1. Introduction

During the past two decades, researchers have begun to use data from several large corpora of historical Spanish to study a number of phenomena that had previously received little or no attention. One of these topics of diachronic syntax has been that of the “raising” constructions, which is the movement of either the subject or object of the embedded clause into either the subject or object position of the main clause\(^1\). In terms of the logical possibilities, there are four types of raising constructions. These are shown in the following sentences, along with references to studies during the past two decades that have dealt with the historical development of these constructions in Spanish:

Subject to Object Raising

\(1\) a. para oír [\textit{ella} cantar]
   b. para oírla [\_ cantar]

Subject to Subject Raising
(Davies 1997a, 1997b)

\(2\) a. parece [María estar enferma]
   b. María parece [\_ estar enferma]

Object to Object Raising\(^2\)

\(3\) a. quiero [PRO hacer\textit{lo}]
   b. lo quiero [PRO hacer \_]

Object to Subject Raising

\(4\) a. es difícil [PRO leer \textit{el libro}]
   b. \textit{el libro} es difícil [de PRO leer \_]

As can be seen, historical “subject to object” and “object to object” raising have

been the focus of several data-oriented studies during the past two decades, and there has also been limited work on “subject to subject” raising.

The one raising construction in Old Spanish and Early Modern Spanish that has received no attention during this time, and which in fact has never been the focus of special attention, is the historical development of “object to subject” raising (OSR). While the OSR construction in Modern Spanish has been the focus of several recent articles (Montalbetti 1983, González 1988, Reider 1993), this is unfortunately not the case for older stages of Spanish. There are only passing references to the construction in standard historical grammars such as Hanssen (1913:256), Menéndez-Pidal (1954:352), and García de Diego (1961:360).

Aside from these historical grammars, there are only two studies that provide more than a passing reference -- Beardsley (1921:143-50) for Old Spanish, and Keniston (1937:530) for the 1500s. In each case, however, the section of the book dealing with the OSR-like construction is a very small part of a larger work on historical Spanish syntax. In addition, in neither of the two books is there any discussion of the syntax of the construction per se, but rather only a simple listing of the adjectives that take [de INF] complements.

One other shortcoming is the fact that in both cases the authors group together all cases of [ADJ de INF], whether the fronted noun phrase is the subject of the embedded clause (5a), or whether it is the object (5b):

(5) a. enojados eran de lo tomar (Crónica general de España, c1270)
   b. grieses son de fazer (Crónica general de España, c1270)

As even the very earliest generative studies have shown, however, cases like John is eager to please [John = lower clause subject] and John is easy to please [John = lower clause object] are quite different in their syntactic behavior (cf. Chomsky 1957). In the present study, we will consider these to be two very different constructions, since only cases like (5b) involve cases of OSR.

This study, then, will be the first to look specifically at OSR-like constructions in Old Spanish and Early Modern Spanish. The first task will be to consider the historical shifts in the governing adjectives that allow OSR in different historical periods. We will then focus on syntactic shifts, such as the decrease in clauses with both synthetic and analytic passives: el libro es fácil de leer, el libro es fácil de ser leído. Most importantly, we will examine the historical frequency of the “OSR” construction (el libro es difícil de leer) and the more basic structure on which it is supposedly based (es difícil leer el libro), and suggest what this might indicate about the correct analysis of the construction.

2. Database

The data for this study are taken from three large corpora of Old Spanish and Early Modern Spanish texts, which represent more than 57 million words in approximately 1100 texts. The composition of these corpora is summarized in Table 1, and a short overview is provided here. The first corpus, which I have created myself, is composed of 3,700,000 words in 74 texts from the 1200s to the 1600s (detailed information can be found online at http://mdavies.for.ilstu.edu/personal/espanol.htm). The second corpus is composed of nearly 33 million words in approximately 800 texts (mainly from the 1400s to the 1600s) from the Biblioteca Virtual (http://cervantesvirtual.com). The third corpus is the ADMYTE collection of medieval Spanish texts (Volumenes 0 and 2; http://www.admyte.com), which includes 20.8 million words in 260 texts from the 1200s-1400s.

<table>
<thead>
<tr>
<th></th>
<th>Original corpus</th>
<th>Biblioteca Virtual</th>
<th>ADMYTE</th>
</tr>
</thead>
<tbody>
<tr>
<td># texts</td>
<td># words</td>
<td># texts</td>
<td># words</td>
</tr>
<tr>
<td>1200</td>
<td>14</td>
<td>776,700</td>
<td>22</td>
</tr>
<tr>
<td>1300</td>
<td>10</td>
<td>744,200</td>
<td>12</td>
</tr>
<tr>
<td>1400</td>
<td>15</td>
<td>765,200</td>
<td>62</td>
</tr>
<tr>
<td>1500</td>
<td>19</td>
<td>745,300</td>
<td>323</td>
</tr>
<tr>
<td>1600</td>
<td>16</td>
<td>701,100</td>
<td>376</td>
</tr>
<tr>
<td>Total</td>
<td>74</td>
<td>3,732,500</td>
<td>795</td>
</tr>
</tbody>
</table>

The large size of these three historical corpora resulted in more than a thousand examples of object to subject raising with nearly 20 separate adjectives (e.g. el libro es duro / ligero / bueno de leer). This rich database – far beyond anything previously collected – will allow us to accurately track several diachronic shifts during the period 1200-1700, and to then use this data to analyze historical shifts in the underlying structure of the clause.

3. Core lexical items and frequency of OSR

The first question that we will consider is whether there has been either an overall increase or decrease in OSR in Old Spanish and Early Modern Spanish. Our initial examination of the historical texts might lead us to believe that OSR was not common in Old Spanish, since there are not examples with either of the two most common adjectives from Modern Spanish (fácil or difícil) until the second half of the 1400s:

(6) a. todos principios son graues & difíciles de aver por esta razón (Tratado de la música, c1440-60)
   b. e por tanto es más fácil de elegir e de obrar (Visión deleitable, c1460)

The absence of cases with fácil and difícil is deceiving, however, since neither of these lexical items were common in any context previous to the 1400s, and we would therefore not expect to find any cases with the OSR construction.
The best approach is to expand our search to consider all of the Old Spanish adjectives that are related to the Modern Spanish fácil and difícil, and perhaps even expand this to the general category of "ease of accomplishment" (fácil / posible / difícil / imposible). Once we do this, we find that indeed OSR was common in Old Spanish. In the ADMYTE collection of Old Spanish texts the adjectives that are related to the Modern Spanish fácil that appear are ligero (112 occurrences), leue (1), and lluiano (45 examples, but 44/45 are from the Sevillana medicina, c1380).

(7) a. ct dio en el arbol ... fue muy ligero de derrirbar (Conde Lucanor, 1330)
b. quien bien quisiere mirar sera leue de concordar (Cura de la piedra, 1498)
c. la carne assada esfuerza mucho y es lluiano de moler (Sevillana medicina, c1380)

Adjectives related to the Modern Spanish difícil that take OSR in the ADMYTE corpus are graue (112 cases), fuerte (36), duro (23), griene (15), dificiloso (12), pesado (7), ásporo (7), caro (6) and trabajoso (5):

(8) a. fazen esso mismo & esto seie graue de prouar (Ordenamiento de Alcalá, c1350)
b. [un pan de ordio] que serie muy fuert de comer (General Estoria, c1270)
c. e el ovogolo mucho a duro de vender (Engaños, 1253)
d. [las despensas] eran grandes a desmesura & grienes de complir (Crónica general de España, c1270)
e. pero sus desembalsuras eran bien dificilosas de hurtar (Grimaite y Gradisida, 1495)
f. porque son cartellinigas y pesadas de moler (Sevillana medicina, c1380)
g. cuyo comienço es duro & aspero de soportar (Eopete ystoriado, 1488)
h. la qual demanda me es muy cara de fazer (Crónica general de España, c1270)
i. si quebrare con los dientes cosa trabajosa de quebrar (Lilio de medicina, 1495)

There are also many cases with largo (9), though most of these are with the set phrase largo de contar (9b):

(9) a. Aquesti emperador era muy largo de dar sus donos (Crónica de los emperadores, tr. 1377)
b. muchas Razones que quieren que serian largas de contar (Atalaya de las Crónicas, c1450)

Related to the "synonyms" of the Modern Spanish fácil and difícil are the adjectives pos(s)ible and impos(s)ible, both of which also refer to "ease of accomplishment" in the continuum "fácil / posible / difícil / imposible". As soon as pos(s)ible and impos(s)ible enter the language as common lexical items (early 1400s), we find examples with the OSR construction. There are eight examples in the 1400s, and this increases to 16 examples in the 1500s, e.g.:

(10) a. la mejor orden e manera que a ellas fue posible de reçebir (Visón deleytable, c1460)
b. Lo cuarto, si fuese o imposible de hacer o dificultosísimo (Summa de tratos y contratos, 1571)

Once we consider the full range of adjectives in the "ease of accomplishment" category (fácil / posible / difícil / imposible) we find that the overall frequency of OSR has not varied much since at least the 1300s. In the original 3,700,000 word historical corpus, there were 19.3 cases of a synonym of fácil and difícil per million words in the 1200s, 56.4 in the 1300s, 31.4 in the 1400s, 43.0 in the 1500s, 37.1 in the 1600s, and 38.8 in the 1700s4. As these data show, while there have been some fluctuations in the frequency of "core" OSR between 1300 and 1700, there has been no general increase or decrease in use during this time.

4. Shifts in the class of governing adjectives

As we have seen, the frequency of OSR with the "ease of accomplishment" adjectives in Old Spanish and Early Modern Spanish is roughly comparable to the situation in later stages of Spanish. When we examine two other categories of adjectives, however, we find that the situation in Old Spanish and Early Modern Spanish is quite different than it is today. First, there are a range of adjectives referring to "evaluation" of the action -- such as agradable, interesante, importante, and divertido -- which allow (at least limited) OSR in Modern Spanish.

(11) a. un espacio distendido ... [que] resulte agradable de ver (España: Baleares-Mundo)
b. cuyos efectos sin duda serian muy interesantes de estudiar (Guatemala: Prensa Libre)
c. lo promuevan como un destino importante de visitar (Nicaragua: Prensa)
d. [el libro] resulta ameno y divertido de leer (España: Barcelona: Vanguardia)

In Old Spanish and Early Modern Spanish, there are almost no examples of OSR with any of these lexical items. In the more than 57 million words of texts
from the 1200s-1600s, there are only three isolated examples with these four adjectives:

(12) a. un estrado bien compuesto, **agradable de mirar** (Poesías [Rodríguez del Padrón], c1440)
   b. la qual bien es de creer que le fuera más **agradable de mirar** (Espejo de príncipes y cavalleros, 1555)
   c. cosas todas muy **importantes de saber** (Predicación del Evangelio en las Indias, 1588)

This is in spite of the fact that these adjectives were quite common as general lexical items during this period. There are frequent examples of **agradable** from the 1300s, **importante** and **divertido** from the 1400s, and **interesante** from the 1500s. The data clearly show, then, that even by the latter part of Early Modern Spanish, OSR had not yet extended in any meaningful way to the "evaluative" adverbs.

Just as there is a category of adjectives that allow OSR in Modern Spanish but did not in earlier stages, the converse is also true -- there is an important category of adjectives that allowed OSR in Old Spanish and Early Middle Spanish, but which do not in Modern Spanish. This is with the category of **bueno/malo** and related adjectives (**delectoso, dulce, fermoso/hermoso** and **sabroso** (positive) and **feo y horrible** (negative)), e.g.:

(13) a. **estas quemas son buenas de fazer** en el mes de marzo (Animales, 1250)
   b. era logar muy sabroso & muy **delectoso de comer** & de beuer (Primera Crónica General, c1270)
   c. & son muy fermosos & muy **dulces de fazer** (Libro de la caza de las aves, c1385)
   d. fallaron muchas tiendas bien **fermosas de ber** (Embajada a Tamarlán, c1403)
   e. Entra el mal camino, y **malo de construir** por lo pendiente y deleznable del terreno (Diario [Joveillanos], c1780)
   f. Los xipisitanos non tienen ley & esto es vna cosa fea de dezir (Cancionero castellano de París, 1444)
   g. **moraron gigantes grandes & horribles de mirar** (Propiedades de las cosas, 1494)

Looking at the core cases **bueno** and **malo**, we find that in the Biblioteca Virtual corpus there are 7 examples of OSR with a form of **bueno** or **malo** in the 1300s (8.8 cases per million words), and the figures for subsequent centuries are 1400s: 6.8 per million words (16 examples), 1500s: 4.9 (90), 1600s: 1.4 (15), and 1700s: 0.8 (7). As can be seen, the frequency of OSR with **bueno** and **malo** decreases slightly from the 1300s through the 1500s, and then decreases markedly after the 1500s. This fits in well with the situation in Modern Spanish, in which OSR with **bueno/malo** is unacceptable for most speakers (### estas manzanas son buenas de comer). Although the data are clear in showing the demise of OSR with **bueno/malo** after the 1500s, it is less clear why this happened. It is doubtful that there is any general cross-linguistic universal tendency away from OSR with these adjectives, as evidenced by the still-acceptable "**these vegetables are good / bad to eat**" in Modern English.

5. Passives

Let us now turn from an analysis of lexical issues to a consideration of the syntactic changes with the OSR construction in earlier stages of the language. The first issue that we will consider concerns sentences like (20)-(21):

(14) a. el liuro es muy ligero :: de se leer
   b. el liuro es muy ligero :: de **ser levido**

(15) a. ?? el libro es muy fácil [[ de leerse ]
   b. ?? el libro es muy fácil [[ de **ser leído** ]

What we will show is that cases like (20) were much more frequent in earlier stages of Spanish, and that they are highly marginal in Modern Spanish. We will postpone the discussion of why this might have changed until we discuss the nature of the underlying clauses in Section 6.

As mentioned, most speakers of Modern Spanish reject examples like those in (21) above, and these intuitions are supported by the data from a large corpus of Modern Spanish (see http://mdavies.for.ilstu.edu/personal/texts.htm). In this corpus, which comprises 43 million words of text from the 1900s, there are only two cases of **se** with **fácil** and **dificil** (and its "synonyms"), or one in every 21.5 million words of text:

(16) a. un desorden institucional **dificil de solucionarse** (Venezuela, Universal)
   b. la verdad es muy **dificil de aprenderse** (Chile, Hoy)

Likewise, there are only five cases of **ser** + PP with **fácil** and **dificil** (and its "synonyms") in the 43 million word corpus, or one in every 8.6 million words of text); e.g.:

(17) a. **las propiedades eran difíciles de ser estudiadas** experimentalmente (España, El País)
   b. un cúmulo de intuiciones **fáciles de ser intuidas** (Argentina, Sobre heroes y tumbas)

In Old Spanish and Early Modern Spanish, however, these constructions are much more common. In the ADMYTE corpus of Old Spanish, for example, there are twelve examples of the **se** construction with synonyms of **fácil** and
6. Clause structure

The final issue that we will consider deals with two competing syntactic analyses of the OSR construction. One analysis, which we have assumed to this point, is that there was actually movement from the subordinate clause to the main clause, as seen in the following:

(22a) _es difícil [ PRO leer el libro ]
    b. el libro es difícil [ de PRO leer_ ]

However, an alternative analysis is that cases like (23b) are simply a variant of more common sentences like (23a), in which the subject (el libro) is base-generated in the main clause subject position, and there is no raising or movement at all.

(23a) el libro es fácil
    b. el libro es fácil [ de leer ]

How can we best decide between these two analyses? As Reider (1993) explains, the OSR model (22) suggests that raising (22b) only occurs if the more basic non-OSR structure is also possible (22a). In other words, we will not find cases of “el libro es fácil de leer” unless there is an even higher frequency of the more basic “es fácil leer el libro”. In fact, from the point of view of Modern Spanish (and Modern English as well), we would find it surprising if there was an almost complete absence of basic sentences of the type “es ADJ leer el libro”.

Yet this strange state of affairs is precisely what we do find in older stages of Spanish, and this suggests that perhaps there was a difference in clause structure at that time. In older stages of Spanish, there were many adjectives that had a high frequency of constructions like “el libro es difícil de leer” (hereafter “NP ser ADJ de V”), without having any cases of the more “basic” constructions es difícil leer el libro (hereafter “ser ADJ V NP”).

For example, there are 25 cases of [NP ser ADJ de V] with fuerte from the ADMYTE corpus, but no cases at all of [ser ADJ V NP] (e.g. **es fuerte matar al rey)

(24a) El puercos fresco porque es fuerte de disimular (Arte cismia, 1423)
    b. serie después muy cara & muy fuerte de desfazer (Castigos y documentos, 1293)

With duro there are 24 cases of [NP ser ADJ de V], but again no cases of [ser ADJ V NP] (e.g. **es duro comprar pan).

(25a) porque avn vuestras cartas son tan duras de auer (Claro varones de Castilla, 1485)
b. en el comienzo son los omnes **duros de meter** al trabajo (General estoria, 1272)

With **ligero** there are 46 cases of [NP ser ADJ de V], but no cases of [ser ADJ V NP] (e.g. **es ligero leer ese libro**):

(26)a. **cuando la yra ha razon sabida es ligera de faze** (Bocados de oro, c1253)

b. **no es menester de dezir como se faze; por que es ligero de faze** (Libro del saber de astrologia, 1276)

Finally, with **liviano** there are 24 cases of [NP ser ADJ de V], but once again no cases of [ser ADJ V NP] (e.g. **es liviano escriuer la carta**):

(27)a. **gallina cozida es de gran esfuerzo y mas liuanova de moler que la assada** (Sevillana medicina, c1380)

b. **los que son liuanos de cozer son mejores que los otros** (Sevillana medicina, c1380)

Overall, if we look at all of the “synonyms” of fácil and dificil in the ADMYTE corpus, we find 276 cases of [NP ser ADJ de V], but only 10 cases of [ser ADJ V NP]. What this suggests is that in Old Spanish, cases like el libro es ligero [de leer] are simply a more “elaborate” version of el libro es ligero, and that the NP is base-generated in subject position, rather than being derived from something like es ligero [leer el libro].

What is interesting is that it appears that in the 1500s-1600s, there was a shift from the “base-generated” structure to the “raising” construction. Again, the evidence for this comes from a comparison of the frequency of the [ser ADJ V NP] construction with the frequency of the [NP ser ADJ de V] construction during this period. Looking at the aggregate total for all of the “synonyms” of fácil and dificil, as well as (im)possible, in the 3,700,000 word corpus, we find that in the 1400s, 78% (25/32 cases) were the [NP ser ADJ de V] construction (el libro es fácil de leer), as opposed to the much less common [ser ADJ V NP] (es fácil leer el libro). In the 1500s this had decreased to 51% (32/63 cases), and finally to 25% (26/102 cases) in the 1600s.

What all of this indicates is that cases like “es fácil leer el libro”, which are so common in Modern Spanish (and Modern English) were relatively uncommon in Old Spanish, and apparently had little to do with constructions like el libro es fácil de leer. Over time, however, the relative importance of the [ser ADJ V NP] construction increases, to the point in the 1600s where it accounts for 75% of all cases. The prominence of the [ser ADJ V NP] construction (es fácil leer el libro) suggests that this construction became more tightly linked to and probably influenced the existing [NP ser ADJ de V] “raising” construction in Early Modern Spanish. It can easily be imagined that this would translate into a scenario in which it makes sense to talk of “raising” (el libro es difícil de leer) being a derivation of the more basic es difícil leer el libro, at least after the 1500s or 1600s.

One other piece of evidence to support this scenario is the data from Section 5, in which we saw that passives like “el libro es difícil de leer / ser leído” were quite common in older stages of the languages, but very rare in Modern Spanish. In this section we have proposed that Old and Middle Spanish “el libro es fácil de leer” derives from the simple “el libro es fácil”. Just as it would be common to have a prepositional phrase with a passive in cases of base-generation like:

(28)a. **ella estaba cansada**

b. **ella estaba cansada [de ser castigada]**

(29)a. **las joyas son muy caras**

b. **las joyas son muy caras [como para verse tan feas]**

it would also be common to have a passive with the base-generated NP in cases like:

(30)a. **el libro es fácil**

b. **el libro es fácil [de leerse / ser leído]**

However, once the construction ceases to be seen as a case of base-generation, as we have suggested is this case in this section, then the frequency of passivized modification would decrease somewhat, as in fact has occurred.

**7. Conclusion**

To summarize, this is the first study to provide extensive data on the “object to subject raising” construction in Old Spanish and Early Modern Spanish, and is based on more than one thousand examples in more than 57 million words of text from the 1200s to the 1700s. We have shown that the frequency of “OSR” with the “ease of accomplishment” adjectives (fácil, posible, difícil, imposible) has remained fairly constant over time, but it has decreased significantly with the bueno/malo adjectives, and only began to increase with the “evaluative” adjectives (e.g. interesante/important) in the latter part of the Early Modern Spanish period.

In terms of syntactic change, we have seen that there was a marked decrease in passives during the past 400-500 years. We have also suggested that the Old Spanish construction is best seen as a variant of simple adjectival clauses in Old Spanish (el libro es fácil [de leer]), although this began to change towards true OSR in Early Modern Spanish. Hopefully, all of these findings help to show how a study that relies both on an extensive database, as well as a simple model of clause structure, can shed light on what was previously a very poorly understood aspect of historical Spanish syntax.
Notes

1. In the GB model of syntax (as well as later derivations of this model), "raising" as a specific syntactic rule has been phased out, in large part to satisfy theory-internal constructs such as the "i-heta criterion" (see Reider 1993 and Montalbetti 1983 for a discussion of this applied to Spanish). However, in several other models of syntax, such as functional, typological, and relational grammar, these constructions continue to be analyzed as "raising".
2. It may seem strange to analyze "clitic climbing" sentences such as these as a type of "raising". However, it does share important similarities with the other raising constructions. The object of the embedded infinitive clearly moves into the object position of the main clause, and it obeys many of the semantic and pragmatic constraints of a base-generated main clause object.
3. There is some overlap between the texts in the three different corpora. In subsequent discussions of the data from the corpus, however, I make it clear which of the three corpora provided the data under discussion.
4. Although these frequencies are based on just the smaller 3,700,000 word corpus, the overall frequency holds true for the larger Biblioteca Virtual and ADMYTE corpora as well.

References
