pro-drop in interrogatives and declaratives.
A parallel study of Old High German and Old Italian

Federica Cognola (Venice) and George Walkden (Konstanz)

Abstract
While there has been a substantial body of research on the asymmetry between main and subordinate clauses in terms of the licensing of pro-drop, potential differences between types of unembedded clause have received much less attention – despite the fact that competing theories of pro-drop make strong, clear predictions about the distribution of null subjects across clause types, especially with regard to interrogatives. This paper presents the first in-depth comparative study of pro-drop in both declaratives and interrogatives in two asymmetric pro-drop languages: Old High German and Old Italian. Based on a parallel corpus study using two translations of Tatian’s Diatessaron, we show that there is a clear difference in distribution between interrogatives and declaratives: null subjects are more frequent in declarative clauses than in interrogatives, and these also differ in terms of the persons in which pro-drop is licensed. Our results speak against the V-in-C licensing theory of asymmetric pro-drop of Benincà (1984) and Adams (1987), and in favour of an account based on an Agree relation with left-peripheral operators in the sense of Frascarelli (2007, 2018).

1 Introduction

Not all languages that allow referential null arguments do so under the same conditions. In this paper we examine one of the factors that null subject licensing appears to be sensitive to: clause type. We label as asymmetric pro-drop languages those languages in which the occurrence of null subjects differs according to clause type. A common pattern here is that the occurrence of null subjects is more restricted in subordinate clauses than in main clauses. An influential line of reasoning originating with Benincà (1984) and Adams (1987) has maintained that the licensor of null subjects in such languages is V-to-C movement and that this explains the clause-type asymmetry.

In this paper we take issue with this kind of account, and propose an alternative in which null subjects must enter into a matching relation with a left-peripheral operator (cf. Frascarelli 2007,

---

1 Parts of this work were presented at the Universities of Konstanz and Göttingen in 2015, at the University of Oslo in 2016 and at the DGF-S-Tagung in Bremen in 2019. We thank those audiences, in particular Marco Coniglio, Cecilia Poletto, Christine Meklenborg Salvesen, Michael Zimmermann, and two anonymous reviewers for useful comments and feedback. All errors are our own. For the concerns of the Italian academy, Federica Cognola takes responsibility for sections 4, 5, 6 and George Walkden for sections 1, 2, 3.

Linguistik online 100, 7/19 – http://dx.doi.org/10.13092/lo.100.6020
CC by 3.0
2018). In support of our analysis we present a new empirical investigation of two asymmetric pro-drop languages, Old High German (OHG) and Old Italian (OI), based on parallel translations of the same text, Tatian’s Diatessaron. Crucial evidence in favour of our analysis comes from interrogative clauses, which have so far been largely ignored in the literature, but which clearly show that an account in terms of V-to-C movement cannot be correct.

Our paper is structured as follows. Section 2 outlines the theoretical landscape as regards the licensing of null subjects in general and asymmetric pro-drop in particular. In section 3 we sketch the state of the art in research on OHG and OI themselves. Section 4 presents our methodology and an overview of our quantitative findings. In section 5 we look more closely at the syntactic contexts in which null subjects are found, both quantitatively and qualitatively. Section 6 concludes.

2 Theoretical approaches to null subject licensing

2.1 Types of null subject languages

The null-subject phenomenon has long been recognized in traditional descriptive grammars, but was first described in generative grammar by Perlmutter (1971). As Roberts/Holmberg (2010: 4) put it, “some languages require finite clauses to overtly express a definite, referential, pronominal subject, while others do not.” Since Rizzi (1982, 1986), the leading idea in generative linguistics has been that this cross-linguistic difference among languages should be captured in terms of different settings of a parameter, the pro-drop parameter. For historical reasons, the prototypical example of a null subject language has been Italian, as exemplified by (1), with English being the prototypical non-null subject language. In Italian, definite, referential, pronominal subjects can be null in all persons, in all tenses and in all syntactic configurations; overt pronouns are only obligatory when emphatic/focalized.

(1)  (Lei, la mamma) parla inglese
she, the mum speak.3SG English
‘She/the mother speaks English.’

Rizzi’s (1986) theory of pro-drop, which takes Italian as its starting point, specifies two requirements for pro: (i) it must be licensed (via government, by a case-assigning head) and (ii) the null element’s feature specification must be recoverable for interpretation.

Research into null subjects in the twenty-first century, however, has by and large moved away from the idea that null subjects are regulated by a single parameter, instead focusing on the different mechanisms and configurations that license them: see Roberts/Holmberg (2010), Barbosa (2011a,b) and D’Alessandro (2015) for overviews.2 The “canonical” or “well-behaved” null subject languages like Italian are recognized as just one type among several, without enjoying any privileged status. Roberts/Holmberg (2010), for instance, propose two further types: “radical” or “discourse” null subject languages such as Chinese and Japanese, in which verb morphology plays no role in licensing null subjects and arguments other than subjects may also

2 In fact, it is suggested as early as Rizzi (1986: 547) that both licensing and recovery are separately parameterized, and hence that the “null subject parameter” is a cluster of at least two independent properties even as regards referential null subject languages
be dropped, and “partial” null subject languages such as Finnish and Hebrew, in which only certain persons and tenses permit null subjects. In this paper we focus on referential null subjects, abstracting away from expletive, arbitrary and generic subjects.

Descriptively, in research on null subject languages, clause type, morphological properties of the verb, person, and discourse context have all been found to exercise an influence on whether a subject may remain unexpressed in a given context or not. Theoretically, the move to Minimalism has eliminated relations such as government that were previously important for analyses of null subject licensing, and rethought locality conditions on syntactic operations, making it necessary to account for the observed phenomena in terms of Agree, Move (internal Merge), Relativized Minimality and phase theory, and encouraging linguists to explore the predictions of these notions in this domain. Two further theoretical developments are worthy of note. First, the dominant conception of inflectional morphology is now that it is post-syntactic (Anderson 1992; Halle/Marantz 1993). From this perspective, there is little theory-internal motivation for pursuing an account of null subjects that attributes a causal role to the morphological “richness” of verbs in synchronic grammar, as proposed by Taraldsen (1978), Rizzi (1982, 1986) and Jaeggli/Safir (1989). At the same time, attempts to make the intuition precise by formulating an explicit and predictive theory of morphological richness (Rohrbacher 1999, Müller 2005, Tamburelli 2006) have not succeeded in capturing the diversity that is observed cross-linguistically (though see Rosenkvist 2018). This has led some authors to suggest that the connection between null subjects and rich verbal morphology is a matter of processing (e. g. Holmberg 2005, Ackema/Neeleman 2007) or an artefact of historical change (e. g. Fuß 2011), with no place in the theory of Universal Grammar.

Another development that has influenced theorizing on null subjects during the same period is what Haegeman/Hill (2013) label the “syntactization of discourse”: the incorporation of information-structural features, positions and entities into narrow syntax (Rizzi 1997, 2001, É. Kiss 1998, Benincà/Poletto 2004, Frascarelli/Hinterhölzl 2007, Cruschina 2009, Aboh 2010). Since the earliest research it has been observed that discourse conditions have a crucial role to play in the licensing of null subjects in at least some languages (e. g. Huang 1984 on Chinese). Against this backdrop, the approach to Italian null subjects in Frascarelli (2007) is of particular importance, and we present it in detail in section 2.3, as it will form the basis of our own account. First, however, we turn to the phenomenon of asymmetric pro-drop and the analysis that has been put forward to account for it.

2.2 Asymmetric pro-drop: the V-to-C analysis

The existence of “asymmetric” null subject languages, in the sense that null subjects in these languages seemed to be more robustly permitted in main than in embedded clauses, was brought to light in generative research by Benincà (1984) and Vanelli et al. (1986), who noted that a subset of the medieval Romance languages seemed to exhibit such behaviour. What is particularly striking about such languages is that canonical null subject languages such as Italian come very close to showing the opposite distribution: the standard view (e. g. Filiaci et al. 2013) is that overt pronouns are strongly disfavoured in embedded clauses when coreferential with the matrix subject, as shown in (2) (though see Frascarelli 2018: 225).
The classic analysis of asymmetric null subject languages was provided by Adams (1987) in a detailed discussion of Old French. Adams (1987) draws on previous literature showing that Old French permits null subjects much more freely in main clauses than in subordinate clauses. Her analysis is based on the claim that Old French is a V2 language, which Adams analyses as involving V-to-C movement via INFL (I or T in modern terms). Following den Besten’s (1983) intuition that V-to-C is blocked when a complementizer occupies C, Adams argues that the asymmetry of pro-drop in Old French follows from the fact that pro must be governed by INFL, a variant of Rizzi’s (1986) licensing requirement. She furthermore argues that the direction of government in Old French is consistently to the right. Since pro is in SpecIP, INFL can only govern it if it is moved to C, which it can only do when C is not occupied by a complementizer. The clause-type asymmetry is thus derived.

Adams observes (1987: 9–10, footnote 11) that some apparently embedded clauses in Old French feature both V2 and pro. These clauses, she claims, are actually main clauses to all intents and purposes, in a paratactic relationship with the apparent embedding clause. The complementizer is thus above C in such clauses, and so both V-to-C and pro are possible.3

Adams’s analysis is stated in terms of government, a relation which is deprecated in current Minimalist theory. Nevertheless, it is easy to reconstruct the core of Adams’s analysis in terms of Agree. Assume that T in null subject languages is fully specified for phi-features, [iφ] (cf. Barbosa 1995, Alexiadou/Anagnostopoulou 1998), and that pro is some kind of null nominal bearing uninterpretable phi-features, [uφ]. Assume furthermore that interpretable features must c-command uninterpretable features in order for the latter to be valued/checked via Agree (Wurmbrand 2012, Zeijlstra 2012), and that this relation must obtain clause-internally for reasons of locality. The only way for this configuration to obtain is for T to move to C, as illustrated in (3). In all other cases, pro is not licensed, as its uninterpretable features cannot be valued/checked. This account maintains the predictions of Adams’s (1987) analysis in a Minimalist framework, as far as we can tell, using only fairly standard ingredients.4 The question is whether it is correct.

(3)  [CP C+T[iφ] [TP pro[uφ] [Γ ... ]]]

Crucially, a prediction of this account is that imperatives and V1 questions in Old French should also robustly permit pro-drop, since these also have V-to-C movement (see Rizzi 1991 on residual V2). Adams (1987: 15–16) presents examples suggesting that this is borne out. Adams’s

3 Adams suggests that all such examples involve “bridge” verbs and the complementizer que (or a null version of it). For V2 in Old French this seems to be broadly correct, and the facts fall out neatly from a split-CP model permitting both high (Force) and low (Fin) complementizers; see most recently Salvesen/Walkden (2017).

4 With a few additional assumptions/stipulations: for instance, one must assume that pro is base-generated in SpecTP, since otherwise its features could be checked/valued in its base position. One also needs something to rule out the valuation of pro’s phi-features by (for instance) a fronted object in SpecCP. Furthermore, a syntax-internal operation of head movement is questionable under Minimalist assumptions (see e. g. Matushansky 2006, Roberts 2010, 2011). Since we will not ultimately be pursuing the V-to-C-licensing analysis, however, we will simply assume that these issues are in principle solvable.
characterization of the Old French facts has been revisited and disputed in various works (Roberts 1993; Vance 1997; Salvesen 2014; Zimmermann 2014, 2018; Simonenko et al. 2018), a debate which lies outside the focus of the present paper. Instead we will focus on the analysis and its cross-linguistic applicability. We claim that for both OI and OHG the V-to-C-licensing analysis is inadequate.

2.3 An alternative: the Topic-matching analysis

Frascarelli (2007, 2018) presents an analysis of present-day Italian in which discourse context plays a crucial role. She works with the standard Minimalist tools of Merge and Agree, along with the cartography of the clausal left periphery given in (4) (from Frascarelli/Hinterhölzl 2007; cf. Frascarelli 2018: 213).

\[(\text{ShiftP[+aboutness]} [\text{ContrP [FocP [FamP* [IP ... ]]]}]\] \]

The innovation in (4) is the distinction between three different types of topic. Familiar topics, which occur below FocP, are simply elements referring to discourse-given entities; like Rizzi’s (1997) TopP, this projection may occur more than once. In Italian, familiar topics are phonologically marked with a low (L*) tone. Contrastive topics are elements that induce alternatives with no impact on the focus value and set up oppositions with respect to other topics (Büring 1999); there is only one contrastive topic projection per clause, and it is above FocP. In Italian, contrastive topics bear a H* tone. Finally, and most importantly for the purposes of this paper, aboutness topics are “what the sentence is about” (Reinhart 1981, Lambrecht 1994). Shifting aboutness topics (newly introduced or reintroduced; Givón 1983) are the structurally highest of the three types of topic projection, and there is only one such projection per clause; shifting topics in Italian are marked by an L*+H contour. Continuing aboutness topics are realized in the same way and in the same position as familiar topics. All three types of topic are assumed to be first Merged in the left periphery in Italian rather than moved there (hence corresponding to the CLLD topics of Cinque 1990).

Frascarelli/Hinterhölzl (2007) make the case that the phonological, information-structural and syntactic properties of these types of topic are closely tied together in the way that would be expected given cartographic assumptions about structure and interpretation. Against this background, Frascarelli (2007) puts forward and assesses the conjecture in (5).

\[(5) \quad \text{A thematic N[ull ]S[ubject] is a pronominal variable, the features of which are valued (i.e., “copied through matching”)} \text{by the local Aboutness-shift Topic.}\]

Frascarelli’s corpus study of Italian shows that null subjects are consistently interpreted in relation to the closest aboutness topic. This aboutness topic must be local, and can be overt or silent; when silent, it is possible for it to also be resumed by a familiar topic lower in the left periphery. Crucial evidence for the proposal is that, if an overt aboutness topic and a null subject are both present, the two must necessarily corefer.

Formally, this idea is cashed out as follows: the Shift head bears a feature which acts as a probe, and hence enters into an Agree relation with a pro at the edge of the vP phase, as in the
configuration in (6) (Frascarelli 2007: 718, her (30)). This operation collapses Rizzi’s (1986) notions of formal licensing and identification into a single relation.\(^{6}\)

\[(\text{ShiftP} \text{DP}_{\text{αPn}} [ \text{Shift}^\circ [ \ldots [\text{Agr}\text{SP} [ \text{Agr}^\circ [\text{VP} \text{pro}] \text{VP} ] \ldots ]]])\]

Frascarelli further assumes that ShiftP is a criterial position, at least in predicational sentences, and that a topic (possibly silent) must be present in the specifier of ShiftP. This “Topic Criterion” is given in full in (7).

\[(\text{Agr}^\circ \text{AgrSP} [ \text{VP} \text{pro}] [\text{αPn}] [\text{Shift}^\circ [ \ldots ]]])\]

The head of a new topic chain can only be established in a clause that is capable of bearing illocutionary force. First and second person null subjects work differently: Frascarelli (2018: 219–222) argues that these do not interfere in topic chains and are not licensed by the same mechanism as third person null subjects. Instead, first and second person null subjects enter into an Agree relation with a logophoric agent (Λ\text{A}) or logophoric patient (Λ\text{P}), syntactically present in the left periphery. Like the [+aboutness] Topic feature, the features of Λ\text{A} and Λ\text{P} are C/edge linkers (CLn) in the sense of Sigurðsson (2011).

\[\text{C/Edge-Linking Generalization} \ (\text{Sigurðsson 2011: 282})\]

Any definite argument, overt or silent, positively matches at least one CLn in its local C-domain, where CLn is an element of the set \{Λ\text{A}, Λ\text{P}, Top \ldots \}.

In sum, Frascarelli’s theory directly incorporates discourse considerations, whereas morphological effects are not taken to be a part of the synchronic grammatical analysis. In this respect, things have come full circle from the earliest generative work on null subjects in the 1970s and 1980s, in which precisely the opposite was true. The analysis that we develop will be based on Frascarelli’s insights, though morphological facts will also play a role in accounting for the special behaviour of particular endings.

3 Existing research on Old High German and Old Italian

In support of our theory of asymmetric \textit{pro}-drop we draw on OHG and OI, two languages attested in historical texts. In this section we briefly present the state of the art in research on these two languages.

3.1 Old High German

Modern German is generally known as a non-\textit{pro}-drop language: expletive, but not referential, subjects can be null (Rizzi 1982; Cardinaletti 1990; Roberts/Holmberg 2010). Modern German

\(^{5}\) The \[\text{αPn}\] bundles are person features in the sense of Sigurðsson (2004, 2011). As the left-peripheral head is the one bearing the interpretable feature, this requires that Agree take place in a configuration in which the valuer \text{c}-commands the valuee, as in Wurmbrand (2012) and Zeijlstra (2012), though Frascarelli does not discuss this.

\(^{6}\) As observed by a reviewer, the essence of Frascarelli’s (2007) proposal thus has antecedents in much earlier work: as well as reflecting Rizzi’s (1986) identification requirement, it is also reminiscent of the proposal in Calabrese (1986) that null subjects must be coreferential to the current topic, albeit cashed out syntactically using modern machinery.
also exhibits a process of “topic drop” or “pronoun zap”: see Ross (1982), Huang (1984), Haider (2010), and Trutkowski (2011, 2016). This process is exemplified in (9).

(9) (In answer to the question: ‘What about Hans?’)
    Hab’ ich heute getroffen
    ‘I met him today.’

Characteristics of ‘topic drop’ are that i) it affects elements in the preverbal position (SpecCP) only, ii) it may occur only in root clauses, and iii) it may affect elements other than subjects (as in (9) above). In general, topic drop provides the only possibility for referential subjects to be null.\footnote{Some German dialects behave differently: see Weiß (2005) and Weiß/Volodina (2018) for an overview, and Bohnacker (2013) for a corpus-based study of spoken Swabian.}


(10) a. Latin: et ex illis ... crucifigetis
    Some hahat pro in cruci
    ‘some of them you will crucify’ (MF XVIII, 17; Mt 23:24)

b. Latin: In persona enim domini patrem accipimus
    In thehmu druhtines nemni archennemes ... pro fater
    ‘in the name of the Lord we recognise … the Father’ (I 279)

c. Latin: Et ascendens in nauicula
    pro steig tho in skifilin
    ‘he then stepped into the boat’ (T 193,1)

Examples like (10a) and (10b), in which SpecCP is filled and the subject is nevertheless null, suggest that OHG exhibited a different pro-licensing mechanism, since in these examples the null subject cannot be a case of topic drop. There is, however, a clear clause-type asymmetry: according to the figures in Eggenberger (1961), between 40% and 64% of pronominal subjects are null in main clauses, compared to between 7% and 16% in embedded clauses. A further asymmetry is found in the person of the pronominal subject: third person subjects are dropped in between 54% and 81% of examples, while first and second person subjects are generally dropped in no more than 40% of examples.
In traditional studies (Eggenberger 1961; Hopper 1975), the possibility of null subjects is ascribed to an effect of loan syntax: slavish translation from Latin is responsible for their occurrence. Axel (2007: 306) takes issue with this, since the clause type and person asymmetries found in the distribution of OHG null subjects remain entirely mysterious under this account. Moreover, she points out that the Hildebrandslied, an entirely autochthonous text, features five instances of null subjects as opposed to 29 overt pronouns. The other early West Germanic languages, Old English and Old Saxon, also display similar distributions of null subjects in both autochthonous and translated texts (Rosenkvist 2009; van Gelderen 2013; Rusten 2013, 2015, 2019; Walkden 2013, 2014: chapter 5). Though translation is likely to have had an effect on the frequency of null subjects, then, it cannot be the whole story.\(^8\)

Axel (2005, 2007), Axel/Weiβ (2011), Volodina/Weiβ (2016) and Weiβ/Volodina (2018) adopt versions of the V-to-C-licensing analysis of Adams (1987), discussed above in section 2.2. According to these authors, null subjects are a genuine syntactic phenomenon, and the features of the finite verb in C are what licenses pro, predicting an asymmetry between clause types. Schlachter (2010, 2012), in a study of the texts of the Isidor group, takes issue with this conclusion. She points out that there are a number of examples of pro-drop in embedded clauses without V-to-C movement, such as (11).

(11) Latin: *sic in consequentibus dicit:*  
so sama so pro hear after quhidhit:  
so same so here after says

‘As he says after this …’  
(Is. IX.11, Eg.703–704, He 43,17; Schlachter 2010: 162)

Axel is aware of such examples, and appeals to Latin translation influence to account for them (2007: 311), which Schlachter (2010, 2012) argues is untenable. Pursuing this further, Walkden (2014: 186–187) looks at the eight examples that Eggenberger (1961) gives of null subjects in subordinate clauses in *Isidor*, including (12). In only four of these is an analysis involving V-to-C movement tenable; (12) can clearly not be analysed in such a way.

(12) Latin: *nisi ex duobus nascatur*  
nibu pro fona zuuem chiboran uuerdhe

NEG-if from two born become-3SG.SBJV

‘if he is not born of two people’  
(*Isidor* 3.15; Walkden 2014: 187)

Moreover, Walkden (2012, 2014) points out that the V-to-C-licensing account does not predict the asymmetry between persons, though this asymmetry appears to be purely quantitative rather than categorical. Both Schlachter and Walkden suggest a discourse-driven account of the type suggested for modern Italian in Frascarelli (2007) and discussed in section 2.3; the relevance of this to the clause type asymmetry will be further elaborated on in section 5.

### 3.2 Old Italian

Like Modern Italian, OI is considered to be a pro-drop language; see Benincà (1984), Salvi/Renzi (2010), and many other works. The main difference, however, is that OI is an

---

\(^8\) See in particular Walkden (2016) for the case against pure loan syntax in Old English.
asymmetric pro-drop language, i. e. null subjects appear to be possible in main but not in embedded clauses. Examples (13) and (14) illustrate.

(13) Quand tu veniss al mondo, se tu voliss
When you came to the world, if you wanted
pensar, negota ge portassi _, negota n
to-think-about-it, nothing there brought.2SG nothing from-there
poi _ portar
can.2SG take
‘When you came into the world, if you think about it, you didn’t bring anything, and nothing you can take away.’

(Old Milanese, Benincà 2006: 68; Bonvesin, 179)

(14) E così ne provò _ de’ più cari ch’elli avea.
and so of-it tested.3SG of-the most dear that-he had
‘So he tested some of the best friends he had.’

(Old Florentine, Benincà 2006: 68; Testi fiorentini, 74)

Importantly, overt subject pronouns in embedded clauses very frequently do not have an emphatic value, as the examples in (15) show (examples are taken from Benincà 2010: 43; glosses and translations are our own). In this respect, OI is different from Modern Italian (on which see above, section 2.2).

(15) a. E certo quando tu2 il voli fare
and obviously when you Him want-2SG make
docile conviene che tu2 insieme lo facci
docile is-advisable that you together him make-2SG
attento
alert
‘If you want to make him docile, you should also make him alert’

(Brunetto Latini, Rettorica, p. 192, rr. 4–5)

b. se·lla naturaj domanda ciò ch’ella2 ha perduto
if.the nature asks what that-she has lost
‘If Nature asks back what she has lost’

(Novellino, 4, rr. 31-32)

c. La formica2 è più savia di te e
the ant is more wise than you and
ongn’ altro animale, inperò k’ella2 raguna la
any other animal since that-she gathers the
state dond’ella2 vive di verno
summer where-she lives in winter
‘The ant is wiser than you or any other animal, since she provides the supplies in summer, on which she lives in winter’

(Disciplina clericalis, p. 74, rr. 5–7)

The standard account of this behaviour has been the V-to-C-licensing analysis: pro-drop is a syntactic phenomenon, dependent on the finite verb occupying C, and asymmetric pro-drop is parasitic on asymmetric V2. In this paper we stick to most of analyses of Old Italian and Old Romance, according to which Old Romance varieties were relaxed V2 languages (Adams 1987;
Benincà 1984, 1995, 2006; Benincà/Poletto 2004; Fontana 1993; Ledgeway 2005, 2007, 2008, 2012; Poletto 2002, 2014; Roberts 1996, 2004; Salvesen 2013, Vance 1987, Wolfe 2015, 2018 among others, and Kaiser 2002; Kaiser/Zimmermann 2011; Zimmermann 2015 for the view that Old Romance was not V2). According to these analyses, which we think are correct, Old Romance languages instantiated another subtype of V2 rule (beside the strict Germanic V2) characterized by obligatory V-to-C in all main clauses (see Benincà 2006; Holmberg 2015). Therefore, in Old Romance V2 was not to be understood as a constraint on linearisation, but as an abstract property involving the movement of the finite verb to a C head in all main clauses. Note, that the V2 subtype instantiated by Old Romance is also found across (mostly non-standard) Germanic varieties (Walkden 2015, Cognola 2013, 2019) and present-day conservative Romance varieties like Rhaeto-Romance (Casalicchio/Cognola 2018). So far, there have been no quantitative studies on the asymmetric pro-drop phenomenon in OI, even though it is uncontroversial that the phenomenon exists. The asymmetric pro-drop system of OI was lost during the 14th century together with the V2 constraint.

3.3 Interim summary

Both OHG and OI were clearly languages in which referential null subjects were possible. In both there is an asymmetry between main and embedded clauses such that pro-drop seems to be a root phenomenon. For both a purely syntactic account has been proposed, in which the finite verb in C licenses the null subject. For OI this remains the state of the art. For OHG, this account has been challenged by Schlachter (2010, 2012) and Walkden (2012, 2014), who propose a discourse-driven account. In neither case is the asymmetry particularly well understood theoretically or empirically, however: for OI there has been no thorough study comparing different clause types quantitatively, while for OHG the rather coarse-grained quantitative study of Eggenberger (1961) has formed the basis for most generalizations. Interrogative clauses have not been considered separately in either case (though see the brief discussion in Schlachter 2010, 2012). In the rest of this paper we present new data supporting a discourse-driven account, and provide a formalization of this account broadly following Frascarelli (2007, 2018).

4 A new investigation

4.1 Sources and methodology

4.1.1 The Diatessaron

In order to investigate the distribution of null subjects in partial pro-drop languages we carried out a parallel study on two translations of the same text in OI and in OHG. The text we

---

10 In this paper OI is taken, as is standard, to refer to the variety represented by the Old Tuscan texts transmitted from the Late Middle Ages (12th-15th centuries). OHG refers to all German texts from what is now central and southern Germany, 750–1250. The OHG texts are not homogeneous; for us, the East Franconian dialect of the monastery at Fulda plays a central role.
focussed on is Tatian’s *Diatessaron* in its OHG and Old Tuscan translations.\(^{11}\) As intensively discussed by Petersen (1997), both texts belong to the Eastern tradition of the translation history of the *Diatessaron* and are translations of a version of a Latin translation stemming from the so-called Codex Fuldensis ordered by San Vittore, Bishop of Capua and finished in 546. This version, which relies on a previous Latin translation of the text (possibly the *Vetus Latina*), is interpolated with passages from the *Vulgata* and is the basis of the Western tradition of the text. The Eastern translations, on the other hand, constitute an independent tradition because they exhibit the heretical and more conservative passages which were removed or changed on the basis of the *Vulgata*.

### 4.1.2 The OHG translation

The only OHG version of the *Diatessaron* that has survived is that of the Codex Sangallensis 56 (ms Sankt Gallen, Stiftbibliothek 56). It is written in East Franconian dialect and dates back to around 830 CE (cf. Petersen 1997). The translators are unknown. The OHG *Diatessaron* is transmitted in a bilingual Latin-OHG code (*lateinisch-alt hochdeutsch Tatianbilingue*).

Whether the Latin in the Codex is the source for the OHG translation is debated in the literature. Masser (1994) considers that it is, whereas Baumstark (1964), Wissmann (1960), and Petersen (1997) suggest that the OHG version is the translation of an unknown Latin version. What appears to be relatively uncontroversial is that the Latin of the Codex Sangallensis can be used for syntactic investigation (Lippert 1974), since the differences from the unknown Latin version are not so dramatic (cf. Petersen 1997).

While traditionally the OHG translation has been considered to be an interlinear translation, which does not say much about OHG and cannot be used for linguistic research on OHG (see among others Sievers 1892, Lippert 1974, Masser 1991, Sonderegger 2003), more recent works have shown that the OHG translators created an independent text, which deviates in many respects from the Latin, and is thus a reliable source that can be used for syntactic research (see Dentschewa 1987, Dittmer/Dittmer 1998, Axel 2007, and Fleischer, Hinterhölzl/Solf 2008).

### 4.1.3 The Tuscan translation

The Tuscan *Diatessaron* is transmitted in 25 manuscripts dating from 1300 to 1500 and has been published in the critical edition by Vaccari/Vatasso (1938), which is also included in the *Opera del Vocabolario Italiano* (OVI) database.

---

\(^{11}\) The *Diatessaron* (from Greek, “one from four”) is a Gospel Harmony, i. e. a type of text which tries to unify the facts told in the four Gospels in a coherent narrative (no omissions, repetitions, inconsistencies etc.), written in Rome by Tatian the Assyrian around 170. Tatian was a pupil of Justin Martyr (a Christian Apologist philosopher) and a theologian, apologist and philosopher himself who developed a new theology reflected in the *Diatessaron*. The original, which was very likely written in Syrian or in Greek, has been lost, but we know of its existence because some parts of it are mentioned in other works (Ephrem the Assyrian’s *Comment*, 4th century). The *Diatessaron* was a “bestseller” in the ancient world, and was translated in many languages. Tatian’s *Diatessaron* was the official text of the Syrian church until the 5th century, when all copies were burned because the *Diatessaron* was then considered to be heretical, and from this Eastern tradition stem the translations in Persian, Armenian, Arabic etc. From a Latin translation (*Vetus Latina*) stem the translations in Old Dutch (Codex of Liege, XIII cent) and Venetian dialect (Codex Marciano 4975, sec. XIV) and in Old High German. See Petersen (1997) and Gambino (2001) for more details of the history of the text.
According to Vaccari/Vatasso (1938), all manuscripts derive from a Latin translation in OI dating back to 1200 (see the specific and very archaic lexical, morphological and syntactic features, Vaccari/Vatasso 1938: 184-190). This translation in OI, which is considered the archetype, stems from a Latin text which was very close to the Fulda manuscript (according to the critical edition of Ranke 1868 used by the authors) but with some differences (e.g. in the interpolation of the material; chapters are not numbered). According to Vaccari/Vatasso (1938), three of the oldest manuscripts from 1300 are copies of the archetype: S = Codice Senese (I.V.9) (Siena), P = Palatino Latino 56 – Biblioteca Apostolica Vaticana (Rome), and L = Riccardiana 2335 (Florence); the later manuscripts are copies of these three texts. Dating is reached through palaeographical considerations.

The critical edition by Vaccari/Vatasso (1938) is based on the Codice Senese (I.V.9) with small corrections at the level of lexical emendations.

4.1.4 Methodology

The data discussed in this paper rely on a qualitative and quantitative study of Tatian’s Diatessaron in OHG and OI. The focus on novel quantitative data is an innovation in our work, since all previous work on the OHG Tatian relies on Eggenberger’s (1961) quantitative data, and there is no published quantitative work on OI at all.

The corpus comprises the first ten chapters of the text as well as all interrogative clauses of the texts for both OHG and OI. The Latin and OHG sentences are taken from the electronic editions available in the TITUS Database, whereas the OI sentences were transcribed manually into the corpus from the critical edition of Vaccari/Vatasso (1938).

The inclusion of interrogative clauses in the corpus is a novelty of this study, since in previous works little attention was devoted to clause type. We are convinced that main interrogative clauses, in the light of the theoretical discussion in sections 2 and 3 above, are a key environment to test the validity of both the syntactic and the discourse accounts for null subjects in OHG and OI. The syntactic hypothesis (V-to-C licensing) predicts that subjects should be mostly null in this syntactic environment, since main interrogative clauses are the prototypical V2 environment (cf. residual V2 in the sense of Rizzi 1991 and Kiparsky 1995), whereas the discourse hypothesis predicts null subjects to be mostly overt in this environment because no topic is present (see sections 2.3 above and 5 below).\(^\text{12}\)

---

\(^\text{12}\) An anonymous reviewer casts doubt upon the idea that interrogative clauses should be considered residual V2 clauses because in present-day Italian no Germanic inversion (Wh-element-AUX-DP subject-lexical VERB) is found. As intensively discussed in the literature (Rizzi 2005, 2006, Cardinaletti 2004, 2010 among others) the ungrammaticality of Germanic inversion in interrogative clauses in present-day Italian depends on the realization of DP subjects in the language, on the relationship between CP and IP and on the interplay between syntax and information structure (DP subjects are only realized when they are topicalized or focussed). Therefore, the absence of Germanic inversion in interrogative clauses in present-day Italian is not fed by the lack of V-to-C movement, but by independent properties of DP-subject syntax.

The same reviewer also casts doubt on the V2 nature of interrogative clauses in OI by providing as an argument the fact that Germanic inversion was not very common in OI in wh-interrogative clauses. We do not think that there should be any doubt about the fact that OI interrogative clauses exhibited the V2 property precisely because of the presence of Germanic inversion – a construction ruled out in present-day Italian. The fact that Germanic
4.2 The data: overview and first generalisations

In this section, we present our corpus data and compare the two translations of the *Diatessaron*. The Latin we use for the comparison is that of the Codex Sankallensis 56 (ms Sankt Gallen, Stiftbibliothek 56) which we are aware is not the source used for the translations, but it is very close to the original (see section 4.1 above).

In Table 1 we consider the translation of all sentences featuring an overt subject in the Latin divided according to clause type. We see that in all the sentences in which an overt subject is present in Latin, an overt subject will also be present in the translations, with virtually no exceptions.

<table>
<thead>
<tr>
<th>Clause Type</th>
<th>Main declaratives</th>
<th>Main interrogatives</th>
<th>Embedded clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latin</td>
<td>101</td>
<td>56</td>
<td>35</td>
</tr>
<tr>
<td>Overt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OHG</td>
<td>101/101 (100%)</td>
<td>56/56 (100%)</td>
<td>35/35 (100%)</td>
</tr>
<tr>
<td>(0%)</td>
<td>(0%)</td>
<td>(0%)</td>
<td>(0%)</td>
</tr>
<tr>
<td>OI</td>
<td>99/101 (98%)</td>
<td>51/56 (91%)</td>
<td>35/35 (100%)</td>
</tr>
<tr>
<td>(2%)</td>
<td>(9%)</td>
<td>(0%)</td>
<td>(0%)</td>
</tr>
<tr>
<td>Total</td>
<td>200/202 (100%)</td>
<td>107/112 (96%)</td>
<td>70/70 (100%)</td>
</tr>
</tbody>
</table>

Table 1: Translations of all sentences with a DP subject in the Latin

Before commenting on the results in Table 1, let us consider the translation of Latin sentences featuring a null subject in Table 2. We see that i) about 90% of main declarative clauses with a null subject in the Latin also feature a null subject in OHG and OI; ii) about 40% of the translated main interrogatives feature a null subject; and iii) null subjects only appear in a minority of embedded clauses in both translations.

Inversion was not obligatory is due to the fact that it coexisted with other syntactic options, like Romance inversion (Wh-element-AUX-lexical VERB-DP subject) or pro-drop (Munaro 2010), simply indicates that DP subjects do not share an identical syntax with present-day strict V2 languages, like German, not that V-to-C did not take place. Interestingly, in OI Germanic inversion was obligatory when the finite verb is followed by a clitic pronoun (Munaro 2010: 1159) - a fact which is very reminiscent of present-day Northern Italian varieties exhibiting subject clitics (Poletto 2000) and some Germanic varieties (Cognola 2019).
Table 2: Translations of all sentences with a null subject in the Latin

<table>
<thead>
<tr>
<th></th>
<th>Main declaratives</th>
<th>Main interrogatives</th>
<th>Embedded clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latin</td>
<td>60</td>
<td>183</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>Overt</td>
<td>Null</td>
<td>Overt</td>
</tr>
<tr>
<td>OHG</td>
<td>5/60</td>
<td>55/60</td>
<td>125/183</td>
</tr>
<tr>
<td></td>
<td>(8%)</td>
<td>(92%)</td>
<td>(69%)</td>
</tr>
<tr>
<td>OI</td>
<td>6/60</td>
<td>54/60</td>
<td>102/183</td>
</tr>
<tr>
<td></td>
<td>(10%)</td>
<td>(90%)</td>
<td>(56%)</td>
</tr>
<tr>
<td>Total</td>
<td>11/120</td>
<td>109/120</td>
<td>227/366</td>
</tr>
</tbody>
</table>

The data presented in Tables 1 and 2 allow us to draw the following descriptive generalisations about the distribution of overt and null subjects.

(16) a. Null subjects are only possible in OHG and OI when the Latin features a null subject and are virtually absent when an overt subject is present in the original;
b. Null subjects are restricted, though with different percentages, to main declarative clauses and main interrogative clauses, and are almost fully excluded from embedded clauses;
c. OHG and OI pattern together in all clause types, but in wh- main interrogative clauses and in embedded clauses null subjects are slightly (about 10%) more frequent in OI than in OHG.\(^{15}\)

The descriptive generalisations above are already clear evidence against the idea that the syntax featured in the OHG and OI texts should be considered loan-syntax on the basis of the Latin model, proposed among others by Eggenberger (1961) based on generalisation (16a). Alongside the convergence between the Latin and the OHG and OI translations, in fact, we see that the translators consistently pattern against the Latin source in all three clause types. In embedded clauses, null subjects are a minority of cases in comparison to the Latin; in main declarative clauses null subjects appear in 80% rather than 100% of cases, which speaks against a slavish translation. This claim is made even stronger by the distribution of null subjects in interrogative clauses, which are null in a minority of wh-interrogative clauses (28% in OHG and 40% in OI) against the Latin.

The generalisations in (16) also provide strong evidence against the V-to-C licensing hypothesis discussed in section 2.2 above. This approach to null subjects in asymmetric pro-drop languages predicts the distribution of null subjects to be determined by the syntactic position of the finite verb in the C head position. Therefore, its main prediction is that null subjects are more restricted in non-root embedded clauses – a prediction which is borne out by our data (Table 2).

\(^{13}\) In one case the null subject involves a passive construction.

\(^{14}\) Out of these 15 sentences 5 involve an impersonal construction (2/5) or an infinitive sentence (3/5).

\(^{15}\) These numbers do not indicate that OI patterns with present-day Italian. The distribution of null subjects in the OI translation, in fact, fully diverges from that of present-day Italian where null subjects are much more frequent across all clause types. Therefore, the slight differences we detect in the OI system are not imputable to the instantiation of a consistent or “canonical” pro-drop system.
This approach also makes two predictions about the syntax of main clauses. The first is that interrogative clauses should exhibit a high percentage of null subjects, given that they are the prototypical context for V-to-C movement, also in residual V2 languages such as English (Rizzi 1991). Our data indicate that this prediction is not borne out, since interrogative clauses show a much lower percentage of null subjects compared to main declarative clauses (around 40% vs 80%) – exactly the opposite of what the V-to-C approach predicts. The second prediction made for main clauses by the V-to-C approach is that null subjects should be possible, to some extent at least, in the translation irrespective of the Latin model. If the licensing of the null element were simply dependent on the position of the finite verb in C, one would expect it to be possible in contexts where this licensing condition is met – regardless of the Latin. Again, this prediction is not borne out, since in our data null subjects are virtually absent when the Latin has an overt subject (see Table 1). We believe that these data call for a new explanation which we will pursue in section 5. First, however, we will take a closer look at the individual contexts where pro-drop is found.

4.3 Analysis of the contexts

4.3.1 On the expression of the subject in main declarative clauses

As summarised in Table 3, the subject is almost always realised in sentences in which it is also present in the Latin, whereas it is mostly (about 80% of the cases) absent in the case in which it is null in the original.

<table>
<thead>
<tr>
<th></th>
<th>OHG</th>
<th>Overt</th>
<th>OI</th>
<th>Overt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null subject in the Latin</td>
<td>55/60</td>
<td>5/60</td>
<td>54/60</td>
<td>6/60</td>
</tr>
<tr>
<td>Overt subject in the Latin</td>
<td>0/56</td>
<td>56/56</td>
<td>5/56</td>
<td>51/56</td>
</tr>
</tbody>
</table>

Table 3: Distribution of overt and null subjects in main declarative clauses

As discussed in section 4.2 above, we suggest that the correspondence in the distribution of overt subjects in the three languages is not to be taken as evidence supporting the loan-syntax hypothesis, but simply indicates that all three languages require an overt subject to be present for pragmatic reasons. More specifically, the overt subject has a crucial function in the narration, since it introduces aboutness/shift topics or bears focus accent – and cannot thus be null (Frascarelli 2007, 2018).

In order to show how overt subjects have a narrative function and can thus not be omitted even in a consistent null-subject language like Latin, let us consider the first three sentences of the Diatessaron’s chapter 2. Let us consider first this passage in English (17). In bold we highlight the overt nominal subjects (a priest, his wife Elisabeth, Elisabeth, they both). From the point of view of information structure, the first two overt subjects realise two new-information foci, whereas the overt subjects Elisabeth and they both in the final clause can be analysed as two contrastive topics (see section 2.3 above, Frascarelli/Hinterhölzl 2007, and Cruschina 2009 on the definition of this topic class). In the second and third sentences, where information is added in the narration, the two new-information foci a priest and his wife are co-indexed with the overt subject pronouns their (English being a non-null subject language, see Roberts/Holmberg 2010: 4); in one case, the pronoun can be left null due to coordination (see below, section 4.3.3).
“In the time in which Herod was king of Judea, there lived a priest, whose name was Zacharia, of the family of Abia, and his wife Elisabeth, one of the daughters of Aaron. They were both righteous before God, and followed all God’s commands and justifications without reproach. They had no son, since Elisabeth was barren and they both old.”

In the above passage we see that overt DP subjects have a clear discourse-informational function in the narration, since they introduce new-information foci and topics, which are then referred to with overt pronouns in further points of the texts.

Let us consider this passage in Latin, OHG and OI (17b). In all three languages, the following are overt subjects: the new-information focus “a priest” (Latin: *quidam sacerdos*; OHG: *sumer biscof*; OI: *uno sacerdote*); “Elisabeth”; “both” (Latin: *autem*; OHG: *siu beidu*; OI: *amendue*); and the contrastive topics “Elisabeth” and “both” (Latin: *ambo*; OHG: *beidu*; OI: *amendue*).

Except for *siu* there are no pronouns in the texts.16

What we see is that all the overt subjects appearing in the texts are actually crucial to narrative function since they all play a discourse-informational role and can thus not be omitted in the OHG and OI translations: they must be expressed overtly.

---

16 In two cases, the construction of the sentence in OI slightly differs from the Latin and OHG, such that the Latin/OHG subject corresponds to a non-subject constituent in OI. For the new-information focus ‘his wife’ (Latin: *uxor illi*, OHG: *quena imo*), the corresponding referent is introduced as part of a PP (*per moglie*, literally ‘for wife’), but is nevertheless overt. In another case the overt subject is within a so-called “dative of possession” construction (Latin: *et non erat illis filius*; OHG: *inti uuard iu sun*), and is rendered as an overt object in OI (*figliuolo*) with the (continuing) subject remaining null.
4.3.2 Main interrogative clauses

Let us consider the realization of the subject in main interrogative clauses.

We extracted all interrogative clauses appearing in the manuscripts and divided them according to the presence or absence of an overt subject in the Latin.

There are 191 main interrogative clauses lacking the subject in the Latin but we excluded 8 sentences from the analysis because the three languages do not pattern together. More specifically, in some contexts the OI translator uses a personal construction to translate an impersonal construction in the Latin and in OHG (3 cases). In the example in (18) we see that the Latin and OHG dative of possession constructions (non est tibi curae / nist thir iz sorga “it is not to you care” = “you do not care”) correspond to a personal construction with the auxiliary to have in OI (non hai tu cura “you do not care”).

(18) a. Quæ stetit et ait: domine, non est tibi curae
who stayed.3SG and said.3SG: Lord NEG is.3SG to you care
quod soror mea reliquit me solam ministrare?
that sister my leaves.3SG me alone serve

b. Thiu stuont tho inti quad: truhin, nist thir iz sorga
She stayed.3SG there and said.3SG: Lord NEG.is.3SG to you it care
thaz min suester liez mih enum embahten? [63,3]
that my sister lets me alone serve

In two cases, it is the OI translator who uses the impersonal construction. In (19) we see that in the OI text the interrogative clause appears with the impersonal verb parere “to seem”, whereas the interrogative features the personal verb to think in the Latin and in OHG texts. Therefore, only in Latin and OHG is there a null referential subject.

(19) a. Quid putatis, quia non veniat ad diem festum?
what think.2PL, that NEG came.3SG to the day festive

b. uuaz uuanet ir, bithius her ni cumit ci themo itmalen tage? [135,34]
what think.2PL you that he came.NEG 3SG to the festive day

c. Che vi pare, ch’egli non è venuto al di della festa?
what to you seem.3SG that he NEG is come to the day of festive

In one paragraph, two interrogative clauses appearing in the Latin and in the OHG are missing in the OI translation. In one case, given in (20), a direct interrogative clause featuring in the
Latin (quid putas “who do you think”) is not translated in OHG: therefore this example was thus excluded from the corpus.  

(20) a. Et accesserunt discipuli ad Ihesum dicentes: and came.3PL pupils to Jesus saying quis putas maior est in regno caelorum? who think.2SG greatest is in realm heaven  

b. Tho giengun the iungoron zi imo quedente: EXPL went.3PL the youngs to him saying uuer ist mero in himilo riche? [94,2] who is most in heaven realm  
c. E vennero i discipoli a llui e dissero: quale ti pensi tu and came.3PL the pupils to him and said: which REF.PRON think.2SG you che sia maggiore nel regno de’ cieli? [95, 273, 16-17] that is.CONJ biggest in the realm of heaven  

‘And the disciples came to Jesus and said: who do you think is greatest in the realm of heaven?’

In Table 4 we provide an overview of the sentences considered in the corpus. We see that in almost all cases in which there is a pronoun in the Latin model, there is one in the OHG and OI translations; conversely, when the subject pronoun is null, in the majority of cases a pronoun is inserted in the translations, contrary to the Latin.

<table>
<thead>
<tr>
<th>Null</th>
<th>Overt</th>
<th>Overt subject in the Latin</th>
</tr>
</thead>
<tbody>
<tr>
<td>OHG</td>
<td>0/56 (0%)</td>
<td>56/56 (100%)</td>
</tr>
<tr>
<td>OI</td>
<td>5/56 (9%)</td>
<td>51/56 (91%)</td>
</tr>
</tbody>
</table>

*Table 4: Translations of interrogative clauses with a DP subject in the Latin*

In all cases in which a DP subject or an overt subject pronoun is present in the Latin (and thus in the OHG and OI translations) it has a discourse/pragmatic function. In the following example, for instance, the DP subject “somebody from Nazareth” is a new-information focus:

(21) a. Et dixit ei Nathanahel: a Nazareth potest aliquid boni esse? And said to him Nathaniel: in Nazareth can.3SG somebody good be Dicit ei Philippus: veni et vide. Said to him Philipp: come and see

b. Thó quad imo Nathanahel: fón Nazareth mág sihuaz quotes uuesan? EXPL said to him Nathaniel: from Nazareth can somebody good be Thó quad imo Philippus: quim inti gish. [17,3] EXPL said to him Philipp: come and see.

---

17 These asymmetries indicate that the OHG and OI translations appear to be autonomous texts from both the Latin, on the one hand, and from each other, on the other. This latter observation speaks in favour of the fact that the OI translation was possibly made from a Latin version of the Diatessaron, as proposed by Vaccari/Vatasso (1938), and makes it highly implausible that the OI translators had access to an OHG version. (A more plausible alternative might be that they had access to an Old French copy (as pointed out to us by Marco Infurna); we will not take a stance on this here.)
c. E disse a llui Nathanael: puote essere da Nazzaret
And said.3SG to him Nathaniel: can.3SG be from Nazareth
alcuno bene?
somebody good
E Filippo disse: vieni a vederlo. [17, 218, 18-20]
and Philip said.3SG: come and see-it
‘And Nathaniel said to him: Can someone from Nazareth be good? And Philip said: come and see.’

In (22), the overt subject pronoun “I” (ego, ih, io) has a contrastive function:

(22) a. Respondit Pilatus: numquid ego Iudeus sum?
Answered Pilatus: is it possible I Jew be?
Gens tua et pontifices tradiderunt te mihi:
People your and pontiffs brought.3PL you to me:
quid fecisti?
what did.2SG
b. Tho antligita Pilatus: eno bin ih Iudeus?
EXPL answered.3SG Pilatus: ENO am I Jew
Thin thiota inti bisgoffa saltun thih mir: uuas
Your people and pontiffs brought.3PL you to me: what
tati thu? [195,3]
did.3SG you
c. Rispuose Pilato, e disse: or sono io giudeo?
Answered Pilatus, and said.3SG: now am I a jude?
La gente tua e i pontefici mi t’anno dato.
the people your and the pontiffs me you-have given
Dunque: che facesti? [168, 350, 18–20]
Then: what did.2SG
‘Pilate answered: Am I a Jew?!! Your people and priests brought you to me. What have you done?’

In both translations there is a strong tendency to insert a subject pronoun (or a DP subject, see section 4.3.3 below) contrary to the Latin (in 70% of the cases in OHG and 56% of cases in OI): see Table 2 above. Here we provide some examples of these cases. All of (23)–(28) involve insertion of a subject pronoun contrary to the Latin. In (24) the subject is overt in OHG but not in OI; in (25) it is the other way round. In the other examples, the subject pronoun is inserted in both OHG and OI.

(23) a. Dicit ei Ihesus: nonne dixi tibi quoniam si
Told to her Jesus: is it not true told you when if
credideris, videbis gloriam dei?
believe.FUT.2SG see.FUT.2SG glory of.God
b. Tho quad iru ther heilant: ia quad ih thir, oba thu
EXPL said.3SG her the saviour IA told I to you if you
giloubist, gishist gotes diurida? [135,24]
believe.2SG see.2SG of.God glory
The first issue to be addressed is whether variation in the distribution of null/overt subjects across interrogative types (yes/no vs wh-interrogatives) is to be detected. In Table 5 we consider
the distribution of null subjects across main interrogative types. We see that in both languages the rate of null subjects in yes-no questions is around 30%. If we consider the percentage of null subjects in wh-interrogative clauses, we see that in OHG it is around 30%, as in yes-no questions. This means that there are no differences in the distribution of null subjects between the two types of interrogative clauses in this language (Fisher’s exact test, $p = 0.8683$). In OI, on the other hand, the percentage of null subjects in wh-interrogative clauses is 50.4% – which indicates that null subjects appear to be favoured in wh-interrogative clauses (Fisher’s exact test, $p = 0.0126$).

<table>
<thead>
<tr>
<th></th>
<th>OHG</th>
<th>OI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes-no interrogatives</td>
<td>Null: 21/71 (30%)</td>
<td>Null: 22/71 (31%)</td>
</tr>
<tr>
<td></td>
<td>Overt: 50/71 (70%)</td>
<td>Overt: 49/71 (69%)</td>
</tr>
<tr>
<td>Wh-interrogatives</td>
<td>Null: 32/103 (31%)</td>
<td>Null: 52/103 (50%)</td>
</tr>
<tr>
<td></td>
<td>Overt: 71/103 (69%)</td>
<td>Overt: 51/103 (50%)</td>
</tr>
</tbody>
</table>

Table 5: Distribution of null/overt subjects in yes/no interrogative clauses

In table 6 we consider the distribution of null subjects across different types of main wh-interrogative clauses. With the exception of interrogative clauses introduced by where (which all feature the 2nd person singular, known to favour null subjects; see below), all types of wh-interrogative clauses feature about 30% of null subjects. We thus conclude that null subjects are not favoured by any type of wh-interrogative element in OHG. In OI, on the other hand, we see that why- and what-interrogative clauses, which are the most numerous in the sample, feature a null subject contrary to the Latin in 58% and in 45% of the cases. We thus seem to observe that wh-clause type appears to be a factor affecting the distribution of null subjects in OI.

<table>
<thead>
<tr>
<th></th>
<th>OHG</th>
<th>OI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wh-element: why</td>
<td>Null: 8/34 (23%)</td>
<td>Null: 20/34 (58%)</td>
</tr>
<tr>
<td></td>
<td>Overt: 26/34</td>
<td>Overt: 14/34</td>
</tr>
<tr>
<td>What (object)</td>
<td>Null: 12/35 (34%)</td>
<td>Null: 16/35 (45%)</td>
</tr>
<tr>
<td></td>
<td>Overt: 23/35</td>
<td>Overt: 19/35</td>
</tr>
<tr>
<td>Whom (indirect object)</td>
<td>Null: 2/7 (28%)</td>
<td>Null: 1/7 (14%)</td>
</tr>
<tr>
<td></td>
<td>Overt: 5/7</td>
<td>Overt: 6/7</td>
</tr>
<tr>
<td>Wh-phrase (how much/many)</td>
<td>Null: 2/8 (25%)</td>
<td>Null: 5/8 (62%)</td>
</tr>
<tr>
<td></td>
<td>Overt: 6/8</td>
<td>Overt: 3/8</td>
</tr>
<tr>
<td>Where</td>
<td>Null: 4/5 (80%)</td>
<td>Null: 1/5 (20%)</td>
</tr>
<tr>
<td></td>
<td>Overt: 1/5</td>
<td>Overt: 4/5</td>
</tr>
<tr>
<td>When</td>
<td>Null: 1/5 (20%)</td>
<td>Null: 2/5 (40%)</td>
</tr>
<tr>
<td></td>
<td>Overt: 4/5</td>
<td>Overt: 3/5</td>
</tr>
<tr>
<td>How</td>
<td>Null: 3/9 (33%)</td>
<td>Null: 7/9 (77%)</td>
</tr>
<tr>
<td></td>
<td>Overt: 6/9</td>
<td>Overt: 2/9</td>
</tr>
<tr>
<td>Total</td>
<td>Null: 32/103 (31%)</td>
<td>Null: 52/103 (50%)</td>
</tr>
<tr>
<td></td>
<td>Overt: 71/103</td>
<td>Overt: 51/103</td>
</tr>
</tbody>
</table>

Table 6: Distribution of null/overt subjects in wh-interrogative clauses

The last variable which has been shown in the literature to play a role in the distribution of null subjects (see section 3.1 above) is person type. The data in Table 7 indicate that the claim that

---

18 We exclude 9 examples featuring a wh-element that appears less than 5 times in the corpus.
person type favours null subjects is confirmed by our data. More specifically, we find that null subjects are more frequent with the first person singular (39%), plural (53%) and second person singular (42%) in OHG, and with the first person plural in OI (94%).

<table>
<thead>
<tr>
<th>Type</th>
<th>OHG</th>
<th>OI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Null</td>
<td>Overt</td>
</tr>
<tr>
<td>1 SG</td>
<td>7/18 (39%)</td>
<td>11/18</td>
</tr>
<tr>
<td>1PL</td>
<td>9/17 (53%)</td>
<td>8/17</td>
</tr>
<tr>
<td>2 SG (enclitic <em>tu</em> excluded)</td>
<td>27/64 (42%)</td>
<td>37/64</td>
</tr>
<tr>
<td>2 PL</td>
<td>8/62 (13%)</td>
<td>54/62</td>
</tr>
<tr>
<td>3 SG</td>
<td>5/18 (28%)</td>
<td>23/18</td>
</tr>
<tr>
<td>3 PL</td>
<td>1/4 (25%)</td>
<td>3/4</td>
</tr>
<tr>
<td>Total</td>
<td>58/183</td>
<td>125/183</td>
</tr>
</tbody>
</table>

Table 7: Distribution of null/overt subjects across persons and numbers

To sum up, we have found that in main interrogative clauses, person appears to favour null subjects in OHG (1st and 2nd singular; 1st plural) but the type of interrogative clause has no visible effect. On the other hand, both person (1st plural) and type of interrogative clause (*wh*-interrogatives) favour null subjects in OI.

4.3.3 Null subjects in main declarative clauses

In this section we examine the distribution of null subjects in main declarative clauses. As shown in Table 8, null subjects appear in the majority (around 90%) of main declarative clauses translating a Latin main clause featuring a null subject in both OHG and OI.

<table>
<thead>
<tr>
<th>Type</th>
<th>OHG</th>
<th>OI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Null</td>
<td>Overt</td>
</tr>
<tr>
<td>Null subject in the Latin</td>
<td>55/60</td>
<td>5/60</td>
</tr>
<tr>
<td>Overt subject in the Latin</td>
<td>0/56 (0%)</td>
<td>56/56 (100%)</td>
</tr>
</tbody>
</table>

Table 8: Distribution of null/overt subjects in main declarative clauses

In our corpus, we considered all finite main declarative clauses lacking an overt subject appearing in the first 10 chapters of the OHG *Diatessaron*. We excluded all impersonal constructions as in (29)–(31), irrespective of their translations in OHG and OI.¹⁹

(29) a. *Sic enim scriptum est per prophetam*
    so therefore written is by prophet
    
(29) b. *Só ist giscriban thuru then uuizagon [8,3]*
    so is written through the prophet

¹⁹ In one case each, OHG and OI translate an impersonal construction with a personal construction featuring an overt DP subject. In the other examples no overt pronoun appears in OI, whereas the overt pronouns *her* ‘he’ and the expletives *thô* and *iz* appear in OHG. In one case no overt expletive appears in OHG.
c. imperciò che ‘l profeta scrisse così [7(5)]
   therefore that the prophet wrote.3SG so
   ‘So it is written by the prophet …’

(30)
  a. sicut scriptum est in lege domini
     so how written is in law of.the.Lord
  b. Só iz gisriban ist in gotes euuu [7,2]
     so how EXPL written is in God’s rules
  c. secondo ch’è scritto nella legge del Signore [8(23)]
     according what-is written in-the law of.the Lord
   ‘As is written in the law of the Lord.’

(31)
  a. Et factum est in die octava
     and done is in day eight
  b. Uuard thô in themo ahtuden tage [4,11]
     became EXPL in that eight day
  c. E adivenne che nell’ ottavo di [4(59)]
     and happened.3SG that in the eight day
   ‘And on the eighth day it happened that …’

We also excluded sentences which were not comparable because they have been translated through different strategies influencing the realization of subjects; see for instance cases like (32) and (33) in which one of the two languages (OI in these cases) translates a main clause with a discourse marker with a complementiser which we know is a context favouring the presence of the overt subject).

(32)
  a. ecce concipies in utero
     here conceive.FUT.2SG in uterus
  b. seno nu imphahis in reue [3,3]
     you will see conceive.FUT.2SG in uterus
  c. Ecco che tu conceperai [3(31)]
     here is that you conceive.FUT.2SG
     ‘You will see that then you will conceive a baby.’

(33)
  a. vidimus enim stellam eius in oriente
     saw.1PL then star his in eastern
  b. uuir gisahumes sinan stellos in ostarlante [8,1]
     we saw.1PL his star in eastern.land
  c. Imperciò che noi vedemmo la stella sua in oriente [7(2)]
     then that we saw.1PL the star his in eastern
     ‘And we saw his star in the East.’

Let us now consider the distribution of null subjects according to person in the 60 relevant sentences. As shown in Table 9 the great majority of sentences in the corpus feature a third person, which is mostly null.
Table 9: Distribution of null/overt subjects across persons and numbers

<table>
<thead>
<tr>
<th></th>
<th>OHG</th>
<th></th>
<th>OI</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Null</td>
<td>Overt</td>
<td>Null</td>
<td>Overt</td>
</tr>
<tr>
<td>3.SG</td>
<td>33/38</td>
<td>5/38</td>
<td>33/38</td>
<td>5/38</td>
</tr>
<tr>
<td>3.PL</td>
<td>16/17</td>
<td>1/17</td>
<td>16/17</td>
<td>1/17</td>
</tr>
<tr>
<td>1.SG</td>
<td>0/1</td>
<td>1/1</td>
<td>0/1</td>
<td>1/1</td>
</tr>
<tr>
<td>1.PL</td>
<td>0/3</td>
<td>3/3</td>
<td>0/3</td>
<td>3/3</td>
</tr>
<tr>
<td>2.SG</td>
<td>1/1</td>
<td>0/1</td>
<td>0/1</td>
<td>1/1</td>
</tr>
<tr>
<td>2.PL</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>50/60</td>
<td>10/60</td>
<td>49/60</td>
<td>11/60</td>
</tr>
</tbody>
</table>

The distribution of null subjects in Latin (and therefore in the OHG and OI translations) is restricted to two recurring contexts. The first, which we label *paragraph-beginning*, is illustrated in the examples in (34) and (35). In this configuration, a new-information DP subject is introduced in the narration in the first sentence (typically a non-finite clause) of the new paragraph and is then left unexpressed in the immediately following sentence.

(34)  a. Audiens autem Herodes rex, turbatus est et omnis Hierusolima cum illo in hearing then Herod king upset is and whole Jerusalem with him
     b. Thò thaz gihorta Herodes ther cuning, uuard pro, gitruobit [8,2] as this heard.3SG Herod the king became upset
     c. Udendo ciò il re Erode, turbossi pro, e tutta Gerusalem co llui. [7(3)] in hearing this the king Herod upset.REFL and whole Jerusalem with him

‘On hearing this, Herod was upset and the whole of Jerusalem with him.’

(35)  a. Exsurgens autem Joseph, a somno emerging then Joseph from sleep
     fecit sicut precepit ei angelus domini did.3SG how commanded.3SG him the angel of.Lord
     b. Arstantanti thò Joseph fon slafe teta pro, emerging EXPL Joseph from sleep did.3SG só imo giböt truhtines engil [5,10] how him commanded.3SG of.Lord angel
     c. Levandosi Gioseppo, fece pro come gli comandò l'angelo [5(25)] getting up Joseph did.3SG how him commanded the angel

‘Once Joseph got up, he did as commanded by the angel.’

The second context, which we label *coordination*, involves the coordination of two main declarative clauses featuring the same abstract referential subject.

(36)  a. Exsurgens autem Joseph, a somno emerging then Joseph from sleep
     fecit sicut precepit ei angelus domini did.3SG how commanded.3SG him the angel of.Lord
     et accepit pro conjugem suam, et pro non
In Table 10 we illustrate the distribution of null/overt subjects in these two contexts in OHG and OI. The data indicate that contexts favours null subject.

<table>
<thead>
<tr>
<th></th>
<th>OHG</th>
<th>OI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Null</td>
<td>Overt</td>
</tr>
<tr>
<td>Paragraph-beginning</td>
<td>19/23</td>
<td>4/23</td>
</tr>
<tr>
<td>Coordination</td>
<td>35/37</td>
<td>2/37</td>
</tr>
<tr>
<td>Total</td>
<td>54/60</td>
<td>6/60</td>
</tr>
</tbody>
</table>

Table 10: Distribution of null/overt subjects across discourse contexts

In nearly all cases in which a subject is inserted contrary to the Latin, we have to do with contrast on the subject, typically of a contrastive topic or focus, as in (37).

(37) a. Parīet autem filium, et vocabis nomen eius Ihesum
give.birth.FUT then son and call.FUT.2SG name his Jesus

b. Siu gibirit sun, inti thū
she give birth.FUT.3SG son and you
ginemnis sinan namon Heilant [5,8]
call.2SG his name saviour

c. Ella partorirà uno figliuolo e
and she give.birth.FUT.3SG a baby.boy and
tu chiamerai il nome suo Gesù [5(21)]
you call.FUT.2SG the name his Jesus
‘And she will give birth to a son, and you will call his name Jesus.’

4.3.4 Summary

As Table 2 shows, when the Latin has a null subject, OHG and OI null subjects are robustly possible in main declaratives, less robustly in interrogatives, and not found at all in embedded clauses, at least in our sample. Where OHG and OI translate a (Latin) overt subject with an overt subject of their own, this is not simply because of slavish translation but rather because of the discourse requirements of the narrative: the overt subjects play a crucial informational role.

In main interrogative clauses, person appears to favour null subjects in OHG (2nd singular) but the type of interrogative clause has no visible effect. On the other hand, both person (1st plural) and type of interrogative clause (wh-interrogatives) favour null subjects in OI. In main declarative clauses, third person appears to play a role, since third person subjects are mostly null; this effect is not found in interrogatives. There are two major contexts for subject omission in declaratives, which we label paragraph-beginning and coordination.

5 Analysis

In this section we outline our analysis of the facts just introduced. Our main claim is that there are two mechanisms ruling the distribution of null subjects in the two languages: a) subject sharing under coordination,20 and b) matching with a left-peripheral topic or logophoric operator. In addition, the special properties of certain verb forms are explicable in morphological terms.

Both of these mechanisms are available in both OI and OHG. However, we claim that there are slight differences between them as regards the availability of the second mechanism. Though synchronically very minor, these differences form the starting point for the diverging diachronic developments in the two languages: while Italian goes on to become a canonical/consistent pro-drop language, German develops into a language in which pro-drop is not available.

We start by outlining our assumptions about clause structure in OHG and OI. We assume the following cartography of the left periphery (cf. (4) above, from Frascarelli/Hinterhölzl 2007):

\[(38) \quad \text{[\text{ForceP} [\text{ShiftP}[\text{aboutness}]] [\text{ContrP} [\text{FocP} [\text{FamP*} [\text{FinP} [\text{TP} \ldots]]]]]]}\]

On the basis of the robust availability of subject-verb inversion and the clear asymmetry between main and subordinate clauses in terms of verb position, we assume that both OHG and OI are varieties of V2 language (for OHG, Lenerz 1984, Axel 2007; for OI, Benincà 1984, 2006, Ledgeway 2008, 2012, Wolfe 2015, 2018). Following Holmberg (2015: 375, his (77)), we assume the following definition of V2:

\[(39) \quad \begin{array}{ll}
  \text{a.} & \text{A functional head in the left periphery attracts the finite verb.} \\
  \text{b.} & \text{This functional head wants a constituent moved to its specifier position.}
\end{array}\]

\[20\] Even though many of these cases involve null topics, we avoid the term “topic drop” in this section, as it is often taken to refer to the specific type of (limited) argument omission found in present-day standard Germanic languages such as German (see e.g. Trutkowski 2011, 2016).
We adopt the recent proposal by Poletto (2002, 2013, 2014), Roberts (2012), Walkden (2015), Wolfe (2015, 2018), and Cognola (2013, 2019) that variation between languages exhibiting V2 structures can be captured in terms of different landing sites in the C-domain for the finite verb to move to. Specifically, in “strict” V2 languages, the finite verb occupies Force, which drastically restricts what can occur preverbally. In “relaxed” V2 languages, on the other hand, the finite verb is in Fin, which allows for V3, V4 etc. orders if enough left-peripheral specifiers are filled, yet still predicts liberal subject-verb inversion on the assumption that the canonical position of the subject is within the TP-domain. In each case, the head that is the target of verb movement also bears an EPP feature. Following Wolfe (2015, 2018), we can label these two possibilities “Force-V2” and “Fin-V2” languages respectively.

In the case of OHG, examples of V3 are documented (Tomaselli 1995, Axel 2007, Walkden 2014, 2015), but – with the exception of a few examples of XP-pronoun orders in two texts, Isidor and the Monsee Fragments – verb-third orders in main clauses in earlier German involve left-peripheral subordinate clauses (Axel 2002, Axel-Tober 2012), which are syntactically un-integrated. We thus take OHG, at least in the variety attested in the Diatessaron translation, to be a Force-V2 language.

OI, on the other hand, is a good candidate for a Fin-V2 language. It can be shown that the verb moves to a low head position, presumably Fin, within a split CP (see e. g. Benincà 2006 among many others). Moreover, the syntax of topics is very liberal, and V3, V4, and V5 word orders are possible in OI. Benincà (2006) takes this to indicate that the left periphery in OI has the same structure, and that the same positions are available, as in present-day Italian.

5.1 Null subjects in coordination structures

A first context in which null subjects are found in OHG is the one labelled antecedent-linked subject drop in the literature (Volodina/Weiß 2016, Weiß/Volodina 2018) and is subsumed to topic drop, i. e. to deletion of the subject in the sentence-initial position (Ross 1982; Trutkowski 2011, Haider 2010). As discussed in Weiß/Volodina (2018:264f), there are two types of antecedent-linked subject drop: Coordination ellipsis, in which the subject in the first clause precedes the finite verb, and subject gap constructions (see also Heycock/Kroch 1994), in which the subject of the first clause appears in the inversion construction after the finite verb.

In (40) we exemplify the subject-gap construction with examples from Diatessaron. In the case at hand the subject <the shepherds> appears after the finite verb in first sentence and remains silent in the second sentence.

(40a) a. Et factum ut discensserunt ab eis angeli in caelum, pastores loquebantur ad invicem: transeamus usque in Bethleem et videamus hoc verbum quod factum est, quod dominus ostendit nobis. And came then in haste and found.3PL Mariam et Ioseph et infantem positum in presepio Mary and Joseph and child put in crib

21 Cognola (2019) shows that multiple topics are allowed with no ordering restrictions among them, contrary to other relaxed V2 languages where multiple topics are highly restricted.
b. Uuard thô thaz arfuorun fon in thie engila in himil; thô sprachun thie hirta untar in zuisgen: farames zi Bethleem int gisêmes thaz uuort thaz thar gitân ist, thaz truhtin uns araугta.

Inti quamun thô ilente inti fundun and came there in haste and found.3pl
Mariun inti Ioseben inti thaz kind gilegitaz in crippeaz [6,4]
Mary and Joseph and the child laying in the crib
‘< The magi > rushed and they found Mary and Joseph and the child lying in the crib.’

(40b) a. Vedendo i magi la stella, ebbero grandissima allegrezza [7(10)]
Seeing the kings the star, had great happiness
E intrando in casa, trovarono il fanciullo con Maria sua madre; and coming in the house found the child with Mary his mother
ed inchinandosi adorarono lui [7(10)(11)]
and bowing adored him
‘The magi were very happy in seeing the star. And once they entered the house, they found the child with his mother Mary. They bowed and adored him.’

In (41) we have an example in which the subject appears before the finite verb in the first sentence and remains unpronounced in the coordinated clause.

(41) a. Tunc Herodes videns quoniam illusus esset a magis, iratus est valde et mittens occidit omnes pueros ,

b. Thô <Herodes> gisah uuanta her bitrogan uuas fon then magin,
EXPL Herodes saw that he fooled was by the kings,
balg sîh hartô
got.angry REFL.PRON a lot
inti ___ sententi arsîuog alle thie knehta [10,1]
and sent kill all the children [10,1]

c. Allora <Erode> vedendo ch’èra beffato da’ magi ù fu molto adirato;
then Herodes seeing that was fooled by the magi was very angry
e ____ fece uccidere tutti i fanciulli [10(1,2)]
and made kill all the children

Following Volodina/Weiβ (2016) we propose that the coordination ellipsis construction (41) should be analyzed as a case of Topic drop, in which the preverbal subject of the first clause can license a null topic in the following coordinated clause.

For cases of the subject-gap construction (40a, b) we suggest that they are instances of conjunction reduction involving subject sharing (e. g. Heycock/Kroch 1994), as found in non-pro-drop languages such as English and present-day German. For an English clause such as Old Italian has pro-drop and is a V2 language, one possible assumption (following Heycock/Kroch 1994) is that two T’ elements are conjoined such that they share a specifier. Another option is to assume a PF process of conjunction reduction (see e.g. Rögnvaldsson 1982) which deletes the subject of a second conjunct TP under identity. In either case, for OHG and OI, the level of coordination must be higher, as the finite verb is distinct in the two conjuncts: in OHG examples
like (40b) we must be dealing with coordination of at least Force’, and in similar OI examples
the conjuncts can be no smaller than Fin’. Regardless, we assume that whatever strategy is used
to derive the present-day English facts is also active in OHG and OI, and so we set these aside
in what follows.

5.2 The paragraph-beginning context

In the vast majority of our remaining main clause declarative examples, the apparent antecedent
of the null subject is introduced by a fronted (full or reduced) adverbial clause. Intriguingly,
there are no cases in which the fronted clause is non-adverbial (relative, interrogative etc). This
type of null subject is always third person. The (otherwise) new-information subject remains
null in the following main clause; the finite verb appears in linear second position.

(42) a. Audiens autem Herodes rex,
in hearing then king
   turbatus est et omnis Hierusolima cum illo
   upset is and whole Jerusalem with him
b. Thô thaz gihorta Herodes ther cuning,
as this heard.3SG Herod the king
   uuard proj. gitruobit [8,2]
   became upset
 c. Udendo ciò il re Erode
in hearing this the king Herodes
   turbossi proj. e tutta Gerusalem co llui. [7(3)]
   upset.REFL and whole Jerusalem with him

‘On hearing this, Herod was upset and the whole of Jerusalem with him.’

In OI, such examples always involve the presence of an enclitic pronoun following the finite
verb (si in dipartirsi), which is generally agreed (e. g. Benincà 1995, 2006, Poletto 2014) to
show that a topic is present before the finite verb.22

(43) a. Qui cum audissent regem abierunt
   who.REL as heard.3PL king left.3PL
b. Thô <sie> gihortun then cuning, <proj> fuorun [8,5]
as they had.3PL the king left.3PL
c. <Li magi> quando ebbero udito il re, <proj> dipartirsi [7(9)]
   the magi when had.3PL heard the king left.3PL.REFL.CL

“When the three magi heard the king, they left”

We note here an interesting parallel with 13th-century French (Steiner 2015), which will be
relevant for our own analysis. 13th-century French robustly, though not categorically, shows
strict verb-second word order. The main class of exceptions have the verb in third position, and
of these 75% involve a fronted adverbial clause. Compare (44) from Steiner (2015: 23 her (32)).
(On the Old French facts, see also Labelle 2007; Vance et al. 2010; Donaldson 2012; Elsig
2012; Mathieu 2013; Salvesen 2013; Wolfe 2015, 2018; Labelle/Hirschbühler 2018.)

22 Rather than a focus, which would imply obligatory proclisis of the pronoun. See Benincà (1995, 2006), Poletto
(2014), and much other work.
When they had some time been, SI said the king to one of his knights,

‘When they had been some time, the king said to one of his knights…’

(Merlin en prose l. 38.25-27)

We hypothesize that the adverbial clause constituent here is in a very high position in the CP, plausibly “clause-external” in the sense of Broekhuis/Corver (2016). Following e.g. Poletto (2002), Labelle (2007), Mathieu (2013), Wolfe (2015, 2018) and Haegeman/Greco (2018), we propose that this is a FrameP specialized for scene-setting elements, above ForceP. Crucially, as a clause-external projection, these fronted constituents – just like coordinating conjunctions – simply “do not count” for the purposes of verb-second, i.e. they do not satisfy the EPP feature associated with either Fin\(^0\) or Force\(^0\). Since the fronted adverbial clause does not count for V2, an additional position is made available within the CP-domain. In OI we can diagnose this as a specifier of TopicP: the presence of enclisis of the object pronouns indicates that the specifier of FocusP is empty and the specifier of TopicP hosts an XP, as shown in (45) (see Benincà 2006).

\[\text{(45)} \quad [\text{FrameP} \quad \text{[Udendo ciò < il re Erode\(_k\)]} \quad [\text{TOPICP} < \text{il re Erode}\(_k\) > [\text{FocusP} \quad [\text{FinP} \quad \text{turbossi [TP prok turbossi]]}]])\]

In OHG, on the other hand, ForceP is available for a further constituent by virtue of the fact that OHG (like present-day German) was a Force-V2 language. We propose in the structure in (46) that in OHG the new-information nominal subject *Herodes ther cuning* is introduced within an embedded adverbial clause in FrameP (like in OI) and then copied to Spec,ForceP to satisfy the requirements of the V2 rule (the finite verb moves to Force\(^0\)). The DP subject in Spec,ForceP can license a corefential null Topic in TopicP which is vital for the creation of the chain for the licensing of *pro* in the IP area.\(^{23}\)

\[\text{(46)} \quad [\text{FrameP} \quad \text{[Thô thatz gihorta <Herodes ther cuning\(_k\)>] [ForceP <Herodes ther cuning\(_k\)> [Force\(^0\) uuard [TOPICP <Herodes ther cuning\(_k\)> [FinP \quad \text{uard projrituobit]]}]})]

In both languages, the available left-peripheral position is filled by a null topic, which then licenses a null subject in the main clause in the manner suggested by Frascarelli (2007, 2018), as discussed in section 2.3 of this paper.

The relative rarity of null subjects in subordinate clauses in both languages can be accounted for, following Walkden (2014: 213), by assuming, first, that the Agree relation between the left-peripheral topic and null subject is subject to standard locality restrictions such that it cannot apply across the boundary of a finite clause unless there is a TopicP in the embedded clause in

\(^{23}\) One crucial question is why the use of SpecFrameP is so much more restricted in present-day German than in OHG, given that all stages of the language are Force-V2 languages in the Poletto-Wolfe typology. We have no complete answer to this at present, but simply note that V3 is permitted as one grammatical possibility at least in cases of “biscuit” conditionals (Scheffler 2008, Csipak 2015), as in (i). It could thus be the case that the position has always been present in the history of German, with independent pragmatic or prosodic factors interfering with its availability in the modern language.

\[\text{(i)} \quad \text{Wenn du durstig bist, es gibt Bier im Kühlschrank.} \quad \text{if you’re thirsty, there’s beer in the fridge.}\]
which a null topic can potentially be hosted and, secondly, that operators cannot be hosted in all embedded clauses since they are in general structurally deficient (“truncated” in the sense of Grewendorf 2002, Haegeman 2006 and de Cuba 2007, 2014) and do not project a full CP-domain. Null subjects hence cannot be licensed in normal embedded clauses, as no local topic is available for them to Agree with. The prediction then is that null subjects should be able to appear only in those embedded clauses that permit embedded main clause phenomena in the sense of Green (1976) and Aelbrecht, Haegeman/Nye (2012).

Let us consider the data from our corpus in support of this hypothesis. In Table 11 we provide the contexts in which referential null subjects are found in OI (10/46). We see that these contexts share two things: the embedded clauses are i) mostly introduced by that or by if/why and ii) they involve a third person.

<table>
<thead>
<tr>
<th>Introductory element</th>
<th>Type of embedded clause</th>
<th>Person</th>
</tr>
</thead>
<tbody>
<tr>
<td>che (that)</td>
<td>Objective</td>
<td>3sg</td>
</tr>
<tr>
<td>che (that)</td>
<td>Objective</td>
<td>3pl</td>
</tr>
<tr>
<td>che+N (what +N)</td>
<td>Objective</td>
<td>3sg</td>
</tr>
<tr>
<td>che (that)</td>
<td>Objective</td>
<td>3sg</td>
</tr>
<tr>
<td>se (if)</td>
<td>Conditional</td>
<td>3sg</td>
</tr>
<tr>
<td>che (that)</td>
<td>Relative on the object</td>
<td>3pl</td>
</tr>
<tr>
<td>poi che (after that)</td>
<td>Temporal</td>
<td>3p</td>
</tr>
<tr>
<td>secondo il tempo che (according to the time which)</td>
<td>Relative on the object</td>
<td>3sg</td>
</tr>
<tr>
<td>perché (since)</td>
<td>Causal</td>
<td>3pl</td>
</tr>
<tr>
<td>acciò che (so that)</td>
<td>Final</td>
<td>1pl</td>
</tr>
</tbody>
</table>

Table 11: Referential null subjects in embedded clauses in OI

A similar situation is also found in OHG with reduced numbers (3/46).

<table>
<thead>
<tr>
<th>Introductory element</th>
<th>Type of embedded clause</th>
<th>Person</th>
</tr>
</thead>
<tbody>
<tr>
<td>thaz (that)</td>
<td>Objective</td>
<td>3pl</td>
</tr>
<tr>
<td>Bithiu uuanta (since)</td>
<td>Causal</td>
<td>3sg</td>
</tr>
<tr>
<td>Thaz (so that)</td>
<td>Final</td>
<td>1pl (mes)</td>
</tr>
</tbody>
</table>

Table 12: Referential null subjects in embedded clauses in OHG

24 An anonymous reviewer wonders whether the standard locality restrictions and the deficiency of embedded CPs hold for languages like Italian where null subjects occur in embedded clauses with high frequency. As is well known, present-day standard Italian is a language exhibiting an articulated structure of the left periphery in main and in most embedded clauses. Since a referential null subject is always possible even in those adverbial clauses exhibiting a reduced CP structure (Haegeman 2006), a TopicP at least is activated in any embedded clause in Italian.
In both languages referential null subjects can appear (but do not have to) in two types of main clauses: those introduced by *that* and those introduced by *why*. Following work on the fine structure of the left periphery, we propose that none of the introductory elements is hosted in the lower portion of CP, but *that* lexicalizes a high FP within the periphery (presumably Force) and *if/why* are hosted in InterrogativeP (Rizzi 2001). Below both FPs hosting *that* and *if/why* a TopicP is found.

(47)  
[ForceP *that* [TopicP Null Topic [InterrP *if/why* [TopicP [FocusP [FinP [TP]]]]]]]

Based on the structure in (47) we propose that referential null subjects can potentially only be licensed in sentences in which the element introducing the embedded clause is hosted either in ForceP or in InterrP and it is able to activate the TopicPs of the left periphery (see Haegeman 2002 for the idea that the presence of the projections TopP and FocusP in root clauses and peripheral adverbial clauses leads to the activation of ForceP (Haegeman 2006)). Since TopicPs are activated, a Topic chain with a preceding topic can be established. In (48) we illustrate this with an example from Italian:

(48)  
E poi che furono compiti gli otto dì, 
accì che si cicuncidesse il fanciullo,
and after that were gone eight days, 
so that IMP.PRON circumcise the child
fu chiamato il nome suo <Gesù>, 
il quale nome era appellato dall'angelo
was called the name his Jesus, 
whose name was mentioned by the angel
*prima che <pro>* fosse conceputo in ventre. [7(1)]
before that were conceived in womb
‘And after eight days had gone, so that the child could be circumcise, he was called Jesus, a name which was mentioned by the angel before he was conceived in his mother’s womb.’

Current topic: Gesù

(49)  
[ForceP *prima che* [TopicP <Null Topic: Gesù> [FinP fosse [TP *pro*]]]]

We assume that in the configuration in (49) referential null subjects can potentially be licensed, but this does not obligatorily take place, for instance due to discourse factors (a subject must be overt because it is focussed). For other clause types, we assume that null subjects can never be licensed because the element introducing the embedded clause sits in FinP and does not activate the TopicPs in the left periphery, i. e. the structure of other clause types is deficient in not activating the TopicP of the left periphery.

We put forth that the mechanism at the basis of the licensing of null referential subjects in embedded clauses is identical in OHG and OI; note, however, that the numbers differ in the two languages. In OHG, in fact, null referential subjects appear in 4/46 embedded clauses (8.6%) whereas in OI they appear in 10/46 (21%). These numbers point to the scarcity of null subjects in embedded clauses in OHG and confirm the data discussed in Weiß/Volodina (2018:278) according to which out of 247 *that*-clauses appearing in the 38 texts collected in Köbler’s (1986) *Sammlung kleineren althochdeutscher Sprachdenkmäler*, a null subject is only found in 8 cases (3.24%). Weiß/Volodina (2018) account for these numbers by assuming that embedded *pro* is indeed a relic from Indo-European – a hypothesis which is in principle not in contrast with our analysis. Crucial for us here is that, despite residual, the distribution of null subjects in OHG embedded clauses appears to be licensed by the same mechanism of OI.
5.3 Interrogative clauses

The analysis developed above for main declarative clauses can be applied as such to yes/no-interrogatives involving a third person null subject. In interrogative clauses, null third person subjects appear in a minority of cases (10 in OI and 6 in OHG), but we think it is not a coincidence that all these examples involve yes/no-interrogative clauses, i.e. there are no third person null subjects in either language.

(50) a. Aut quis est ex vobis homo, quem si petierit filius or who is of you.PL man, who if asks son suus panem, his bread numquid lapidem porrigit ei, aut si piscem instead stone gives him, or if fish petit, numquid asks, instead serpentem porrigit ei, aut si ovum petierit, numquid porrigit snake gives him, or if egg asks, instead gives illi scorpionem?

b. Odo uuer ist fon iu manno, then oba bitit sin sun brotes, or who is of you.PL man who if asks his son bread ia ni gibit her imo stein? Oba her fiskes bitit IA NEG gives he to him stone if he Fish asks ia ni gibit her imo thanne natrun? IA NEG gives he him then serpent Odo oba her eies bitit, ia ni gibit imo thanne or if he eggs asks IA NEG gives him then scorpionem? [40,6]

c. Qual’è quell’ uomo di voi che se’l figliuolo which is that man of you.PL that if-the small son gli chiede pane, him asks bread che gli dia pietre? O vero s’egli that him gives.CONJ stones or true if.he domanda pesce fish <topic/egli> gli dia serpente? O vero him gives.CONJ snake or true s’egli domanda uovo <topic/egli> gli dia if.he asks egg him gives.CONJ scorpione? [41, 229, 20-24]

‘Who of you would give their sons stones when they ask for bread? Or would give them snake when they ask for fish? Or scorpion when they ask for eggs?’
This fact is fully expected within the proposed analysis: only when the relevant left-peripheral specifier position (either Spec,ForceP for OHG or Spec,Fin for OI) does not contain overt material can a null topic be licensed.25

The mechanism developed for third person null subjects in the previous section cannot, however, be applied to the first and second persons, which are most frequently null in interrogative clauses. We thus propose, following Sigurðsson (2011) and Frascarelli (2018), that logophoric operators, agent (Λ^A) or patient (Λ^P), play the crucial role in licensing the null subject in interrogative clauses: see section 2.3.

In OHG, 9 of 17 examples of first person plural subjects in interrogative clauses are null. This, however, correlates strikingly with the use of the long ending -mēs (see Table 13): all null subjects in the first person plural occur with this ending, rather than with the shorter -n ending, which confirms the observation in Axel (2005, 2007). In traditional analyses -mēs is taken to “show[s] evidence of being synchronically pronominal” in the Diatessaron translation (Somers et al. 2018: 243), following a long tradition of research (e. g. Kuhn 1869).26

Setting the first person plural aside, in OHG only the first and second person singular allow null subjects to a certain degree (39% and 42% respectively). We suggest that the special properties of these persons may relate to their morphological expression. Table 13 gives an overview of the verbal endings in OHG.

<table>
<thead>
<tr>
<th>Person</th>
<th>Number</th>
<th>ziohan ‘to pull’</th>
<th>salbōn ‘to anoint’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sg</td>
<td>ziuhu</td>
<td>salbōn, salbōn</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>ziuhis(-t)</td>
<td>salbōs(-t)</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>ziuhit</td>
<td>salbōt</td>
</tr>
<tr>
<td>1</td>
<td>Pl</td>
<td>ziohemēs, ziohen</td>
<td>salbōmēs, salbōn</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>ziohet</td>
<td>salbōt</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>ziohent</td>
<td>salbōnt</td>
</tr>
</tbody>
</table>

Table 13: Verbal endings in the OHG Tatian (cf. Axel 2007: 316)

As can be seen from Table 13, there are two possible endings for regular present tense verbs in the second person singular: -s and -st. A possible hypothesis at this point is that the -t ending is a clitic pronoun, like -mēs; this is suggested (at least as a diachronic origin) by Braune/Reiffenstein (2004) and Somers (2011). However, the data in Table 14 show that as a synchronic analysis this does not help us. Table 14 shows the distribution of null and overt subjects with these two endings in interrogative clauses. We also have a separate column for the enclitic -tu, which never occurs with an overt subject and so can be said (like -mēs) to be a true pronominal clitic.

---

25 If this is correct, then, contra one popular analysis following Grimshaw (1993) and Roberts (1993) for modern English, there can be no null operator in SpecCP in V1 interrogatives at least in Old High German. Since modern German disallows topic drop in V1 interrogatives, this could be a parametric difference (in the sense of lexical variation) between the two stages of the language.

26 For our purposes it does not matter whether -mēs is an inherited pronominal clitic, as originally proposed by Kuhn (1869) and supported by Somers et al. (2018), or was rather reanalysed as pronominal during the OHG period; we take no stance on this issue. What is important is that in OHG this ending was synchronically pronominal.
Table 14: Distribution of 2SG null/overt/clitic subjects in OHG interrogatives

<table>
<thead>
<tr>
<th></th>
<th>Null</th>
<th>Overt</th>
<th>Clitic -tu</th>
</tr>
</thead>
<tbody>
<tr>
<td>-s</td>
<td>24</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>-st</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Other (irregular)²⁷</td>
<td>2</td>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

What emerges as a generalization from Table 14 is a) that the ending -st is very rare overall in this text in all contexts, and b) that the ending -s favours null subjects, whereas with -st null subjects are no more common than in other persons (though the effect is far from categorical); this is the opposite of what we would predict if -t were synchronically a pronominal clitic. Another analysis is needed.

We appeal to the morphological distinctness of the different endings in Table 13 as an explanation for the distribution.²⁸ Starting with -t, this ending is found in various different persons and numbers, including 3rd singular and 2nd and 3rd plural; thus, these person/number combinations are not distinct enough to license null subjects. The -n found in the 1st plural is also not distinctive: at least in the salbôn-paradigm, this ending is found in the 1st singular too. That leaves the distinctive -u and -ôm, which are only found in the 1st singular, and -s(t), which is only found in the 2nd singular. In OHG, these unique exponents are able to license null subjects. In OI, on the other hand, all endings are distinct: therefore, we predict that any person and any number can exhibit a null subject, and this is indeed what we find (see Table 7).

Formally, we cash this idea out as follows. The relation between the logophoric operators agent (Λ_A) and patient (Λ_P) and pro is mediated by verbal endings. Forms with distinct endings bear a strong set of phi-features, which allows them to agree with the operator and form a chain with it (Frascarelli 2018) even in the case in which an intervening wh-element is present in the left periphery.

For interrogative clauses we thus propose that the role of distinct/rich morphology is that of making a relation between pro and a logophoric operator possible in cases in which CP hosts another constituent, i.e. the wh-element (note that when the left periphery is occupied, the licensing of pro is typically blocked; see the discussion of embedded clauses above). Therefore, rich/distinct morphology function as a sort of “repair strategy” for cases in which pro could not be otherwise licensed.

We assume that, in both languages, morphology does not play any relevant role in the case in which a Topic is in the left periphery with which a chain can be established by pro, like in the licensing of null subjects in main declarative clauses (sections 5.1 and 5.2 above). For that clause type, third person null subjects are the most frequent cases of null subjects even though we have shown that third person morphology is not distinct/rich (at least in OHG). We thus propose that in main clauses the mechanism of pro licensing crucially relies on the availability of a topic in CP with which pro can agree independently of morphological agreement.

²⁷ Under “Other” we count all second person singular verb forms not ending in -s or -st. This category includes strong verbs in the past tense, preterite-presents, and subjunctive verbs.

Summing up, we have proposed that in both OHG and OI pro licensing takes place through an Agree relation between the left-peripheral topic and the null subject, independently of the agreement morphology on the finite verb. As the restrictions widely attested in the distribution of pro in embedded clauses in both languages show, the Agree relation between elements in the left periphery and pro is subject to standard locality restrictions such that it cannot apply across the boundary of a finite clause. We claim that the distribution of null subjects in interrogative clauses (second person singular most frequent in OHG) emerges from a type of locality restriction caused by the presence of an operator in the left periphery which interferes with the Agree relation between the logophoric operator and pro. We suggested that the intervening operator can be circumvented only in those persons exhibiting rich morphology, i.e. rich morphology can be seen as repairing strategy to establish an agree relation in cases in which this should not be possible.

What about OI? We have said that this language exhibits distinct/rich agreement in all persons, unlike OHG – this is why null subjects are possible in all persons in interrogative clauses. However, rich/distinct morphology does not appear to play a key role in the case of main declarative clauses (in which pro is licensed through Agree with Topic), where the distribution of null subjects in OI is comparable to that of null subjects in OHG. We thus propose that null subjects are licensed in the same way in both languages in both declarative and wh-interrogative clauses, with the only difference that in OI null subjects are possible in all persons in wh-interrogatives due to rich morphology.

Now, the asymmetry in the distribution of pro across different persons in wh-interrogative clauses is the only relevant difference between OHG and OI, and possibly the origin of their different developments. More specifically, we tentatively propose that the “repair strategy” to license pro available in wh-interrogatives (i.e. licensing through morphology in both languages) was then generalized to main declarative clauses in OI (due to the availability of rich/distinct morphology in the language) but not in OHG.

6 Conclusions and outlook

In this paper we have explored the nuances of null subject licensing in OHG and OI, with a particular focus on the role played by clause type. Descriptively, we have shown that this – along with person – is a major factor influencing the possibility of null subjects in both languages. The two languages behave similarly in this respect, though are not identical, and we have shown that the differences between them can be attributed in part to the different structures of the left periphery in the two languages: while OHG is a Force-V2 language, OI is a Fin-V2 language.

The strongest predictor of whether a subject will be null or overt in both languages is the presence of an overt subject in the Latin original: in such cases the subject is virtually always overt in both OHG and OI. In cases where there is a null subject in the Latin, there is variation: main declarative clauses have null subjects (in paragraph-beginning and coordination contexts), embedded clauses almost never have them, and interrogatives behave more variably. Third person null subjects are only found in yes-no interrogatives, not in wh-interrogatives. In addition, person and morphology affect whether subjects are overt or unexpressed in interrogatives.
Our analysis ties the availability of null subjects to an Agree relation with an appropriate operator in the left periphery, following Frascarelli (2007, 2018) and Sigurðsson (2011). Crucially, since interrogatives exhibit V-to-C across the board in both languages, a V-in-C-licensing account of null subjects such as that of Adams (1987) cannot be upheld for our dataset. Another crucial aspect of our analysis is that multiple distinct strategies for the licensing of null subjects may be operative in one and the same language; see also Rezac (2017), who reaches the same conclusion for Old Icelandic. In this respect we have come a long way from the days of a single null subject parameter, globally governing the availability of referential pro. But what we have lost in theoretical parsimony we have gained in empirical coverage and in understanding of the principled discourse basis of phenomena previously thought to be irreducibly syntactic.

References

Primary sources


Secondary sources

Aboh, Enoch (2010): “Information structuring begins with the numeration”. Iberia 2: 12–42.


Ledgeway, Adam (2007): “Old Neapolitan word order: some initial observations”. In: Lepschy, Anna Laura/Tosi, Arturo (eds.): Histories and Dictionaries of the Languages of Italy. Rovenna, Longo: 121–149.


Müller, Gereon (2005): “Pro-drop and impoverishment”. In: Brandt, Patrick/Fuß, Eric (eds.): *Form, structure and grammar: a festschrift presented to Günther Grewendorf on the occasion of his 60th birthday*. Tübingen, Narr: 93–115.


Rezac, Milan (2017): *Argumental null subjects in Old Icelandic*. Ms., IKER/CNRS.


ISSN 1615-3014


