**Tornar and volver: The interplay of frequency and semantics in compound tense auxiliary selection in Medieval and Classical Spanish**

Malte Rosemeyer, Albert-Ludwigs-Universität Freiburg

This paper examines how the interplay of frequency and semantics may have influenced the directionality of the change of the compound tense auxiliary system in Spanish, taking the near-synonymous verbs *tornar* and *volver* ('to return') in 16th century as examples. There is a significant contrast in the auxiliary selection of the two verbs that can be explained by taking into account the differences in semantic usage of the two verbs, as well as frequency effects. On the one hand, the higher degree of motional usage of *volver* leads to the more frequent selection of *ser*. On the other hand, the higher usage frequency of *volver* led to a conservation of the usage of the *ser*-auxiliary with this verb. In consequence, the article argues for a joint influence of semantic usage and frequency effects on the development of auxiliary selection in Spanish.

1. Introduction

Ever since Perlmutter’s (1978) Unaccusative Hypothesis, intransitive compound tense auxiliary selection has received much attention in linguistic research. This is due to its status as a heuristic of split intransitivity. In the last decade, however, researchers have started to highlight the fact that – both cross-linguistically and within a single language – there is much variation in the use of HAVE or BE as compound tense auxiliaries. Hence, while many authors agree that the semantics of the verb occurring as the construction’s participle is the main predictor of auxiliary selection, it has been acknowledged that verb semantics can probably not account for all of the variation regarding auxiliary selection (e.g., Aranovich 2003; Smith 2007).

A possible solution for this problem is to analyse auxiliary selection from a diachronic perspective. Synchronic gradience is often determined by diachronic “gradualness” (cf. Traugott & Trousdale 2010). It has been proposed that the instability of the auxiliary selection behaviour of verbs from certain semantic classes is a result of the directionality of the analogical spread of HAVE to contexts associated with BE (Shannon 1990; 1996). This change is well documented in Spanish, where *ser* (BE) as a compound tense auxiliary was gradually replaced by *aver* (HAVE) until it eventually disappeared in the 17th century. While verbs of continuation of a pre-existing state like *quedar* ('to stay')

---

1 The term “compound tense” is used here for both *aver* + PP and *ser* + PP-constructions, even though in many cases the latter have to be analysed as resultative constructions (cf. Mackenzie 2005).

2 In the following HAVE and BE in capital letters will be used to denote the corresponding lexemes, whatever the language debated, while italised small letters will be used to denote the language-specific lexeme (e.g., *aver*).
and *fincar* (‘to stay’) show some variation in terms of auxiliary selection in the 13\textsuperscript{th} and 14\textsuperscript{th} century, they categorically select *aver* after the middle of the 15\textsuperscript{th} century. This development coincides with the spread of the *aver*-auxiliary to the semantic classes of change of state and motion verbs (Benzing 1931; Aranovich 2003).

If synchronic variation is determined by diachronic development, a model of auxiliary selection has to explore factors that are not obvious in synchrony. In particular, the token frequency of individual lexemes has been argued to have played a decisive role in the diachronic pathway of the development of auxiliary selection in English (Smith 2007). Verbs that appear very often with *BE* are less likely to start preferring *HAVE* than less frequent verbs. Auxiliary selection can thus be compared to multiple argument realisation in that seemingly idiosyncratic factors affect individual lexemes, while a global principle of the mapping of semantic to syntactic structure influences auxiliary selection (Rappaport Hovav 2008). Therefore, variation regarding auxiliary selection within the same semantic verb classes, and even between verbs with nearly synonymous meanings, should be expected. This type of variation can be modelled by quantitative approaches to language change. The goal of such an approach to auxiliary selection would be to clarify to what degree variation can be explained by verb semantics (as a global factor) and verb lexeme token frequency (as a specific factor), respectively.

In this paper, the influence of the interplay of verb semantics and lexeme frequency on the development of Spanish auxiliary selection is illustrated using a corpus-based case study on the two near-synonymous Spanish verbs *tornar* and *volver*. In the next section, an introduction to auxiliary selection in Medieval and Classical Spanish is given. In section three, the respective auxiliary selection behaviour of *tornar* and *volver* in 16\textsuperscript{th} century historiographical writing is compared. *Tornar* appears to select *aver* significantly more often than *volver*. In section four, a statistical analysis of the differences in semantic usage of the two verbs is conducted. In section five, the domain of investigation is expanded to include tokens from historiographical texts between the 13\textsuperscript{th} and 17\textsuperscript{th} century in order to evaluate the interplay of semantic usage and token frequency in the development of the auxiliary selection behaviour of *tornar* and *volver*. The differences in semantic usage, as well as usage frequency of the two verbs affect their degree of selection of *HAVE*. These results are summarised in the last section.
2. Compound tense auxiliary selection in Spanish

The Romance languages are an interesting field for the study of compound tense auxiliary selection since they simultaneously exhibit both divergent and uniform behaviour. Contemporary French and Italian display split auxiliary selection for intransitive verbs: the auxiliary HAVE is used for verbs denoting controlled activities, and BE for telic verbs of motion and change of state. In the Iberian languages and Romanian, however, HAVE is used as the only perfect auxiliary with few exceptions, even though in earlier stages of these languages the auxiliary BE was used for telic verbs (e.g., for Spanish, Benzing 1931; Keniston 1937; for Catalan, Batlle 2002). With regard to Spanish, Benzing (1931) notes that in the earliest source texts from the 12th and 13th century, a gradual reduction in the use of ser (BE) as a compound tense auxiliary in favour of aver (HAVE) can already be observed. By the late 17th century, ser had lost all of its productivity as a compound tense auxiliary. Later examples of ser + PP have been analysed as “frozen” discourse fragments (Stolova 2009).

Based on Benzing’s (1931) findings, a number of researchers have argued for the lexical aspect of the verb used for the participle as the main predictor of Medieval and Classical Spanish compound tense auxiliary selection (Elvira González 2001; Aranovich 2003; Mackenzie 2005; Castillo Herrero 2006; Mackenzie 2006; Mateu 2009). For example, Castillo Herrero (2006) claims that compound tense auxiliary selection in Medieval Spanish is determined by the dynamity [+/- DYN] and delimitation [+/- DEL] of the event expressed by the participle. Hence, stative verbs (which are [- DYN], [- DEL]) usually select aver, while change of state verbs (which are [+ DYN], [+ DEL]) select ser. Such approaches are based on the projectionist assumption that once the grammatically relevant facets of verb meanings have been extrapolated, mapping rules from semantics to syntax can be established to fully explain phenomena like auxiliary selection (Levin & Rappaport Hovav 2005: 7-9).

However, projectionist approaches face various problems when dealing with linguistic variation. In particular, assuming mechanistic mapping from semantics to syntax does not fully explain the variation in auxiliary selection for one and the same lexeme. For instance, while verbs denoting controlled non-motional activities like trabajar (‘to work’) appear to categorically select aver in Medieval and Classical Spanish, the picture

---

3 A number of Italian dialects contradict this generalisation, displaying a spread of essere in certain paradigms and with activity verbs (Tuttle 1986; Bentley 2006).
becomes more complicated when taking into account verbs like *quedar* (‘to stay’) which denote the continuation of a pre-existing state. Thus, in the following example *quedar* is used with both HAVE and BE:

(1) E asy andudo souetyendo a todas quantas mugeres de altos omnes ouo en grecia. otrosy a algunos que noble men have-PST.MSG in Greece. also to some that aujan quedado en grecia en tal manera que todas las mas delas grandes mugeres de grecia se leuantaron contra sus maridos con algunos delos que enla tierra eran quedados [SUM] have-PST.MSG stay-PTCP.M.PL 'And so he went about subverting all of the women of noble men that there were in Greece, and also some (of the noble men) who had stayed in Greece rose against their husbands with some of those (noble men) who had stayed in the country'

Following Levin & Rappaport Hovav (2005), in this case the variation may have two sources, depending on the lexemes in question. First, the lexeme may be used with different meanings, in which case different projection rules apply. For instance, when used with an adjectival complement, *quedar* could mean ‘to become (something)’ already in Medieval Spanish, as illustrated by the following example from the *Gran Conquista de Ultramar* (c. 1293), cited by Vergara Wilson (2009: 274):

(2) E en aquella primera noche delas bodas que el conde & la condesa durieron queda ella preñada count and the countess sleep-PST.MSG she pregnant 'And when the count and the countess slept (together) in the first night of the wedding, she became pregnant'

Since change-of-state verbs display a higher degree of telicity than continuation-of-state verbs, *quedar* in the change-of-state meaning would likely be more prone to auxiliation with *ser* than *quedar* in the continuation-of-state meaning. However, in example (1) both tokens of *quedar* are apparently used with the same meaning. Levin & Rappaport go on to argue that, if the different usage of the lexeme cannot be explained by a difference in meaning, pragmatic factors like information structure may determine the variation (Levin & Rappaport Hovav 2005: 216-219). From this standpoint, the variation regarding auxiliary selection as shown by example (1) could be confined to differences in the contexts in which the two tokens appear. It should be noted, however, that the contexts of the two tokens of *quedar* in this example are actually quite similar, both
involving a relative clause, a locative adverbial, and even the same lexeme for their subject referents (algunos).

Sorace’s (2000) Auxiliary Selection Hierarchy (ASH) accounts for the different degrees of variation between and within certain semantic verb classes with regard to auxiliary selection. She claims that, cross-linguistically, stative verbs display a greater degree of variation in auxiliary selection than controlled activity verbs and change of state or motion verbs (2000: 869-870). In particular, controlled activity verbs almost exclusively select HAVE, while change of state and motion verbs show a relatively consistent selection of BE.

According to Sorace, the greater degree of variation found with stative verbs can be explained by the fact that these verbs are aspectually underspecified and thus show a greater sensitivity to syntactic and pragmatic contexts than the other verb classes mentioned above (2000: 870). For this reason, stative verbs are more susceptible to processes of “template augmentation” (Rappaport Hovav & Levin 1998). The meaning of a stative verb is easily altered when introducing linguistic elements like adverbials that affect meaning at the sentence level. For instance, the auxiliary selection of Italian stative verbs differs according to the telicity imposed by the adverbials in a given clause. In example (4), the aletic verb durare is telicised in a template augmentation process involving the adverbial fino al 1974 (‘until 1974’), resulting in the selection of essere over avere:

\[
\begin{align*}
(3) & \quad \text{Lo spettacolo ha durato essatamente due ore senza pausa.} \\
& \quad \text{The show has lasted exactly two hours without pause.} \\
& \quad \text{(http://lifeasorkish.blogspot.com/2010_11_23_archive.html, last access 21.2.2012)}
\end{align*}
\]

\[
\begin{align*}
(4) & \quad \text{Lo spettacolo è durato fino al 1974 senza pausa.} \\
& \quad \text{The show has lasted until 1974’ (http://it.wikipedia.org/wiki/Larry_Wilcox, last access 21.2.2012)}
\end{align*}
\]

As noted above, in Spanish, the process by which aver was generalised appears to have affected stative verbs first, and only afterwards change-of-state and motion verbs\(^4\). However, a closer look at the historical data reveals variation in auxiliary selection in

\(^4\) There is evidence for similar processes in Old Neapolitan (Cennamo 2008), Early Modern Catalan (Batlle 2002) and Canadian French (Sankoff & Thibault 1977). The similarity of the process of substitution of BE by HAVE in various Romance languages points to its being conditioned by macro-factors and its origin in Late Latin configurations (Cennamo 2008).
the earliest texts even within the seemingly uniform telic semantic classes. While the
telic verb *venir* was usually auxiliated with *ser* in the 13th century (5), examples of *aver*
*venido* like (6) can be found:

(5) Ellos respusiéronle que pues que en aquel logar

*que* que *en* *aquel* logar

(6) Orosio [...] diz que non cuentan aquel Pompeo e

*que* *que* *aquel* Pompeo e

Moreover, while the general trend of the actualisation process follows the predictions of
the Auxiliary Selection Hierarchy, there are a number of exceptions. For instance, in
Benzing’s (1931) list of the dates of the latest occurrences of selected verbs with the *ser*
 auxiliary, motion verbs like *descender* (‘to descend’) selected *aver* categorically as early as
in the 15th century, while other motion verbs like *pasar* (‘to pass by’) only began to
exclusively select *aver* in the 17th century (1931: 443).

As argued above, variation within a given semantic verb class may be explored by
differentiating semantic usage of the verbs. However, it has been argued that factors
related to the historical process of substitution of *ser* by *aver* influenced gradient
auxiliary selection of intransitive verbs at synchronic points in time. In particular,
Aranovich (2003) hypothesises that frequency of occurrence influenced the spread of
*aver*, in that “the most frequent verbs are the last ones to lose their ability to combine
with *ser*” (30). Moreover, he claims an influence of the interplay between frequency and
semantics in the development of the auxiliary selection in Spanish: “A possible scenario
is one in which the two approaches complement each other, with the most frequent
verbs resisting the expansion of *haber* the longest within each lexical semantic class”
(ibid.).

Aranovich’s hypothesis is implicitly based on the assumption that high absolute
frequency leads to “conserving effects” as described by Bybee (2006; 2010) and Bybee &
Hopper (2007): High token frequency is known to result in a stronger cognitive representation of a specific complex linguistic item because that item is repeatedly accessed holistically. The stronger cognitive representation of that element leads to “entrenchment”: its morphosyntactic form is less likely to be altered. Consequently, the linguistic item grows more and more autonomous from the construction to which the item belonged. If the item’s original construction is altered by a grammatical change, the linguistic item is less affected by that change due to its independence from the construction:

...frequent forms resist regularizing or other morphological change with the well-known result that irregular inflectional forms tend to be of high frequency. Assuming that regularization occurs when an irregular form is not accessed and instead the regular process is used, it is less likely that high-frequency inflected forms would be subject to regularization. (Bybee 2010: 25)

In Aranovich’s (2003) scenario, the asymmetries in auxiliary selection within one and the same verb class would appear to be the result of frequency-induced micro-changes that affect to a certain degree the directionality of the macro-change (the replacement of *ser* by *aver*). In order to test this hypothesis, the auxiliary selection behaviour of verbs within one and the same semantic class would have to be investigated. The behaviour of single verbs would have to be compared to the general trend in order to observe frequency effects. In the following section, an analysis of the two Spanish verbs *tornar* and *volver* is conducted that demonstrates the joint influence of verb semantics and usage frequency on the development of auxiliary selection in Spanish.

3. Variation in auxiliary selection between *tornar* and *volver*

In a first step of the study, the auxiliary selection behaviour of the verbs *tornar* and *volver* in 16th century historiographical texts was compared. *Tornar* and *volver* have essentially the same semantic value. Both verbs can carry the meaning ‘to come back’, as in (7) and (8). They can also appear with an adjectival or subject complement, thus taking on the meaning ‘to turn into (something)’, as in (9) and (10). In addition, they can be used as an iterative periphrasis, as in (11) and (12). Both *tornar* and *volver* were originally used as verbs of change of position (Melis 2006: 883), from which the three meanings examined here were derived. The periphrastic forms *tornar + a + infinitive* and *volver + a + infinitive* are the result of a grammaticalisation process in which the original motional meaning of these verbs was lost.

---

5 “Iterativity” is defined here as the sum of repetitive and restitutive “again”-meanings (cf. Wälchli 2006).
(7) y así, como creyó que Piñarro se avía
...and so, since he believed that Pizarro had returned
return-PTCP.MSG to Los Reyes, himself believe-PTCP.MSG in order to enter
al Quito [NAT]
to the Quito
'...And so, since he believed that Pizarro had returned to Los Reyes, he prepared to enter Quito'

(8) Estando el rey en Perpinian después de ser vuelto
be-PROG the king in Perpignan after be return-PTCP.MSG
de Monpellier... [ANA]
from Montpellier
'While the king was in Perpignan after having returned from Montpellier,...'

(9) hacia el suro, donde ha habido grandes contrataciones y se han tornado muchos negros cristianos...
...to the south, where have been great recruitments and many Blacks have become Christians...

(10) ...y a los judíos y moros que se habían vuelto
...and to the Jews and Moors that themselves had become Christians
return-PTCP.MSG Christians much time have-PTCP.MSG them
hacían volver á ser moros... [CEC]
make-PTCP.MSP return to be Moors...
'...and they forced the Jews and Moors that had become Christians to become Moors again...

(11) Haviste arrojado en el infierno, y agora
have-yourself-PTCP.MSG throw-PTCP.MSG in the hell, and now
has vuelto a resucitar en este mundo... [NUE]
return-PTCP.MSG to be reborn in this world
'You had thrown yourself in hell, and now you have been born again into this world...

(12) su hija madama Margarita, que avía otra vez
...his daughter Madame Margarita, who had another time
return-PTCP.MSG to become widow of the duke of Saboya [CAT]
tornado a enviadir del duque de Saboya [CAT]
return-PTCP.MSG to become widow of the duke of Saboya
'...his daughter Madame Margarita who again had become a widow (the widow) of the Duke of Saboya'

The 16th century is an interesting period for the comparative study of the auxiliary selection behaviour of tornar and volver due to the diverging historical paths of the two verbs. While both tornar and volver can be found in the oldest Spanish texts, tornar was used with a much higher frequency until the first half of the 16th century. Volver, conversely, had a very low usage frequency until the beginning of the 16th century. It then started gaining ground, eventually replacing tornar by the first half of the 17th century (Stolova 2005: 72; Melis 2006: 912). The following figure illustrates this development6:

---

6 Since the CORDE is not annotated for verb lemmata, it is difficult to obtain frequency information for verb lemmata from it. Therefore, the frequency information for tornar and volver was extracted from the annotated 100 million-token diachronic Corpus del Español (Davies 2002-).
Comparing the token frequencies of *volver* to other Spanish motion verbs after the 17th century, *volver* is one of the highest-ranking motion verbs in terms of absolute token frequency. *Tornar* and *volver* thus present an ideal testing ground for the hypothesis that token frequency influenced the development of auxiliary selection in Spanish.

The present study was conducted taking into account only historiographical texts in order to avoid genre-specific skewing of the data. While this method only allows for limited generalisations for the entire language, it enhances the comparability of the tokens extracted from the corpus.

All of the 16th century tokens of *volver*- and *tornar*- participles in historiographical texts in the Corpus diacrónico del español (CORDE, Real Academia Española 2010) were extracted. The enormous orthographic variation (inflections, capital letters, <b> and <v>) was accounted for. After manually excluding passives, nominalisations and transitive tokens, the operation yielded 67 tokens for *tornar* and 137 tokens for *volver*. The following figure shows the auxiliary selection behaviour of the two verbs in the given time frame:

![Graph showing the development of the usage frequencies of tornar and volver (lemma frequencies / million tokens)](image)

Figure 1: Development of the usage frequencies of *tornar* and *volver* (lemma frequencies / million tokens)

![Graph showing auxiliary selection of tornar and volver in 16th century historiographical texts](image)

Figure 2: Auxiliary selection of *tornar* and *volver* in 16th century historiographical texts
Figure 2 demonstrates a highly significant difference in the behaviour of tornar and volver with regard to auxiliary selection in the 16th century ($\chi^2(1) = 15.3$, p < .001). While tornar shows a clear preference for aver, volver appears to select aver with approximately the same frequency as ser.

The different auxiliary selection of tornar and volver is surprising given that the verbs are close synonyms. Hence, the phenomena discussed in section two will have to be accounted for. On the one hand, the lexical relation between the two verbs needs to be examined in order to ascertain to what extent the two verbs are truly synonymous, or whether there are differences in their semantic usage. This question will be dealt with in section four. On the other hand, it may be that the different auxiliary selection of the two verbs is a conserving effect resulting from the difference in their usage frequencies, which will be argued in section five.

4. The semantics of tornar and volver

In this section, it is argued that there are subtle differences between the use of tornar and volver with respect to their three core meanings: motion, iterativity, and change of state. In order to trace possible semantic differences between volver and tornar, all 204 tokens of volver and tornar compound tenses were coded with respect to these three meanings. While all three meanings are present in the use of both verbs, the following graphs illustrate vast differences in the two verbs in the use of these meanings:

Figure 3: Distribution of meanings of tornar and volver in the 16th century

---

7 All of the statistical tests cited in this article were conducted using the open source statistical software R (R Development Core Team 2012).
*Volver* is used more often with a motional semantics than *tornar*. By contrast, *tornar* is used more often with an iterative or change-of-state meaning. These findings were tested using a “mixed effects” logistic regression analysis (Pinheiro et al. 2009). Logistic regression analysis calculates the influence of multiple predictor variables on a binary dependent variable. However, depending on the design of the study, variables are often nested. In the case of this study, usually several tokens were extracted from one book. While the variable “verb lemma” (*tornar* or *volver*) is, in fact, variable for each of these tokens, the variable “author” is (usually) not. It stands to reason that authors display personal preferences regarding the question of whether to use *tornar* and *volver*. Mixed effects regression models can capture exactly this type of non-variation since they modulate the intercepts of all tokens by any one random factor (Baayen 2008: 263-264). Accounting for individual preferences of the authors regarding the two verbs leads to a greater accuracy of the analysis.

The dependent variable in the mixed model analysis used for this study was the binary choice between the lexical items *volver* and *tornar*. The predictor to be tested was the tokens’ semantic usage. Not all of the three semantic factors (“iterativity”, “motion” and “change of state”) could be included in the model as they are mutually exclusive: in the data analysed, a motional interpretation excludes an iterative or change-of-state interpretation. Similarly, an iterative interpretation excludes the other interpretations\(^8\). Consequently, the three factors are collinear: they explain the same part of the variance encountered in the data (Baayen 2008: 181). Therefore, only “motion” vs. “non-motion” was included as a predictor in the model. The results of the analysis are summarised in the following table\(^9\):

|                | Estimate\(^10\) | Std. Error | z value | Pr(>|z|)  |
|----------------|-----------------|------------|---------|-----------|
| (Intercept)    | -3.35           | 0.99       | -3.39   | 0.000692  ***|
| motionTRUE     | 2.67            | 0.58       | 4.61    | 0.000004  ***|
| ---             |                 |            |         |           |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ’ 1

**Figure 4: Mixed-effects model for the difference in use of ‘tornar’ and ‘volver’ in the 16th century**

\(^8\) There are two cases in which *tornar* is used iteratively with *volver* as its verbal complement. These cases were coded as [- MO], since the motional semantics in these cases is contributed not by the auxiliary verb *tornar*, but by *volver*.

\(^9\) The formula for the lmer-model was *verb_lemma ~ motion + (1 | author)*. The model was set to expect a binomial distribution since the response variable is binary (Crawley 2007: 593-594; Gries 2009: 294-295). It was evaluated with the *somers2*-function, which can also be used to measure the explained degree of variance. The relevant C-value shows the model to explain a high degree of variance (0.96 of 1).

\(^{10}\) The “estimate”-values should be interpreted as follows: the higher the value, the more likely the preference for *volver* over *tornar*. 

11
The analysis confirms that the use of *volver* and *tornar* differ precisely in their motional reading: As expected, *volver* is used more often with a motional semantics than *tornar*. This effect is highly significant (p < .001). Conversely, this analysis predicts that non-motional meanings, i.e. “iterativity” and “change-of-state”, favour the use of *tornar*. In order to find out whether “iterativity” and “change-of-state” differ in their influence on the selection of *volver* or *tornar*, a second logistic regression analysis with mixed effects was conducted. To avoid the collinearity effect mentioned above, the analysis was performed only on tokens without a motional meaning. This mixed-effects model\(^\text{11}\) shows a slight distinction (p < .1) between change-of-state and iterative interpretations for the selection of the verb lemma, with the change-of-state interpretation favouring *volver* and the iterative interpretation favouring *tornar*. As this effect is not significant, it can be concluded that from a synchronic point of view the “motion” / “non-motion” – distinction is a reliable predictor of the use of *volver* and *tornar* in 16\(^{th}\) century historiographical texts.

5. Semantics and frequency as predictors as auxiliary selection

Based on the previous section, it can be assumed that the different auxiliary selection of *tornar* and *volver* in the 16\(^{th}\) century is a function of their semantic differences. A third mixed-effects regression analysis was conducted. The dependent binary variable was the auxiliary (*aver* vs. *ser*). This time two binary predictors were included: “motion” (motional or non-motional interpretation) and “verb lemma” (*tornar* or *volver*). The reasoning was that if both verb lemma and semantic factors are included, the model should show the verb semantics to be a more accurate predictor of auxiliary selection than the verb lemmata themselves. The results of the analysis are illustrated below\(^\text{12}\).

|             | Estimate | Std. Error | z value | Pr(>|z|)   |
|-------------|----------|------------|---------|------------|
| (Intercept) | 4.31     | 1.16       | 3.73    | 0.000195 ***|
| verblemmavolver | -0.58  | 0.57       | -1.02   | 0.309947   |
| motionTRUE  | -3.75    | 1.15       | -3.25   | 0.001152 ** |
| ---         |          |            |         |            |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ’ ’ 1

**Figure 6: Mixed-effects model of the auxiliary selection behaviour of *tornar* and *volver* in the 16\(^{th}\) century**

\(^{11}\) For convenience, the model is not reproduced here. It was set to expect a binomial distribution since the response variable is binary. The somers2- evaluation method (cf. fn. 10) shows the model to explain a high degree of variance (C=0.93 of 1).

\(^{12}\) The formula for this model was aver ~ verb lemma + motion + (1 | book) + (1 | author). It was set to expect a binomial distribution since the response variable is binary. The C-value obtained by the somers2-function is lower than with the first two models, but still good (C=0.83 of 1).
Comparing the mixed-effects model to the chi-square test done in section three (figure 2), it appears that an inclusion of the semantic annotation of the tokens explains a great deal of the variation in auxiliary selection of *tornar* and *volver*. The factor “verb lemma” does not significantly influence the auxiliary selection found in the sample if the semantic factor “motion” is included. By contrast, it appears to be the factor “motion” that determines whether *aver* or *ser* is selected, with *ser* being favoured in motional contexts. Hence, *volver* selects *ser* significantly more often in the 16th century than *tornar* does because it is used much more often as a motion verb. This is also illustrated in the following graph:

![Graph showing auxiliary selection of tornar and volver](image)

**Figure 7: Auxiliary selection of tornar and volver in the 16th century as a function of their semantic usage**

As the graph shows, there conversely exists a strong bias towards *aver*-selection for both the change-of-state (COS) and iterative (IT) readings: Of the 19 cases of *tornar* and *volver* in compound tenses with change-of-state-readings in the 16th century, *aver* is used in 18. Moreover, in all 30 iterative tokens of *tornar* and *volver* in compound tenses *aver* is selected. Since both the iterative and change-of-state readings favour *aver*-selection, the mixed-effects regression model selects “motion” as the best predictor of the auxiliary selection behaviour of the verbs.

The finding that the change-of-state and motion verb readings of *tornar* and *volver* differ with regard to auxiliary selection is in line with the predictions of Sorace’s (2000) Auxiliary Selection Hierarchy. Thus, while motion verbs display the most categorical degree of selection of BE in the hierarchy because of their strong inherent telic meaning, change-of-state verbs “can [...] be regarded as telic, but to a lesser extent than verbs of
change of location” (865). Consequently, they display more aver-selection than motion verbs.

It has been shown that the difference in auxiliary selection of tornar and volver in the 16th century is, to a certain extent, the result of the different usage of the verbs’ meanings. However, this is not a comprehensive explanation. In section two, frequency was hypothesised to play a role in auxiliary selection. In order to determine a possible influence of conserving effects on auxiliary selection, a diachronic perspective has to be taken. Thus, the domain of investigation of the behaviour of tornar and volver was expanded to include data from the 13th to the 17th century. The data was extracted from the CORDE in the same manner as the data from the 16th century, with the exception that, due to the very high number of tokens of volver in the 17th century, only a randomised selection of 50 volver-tokens from that century was extracted. Only tokens from manuscripts that Fernández-Ordoñez (2006) regards as original or at least contemporary copies of the original were included in order to ensure maximal authenticity of the data (cf. the criticism by Rodríguez Molina 2006). The new data was annotated in the same way as the 16th century tokens.

The following graph illustrates the development of the selection of aver over ser for both tornar and volver\textsuperscript{13}:

![Graph showing development of aver-selection for tornar and volver](image)

Figure 8: Development of the aver-selection of tornar and volver

Comparing the development of aver-selection of tornar and volver, it is evident that for both verbs, ser is being replaced by aver. However, tornar succumbs to the analogical

\textsuperscript{13} Due to the scarcity of examples for volver in compound tenses before the 15th century, the auxiliary selection behaviour of tornar and volver before the 15th century cannot be compared quantitatively.
pressure exerted by other verbs that tend towards selecting *aver* at an earlier point in history than *volver*. Between the 15th and the 17th century, the rate of *aver*-selection for *tornar* rises rapidly, from about 5% in the 15th century to almost 100% in the 17th century. By contrast, *aver*-selection appears to advance more slowly for *volver*.

In the light of the results of the preceding section, the different speed of the extension of the *aver*-auxiliary to the two verbs leads to the expectation that this development is bound to a change of the semantic usage of the two verbs. Since it has been shown that the motional usage of the two verbs favours the selection of *ser*, it is expected that the faster increase of *aver*-selection of *tornar* compared to *volver* between the 15th and the 16th century coincides with a faster decrease of its motional usage – an expectation that has been confirmed by the data. Between the 15th and the 16th century, the motional usage of *tornar* decreased from 62.9 to 41.8% (from 39/62 to 28/67 compound tense tokens), while the motional usage of *volver* actually increases from 90.8 to 92.7% (from 24/28 to 127/137 tokens). *Tornar* was already losing its function as a motion verb before disappearing. Therefore, it was auxiliated more frequently with *aver*.

However, this view is challenged by the development of the auxiliary selection behaviour of only the motional variants of *tornar* and *volver*, as illustrated by the following table:

<table>
<thead>
<tr>
<th></th>
<th>tornar (motion)</th>
<th><em>volver</em> (motion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15th century</td>
<td>2.6% (1/39)</td>
<td>20.8% (5/24)</td>
</tr>
<tr>
<td>16th century</td>
<td>57.1% (16/28)</td>
<td>51.2% (65/127)</td>
</tr>
</tbody>
</table>

*Figure 9: Degree of *aver*-selection of *tornar* and *volver* in motional usages*

While the degree of *aver*-selection of motional *tornar* in the 16th century is about 22 times higher than in the 15th century, the degree of *aver*-selection of motional *volver* in the 16th century is only about 2.5 times higher than in the 15th century. Note that this development reverses the auxiliary selection behaviour of the two verbs. Since the development of the semantic usage of the verbs is not a valid point of argument here, a different kind of explanation has to be looked for.

It is proposed here that this finding is the result of the interplay of token frequency and type frequency. By the 15th century, *aver* + PP already had a high type frequency: all atelic verbs (controlled and uncontrolled activities, states, continuation of states) displayed near-categorical selection of *aver*. Moreover, *aver* had already started to be preferred for some telic verbs (change of state, telic motion). Following Barðdal (2008),
the productivity of a construction can be predicted by its type frequency and semantic 
coherence (27). Hence, the high type frequency of the *aver*-compound tense leads to a 
high degree of productivity of that pattern.

In contrast, the high absolute token frequency of *tornar* until the end of the 15th century 
leads to an entrenchment process: the high token frequency of *tornar* has increased the 
strength of the association between *tornar* and *ser* in compound tenses and thus 
disfavours the use of *aver* with this specific lexeme. On the contrary, in the 15th century, 
*volver* was still used infrequently, so that no comparable entrenchment effect could have 
taken place. Therefore, it is not surprising to find a higher rate of *aver*-selection for 
*volver* than *tornar*.

Assuming that these results are accurate, it is expected that the subsequent reversal of 
usage frequencies of *tornar* and *volver* will have an influence on the degree of 
entrenchment of *tornar* + PP and *volver* + PP. Indeed, the increase in absolute usage 
frequency of *volver* between the 15th and the 16th century appears to have led to a higher 
degree of entrenchment of *ser*-selection of this verb. In contrast, the decrease in the 
frequency of *tornar* in the same period correlates with the comparatively sharper 
increase of *aver*-selection. In this manner, the development of the auxiliary selection 
behaviour of *tornar* and *volver* in their motional senses can be explained by taking into 
account entrenchment processes conditioned by usage frequency.

In order to give statistical evidence for frequency effects on the auxiliary selection 
behaviour of the two verbs, a regression analysis of the development of all semantic 
variants of the two verbs over the centuries was conducted. As in the earlier mixed-
effects regression model, the binary dependent variable was the selection of *aver* or *ser*. 
However, this model analyses the entire time period from the 13th to the 17th century. 
For this reason, it would not have been sensible to include “author” or “book” as random 
effects. Trivially, the points in time analysed in the statistical model correspond exactly 
to the variables “author” and “book”. An inclusion of these factors would have produced 
a collinearity effect. Instead, a binomial logistic regression analysis without random 
effects was conducted using the function *glm* in R.

The following predictors were included: First, since in the previous discussion “motion” 
has been shown to be an important predictor of auxiliary selection, this factor is
expected to continue playing a role in this analysis. Second, the average date\textsuperscript{14} of the
sources of the examples was included as a numerical predictor. Hence, “date\_average” is
expected to mirror the historical spread of aver (i.e. the increase of aver-selection over
time). Third, the average lemma frequency of the two verbs at each data point was
calculated\textsuperscript{15} and included in the analysis both as a main effect and in an “interaction”
with “date\_average”. In statistics, interactions denote the joint influence of two predictor
variables on a dependent variable where this influence is partly independent of the
contribution of either single variable. Here, a “moderator approach” to the analysis of
interaction effects has been taken, assuming that “the effect of an independent variable
on a dependent variable differs depending on the value of a third variable, commonly
called a ‘moderator variable’” (Jaccard 2001: 12; cf. also Baayen 2008: 251). This
interaction effect is to be expected after the discussion of conserving effects in this
section: While a general trend towards aver-selection exists, this trend is slower when
the relative token frequency of a linguistic item is high. The results of the analysis are
shown below\textsuperscript{16}:

| Coefficients: | Estimate | Std. Error | z value | Pr(>|z|) |
|---------------|----------|------------|---------|----------|
| (Intercept)   | -30.97   | 3.43       | -9.02   | < 2e-16 *** |
| date\_average | 0.02     | 0.002      | 9.06    | < 2e-16 *** |
| MOTRUE        | -1.57    | 0.36       | -4.41   | 0.0000106*** |
| date\_average:freq | -8e-07 | -3e-07 | -2.56 | 0.0106 * |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ’ ’ 1

**Figure 10:** Binomial logistic regression model of the auxiliary selection behaviour of tornar and volver from the 13th to the 17th century

\textsuperscript{14} The sources were dated according to the information given in the CORDE. When the source was dated
by a time-span-approximation, the average date was calculated as the statistical mean of the end points of
the time span.

\textsuperscript{15} Again, the lemma frequency information from the Corpus del Español was used (cf. fn 6). Since the
Corpus del Español only delivers frequency information for centuries, an extrapolation method was used
to calculate a frequency value for each of the time points represented by tokens in the data in a linear
fashion.

An obvious problem with this frequency measure is that it is assumed that the overall usage frequencies of
the verb lemmata in the corpus are mirrored in the relative frequency of the verb lemmata within the
construction. In this case, however, this problem does not appear to skew the results: a preliminary
analysis using the token-counts of the two verbs per century and the overall token counts of the source
texts per century has shown that the curves of the two frequency measures do indeed appear largely
aligned for the two verbs. It might be more accurate to compute the absolute token frequency of a verb in
the aver- and ser-constructions.

\textsuperscript{16} The formula for this model was aver ~ date\_average + MO + date\_average:freq. It was set to expect a
binomial distribution since the response variable is binary. Somers\textsuperscript{2} shows a high degree of variance
explained (C=0.89 of 1). Note that if freq is included as a main effect, the interaction date\_average:freq is
no longer significant. However, a comparison of the two models using anova favours the model as
presented here, thus justifying the exclusion of freq as a main effect.
The main effects for “motion” and “date_average” both have high statistical significance. As expected after the discussion in the previous sections, the verbs’ motional acceptation disfavours aver. The positive value of “date_average” is equally unsurprising: as time progresses, aver is selected significantly more frequently than ser.

In the last line of the model, a small but significant effect of the interaction between the date of the source book of the token and the usage frequency of the verb lemma is indicated. As such, the effect of “date_average” on the auxiliary selection of tornar and volver, in this case, is moderated by the influence of the verbs’ usage frequencies. As indicated in the second column, the effect value is negative. Hence, the positive effect of “data_average” is attenuated to some degree by “freq”: While aver-selection increases over time, it increases less with the more common verb (after the 15th century, volver). Consequently, this analysis corroborates the hypothesis that conserving effects influenced the development of the auxiliary selection of tornar and volver: The high usage frequency of volver after the 15th century inhibited, to a certain degree, the use of aver with that verb.

In summary, the regression model demonstrates the co-existence of influence of both the semantics of tornar and volver and their usage frequency on auxiliary selection. In general, the effect of verb semantics seems to be more easily traceable and more stable than the conserving effect. This observation is in line with Aranovich’s (2003) hypothesis that verb semantics can be identified as the main determinant of the development of auxiliary selection in Spanish, while conserving effects show a smaller influence.

6. Conclusion

In conclusion, the analyses conducted in this study have shown that both the semantic usage and the relative usage frequencies of tornar and volver appear to influence the verbs’ auxiliary selection behaviour. First, selection of ser is preferred for motional tornar and volver when compared to their change-of-state and iterative readings. Since volver is used much more frequently in its motional meaning in the 16th century than

---

17 The small effect size is due to the scale selected for time. Indeed, the interaction shows the effect of frequency on the development of auxiliary selection for each year in the investigated time period. Since the main effect of time on the development of auxiliary is already rather small, the interaction effect has to be even smaller. The small effect sizes of the genuinely diachronic factors evince the slow pace of language change.
tornar, in compound tenses it consequently occurs more often with the auxiliary ser than tornar. As the regression analysis in figure 10 suggests, this effect is evident not only in the 16th century, but in the entire dataset.

Second, a conserving effect as put forward by Bybee (2006) and Bybee & Hopper (2007) has been uncovered: The process by which aver was generalised to the single auxiliary of Spanish was slower with the verb volver than with tornar due to its relatively higher usage frequency after the 15th century. This suggests that due to the high absolute frequency of syntagms of the type ser vuelto, many speakers view this syntagm as autonomous from the ser + PP construction. Consequently, the replacement of ser + PP by aver + PP affected ser vuelto in less cases than ser tornado. This study thus illustrates the importance of taking into account diachronic gradualness when analysing synchronic variation in auxiliary selection at particular time periods.

Note that a greater number of verbs, including other semantic classes, would have to be analysed in order to demonstrate the pervasiveness, as well as the limits, of conserving effects in the development of the auxiliary selection system in Spanish. Above all, the impact of usage frequency would have to be measured on different levels of structural complexity, including not only verb lemma frequencies, but also a measure of the frequency of verb lemmata within the compound tenses. This issue is left to further investigation.

7. Bibliography

7.1 Documents cited


### 7.2 References


Caballero Martínez, Irma (2000): *La primera parte de la Historia natural de las Indias*. Salamanca: CILUS.

Carriazo, Juan de Mata (1951): *Cronica de los reyes católicos (Hasta ahora inédita).* Sevilla: Escuela de Estudios Hispano-Americanos.


Keniston, Hayward (1937): *The Syntax of Castilian Prose (The Sixteenth Century).* Chicago, IL.


