

# **Register Variation and Anglo-Norman Syntax in MS Harley 2253: Null Subjects, Generic Subjects, and V2 Word Order**

Matthew Maddox

DRAFT - Please do not circulate

## **0. Introduction**

Of the three languages that constitute the textual material in MS Harley 2253, Anglo-Norman (AN) is the most representative and the most understudied, at least in the field of linguistics. While a growing body of work has been produced over the years, it has focused primarily on questions related to the external history and use of the language (Tyler 2011). Meanwhile, little work has been done that addresses the internal linguistic system, or the "language faculty," as it is known in generative linguistics. This study is an attempt to examine how the external (performance) interacts with the internal (competence<sup>1</sup>).

More specifically, the issue being investigated is the following: to what extent are internal syntactic features of AN texts in MS Harley 2253 conditioned by external factors such as register choice? The syntactic features under consideration are null subjects, generic subject expression, and V2 word order, all of which will be further explained in Section 2. Register, also referred to as genre or style, can be defined as a variety of language "associated with a particular configuration of situational characteristics and purposes" (Biber & Conrad 2001:175). As Romaine (1982:117) points out, a register is essentially a dialect. Just as we study the variation of linguistic features from dialect-to-dialect, we can study variation among different registers as well. In this study I investigate how the aforementioned linguistic features vary in two registers: AN poetry and prose. I hypothesize that choice of register will be a significant factor in the

---

<sup>1</sup> Since Chomsky (1965:4), it has been customary to distinguish between these two. Competence is a speaker's innate knowledge of language, Universal Grammar. Performance is actual use, which may be subject to external factors such as social roles, class, gender, etc.

licensing of null subjects, V2 word order, and generic subjects. I predict that there will be a higher rate of null subjects and a lower rate of V2 word order in poetry due to this register's need to be economical and its ability to be more flexible syntactically. Since generic subject expression is related to whether a language can or cannot have null subjects (Belletti 1982; Holmberg 2005, 2010), I also predict that a pattern might emerge in this respect as well.

The amount of work that has been done on the syntactic features being studied here is considerable, at least within the generative framework. Many have examined how French changed diachronically from allowing null subjects in main clauses, as in Old French, to disallowing null subjects everywhere, as in Modern French, and relating it to V2 order (Adams 1987; Vance 1997). However, only one study, Ingham (2012) has looked at these phenomena in AN. None of these has approached the topic from the perspective of register variation. There are studies of this kind that have investigated null subjects (Haegeman & Ihsane 2001; Haegeman 2013), but they are focused on diary dialects in the Modern English and French. The findings of this current study are intended to contribute to this field of inquiry by expanding the body of data, looking at medieval rather than modern language, and to enhance what we know not only about the language of MS Harley 2253, but about its textual content as well.

The format of this paper is as follows: in Section 1 I present the materials that make up the corpus and the type of methodology being implemented; in Section 2, the phenomena being investigated are explained thoroughly, though with as little theoretical background as possible; in Section 3, I present and analyze the results of the investigation; in Section 4, I conclude and remark upon possible implications for what we know about these linguistic features and discuss what the next step may be in this ongoing project.

## 1. Methodology and Corpus

The corpus is comprised of three prose texts and three poetic texts, as shown in Table 1:

Article Number	Title	Theme	Register
Article 8	<i>ABC a femmes</i>	Religious	Poetry
Article 20	<i>Quant voy la revenue d'hyver</i>	Secular	Poetry
Article 24	<i>Chaunter m'estoit</i>	Political	Poetry
Article 38	<i>Les pelerinages communes que crestien fount en la Seinte Terre</i>	Religious	Prose
Article 91	<i>La destinccioun de la estature Jesu Crist Nostre Seigneur</i>	Religious	Prose
Article 95	<i>L'enqueste que la patriarche de Jerusalem fist</i>	Religious	Prose

Table 1: Corpus

Each of the individual texts has certain traits that may be relevant with respect to the linguistic features being studied. Article 8, for example, is an acrostic poem in which the first word of each stanza, after the first two introductory stanzas, is taken from the order of the alphabet. Additionally, R. Dean (1999) suggests that it was written to be sung and accompanied by a stringed instrument. How these factors correlate with null subjects, for example, may not be immediately clear, but it is nevertheless illustrative of the fact that this is not spontaneous language but rather a constructed piece of literature. Given that this type of textual information may help inform the linguistic analysis, it will be brought up when and if relevant in Section 3.

Verb tokens were extracted from the corpus and coded for the following characteristics: 1) null or overt subject, 2) V2 or non-V2 position in a matrix clause. Given that this is a pilot study, I did not code for other factors generally taken into account in subject expression studies such as verb type, person/number morphology, tense/mood/aspect, etc. Regarding V2, only matrix verbs were extracted based on the observation in Ingham (2012) that AN allowed V2 primarily in matrix clauses. It should be pointed out that AN is also thought, like continental Medieval French, to allow null subjects primarily in matrix clauses as well. However, when

coding for null and overt subjects, I extracted all verbs from all clauses. This presents a greater picture of the overall phenomenon of null subjects, rather than restricting the study to null subjects in matrix clauses.

After extracting and coding, a descriptive statistical analysis was done in order to determine the frequency of null subjects and V2 order. Given the low amount of tokens of generic subjects, no statistical analysis was necessary for this feature. Instances of generic subjects were analyzed qualitatively, within their individual clausal contexts. This study is based on the generative model of linguistic theory, specifically the Minimalist Program (Chomsky 1995). The theoretical terminology will be kept to a minimum.

## **2. Null Subjects, V2 Word Order, and Generic Subject Expression**

### 2.1 Null Subjects

The observation that some languages allow a phonetically unrealized (null) referential subjects in finite clauses goes back in the generative literature to Perlmutter (1971). (1) is Spanish, a consistent null subject language (cNSL), where a tensed sentence is grammatical both with and without the overt subject. In French, a non-null subject language (nNSL), an overt subject must be expressed (2a) or the sentence will be ungrammatical (2b).

- |     |                        |     |                         |
|-----|------------------------|-----|-------------------------|
| (1) | a. Juan habla español. | (2) | a. Jean parle français. |
|     | b. Habla español.      |     | b. *Parle français.     |

The intuitive explanation for why certain languages allow a null subject is that they tend to have "rich" morphology, so that the subject can be identified by correlation with verbal endings. This explanation is a traditional one, and was stated as early as the second-century C.E. by Apollonius Dyscolus in reference to Ancient Greek (Roberts & Holmberg 2010). Within generative linguistics, this explanation continues to guide the theoretical modelling of what is referred to as the Null Subject Parameter, a formal property the language faculty which is "set" during native

language acquisition. If the parameter is set positive, the language allows null subjects; if negative, null subjects are disallowed.

Of course, the story is much more complicated than this and the issue is continuing to be studied and debated. One of the complications is that there are certain languages that seem to be in-between cNSLs and nNSLs, known as partial null-subject languages (pNSL). These are languages where null subjects are allowed freely in some contexts but not in others. For example, Finnish allows null first- and second-person referential subjects, but not third-person referential null subjects. If a third-person form has no overt subject in this language, it is interpreted as referring to a "generic" subject. In (3) and (4) there is no overt pronoun equivalent to English "one;" the verb is without any overt subject and thus receives a generic interpretation.

- |     |   |     |   |
|-----|---|-----|---|
| (3) | Tässä istuu mukavasti<br>here sit.3S comfortably<br>"One can sit comfortably here." | (4) | Opettajana odottaisi vähän kunnioitusta.<br>teacher.ESS expect.3S some respect<br>"As a teacher one would expect a bit of respect." |
|-----|---|-----|---|

According to Holmberg (2005, 2010), this is characteristic of pNSLs; a point that will be elaborated upon in the Section 2.3<sup>2</sup>.

An additional complicating point is that even in nNSLs such as English and French, subjects can be null in certain pragmatic contexts:

- (5) Cried yesterday morning: as if it were an hour for keening: why is crying so pleasureable? (Plath; taken from Haegeman & Ihsane 2001)

How do we explain why Sylvia Plath can omit subject pronouns in her diary, but in most other contexts we, as English-speakers, can not do the same? Most generative analyses of null subjects attribute its licensing to the presence of a specific type of formal feature on a functional head (Rizzi 1982, 1986; Alexiadiou & Anagnostopoulou 1998; Roberts 2010). The abstruse

---

<sup>2</sup> This generalization, known as Holmberg's Generalization, has actually been questioned by Sigurðsson & Egerland (2010). Furthermore, data that I have obtained through my own research supports this criticism. However, the validity of the generalization is not important for the purposes of this paper.

theoretical details are not important here, but rather the question that examples like (5) raise; namely, if what licenses null subjects is "rich" morphology, then how do we explain null subjects in English, a language with very "poor" morphology?

The most obvious explanation is that when null subjects occur in nNSLs it is not that they are licensed by the grammatical system; rather what happens is that pragmatics overrides syntax in these cases. In other words, the "rules" can be broken when language is being used in a certain domain; i.e. registers. For example, in English and French, null subjects can occur not only in diary dialects (Haegeman & Ihsane 2001; Haegeman 2013), but also in text messages, status updates on social media websites, etc. This immediately raises others questions. Are null subjects in these registers completely unregulated by grammar? If you can have null subjects in these registers, why not in everyday spoken English? The answer to the first question appears to be negative. Haegeman (2013) demonstrated that null subjects are almost exclusively limited to matrix clauses in these diary dialects. Thus, while it may seem that syntax is being overridden, it cannot be completely ignored.

Interestingly, the previously mentioned pattern of null subjects in diary English also occurs in Old French (6): null subjects tend to occur in matrix rather than embedded clauses. However, this property was being lost in both continental (7) and Anglo-Norman (8) French during the period which concerns us, the fourteenth century:

(6) Tresqu'en la mer cunquist la tere altaigne.  
 until the sea conquered.3S the land high  
 "He conquered the high land all the way to the sea."

(7) Et ly direz que je me recommande humblement a elle.  
 and her will-say-2P that I myself recommend.1S humbly to her  
 "And you will say to her that I humbly ask her good will."

- (8) Quant voy la revenue d'yver qe si me argue qe ly temps se  
 when see.1S the return of-winter that so me afflicts as the weather SE  
 remue lors aym buche fendue.  
 changes then love.1S log split  
 "When I see the return of winter that so afflicts me, as the weather  
 changes, then I love a split log." (Harley 2253, art. 20:1-4)

Why did French lose the ability to have null subjects? Was it due to contact with English? This question is not the primary concern here, but Ingham (2012) concludes that non-contact French and AN lost null subjects at the same rate over the same period of time. If it was not due to contact then the most obvious explanation would be that phonological erosion of verbal endings forced speakers to express overt subjects so that the referent would be understood. Even if this is determined to be the ultimate source of loss of null subjects in French, it very well may not be, what is important for this study is that a change was going on in AN during the period in which the Harley manuscript was compiled. Consequently, we can expect to see a degree of variation in null subject expression in this material. If we do, what can it tell us about the process of loss, or about AN grammar in general? An analysis of this variation might also serve as an additional clue as to when each piece was composed, though this would be outside the scope of this study.

## 2.2 V2 Word Order

Since Greenberg (1963), the order of subject, verb, and object within sentences has been seen as a characteristic that can be used to classify the world's languages. For example, Japanese and Korean are known as verb-final (also head-final) languages because the verb almost always appears in the last constituent position of the sentence; i.e., they are SOV languages. Most modern Romance languages are SVO, like Spanish and French. There also exist VSO languages like Welsh and Irish, or Arabic and Hebrew. Of course, these are just the patterns most characteristic of these languages; a certain flexibility may be allowed. This flexibility varies

from language-to-language, but it may also vary language-internally. For example, Spanish is generally an SVO language, but in questions (9b) the pattern is VSO:

- (9) a. Juan corre por el parque.  
b. ¿Corre Juan por el parque?

Thus, as a rule, Spanish allows both orders, though within formal linguistics (9b) is analyzed as being derived from (9a) via what used to be called "transformation." In other words, there is a universal word order determined by Universal Grammar, which is assumed to be SVO.

The type of word order of concern for this study is called Verb-Second order because, regardless of where the subject or object may be, a finite verb must always be in the second constituent position. To be clear, a constituent does not refer to a single word but rather to a whole unit of syntactic structure such as a Prepositional Phrase, Adverb Phrase, Noun Phrase, etc., all of which are single units (a "constituent," sometimes called "syntagms" in earlier structuralist linguistics) but may be composed of more than one word. Thus, in rigidly V2 languages, which may be rare, orders such as SVO and OVS may both occur, but not SOV. German (10) is a classic example of a V2 language, though it still has verb-final position in embedded clauses.

- (10) Ich las schon letztes Jahr diesen Roman.  
I read.PAST already last year this novel  
"I read this novel last year already."

Old French (11), continental Middle French (12), and AN were V2 matrix languages; V2 was the preferred pattern in matrix clauses:

- (11) Par Petit Pont sont en Paris entré.  
by Petit Pont be.PRES.3P in Paris enter.PART  
"They entered Paris by the Petit Pont."  
  
(12) Et aussi fis je de par vous.  
and also make.Past.1S I from by you  
"And I did likewise with respect to you."

- (13) Chescun devereit mettre cure pur l'amour de une dame.  
 each should.3S place.INF care for the-love of a woman  
 "Each of us ought to be careful on account of the love of a lady."  
 (Harley 2253, Art. 8, line 26-27)

This is surprising because Old French is descended from Latin, a language that is generally verb-final, SOV<sup>3</sup>. How did French come to acquire a property not already present in its ancestor? Some have argued that it was due to contact from Germanic languages, though the issue is still being debated.

Whatever its ultimate origin may be, V2 order was being lost in the fourteenth century in both continental and AN French (Ingham 2012). Just as the allowance of null subjects is thought of as an abstract property of Universal Grammar, the Null Subject Parameter, so is V2 order. In fact, some authors consider the "operation" that creates V2, T-to-C movement, also to be a parameter (Roberts 2007). Both parameters were subject to variation in AN at this time. Can this variation also be detected in the AN material of the MS Harley 2253? As with null subjects, an analysis of V2 order in the Harley manuscript can contribute to what we know about how and when this property was lost in the history of French.

### 2.3 Generic Subject Expression

A "generic" subject is one that refers to a non-restricted, [+human], antecedent (Sigurðsson & Egerland 2009:161). It is equivalent to English pronouns *you* (14a), *they* (14b), or *one* (14c) when referring to "people in general:"

- (14) a. You have to work hard in life if you want to be successful.  
 b. They always kick you when you fall down.  
 c. One must work until the age of 65.

Belletti (1982) made the observation that in cNSLs, a generic subject is expressed with the reflexive clitic, a so-called impersonal clitic construction, as in Spanish (15a) and Italian (15b):

---

<sup>3</sup> Though it has been analyzed within contemporary linguistics as a "discourse-configurational language," in which there is no majority word-order pattern (Devine & Stephens 2006).

- (15) a. Aquí no se puede fumar.  
 here NEG SE can.3S smoke.Inf  
 "One cannot smoke here."  
 b. Spesso si e traditi dai falsi amici.  
 often SI is.3S betrayed.Part by false friends  
 "One is often betrayed by false friends."

In some languages, like Brazilian Portuguese, a generic subject can be expressed by a bare verb with no overt subject (16a). In a cNSL like Spanish, the null subject is usually interpreted as referential, not generic (16b).

- (16) a. Aqui não pode nadar.  
 here NEG can.3S swim.Inf  
 "One can't swim here."  
 b. Aquí no puede nadar.  
 here NEG can.3S swim.Inf  
 "He/she/it can't swim here."

Holmberg (2005, 2010) claims that the ability to express a null generic subject is a characteristic of partial null subject languages like Finnish, Marathi, Hebrew, Brazilian Portuguese. He connects it with the setting of the Null Subject Parameter in these languages. In other words, we can use the presence or absence of a null generic subject in a language as a typological tool, in the same way that we use the presence or absence of a null referential subject. Some languages will allow it; others will not. Crucially, according to Holmberg, whether a language allows a null generic subject correlates with whether it allows a null referential subject. Languages will allow one or the other, but not both; i.e. Holmberg's Generalization<sup>4</sup>.

How does this relate to AN poetry and prose in MS Harley 2253? Given that AN and Middle French were undergoing change with respect to the extent to which null subjects were allowed, we can predict, if Holmberg's Generalization holds, that variation will also occur for null generic subjects. AN could be classified as a pNSL because its null subjects tend to occur in

---

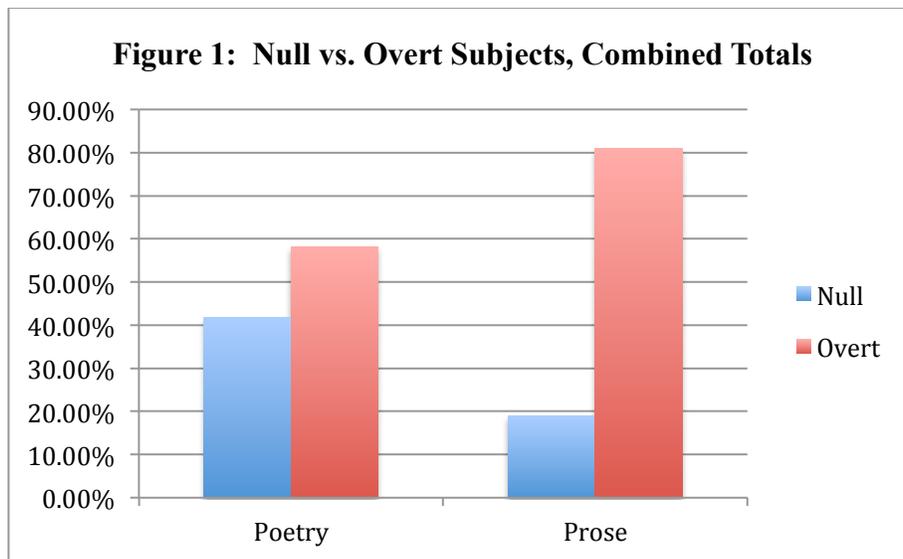
<sup>4</sup> There is much more that could be said here about null generic subjects, but given the intended readership of this paper, much has been omitted. It is hoped that a this brief explanation is sufficient to give the reader a basic understanding of the phenomenon.

matrix clauses rather than embedded clauses. Consequently, like Finnish and Brazilian Portuguese, AN should allow null generic subjects. At this point, there is only one prediction we can formulate about null generic subjects in AN: they should be allowed. Whether this prediction is borne-out by the AN data in MS Harley 2253 will be discussed in Section 3.3.

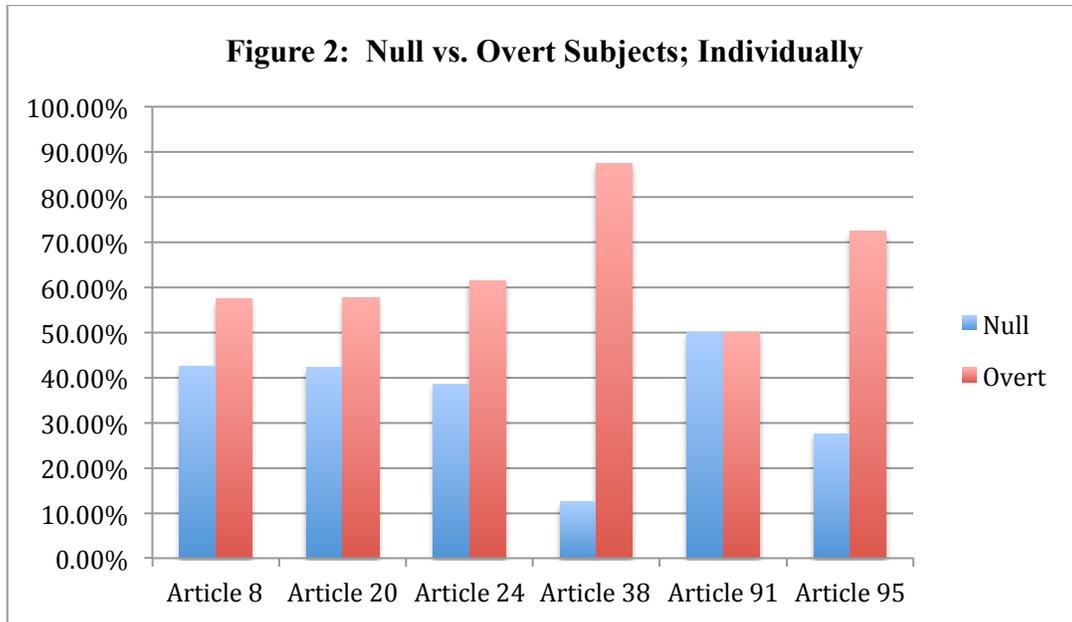
### 3. Results and Discussion

#### 3.1 Null Subjects

A combined total of 801 tokens were extracted and coded for null or overt subject: 240 had null subjects, 561 had overt subjects. When register is taken into account, null subjects are more frequent in poetry than they are in prose. 41.8% of all the subjects in the poetry are null compared to 19% null subjects in prose, as shown below in Figure 1:



In both poetry and prose, overt subjects are still more frequent than are null subjects, but the difference is greatly reduced in poetry. This suggests that in poetry there are less restrictions on when and where a null subject is licensed, in the same way that diary English is less restrictive than spoken English. An even more interesting picture emerges when we look at each Article individually:



In Figure 2, the frequency of null and overt subjects is given with poetry on the left (Articles 8 to 24) and prose on the right (Articles 38-95). While in poetry the rate of null subjects is more or less consistent, ranging from 42.6% to 38.6%, the prose data show a greater range of variation.

For example, Article 91 shows 50/50 ratio of null and overt subject expression. In other words, this text has the same amount of null subjects as it does overt subjects. The total number of tokens for Article 91 is twenty. In fact, it is the shortest of all six selections. Perhaps text-length is a conditioning factor, a register-related factor, on the presence of null subjects. Why this should be the case is not clear; further investigation is required.

Another factor that might be contributing to these results is that, according to Fein's explanatory notes, Article 91 is that it is based off of an apocryphal letter written in Latin, which may have originally been composed in Greek (Volume 2, page 337). Depending upon how closely the putative translator was following the language from which he was translating, it is quite possible that this affected the rate of null subject expression. In order to answer this

question a comparison with extant versions of the "Letulus Letter" would need to be carried, a task which will be left for future work.

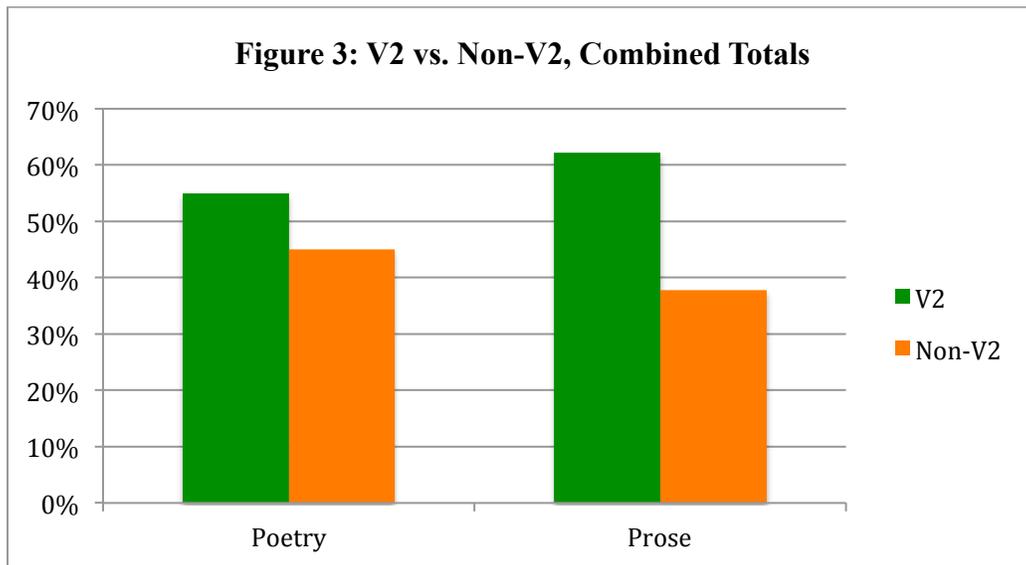
One final observation on Article 91 is that it is a third-person narrated description of the physical appearance of Jesus Christ. Naturally many of the verbs are going to have the same referent. In fact, 9/20 refer to Jesus Christ; 7/9 have a null subject. This is not surprising given that coreferentiality is one of the strongest statistically significant conditioning factors for null subjects in cNSLs (Bayley & Pease-Alvarez, 1996; Bayley et al., 2012; Claes, 2011; Holmquist, 2012; Matos Amaral & Schwenter, 2005; Orozco & Guy, 2008; Otheguy et al., 2007; Silva-Corvalán, 1994). Nevertheless, this is not an effect that can be attributed to register, given that coreferentiality is a language-internal factor.

The other two prose selections, have rather low rates of null subject expression, compared with the other material. Article 38 is 12.5% null; Article 95 is 27.5% null. Compared with the rest of the corpus these two selections have the lowest rate of null subjects; that is, they reflect a greater degree of preference for overt subjects. Why should this be the case? What do these two texts have in common that would explain the low rate of null subjects? There are a couple of possibilities. First, both are prose and, as mentioned above, prose is a closer approximation of spontaneous spoken language than is poetry. In prose the writer is likely less concerned about syllable count and space restrictions, so there is no need to drop subjects. Second, both selections are essentially lists. Article 38 is a travel guide that lists all of the essential tourist attractions when going on pilgrimage to the Holy Land. Article 95 is very similar, as Fein observes in her explanatory notes (Volume 2, page 341). If these works were composed by the same writer, perhaps he spoke a variety of AN in which the change from pNSL

to nNSL was nearly complete. On the other hand, low null subject rate could be a feature that characterizes list language in general; i.e., list dialects.

### 3.2 V2 Word Order

A total of 481 tokens were extracted from main clauses only. V2 is the majority pattern with 59% of verbs being in V2 position. This same pattern is reflected more-or-less consistently when the data is divided into poetry and prose as seen below in Figure 3.

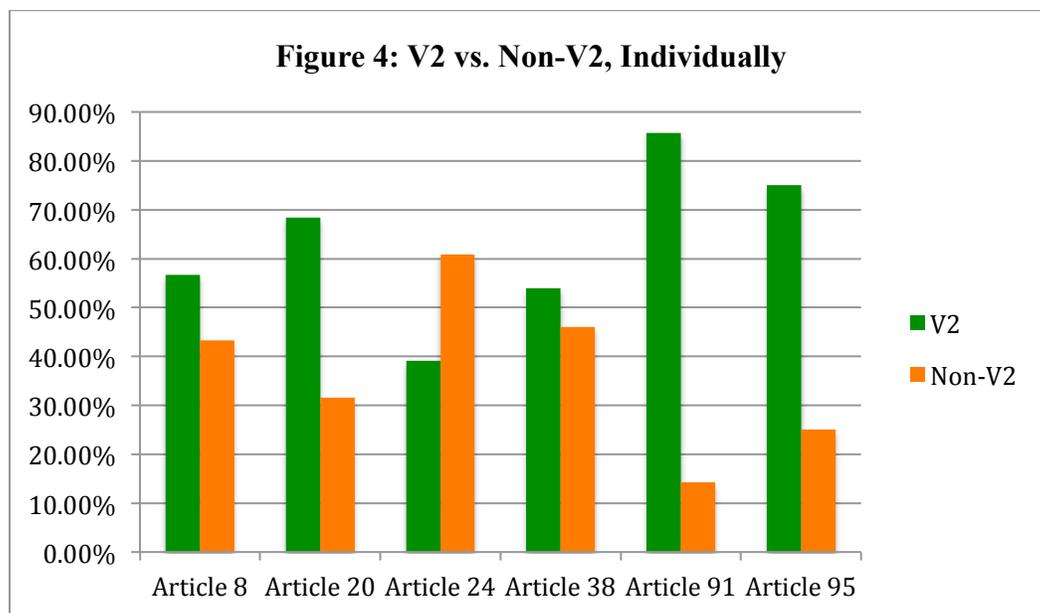


While V2 is favored in both registers, the rate of V2 expression is somewhat higher in prose at 62.2% compared with 55% for poetry. Part of my hypothesis was that in poetry there would be a lower rate of V2 than in prose, due to the more flexible nature of poetic syntax. This prediction is supported by the data. This makes sense if V2 is a syntactic constraint that is only broken in special contexts, like poetry. Both poetry and prose have V2 as a majority pattern. This suggests that the change from V2 to non-V2 in French was not yet complete the varieties of AN from which this data was gathered.

Of course, in Modern French there is still a certain frequency of V2, not because it is still a formal syntactic operation, but because of statistical accident. Many sentences start with a

subject and then a verb without any intervening material, so it would not be surprising to find instances of V2 in Modern French. The question that needs to be answered is: what is the frequency of V2 expression in Modern French? The answer to this question would allow for comparison with the AN data, which would better help situate at what stage in the process of change the data represented by MS Harley 2253 is in. We can predict that it will be lower than what is seen in AN, but further investigation is needed.

In Figure 4, the results for individual articles are given:



As with null subjects, the results are even more interesting when displayed individually. There are several observations to make here. The first is that V2 is more frequent in all of the selections except one, Article 24, which has 39% V2. The only other article that comes close to this low of a rate of V2 is Article 38, though it is still majority V2 at 54%. Is there something unique about Article 24 that might explain this discrepancy? The most immediate explanation would be that Article 24, compared with the other two poems and with the prose selections, is the only text that has a section that repeated more than once, the chorus. In fact, it is repeated

nine times. Thus, the patterns of V2 expression displayed in this one section will be reproduced statistically in a quantitative analysis of this sort and will skew the results.

A close look at the chorus, given below with the matrix verbs in bold, may help illustrate this point better:

- (17) a. Ore **est ocys**, la flur de pris, qe taunt savoit de guerre;  
 "Now he is slain, the flower of fame, who knew so much of war."  
 b. ly quens Montfort, sa dure mort molt en **plorra** la terre.  
 "The Earl of Monfort, his cruel death the land will deeply mourn."

The chorus consist of two matrix clauses. The first clause (17a) is V2; the verb *est ocys* is preceded by one constituent, the Adverbial Phrase *ore*. In the second clause, the surface order is OVS, which can be analyzed as having a fronted/topicalized object phrase and Verb Phrase. However, there are also two other constituents intervening between the object and the verb: *molt*, and Adverbial Phrase and *en*, an object clitic that could be coindexed with the object phrase. This means that the verb in this clause is actually in fourth position. Given that this is repeated in the chorus, it affects the overall frequency of V2 expression in this poem.

Another observation that can be made from the information presented in Figure 4 is that two of the prose selections, Article 91 and Article 95 both have the highest rate of V2 in the entire corpus, 85.7% and 75%, respectively. How do we account for this? As mentioned above, Article 91 is likely a translation either from Latin or Greek. However, neither of these languages has ever been analyzed as having a regular V2 constraint like what is seen in Germanic languages. One striking characteristic of this selection is that it is all very straightforward description about one person, Jesus Christ. For example, there are two verbs more frequently repeated than any others: *avoit*-he had, and *estoit*-he was. This is to be expected in text purporting to be a physical description of an individual. The language used here has a repetitive

feeling to it: he had an honorable face, he had a Greek nose, he had a beard, etc. Since the intention of the author is to say what Christ had and how he was, there will necessarily be a lot of direct objects and adjectives/adverbs with the two previously mentioned verbs. One property that has been identified in previous research on V2 is that it correlates strongly with sentence-initial direct objects and adjectives/adverbs. These preposed (sentence-initial) elements are said to trigger V2. Thus, in a type of writing in which there have to be a lot of direct objects and adjectives, it would not be surprising to find a lot of V2.

With respect to Article 95, the list of all the places conquered by the Saracens, there is also abundant repetition. The most frequent patterns can be seen in the selections below:

- (18) Melchifas, la tierce fitz, **tint** la terre de Caine, plus qe quatre cent cités e chastels fortz. (line 11)
- (19) En Baruch **est** un ymage de Jesu Crist crucifié... (line 45)

The pattern can be summarized as follows: Person X/Place X has/is Y. The verb in this pattern is almost always preceded by a single constituent, a Noun Phrase in (18) or a Prepositional Phrase in (19). This kind of repetition is to be expected in a list; a specific register where no attempt is made to make the prose a pleasant or entertaining read. As with null subjects seen above, it appears that high V2 expression rate is a feature that characterizes the list dialect.

There is another factor that likely contributes to high V2 not only in Article 95 but also Article 38, the two lists, possibly written by the same author: the grammaticalization of *la* as an expletive in existential constructions:

- (20) a. Dames est une bone cité e forte. **La est** une ymage de Nostre Dame...  
"Damascus is a strong and important city. There is an image of Our Lady..."  
(Art. 95, line 32)
- b. E de yleqe a Japhet, xij liwes. **La est** um peron...  
"And from there to Jaffa, twelve leagues. There is a rock..." (Art. 38, line 12)

An expletive is a word that appears subject position but has no real semantic value, as in English constructions such as "There are too many student in this class" or "It is raining." The words *there* and *it* do not refer to anything semantically, rather they are there for structural reasons that are not necessary to go into here. In (20) *la*, which was originally a locative adverb, may have the same function. In other words, the word is losing its semantic value and taking on a purely grammatical function. In some languages, existential constructions require an expletive. In Latin, an expletive was not required:

- (21) quoniam non est herba gregibus servorum tuorum.  
 because NEG is grass flocks.Dat servants.Gen your.Gen  
 "Because there is no grass for the flocks of your servants." (Genesis 47:4)

Thus, while in Latin a copula did not require an expletive in existential constructions, the AN copula did. Whether this is the case or not in these texts and in AN in general requires a much more detailed study than what is intended here, but it explain why the recurrent pattern of adverb plus verb is so frequent.

### 3.3 Generic Subject Expression

How is a generic subject expressed in these selections? It was predicted at the outset of this study that, if AN is a pNSL, it should have null generic subjects. In order to determine whether this prediction was correct or not, a closer look at the data is required. The following sentences are possible examples of a null generic pronoun ("pro"):

- (22) Mout *pro* purrad mener sure vie que de femme *pro* puet aver choys.  
 much *pro*<sup>5</sup> could.3S lead.Inf secure life who of woman *pro* can.3S have.Inf choice  
 "He might lead a very secure life whoever can take his choice of women." (Art.  
 8:104-105)

---

<sup>5</sup> The italicized *pro* is not actually in the written text. It is a convention that stands for a null pronoun, assumed to be in subject position. It is inserted to help demonstrate the existence of a null generic pronoun.

- (23) Korteysie en femme git en lu ou *pro* ad bel desport.  
 courtesy in woman lies.3S in place where *pro* has.3S sweet delight  
 "Kind courtesy lies in a woman in the place where one has sweet delight." (Art.  
 8:122-123)
- (24) E quant ele est destrue, si *pro* seme un tote maneres blees,  
 and when it is.3S receded then *pro* sows.3S a whole manners grains  
 e les *pro* sient en mars.  
 and them.Acc *pro* harvest.3P in March  
 "And when it has receded, all kinds of grains are sowed, and they are harvested in  
 March." (Art. 95:70)

In (22), the subject of both the matrix and embedded clause is a null pronoun with no specific referent. However, gender may be inferred from the context. Null *pro* is generally analyzed as being [+human] but not specified with respect to gender. If *pro* were [+masculine], for example, it would only be interpreted as "men in general" rather than "people in general." On the other hand, there is no other masculine marking that shows agreement with the null *pro* in this clause, and so a null generic pronoun may very well be present. In (23) the subject of the embedded clause is also a null generic referent, translated as "one."

The data in (24) is more complicated. The verb *seme* is third-singular and, though translated with an English passive by Raybin, it is actually an active form. Null generic *pro* generally occurs with third-person singular active verb forms, like English *one*. A more literal translation would be: *then one sows all kinds of grains* (or) *then people sow all kinds of grains*. However, the next verb, *sient*, poses a problem: it is a third-plural form. As with the previous verb, it is an active form, rather than a passive. But if it shares the same antecedent as *seme*, and receives a generic interpretation, why does it show plural agreement? A more in depth analysis is needed, but a possible explanation is that the null generic pronoun in AN is not specified for person in its formal features. Thus, either singular or plural agreement can appear on the verb.

Of course, another possibility is that this was a copy error. Unfortunately, the question has to be left unanswered for now.

Whether the previous examples have a null generic pronoun or not, the most common method of generic subject expression in these texts is the impersonal *on*-construction, as in Modern French, though there is considerable orthographical variation in AN:

(25) Ne vueil qe l'**em** die.  
 NEG wish.1S that it-ON say.3S  
 "I do not wish that people say it." (Art. 24:117)

(26) La est um peron qe **um** apele le Peroun Seint Jake.  
 there is a rock which ON calls.3S the Rock Saint James  
 "There is a rock which people call the Rock of St. James." (Art. 38:12)

That AN uses this construction would be consistent with continental French (Kibler 1984:78), but if a null generic pronoun is equally available for the same interpretation, this may serve as evidence that contradicts Holmberg's Generalization, as discussed in Section 2.3. There is no need to elaborate this point any further here, but it may serve as useful data for future research.

A final observation is in order here prior to the conclusion. One of the predictions made in the introduction was that a pattern might emerge with respect to the distribution of methods of generic expression in the poetry and prose selections. While a thorough quantitative study of this variable was not carried out, I did originally identify nine putative examples of generic subject expression. Two were later ruled-out due to not actually having a generic interpretation. Of the seven remaining tokens, four are impersonal *on*-constructions and three are possibly generic null pronouns. Since the token-count is so low, there is not sufficient data to establish a correlation.

#### 4. Conclusion

In this pilot study, data from six AN texts in MS Harley 2253 was analyzed in order to determine to what extent register was a strong conditioning factor in the variation of null

subjects, V2 order, and generic subject expression. With respect to the first two linguistic variables, choice of register (prose or poetry) does appear to correlate with which variant is used. In prose, null subjects are much less frequent than in poetry, which may be due to the need to be economical in a poetic context. Individual differences in the prose selections also help explain the preference for overt subjects in that register. V2 was found to be less frequent in poetry than in prose. The poetic register allows a greater amount of freedom when it comes to syntax; word order constraints such as V2 are more easily broken in poetry than in prose. As with null subjects, certain differences between the texts also help explain the statistical patterns. The result of the analysis of generic subject expression is that AN, like continental Middle French, also uses the impersonal *on*-construction, though there is evidence that generic null subject may also be used.

The next step in this project is to go back to the data and code for linguistic factors such as coreferentiality, clause type, verbal form, etc. More complex statistical models may also be used to better understand what the most significant factors are in determining which variants are used. Finally, the corpus should be expanded to incorporate more AN material. Syntactic variation can be observed throughout the entirety of MS Harley 2253. The variegated nature of this work provides a wealth of data that has not even been approached within the field of linguistics. Further linguistic investigation of this manuscript and Anglo-Norman in general will greatly enhance our understanding of language change and variation.

## References

- Adams, Marianne. 1987. Old French to the Theory of Pro-Drop. *Natural Language and Linguistic Theory* 5: 1-32.
- Alexiadou, A. & E. Anagnostopoulou. 1998. Parameterizing Agr: Word Order, Verb-Movement and EPP-Checking. *Natural Language and Linguistic Theory* 16:491-539.
- Belletti, Adriana. 1982. Morphological Passive and Pro-Drop: The Impersonal Construction in Italian. *Journal of Linguistic Research* 2: 1-34.
- Biber, Douglas & Susan Conrad. 2001. Register Variation: A Corpus Approach. In *The Handbook of Discourse Analysis*, (eds.) Deborah Schiffren, Deborah Tannen & Heidi E. Hamilton. Malden, Mass.: Blackwell.
- Chomsky, Noam. 1965. *Aspects of a Theory of Syntax*. Cambridge: MIT Press.
- . 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Dean, Ruth J., with Maureen B.M. Boulton. 1999. *Anglo-Norman Literature: A Guide to Texts and Manuscripts*. London: Anglo-Norman Text Society.
- Devine, A.M. & Laurence D. Stephens. 2006. *Latin Word Order: Structured Meaning and Information*. Oxford: UP.
- Fein, Susanna (ed.). 2000. *Studies in the Harley Manuscript: The Scribes, Contents, and Social Contexts of British Library MS Harley 2253*. Kalamazoo: Medieval Institute.
- Fein, Susanna, David Raybin & Jan Ziolkowski (eds., trans.). 2012. *The Complete Harley 2253 Manuscript in Three Volumes*. Kalamazoo: Medieval Institute.
- Greenberg, J. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In *Universals of Grammar*, (ed.) J. Greenberg, 73-113. Cambridge: MIT Press.
- Haegeman, Liliane & Tabea Ihsane. 2001. Adult Null Subjects in the Non-Pro Drop Languages: Two Diary Dialects. *Language Acquisition* 9:329-346.
- Haegeman, Liliane. 2013. The Syntax of Registers: Diary Subject Omission and the Privilege of the Root. *Lingua* 130:88-110.
- Holmberg, Anders. 2005. [Is there a little pro? Evidence from Finnish](#). *Linguistic Inquiry* 36: 533-564.
- . 2010. The Null Generic Pronoun in Finnish: A Case of Incorporation in T. In *Parametric Variation: Null Subjects in Minimalist Theory*, (eds.) Biberauer et al. Cambridge:UP.
- Ingham, Richard. 2012. *The Transmission of Anglo-Norman: Language History and Language Acquisition*. Amsterdam: John Benjamins.
- Kibler, William W. 1984. *An Introduction to Old French*. New York: Modern Language Association of America.
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.
- . Null Objects and the Theory of pro. *Linguistic Inquiry* 17:501-557.
- Roberts, Ian. 2007. *Diachronic Syntax*. Cambridge: UP.
- Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in Minimalist Theory. In *Parametric Variation: Null Subjects in Minimalist Theory*, (eds.) Biberauer et al. Cambridge: UP.
- Romaine, Suzanne. 1982. *Socio-Historical Linguistics*. Cambridge: UP.
- Sigurðsson, Halldór Ármann & Verner Egerland. 2009. Impersonal Null Subjects in Icelandic and Elsewhere. *Studia Linguistica* 63:158-185.
- Tyler, Elizabeth (ed.). 2011. *Conceptualizing Multilingualism in England, C. 800-C. 1250*. Belgium: Brepols.
- Vance, Barbara. 1997. *Syntactic Change in Medieval French: Verb Second and Null Subjects*. Dordrecht: Kluwer Academic Publishers.