STUDYING WORD ORDER DIFFERENCES IN A HISTORICAL PARALLEL CORPUS. AN EXAMPLE FROM OLD SPANISH AND OLD PORTUGUESE

KRISTINE GUNN EIDE
University of Oslo

ABSTRACT

In this article, I use a parallel corpus from the 13th and 14th century to tease out some of the structural differences that existed between Old Spanish and Old Portuguese. While these two related languages were relatively similar in many respects, and the parallel corpus reflects these similarities, differences in syntax and information structure are also apparent. By comparing the syntactic and information structural properties of the sentences that display different word orders, it is possible to pinpoint more exactly what these differences were. The parallel sentences show that information structural properties of the left periphery, where Spanish allows for new information where Portuguese does not, account for differences in both object and subject placement.

[1] INTRODUCTION

Spanish and Portuguese are languages with a relatively free word order, which is regulated by a combination of syntactic restrictions, prosody and information structure (e.g Zubizarreta (1998) for Modern Spanish, Costa (2004) for Modern Portuguese). There are also certain syntactic and information structural differences between the two languages, in particular with regard to subject placement. Portuguese has a strong tendency for preverbal subjects, even when the sentence is introduced by an adverb, as in the two possible translations in (2) of the Old Portuguese sentence (OP) in (1). The post verbal subject in (2a) is only allowed in contexts where the subject contains new information. In sentences introduced by an adverb, Modern Spanish (MSp) either prefers a post verbal subject (4a), similar to Old Spanish (Osp) (3) or has a pause between the adverb and the preverbal subject, indicated by a comma in (4b). The word order in (4a) would be infelicitous in Modern Portuguese (MEP) when *los romanos* is old information.

(1)   E entom sayo do boosco hũa molher nua (OP) (Orto do Esposo 4.70, 155r)
        And then came-out from-the woods a woman naked
        'And then a naked woman emerged from the woods'

(2)   a. E então saiu do bosque uma mulher nua (MEP)
        b. E então uma mulher nua saiu do bosque (MEP)

(3)   Entonces tomaron los romanos la dicha cibdad (Osp)
        Then took the Romans the already-mentioned city
        'Then the Romans took the above mentioned city'

(4)   a. Entonces tomaron los romanos dicha ciudad (MSp)
        b. Entonces, los romanos tomaron dicha ciudad (MSp)

In older stages of these languages, on the one hand, Old Portuguese and Old Spanish word orders are also very similar and seemingly structured according to some of the same principles, and are often (mistakenly) taken for more similar than they were. On the other hand, there are certain syntactic and information structural phenomena that have been described for only one of the languages, and not for the other. Among them are,

(i) Old Spanish has been classified as a VSO language, a claim which has not been proposed for Old Portuguese; and

(ii) the sentence final position in Old Portuguese has been strongly associated with new information while studies on the new information in Old Spanish is less restricted to this one position.

To my knowledge, no systematic comparison has been made between the information structure in these two languages. In this paper, I compare a parallel text from the 13th and the 14th century and show that the difference in subject placement that we find between Modern Spanish and Modern Portuguese, was already systematically present in the old languages and driven by a combination of syntax and information structure. The differences in information structure also affected the objects in a regular way, even though modern traces of this difference are not found.

In section [2], I will outline briefly the syntax and information structure of these languages, in section [3], I will discuss advantages and disadvantages of using a parallel corpus in the study of older texts, in section [4], I present the results of the searches, concluding in section [5].

[2]   OUTLINE OF OLD SPANISH AND OLD PORTUGUESE WORD ORDER

Both Old Spanish and Old Portuguese have been classified as V2 languages by some authors, such as Salvi (2004) for old Romance in general, Fontana (1993) for Old Spanish and Ribeiro (1995) for Old Portuguese. They argue that these languages share a relatively frequent XPVS structure where the verb moves to C and the preverbal positions in CP are reserved for specific information structural categories, such as topic and (contrastive) focus. Others refute this claim (Kaiser (2002) for

Old Romance in General, Rinke (2007) for Old Portuguese, Sitaridou (2012) for Old Spanish) mainly due to the large number of V1 and V3 sentences also found in these languages. This discussion around the V2 structures of old Romance languages has been going on for more than a decade. I will not go further into it here, but concentrate on the similarities and differences between Old Spanish and Old Portuguese.

The similarities are many, and the two languages tend to group together in cross linguistic studies, such as Sitaridou (2012), in that they both have frequent verb-subject inversions, that topics and contrastive foci are in the left periphery, and that information focus can be sentence final. Both Old Spanish and Old Portuguese have all possible combinations of subject, object and verb, even though there is a certain variation with regard to text types. Unlike Portuguese however, Spanish has also been classified as a VSO-language or a language with a mixed SVO/VSO order (Neumann-Holzschuh 1998), while no one has claimed this for Old Portuguese. Old Portuguese is generally known to display a basic SVO order (Martins 2002, 2011).

[2.1]   *Information structure*

Information structure is a common term for several related linguistic and textual phenomena. It includes, among others, *relational* information structure and *referential* information structure. Referential information structure refers to the information value of the referents in a context, i.e. whether a referent is being introduced in a text for the first time or whether it has been mentioned before. In (5), *A mermaid* is new information, whereas in the continuation of the story (6) *she* is old, or given, since this pronoun refers to the aforementioned mermaid who is being 'picked up again'. As for *Neptune* in (6), this has a status somewhere in between given and new. While *Neptune* is being introduced for the first time in this text, and hence qualifies as new information, it is also a well-known Roman god, which in some circles at least is part of a general, shared knowledge of the world.[1]

(5)   Suddenly, a mermaid emerged from the seaweed.

(6)   She was looking for Neptune.

Relational information structure refers to the internal structure of a sentence, where referents are marked, syntactically or otherwise, according to pragmatic notions such as topic, focus and background. There is a universal tendency for topics, generally understood as 'what the sentence is about' to be old informa-

---

[1]   There are several intermediate stages between the core categories 'given' and 'new', which will not be dealt with here. For an overview of the information structural categories used in the ISWOC and PROIEL projects, see Haug et al. (2014) and PROIEL.

tion. The term focus is used in a variety of meanings, roughly corresponding to either 'information focus' or 'contrastive focus'. Similar to the correlation between topics and old information, there is a general correspondence between information focus and new information. The notion of contrast, which will be mentioned as relevant for some of the examples in this article, is more complex. It cannot be reduced to the old-new dichotomy, and has particular phonological traits as well as syntactic reflexes that distinguish it from both information focus and topics. In modern languages, information structural categories are often teased out through question-answer pairs. This is illustrated in (7), (8) and (9) where in (7) the entire VP is new information and in focus, in (8) the object is in focus, whereas in (9) there is no information focus, but a contrastive focus which picks Peter out of a set of two candidates, *Peter* and *Sebastian*.

(7)   What did Peter do?
        Peter [hit John]$_F$

(8)   Whom did Peter hit?
        Peter hit [John]$_F$

(9)   Did Sebastian or Peter hit John?
        PETER hit John.

In English, relational information structure is primarily marked by intonation and sentence stress whereas in Romance languages, relational information structure is primarily marked by word order *in combination* with stress and intonation. As a general rule (although not without exceptions) in English, the word order is kept and the stress patterns moves, while in Romance languages, the stress pattern remains the same and the words move to accommodate the discourse patterns. This is illustrated in the question-answer pairs below, where the Portuguese examples (10) and (11) the object placement changes when the object is given and thus 'escapes' sentence final focus stress, whereas in the English examples (12) and (13), the stress pattern changes and the word order remains the same regardless of information structure.

(10)   O que é que o Paulo fala bem? (MEP)
        O Paulo fala bem francês

(11)   Como é que o Paulo fala francês? (MEP)
        O Paulo fala francês bem (Costa 2004)

(12)   What does John speak well?
        John speaks FRENCH well.

(13)   How is John's French?
        John speaks French WELL.

We do not have direct access to prosodic features in historical corpora, nor do we have the possibility to ask for native speakers' judgement on construed question-answer pairs. One way of working around this problem is to map referential information structure on to syntactic structures, given the universal correlation between syntax, referential and relational information structure, and try to 'reconstruct' the relational information structure through the other two.

The left periphery, in particular, has been subject to scrutiny, since this part of the sentence is where we find and array of positions carrying different information structural values. Schematically, the syntax of the left periphery of Old Spanish and Old Portuguese has been outlined as follows:

- Spanish:
  ForceP > TopP > ContrastiveFocusP > FinP > InfoFocusP > TP > vP
  (Sitaridou 2011, 178)

- Portuguese:
  Comp > topic > focus > clitic > scrambled/interpolated XPs > verb
  (Martins 2011)

The differences in the left periphery can be summed up in the following:

- Old Spanish allows for preverbal information focus (FocusPinfo), but Old Portuguese does not.

- In Old Portuguese, non-focused elements scramble to a non-focused place before the verb (or the non-focused place after the verb).

This type of scrambling to a preverbal position is not found in the Old Spanish scheme. In Old Portuguese the evidence for a non-focused preverbal scrambling position are examples such as (14) where elements are placed between the clitic (*me*) and the verb (*disse*). According to Martins (2011) this position is reserved for elements which need to avoid a sentence final focus stress, normally associated with information focus.

(14)   depois que me o sancto bispo Nono estas cousas disse[2] (OP)
         (Vidas dos Santos, Vida de Santa Pelágia, 177v)
         after that me.CL the holy bishop Nono these things said
         'After the holy Bishop Nono had said these things to me...'

Needless to say, the clitic is not always present, and in most cases we therefore have no diagnostic for determining the exact syntactic position of a preverbal element, ie whether it is in this 'non-focus' position or if it is a topic or a contrastive

---

[2]   In the short parallel text at hand, no examples with a clitic and an interpolated object were found. This example is from *Vidas de Santos de um Manuscrito Alcobacense, Vida de Sta. Pelágia*.

focus. We can only rely on information structural criteria and, as I shall argue in section [4], on the information structural and syntactic exponence of parallel texts. As for Old Spanish, even though interpolated elements occur, it is far less frequent and much more restricted as to what types of elements that can occur here (Castillo Lluch 1998).

In addition to the preverbal 'avoidfocus' position in Old Portuguese, arguments can also scramble to a lower position, just above VP (Martins 2011). This 'avoidfocus' strategy is the same as the one used in the Modern Portuguese example (11). Schematically, the two languages seem to have the following information structural properties, where contrastive focus also incorporates anaphoric, or discourse-linked elements[3]:

- Spanish:
  Topic # Focuscontrast Focusinfo verb non-focus/ Focusinfo ?

- Portuguese:
  Topic # Focuscontrast non-focus verb non-focus VP Focusinfo

While the information structural properties of the sentence final position(s) are relatively clear for Old Portuguese, they have not been properly described for Old Spanish and are therefore left with a question mark.

In this paper I will look at the three differences which we should be able to find evidence for in a parallel corpus:

- Spanish has preverbal information focus (Sitaridou 2011), in addition to, possibly, sentence final information focus. In other words, referents that are introduced as new, can be either pre or post verbal. Preverbal information focus has not been proposed for Portuguese, and given the fact that it has been classified as an SVO language, we would only expect to find preverbal subjects that contain new information, never preverbal objects.

- While both Spanish and Portuguese seem to accept sentence final information focus, and move objects that contain old information away from the sentence final focus position, Portuguese has two 'avoidfocus' positions, one preverbal, one below the verb. Spanish has only the one position below the verb.

- The 'avoidfocus' rule is very strong in Portuguese, and very much linked to the prosodic structure of the language. Although the same sentence final focus stress has been described for Modern Spanish (Zubizarreta 1998), it

---

[3]   Both languages are null subject languages. An expressed subject pronoun is contrastive or, in some cases, serves to disambiguate between referents. Disambiguating subject pronouns are post verbal in Old Spanish, see Eide & Sitaridou (2014) for a further discussion.

does not have a prominent place in the literature on Old Spanish, perhaps not surprising since information focus occurs preverbally as well (Sitaridou 2011).

In the present study, I will check the parallel text for

- what happens to Spanish preverbal objects in the Portuguese translation, in light of the fact that Portuguese has no preverbal information focus,

- given the fact that Portuguese has two 'avoidfocus' positions, what is the distribution of non-focused elements in the translated text and

- whether the strong 'avoidfocus' rule in Portuguese affects the translation. For instance, do more objects that contain old information move from a sentence final focus position in the Portuguese translation?

[3] THE DATA: CORPUS AND METHODOLOGY

These texts are a continuation of the Bible, where we are told what happened to the people who were driven from Babylon, in particular what happened to the ones who ended up in Spain. In other words, we are dealing with a supposedly original Spanish text *Estoria de Espanna* (12th century), chapters 1-12. The Portuguese version, *Crónica Geral de Espanha* (14th century), adds bits and pieces about who went on to colonise Portugal and to found Portuguese cities[4].

Additionally, the Portuguese text is probably not a direct translation of the Spanish one. It is therefore difficult to automatise the searches, and not all sentences are comparable. The criteria for selecting the sentences used for this study was that they, in both texts, are the same type of clause (main clause/subordinate), that they have the same arguments to the verb, and that they carry the same meaning. Given these criteria, the corpus consists of 545 parallel sentences. Differences in word order, including adverbs, adjectives and quantifiers, were found in 104 of the comparable sentences. Of these, only sentences with subjects and/or objects (except clitics) were taken into consideration, leaving 40 sentences.

[3.1] *Using parallel corpora in historical studies*

Even though the texts are written, chances are we may not have the original manuscript upon which the translation is based, unlike modern parallel corpora. The manuscript tradition and translation history of these texts are complex. Although the Spanish text is older than the Portuguese one, it has been proposed that the Spanish text was translated from a presumably lost Portuguese original and that the Portuguese text known today, is an edited version of the first Portuguese original (Cintra 1951). In fact, we are not even sure if the translations

---

[4]   The annotated texts have been parallelised and annotated by the ISWOC project and are available at http://foni.uio.no:3000/.

were directly translated from a written source, or if the original text was read out loud for someone to translate. Even though we know certain things about medieval translations in general, we know very little about the individual translations. In the parallel text we are presented with here, the translator/editor feels free to explain unclear passages in the original, or to even add something to satisfy the local audience, in this case the Portuguese, who wanted to know what went on in Portugal.

Schematically, the differences between modern and older translations are as indicated in Table 1.

TABLE 1: Differences between modern and older translations

MODERN TRANSLATION	OLDER 'TRANSLATIONS'
Original manuscript available	Original manuscript not available/non-existing or uncertain
Written transmission	Written or oral transmission
Translator known	Translator unknown
Clear principles of translation: translate everything, no more, no less	Unclear principles of translation: translate most parts, explain what is unclear, add if something is missing.

[3.2]   *Contamination*

It is not improbable that the language in the translation has been influenced by the original. In the case of word order, this is commonly known to be a factor. Related to the uncertainties regarding the original manuscript upon which the translation is based, is that the differences in word order, theoretically at least, could be due to a translation from a different manuscript than the one we have available. In this paper, I adopt the solution proposed by Petrova & Solf (2009) among others, to only study sentences in which there is a difference in word order between the two texts. This is under the assumption that such sentences are not subject to contamination, rather they reflect a choice made by the translator to give the translated text a proper syntactic and information structural configuration in the target language.

[3.3]   *Advantages of studying information structure in parallel texts*

While the problems regarding word order contamination in old parallel corpora are obvious, there are some great advantages to using these texts, in particular when it comes to information structure. Information structure is context-dependent, and in parallel corpora, the context is presumably the same, and hence the referential information status of the arguments is also the same. Parallel corpora thus give us a unique opportunity to compare sentence structures *in* context.

SYNTACTIC STRUCTURES IN THE PARALLEL TEXTS OF OLD SPANISH AND OLD PORTUGUESE

Turning now to a description of the parallel text at hand, we find that of the 545 of the parallel sentences, more than 450 display the same word order. Typical examples are found of the following parallel word orders: SVO, OVS, VSO as well as VS with unaccusatives. They correspond to distinctive information structural configurations. SVO is the preferred neutral word order in unmarked cases where both subject and object contain new information.

(15)   SVO
        a. Moysen escrivio um libro (OSp)
        'Moses wrote a book'
        b. Moyses fez huu livro (OP)
        'Moses wrote a book'

The object is preposed when it is contrastive or when it serves as a discourse link, which is when it refers back to something previously mentioned in the text. In these cases the preferred position for the subject is post verbal.

(16)   OVS
        a. Tod esto cuenta Moysen (OSp)
        'Moses recounts all this'
        b. todas estas coisas e outras muytas cõta Moyses (OP)
        'Moses recounts all these things as well as many others'

We also find cases of VSO, when there is no such discourse link. For other languages such sentences have been classified as thetic sentences, i.e. sentences without topic (for Old French and Old High German, (Petrova & Rinke 2014), for Old Spanish Neumann-Holzschuh (1998)).

(17)   VSO
        a. cuemo (...) aduxo dios el grand diluuio (OSp)
        'How God brought the great flood'
        b. como (...) adusse Deus o grande deluvio (OP)
        'How God brought the great flood'

Additionally, both languages have SOV word orders, although no examples were found which were both parallel and displayed the same word order. The information structural properties of these sentences will be discussed in section [4.2].

(18)   SOV
        a. yo tal vida fago. (OSp)
        I such life lead

'I lead such a life'
b. Quando Hercoles esto ouvyo (OP)
when Hercules this heard
'When Hercules heard this'

With unaccusatives (19) and passives (20), both word orders VS and SV are possible. As for the information structure of the subjects, *europa* is old information, and can be interpreted as contrastive in the sense that it picks out Europe from the set *Europe and Africa*[5]. The subject of the passive verb *fue destroyda / foi destroida* is also somewhat complex: First of all, the Spanish and the Portuguese context differ: In the Spanish text, Troy is old information, as shown by the context. In the Portuguese text, this is the first time the city of Troy has been mentioned. However, given the context, Troy could be inferred from *Iliones* (another word for Trojans) in the previous sentence. As such it can be thought of as old, rather than new (see also footnote 1).

(19)   VS – unaccusatives/passives
         Context[6] (translated from the Portuguese version): ['Together with the sons of Japheth, some of the sons of Shem inherited all of Europe, from the great sea which encircles all the earth and which in Greek is called *the Ocean* to the other sea which is called the Mediterranean because it lies in the middle of the earth and is the division between Europe and Africa.']
         a. e acabasse europa en un cabo d'Espanha (OSp)
         and end.3SG-REFL europe in one end of'Spain
         'and Europe ends in one corner of Spain'
         b. e acabasse Europa encabo despanna (OP)
         and end.3SG-REFL europe in-end of-Spain
         'and Europe ends in the corner of Spain'

(20)   SV – unaccusatives/passives
         a. ['From Ivan, son of Japheth came the Iliones who populated a part of Greece and later Troy (...)']
         despues que troya fue destroyda (OSp)
         'after Troy was destroyed'
         b. ['And also from Ivan, who we have already mentioned, came the Iliones who populated a part of Greece and Persia (...)']
         despois que Troya foy destroida (OP)
         'after Troy was destroyed'

---

[5]   Alternatively, that we are dealing with a thetic sentence.
[6]   The contexts are inserted when they are relevant for the (referential) information structure, and marked with square brackets []. When the contexts are parallel, only the English translation is given. Whenever the contexts diverge, the context is given for each language, along with the English translation.

In the data set, no parallel examples were found which have OSV or VOS. For the purposes of the present paper, we concentrate on the parallel sentences.

[4.1] *Interpolation*

Turning now to the differences between the two texts, as mentioned in section [2], interpolation in Old Portuguese is far less restricted than in Spanish. Although no examples of interpolated non-focused objects were found in the present corpus, several examples were found with subjects (21), negations (22) and adverbs (23) and (24), where the Spanish and Portuguese text show a distinct pattern of interpolation in Portuguese, yet none in Spanish. (The clitics are italicized, the interpolated element and its Spanish equivalent is in bold.)

(21)   Subjects
        a. y ella diogela por quel criasse aquel moço. atal pleyto que quando fuesse grand que fiziesse del aquello que **ella** *l* mandasse (OSp)
        and she gave it (the land) to him (Eurysthes) in order for him to raise that boy on the condition that when he was grown, that he did to him that which she-CL ordered.3SG
        b. E esto com tal condiçom que, quando elle fosse grande, que fezesse delle aquello que *lhe* **ella** mandasse (OP)
        and this on the on the condition that when he (Hercules) grew up, he would do with him that which CL she ordered.3SG

(22)   Negation
        a. y rogol que **no** *l* fiziesse mal. (OSp)
        and asked-him that not-CL do.3SG ill
        b. e rogouho que *lhe* **nom** fezesse mal. (OP)
        and asked-CL that CL not do.3SG ill
        'and asked him not to do him ill'

(23)   Adverbs
        a. E cuenta lucan que desque *la* ouo **alli** poblada. que fue a caliz (OSp)
        And tells lucan that after CL had.3SG there populated, that went to Caliz
        b. Mas conta dom Luchas que, despois que *a* **ally** ouve poboada, que se foy a Calez (OP)
        But tells dom Luchas that after that CL there had.3SG populated, that CL went to Caliz
        'But D.Lucas recounts that after he had populated/founded it (Ispalis, a city), that he went to Caliz'

(24)   Adverbs
        a. que lo que caçaua traye*gelo* **alli**. (OSp)
        that CL which hunted.3SG brought.3SG-CL.DAT-CL.ACC there

'that what he hunted, he should bring it there to him.'
b. que, de quão caçava, de todo *lhe* **ally** tragia. (OP)
that of all hunted3.SG. of all CL.DAT there brought.3SG
'That all he hunted, he should bring all (of it) there to him.'

In the examples above we observe two things in particular: The Spanish post verbal adverbs in (22a) and (23a) appear preverbally in Portuguese. The subject in (21a) and the negation in (22a) were both preverbal in the Spanish text. If Martins (2011) is correct, that the position between the clitic and the verb is an 'avoidfocus position'[7], adverbs such as *alli*, 'there', which can be argued to contain old information since they refer to a place already mentioned in the text, should not be sentence final in Portuguese. Spanish does not seem to have this restriction. The subject in (21) is similarly interpreted as neutral with regard to information focus in the Portuguese text. The context does give room for a contrastive interpretation of the subject, which would account for the Spanish word order.

[4.2]   *Subjects*

In most sentences, there is no clitic to give a clear diagnostic for the exact position of the preverbal subject in the Portuguese text. However, since VSO order has been proposed for Spanish, and since there is a post verbal non-focus position in Spanish, we would expect to find subjects that are neither topics nor foci in this position, while in Old Portuguese we would also expect to find them preverbally, given that it has been analysed as an SVO language.

Old Spanish: VS_{old}

Old Portuguese: $S_{old}V$

Indeed, in the examples that do not display the same word order, this is by far the most common difference, (25)-(28).

(25)   a. de cuemo crio **dios** el cielo ... (OSp)
of how created God the heaven...
b. como **Deus** criou o mundo ... (OP)
of how God created the world...

(26)   [Hercules is the protagonist of this story, hence old information]
a. quando esto oyo **hercules** (OSp)
when this heard Hercules
b. quando **Hercoles** esto ouvyo (OP)
When Hercules this heard.

(27)   a. lo uno por ques era **ella** muy fermosa (OSp)
the one because was she very beautiful

---

[7]   Although see Eide & Sitaridou (2014) for examples of highlighted elements also occurring here.

  b. a primeira, por que **ella** era muy boa e muy fremosa (OP)
  the first because she was very good and very beautiful

(28) [Rocas has travelled to Troy before the city's first destruction and is talking to the people there.]
  a. Quand esto oyeron **los troyanos**... (OSp)
  When this heard the trojans
  b. quando **os Troyããos** ouvyram o que dezia, ... (OP)
  when the Trojans heard what he said

When the subject patterns in the opposite way, and Spanish SV is rendered VS in the Portuguese text, the subject is either not a topic or it contains new information, i.e., it lacks one or more of the properties that are typical of subjects. In two of these instances, (29) and (30), the subject may be interpreted as contrastive. (*este* contrasts with *Hercules*, *yo* may be interpreted as contrasting with an unexpressed 'tu', referring to *Tarcos*.) Rinke (2007) has argued that postverbal pronominal subjects receive a contrastive interpretation in Old Portuguese. The following two examples seem to confirm to this hypothesis.

(29) a. assi cuemo hercules se apoderaua de la tierra por fuerça assi **este** se apoderaua della por amor. (OSp)
  like Hercules REFL took of the land by force, thus this REFL took of-her by love
  'just like Hercules conquered the land by force, this (man) conquered it through love.'
  b. bem assy como Hercolles se apoderara da terra per força, assi se poderou **este** per afaagos e per ben fazer. (OP)
  well thus as Hercules REFL took of the land by force, thus REFL took this by caresses and by well do.INF
  'just like Hercules conquered the land by force, this (man) conquered it by doing good deeds.'

(30) [Rocas has been living in a cave with a dragon, when Tarcos/Tharcus finds him: 'And while they were talking, the dragon came, and Tarcos, when he saw it, was frightened and wanted to flee. But Rocas told him not to, and promised him that the dragon would do him no harm. And the dragon threw half an ox before Rocas, since it had eaten the oher half. And he asked Tarcos if he wanted to eat from it and Tarcos said no, that he would rather eat with his companions. And Rotas said to him:']
  a. **yo** tal vida fago. (OSp)
  I such life lead
  'I lead such a life'
  b. Pois tal vida faço **eu** (OP)

for such life lead I
'For I lead such a life'

The following examples show that the information structure constraints which are valid for objects also affect the subject positions when the correlation *subject - topic - old information* does not hold, for instance when the subject is new information. This is the case with the following examples: the subject contains new information, whether in transitive (31), unaccusative (32) or predicative (33) structures, and the differences point to the strong sentence final information focus tendency found in Old Portuguese, while Old Spanish on the other hand, allows for sentence initial information focus.

(31)   a. assi cuemo **las montannas e los mares** la cercan. (OSp)
        thus like the mountains and the seas it surrounds
        'like the mountains and the seas surround it'
        b. assy como a cercã **os mares e as serras**. (OP)
        thus like it surrounds the mountains and the hills
        'like the mountains and the hills surround it'

(32)   a. o **el rey Leomedon** estaua (OSp)
        where the king Leomedon was
        b. onde estava **el rey Leomedon** (OP)
        where was the king Leomedon

(33)   a. quando oyo que **Rey era** e noble omne. ouo grand duelo del. (OSp)
        when heard.3SG that king was.3SG and noble man had.3SG great pain of-him
        'When he heard that he was a king and a noble man, he felt very sorry for him.'
        b. quando ouvyo como era rey e nobre homen, ouve delle grande doo (OP)
        when heard.3SG how was.3SG king and noble man had.3SG of-him great pain.

[4.3]   *Objects*

Objects usually contain new information and are usually VO, even in Old Spanish. (or VS if the verb is unaccusative). However, if the word order differs, we would expect Spanish to have OV (or SV with unaccusatives) since Old Spanish allows for preverbal information focus.

Old Spanish: OV (or SV$_{unacc}$)

Old Portuguese: VO (or V$_{unacc}$S)

Old Portuguese has a stronger tendency than Old Spanish for the sentence-final position to contain new information (cf. (24)), and the sentence is more likely to

be reorganised according to the following principle if the object is new information. (No examples were found of the opposite, where a Portuguese object containing old information scrambles and the Spanish does not.)

Old Spanish: V O PP_{old}/Adv_{old}

Old Portuguese: V PP_{old}/Adv_{old} O

Traces of these structural differences are also found in the texts. When the objects display a different word order in the texts, the Spanish preverbal objects may contain new information. The object in (34) *otra tierra* introduces a new referent[8] and *estorias* in (35) is also new information[9].

In Portuguese these objects remain sentence final, and receive sentence final focus stress[10].

(34)   ['And there were other peoples there which were called Vandals and those inhabited Andaluzia. And (this land) runs in width from the river called Guadiana to the Mediterranean sea.']
a. **otra tierra** y ouo que llamaron luzenna (OSp)
other land there was.3SG
b. ainda hi ha **outra terra** a que chamam Luçena (OP)
also there is.3SG other land

(35)   a. e deste fablaron todos los sabios que **estorias** fizieron (OSp)
who stories made.3PL
'who wrote stories'
b. e deste fallarom todos os sabedores que composeron **estorias** (OP)
who wrote stories

The following examples show another one of the most common word order differences, where the Portuguese Adverb (*assim, ally*) or PP (*com ella, per força*) which

---

[8]   Although it can also be argued that it is contrastive. Additionally, the object state of this argument is debatable. It could also be analysed as a subject. It would however, be subject to the same information structural principles mentioned above for subjects.

[9]   Although the following example is puzzling:

(i)   a. **este nombre** ouo por trebeios que mando y fazer hercules (OSp)
this name had.3SG for work which ordered.3SG there make.INF Hercules
'It got this name because of the work Hercules ordered done there'
b. ouve **este nome** por trebelhos que hy mandou fazer Hercolles (OP)
had.3SG this name for work which there ordered.3SG make.INF Hercules
'it got this name because of the work Hercules ordered done there'

[10]   One reviewer pointed out that this could be explained by a syntactic preference for these word orders, independently of discourse structures. While this analysis is not impossible, it does not explain why these differences are only found in particular information structural contexts. However, the corpus is too small to draw conclusions regarding this question.

contains old information, precedes the object containing new information, (or in the case of (36), more prominent information than the adverb) while the Spanish sentences display the opposite pattern.

(36)   ['And some say that the wise men of Greece knew (...) that there would be born a man there who would have the name Hercules and that he would do many and great and marvelous things throughout the world, more than any other man.']
         a. e los dos hercules primeros cuydando que cada uno dellos serie aquel. pusieron se **nombre** assi (OSp)
         put.3PL REFL name thus
         b. E por esta razom os dous Hercolles primeiros, cuydando cada hũũ seer aquelle, poseronse assy **nome** (OP)
         put.3PL-REFL thus name
         'And for this reason, the two first Hercules, each thinking that he was the one, named themselves thus'

(37)   [Rocas went with Tarcos and married one of his daughters and had two sons with her. Then he went back to the cave (where he had lived with the dragon).]
         a. e fizo una torre sobraquella cueua. e moro alli ya quanto. Depues que el murio fincaron **sos fijos** alli (OSp)
         remained.3PL his sons there
         'his sons remained there'
         b. e fez hi hũa torre sobre aquella cova. E despois se foy andar pello mundo como da prymeira. E ficaron ally **aquelles dous seus filhos** (OP)
         remained there those two his sons
         'those two sons of his remained there'

(38)   ['And this Juno, for this reason, got friendly with king Eurystheus who was her neighbour']
         a. Auie **contienda** con ella (OSp)
         had.3SG dispute with her
         'He had a dispute with her'
         b. e avya com ella **contenda** (OP)
         had.3SG with her dispute
         'and he had a dispute with her'

(39)   ['and he buried his brother Priamus in Pádua (...) and after he had buried him, he became very powerful']
         a. e gano **una tierra** per fuerça (OSp)
         won.3SG a land by force
         'and won a land by force'

b. e guaanhou per força **hua terra** (OP)
won.3SG by force a land
'and won a land by force'

[5] CONCLUSION

The differences in information structure that can be derived from the different descriptions of Old Spanish and Old Portuguese are reflected in the parallel corpus. They are captured by three main principles: (a) New information in Spanish can, but need not be, preverbal, and this leads to a larger number of OV sentences. (b) Old Portuguese has sentence final focus stress, thus post verbal Adverbs, PPs and Objects ordered according to this principle and (c) Spanish has a non-focus position below the verb reflected in the VSO order, while Portuguese has a non-focus position preverbally which does not exist in Spanish. This leads to a relatively larger number of SV in Old Portuguese in cases where the subject is old information.

REFERENCES

Castillo Lluch, Mónica. 1998. La interpolación en español antiguo. In *Actas del Cuarto Congreso Internacional de Historia de la Lengua Española* , 409–422. Logroño: Universidad de la Rioja.

Cintra, Luís Filipe Lindley. 1951. *Crónica Geral de Espanha de 1344. Edição Crítica do Texto Português. Vol. I-III*. Lisboa: Imprensa Nacional - Casa da Moeda.

Costa, João. 2004. *Subject Positions and Interfaces. The Case of European Portuguese.* Berlin: Mouton de Gruyter.

Eide, Kristine Gunn & Ioanna Sitaridou. 2014. Contrastivity and Information Structure in the old Ibero-Romance languages. In K. Bech & K. Eide (eds.), *Information Structure and Syntactic Change in Germanic and Romance Languages*, 377–412. Amsterdam/Philadelphia: John Benjamins.

Fontana, Josep. 1993. *Phrase Structure and the Syntax of Clitics in the History of Spanish.* Philadelphia: Univ. of Pennsylvania dissertation.

Haug, Dag, Hanne Martine Eckhoff & Eirik Welo. 2014. The theoretical foundations of givenness annotation. In K. Bech & K. Eide (eds.), *Information Structure and Syntactic Change in Germanic and Romance Languages*, 17–52. Amsterdam/Philadelphia: John Benjamins.

Kaiser, Georg. 2002. *Verbstellung und Verbstellungswandel in den romanischen Sprachen*. Tübingen: Max Niemeyer Verlag.

Martins, Ana Maria. 2002. The loss of IP-Scrambling in Portuguese: Clause structure, word order variation and change. In D. Lightfoot (ed.), *Syntactic Effects of Morphological Change*, 232–248. Oxford University Press.

Martins, Ana Maria. 2011. Scrambling and Information Focus in Old and Contemporary Portuguese. *CatJL* 10. 133–158.

Neumann-Holzschuh, I. 1998. Transitivität als parameter für sprachlichen wandel? In H. Geisler & D. Jacob (eds.), *Transitivität und Diathese in romanischen Sprachen*, 91–103. Tübingen: Max Niemeyer Verlag.

Petrova, Svetlana & Esther Rinke. 2014. Formal properties of event-reporting sentences in Old High German and Old French. In K. Bech & K. Eide (eds.), *Information Structure and Syntactic Change in Germanic and Romance Languages*, 271–294. John Benjamins.

Petrova, Svetlana & Michael Solf. 2009. On the methods of the information-structural analysis of historical texts: a case study on Old High German. In R. Hinterhözl & S. Petrova (eds.), *Information Structure and Language Change. New Approaches to Word Order Variation in Germanic*, 121–160. Berlin: Mouton de Gruyter.

PROIEL. 2011. Guidelines for annotation of Givenness. http://folk.uio.no/daghaug/info_guidelines.pdf.

Ribeiro, Ilza. 1995. Evidence for a verb-second phase in Old Portuguese. In Adrian Battye & Ian Roberts (eds.), *Clause Structure and Langage Change*, 110–139. Oxford: Oxford University Press.

Rinke, Esther. 2007. *Syntaktische Variation aus synchronischer und diachronischer Perspektive*. Frankfurt am Main: Vervuert.

Salvi, Giampaolo. 2004. *La formazione della struttura di frase romanza*. Tübingen: Max Niemeyer.

Sitaridou, Ioanna. 2011. Word order and information structure in Old Spanish. *CatJL* 10. 159–184.

Sitaridou, Ioanna. 2012. A comparative study of Word order in Old Romance. *Folia Linguistica* 46(2). 553–604.

Zubizarreta, Maria Luisa. 1998. *Prosody, Focus and Word Order*. Cambridge MA: MIT Press.

AUTHOR CONTACT INFORMATION

Kristine Gunn Eide
Department of Literature, Area Studies and European Languages
University of Oslo
Norway
k.g.eide@ilos.uio.no