acquisition* 30, 511–538.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

Boutière, Jean. 1964. *Biographies des troubadours: Textes provençaux des XIIIe et XIVe siècles*. Paris: A.-G. Nizet.
Brunel, Clovis. 1926. *Les plus anciennes chartes en langue provençale: Recueil des pièces originales antérieures au XIIIe siècle*. Paris: Auguste Picard.
Brunel, Clovis. 1952. *Les plus anciennes chartes en langue provençale: Recueil des pièces originales antérieures au XIIIe siècle – Supplément*. Paris: Éditions A. & J. Picard.
Cardinaletti, Anna. 2004. Towards a cartography of syntactic positions. In Luigi Rizzi (ed.), *The structure of CP and IP: The cartography of syntactic structures*, 115–165. Oxford: Oxford University Press.
Combettes, Bernard. 1971. Remarques sur la structure grammaticale et l'organisation thématique de la phrase dans les vies de troubadours. *Actes du VIe Congrès International de Langue et Littérature d'Oc et d'Études Franco-Provençales*, 293–308. Montpellier: Centre d'Estudis Occitans.
Donaldson, Bryan. 2012. Initial subordinate clauses in Old French: Syntactic variation and the clausal left periphery. *Lingua* 122, 1021–1046.
Donaldson, Bryan. To appear. Discourse functions of subject left dislocation in Old Occitan. *Journal of Historical Pragmatics*.
Ferraresi, Gisella & Maria Goldbach. 2002. V2 syntax and topicalization in Old French. *Linguistische Berichte* 189, 2–25.
Firbas, Jan. 1964. On defining the theme in functional sentence analysis. *Travaux Linguistiques de Prague* 1, 267–280.
Fleischman, Suzanne. 1990. *Tense and narrativity: From medieval performance to modern fiction*. Austin, TX: University of Texas Press.
Fleischman, Suzanne. 1991. Discourse pragmatics and the grammar of Old French: A functional reinterpretation of *si* and the personal pronouns. *Romance Philology* 44, 251–283.
Fontana, Josep M. 1993. *Phrase structure and the syntax of clitics in the history of Spanish*. Ph.D. dissertation, University of Pennsylvania.
Foulet, Lucien. 1928. *Petite syntaxe de l'ancien français*. Paris: Honoré Champion.
Frascarelli, Mara. 2007. Subjects, topics, and the interpretation of referential *pro*: An interface approach to the linking of (null) pronouns. *Natural Language & Linguistic Theory* 25, 691–734.
Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning, and form*, 87–116. Amsterdam: John Benjamins.
Garay, Kathleen & Madeleine Jeay. 2001. *The life of Saint Douceline: Beguine of Provence*. Rochester, NY: Boydell & Brewer.
Givón, Talmy. 1983. *Topic continuity in discourse: A quantitative cross language study*. Amsterdam: John Benjamins.
Gout, Raoul. 1927. *La vie de Sainte Douceline*. Paris: Ars et Fides.
Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and the composition of the left periphery: The cartography of syntactic structures*, vol. 8. Oxford: Oxford University Press.
Hamlin, Frank R., Peter T. Ricketts & John Hathaway. 1967. *Introduction à l'étude de l'ancien provençal*. Geneva: Librairie Droz.
Henrichsen, Arne-Johan. 1955. *Les phrases hypothétiques en ancien occitan: Etude syntaxique*. Bergen: John Griegs.
Herman, Josef. 1954. Recherches sur l'ordre des mots dans les plus anciens textes français en prose. *Acta Linguistica Academiae Scientiarum Hungaricae* 4, 69–93, 351–379.
Hinzelin, Marc-Olivier. 2007. *Die Stellung der klitischen Objektpronomina in den romanischen Sprachen*. Tübingen: Gunter Narr.
Hinzelin, Marc-Olivier. 2009. La position des pronoms clitiques en ancien occitan. In Guy Latry (ed.), *La voix occitane. Actes du VIIIe Congrès de l'Association Internationale d'Études Occitanes*, 67–81. Bordeaux: Presses Universitaires de Bordeaux.
Holmberg, Anders. 2015. Verb second. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 1, 342–383. Berlin: De Gruyter.
Jensen, Frede. 1972. *From Vulgar Latin to Old Provençal*. Chapel Hill, NC: University of North Carolina Press.
Jensen, Frede. 1986. *The syntax of medieval Occitan*. Tübingen: Max Niemeyer Verlag.
Jensen, Frede. 1994. *Syntaxe de l'ancien occitan*. Tübingen: Max Niemeyer Verlag.
Kaiser, Georg. A. 2002. *Verbstellung und Verbstellungswandel in den romansichen Sprachen*. Tübingen: Max Niemeyer Verlag.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

Kunert, Hans P. 2003. L'ordine degli elementi della frase in Occitano antico. *Romanische Forschungen* 115, 194–209.
Labelle, Marie. 2007. Clausal architecture in Early Old French. *Lingua* 117, 289–316.
Lafont, Robert. 1967. *La phrase occitane: Essai d'analyse systématique*. Montpellier: Presses Universitaires de France.
Lambrecht, Knud. 1981. *Topic, antitopic, and verb agreement in non-standard French*. Amsterdam: John Benjamins.
Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Amsterdam: John Benjamins.
Lambrecht, Knud. 2001. Dislocation. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals*, 1050–1078. Berlin: de Gruyter.
Ledgeway, Adam. 2008. Satisfying V2 in early Romance: Merge vs. Move. *Journal of Linguistics* 44, 437–470.
Ledgeway, Adam. 2010. Subject licensing in CP: The Neapolitan double-subject construction. In Paola Benincà & Nicola Munaro (eds.), 2010. *Mapping the left periphery: The cartography of syntactic structures*, vol. 5, 257–296. Oxford: Oxford University Press.
Lemieux, Monique & Fernande Dupuis. 1995. The locus of verb movement in non-asymmetric verb-second languages: The case of Middle French. In Adrian Battye & Ian G. Roberts (eds.), *Clause structure and language change*, 80–109. Oxford: Oxford University Press.
Marchello-Nizia, Christiane. 1985. *Dire le vrai: L'adverbe "si" en français médiéval – Essai de linguistique historique*. Geneva: Librairie Droz.
Marchello-Nizia, Christiane. 1998. Dislocations en ancien français: Thématisation ou rhématisation? *Cahiers de praxématique* 30, 161–178.
Mathieu, Eric. 2013. The left-periphery in Old French. In Deborah Arteaga (ed.), *Research on Old French: The state of the art*, 327–350. Berlin: Springer.
Mériz, Diana T. 1978. Observations on object pronoun collocation with finite verb-parts in medieval Occitan (to 1300). *Romania* 99, 145–182, 289–310.
Mussafia, Adolf. 1886. Una particolaritá sintattica della lingua italiana dei primi secoli. *Miscellanea di folologica e linguistica: Dedicata alla memoria di Napoleone Caix e Ugo Angelo Canello*, 255–261. Firenze: Le Monnier.
Paden, William D. 1998. *An introduction to Old Occitan*. New York: Modern Language Association of America.
Poletto, Cecilia. 2005. *Sì* and *e* as CP expletives in Old Italian. In Montserrat Batllori, Maria-Lluïsa Hernanz, Carme Picallo & Francesc Roca (eds.), *Grammaticalization and parametric variation*, 206–235. Oxford: Oxford University Press.
Poole, Geoffrey. 2006. Interpolation and the left periphery in Old Spanish. *Newcastle Working Papers in Linguistics* 12–13, 188–216.
Posner, Rebecca. 1996. *The Romance languages*. Cambridge: Cambridge University Press.
Prince, Ellen F. 1997. On the functions of left-dislocation in English discourse. In Akio Kamio (ed.), *Directions in functional linguistics*, 117–144. Amsterdam: John Benjamins.
Ramsden, Herbert. 1963. *Weak-pronoun position in the early Romance languages*. Manchester: Manchester University Press.
Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27, 53–94.
Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer.
Roberts, Ian G. 1993. *Verbs and diachronic syntax: A comparative history of English and French*. Dordrecht: Kluwer.
Roberts, Ian G. 2012. Phases, head movement and second-position effects. In Ángel J. Gallego (ed.), *Phases: Developing the framework*, 385–440. Berlin: de Gruyter Mouton.
Rohlfs, Gerhard. 1970. *From Vulgar Latin to Old French*. Detroit, MI: Wayne State University Press.
Romieu, Maurice & André Bianchi. 2002a. *Iniciacion a l'occitan ancian: Dètz e nòu tèxtes de l'Edat Mejana comentats*. Bordeaux: Presses Universitaires de Bordeaux.
Romieu, Maurice & André Bianchi. 2002b. *La lenga del trobar: Précis de grammatica d'occitan ancian*. Bordeaux: Presses Universitaires de Bordeaux.
Rouveret, Alain. 2004. Les clitiques pronominaux et la périphérie gauche en ancien français. *Bulletin de la Société de linguistique de Paris* 99, 181–237.

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press

Salvi, Giampaolo. 2000. La formazione del sistema V2 delle lingue romanze antiche. *Lingua e stile* 35, 665–692.
Schutz, Alexander H. 1938. Where were the Provençal *vidas* and *razos* written? *Modern Philology* 35, 225–232.
Simonenko, Alexandra & Paul Hirschbühler. 2012. Placement de clitiques dans les propositions V1 et évolution de la structure de la proposition en ancien français. In Monique Dufresne (ed.), *Typologie, ordre des mots et groupe verbal en français médiéval*, 11–53. Laval, Quebec: Presses de L'Université Laval.
Sitaridou, Ioanna. 2005. A corpus-based study of null subjects in Old French and Old Occitan. In Claus D. Pusch, Johannes Kabatek & Wolfgang Raible (eds.), *Romance corpus linguistics II: Corpora and diachronic linguistics*, 359–374. Tübingen: Gunter Narr.
Skårup, Povl. 1975. *Les premières zones de la proposition en ancien français Revue Romane* [special issue] 6.
Skårup, Povl. 1997. *Morphologie élémentaire de l'ancien occitan*. Copenhagen: Museum Tusulanums Forlag.
Smith, Nathaniel B. & Thomas G. Bergin. 1984. *An Old Provençal primer*. New York: Garland.
Steiner, B. Devan. 2014. *The evolution of information structure and verb second in the history of French*. Ph.D. dissertation, Indiana University.
Tobler, Adolf. 1875. Review of Jean J. Le Coultre, *De l'ordre des mots dans Chrétien de Troyes*, 1875. *Göttingische gelehrte Anzeigen* 34, 1057–1082.
van Dijk, Teun A. 1976. Sentence topic and discourse topic. *Papers in Slavic Philology* 1, 49–61.
van Reenen, Pieter & Lene Schøsler. 1992. Ancien et moyen français: *si* 'thématique'. Analyse exhaustive d'une série de textes. *Vox Romanica* 51, 101–127.
Vance, Barbara S. 1997. *Syntactic change in medieval French: Verb-second and null subjects*. Dordrecht: Kluwer.
Vance, Barbara S. 2011. Variation and change in medieval Romance word order: The case of object clitics in Occitan. Presented at Language, Variation, and Change Workshop, The Univeristy of Chicago.
Vance, Barbara S., Bryan Donaldson & B. Devan Steiner. 2010. V2 loss in Old French and Old Occitan: The role of fronted clauses. In Sonia Colina, Antxon Olarrea & Ana Maria Carvalho (eds.), *Romance linguistics 2009: Selected papers from the 39th Linguistic Symposium on Romance Languages*, 301–320. Amsterdam: John Benjamins.
Vanelli, Laura, Lorenzo Renzi & Paola Benincà. 1985. Typologie des pronoms sujets dans les langues romanes. *Actes du XVIIe Congrès International de Linguistique et Philologie Romanes 3*, 163–176. Aix: Université de Provence.
Wanner, Dieter. 2010. Review of M.-O. Hinzelin, *Die Stellung der klitischen Objektpronomina in den romanischen Sprachen*, 2007. *Zeischrift für Katalanistik* 23, 270–280.

Author's address: *Department of Languages and Applied Linguistics,*
University of California, 1156 High Street,
Santa Cruz, CA 95064, USA
bryandonaldson@ucsc.edu

https://doi.org/10.1017/S0022226714000619 Published online by Cambridge University Press