

multi-million word corpus. This data can then be used to create a linguistic "fingerprint" of the text, which the philologist can use to answer a number of questions. For example, this data can be used to verify the dating of a particular text, by comparing the data from that text against a large number of "control" texts from the same period. This data can also be used to compare the text against a source text, provided that one is available.

In this study we will use large computer-based corpora in the analysis of four separate versions of the medieval Spanish bible. There are a number of reasons for examining the bibles. First, they are one of the few texts where we have available a number of computer-based copies of the same text, as well as computer-based versions of the text in the source language. Second, it is only within the past fifteen years that some of the best editions of these bibles have been published; as of yet there have been few studies that have examined these new editions in depth. Third, these texts represent some of the earliest extensive prose available in Spanish and are thus of interest to linguists looking at this early stage of Spanish. While there are other Spanish prose texts from the early 1200s, e.g. *El libro de los buenos proverbios* or *La doncella Teodor*, few of these are as large as the versions of the Bible that we will analyze, which each contain between 200,000 and 300,000 words.

One final reason for studying the bibles has to do with their origin and what they might tell us about the way in which medieval translators carried out their work. Unlike the medieval bible in other countries, many of the Spanish bibles were translated from Hebrew, rather than the Latin Vulgate. In addition, these Hebrew-based bibles were sometimes created for a Christian audience and in other cases for use in the Jewish community, which also affected the translation.

USING LARGE COMPUTER- BASED CORPORA AS A PHILOLOGICAL TOOL: AN ANALYSIS OF FOUR MEDIEVAL SPANISH BIBLES

MARK DAVIES
Illinois State University

1. Introduction

Within the past ten to twenty years, a number of computer-based tools have been developed that have the potential to aid philologists in their analysis of texts. Among these tools are large computer-based corpora, which can be quickly and thoroughly analyzed with text-retrieval software. These computer-based corpora and the related software make it possible to extract and organize, in a very short period of time, thousands of examples of a particular syntactic construction in a

Although there has been some very good work that has looked at how the Latin or Hebrew original and the Christian or Jewish audience might have influenced the resulting translation,¹ there is still much work that remains to be done. Our hope is that a computer corpus-based analysis might provide some insight into this and other questions.

In Section 2 we present the four different versions of the medieval Spanish bible, which comprise a total of 1,250,000 words of text. In Section 3 we discuss the two syntactic features that will give us a "syntactic fingerprint" of the four texts, and we present "control" data for these two syntactic phenomena from a separate 5,300,000 word computer corpus of historical Spanish prose. These two phenomena concern the placement of clitics (unstressed object pronouns) with infinitives, both when the infinitive is governed by a preposition (*para lo fazer / para fazerlo*) or a main verb (*lo quiero fazer / quiero lo fazer / quiero fazerlo*). In Section 4 we analyze the syntax of the two versions of the Bible that are based on the Latin Vulgate. We examine how closely the syntax of these two texts compares with other texts from the period in which the two texts are assumed to have been created. We also compare the syntax of these two texts to the original Vulgate, to see whether these versions are influenced in any way by the Latin source text. In Section 5 we carry out a similar analysis of the two Hebrew-based bibles, and compare them to other texts of the same period as well as to the original Hebrew Masoretic text on which they are based. Finally, in Section 6 we look at which group of texts, the Latin or the Hebrew based bibles, holds closest to the original source text and examine why there should be a difference between the two.

2. The texts

The table below summarizes the information about each of the four versions of the medieval Spanish bible that we will use in this study. Both individually and taken as a whole, these texts are the largest computer-based medieval bibles presently available. The three texts from the Old Testament (E8, E4, and E19) each contain about 275,000 words of text, while the text from the New Testament (E6) contains about 150,000 words, for a total of nearly one million words. Table 1 includes the name of the text, the supposed date of translation from the source language, the date in which the manuscript itself was produced, its source language, the intended audience, the portions of the manuscript used in this study, and publication information. The name of each of the texts (E6, E8, E4, E19) is an abbreviation for a particular manuscript at the *Escorial*, where the manuscripts are located.

Table 1. Medieval Spanish bibles used in the study²

Text	Translated	Manuscript	Source	Audience	Portions of text used	Publication
E6	1254-1270	1254-1270	Latin	Christian	(NT) Luke, Acts-Rev	Montgomery (1970)
E8	1200-1300	1200-1300	Latin	Christian	(OT) Lev 6-Job	Littlefield (1983)
E4	1200-1400	1400-1440	Hebrew	Christian	(OT) Joshua-Job	Littlefield (1987)
E19	1200-1400	cl 420	Hebrew	Jewish	(OT) Lev 6-2 Kings	Littlefield (1992)

For each of the five Old Spanish versions, we acquired an electronic copy of the text either from Mark Littlefield (E19), the Hispanic Seminary of Medieval

Studies at the Univ. of Wisconsin-Madison (E8 and E4), or by scanning the text into the computer (E6). In addition we downloaded from the Internet copies of the relevant portions of the Vulgate Latin text (from the Online Book Initiative), the Hebrew text (the consonantal text of the *Biblia Hebraica*), and the "Rey Valera" version of Modern Spanish (both from the *Online Bible*)³. Once we had the five medieval bibles and the corresponding sections in Latin, Hebrew, and Modern Spanish, we created a verse-by-verse "parallel version" of the texts by placing the Latin text next to E8, and E4 and E19 next to the Hebrew, along with the Modern Spanish. This was done for each of the 9300+ verses of the Old Testament that we used in the study (Leviticus 6-Job 42, excluding 1-2 Chronicles). In the case of E19 the books of Ruth and Ezra-Job were missing, as was Leviticus 6-Deuteronomy for E4. The 5300+ verses from the New Testament covered Luke and Acts-Revelation, and included E6, the Latin Vulgate, and the Modern Spanish. Below is an example of sample verses from the Old and New Testaments (1 Samuel 7:5 and Acts 15:3):

LAT:	dixit autem Samuēl congregare universum Israel in Masphat ut orem pro vobis Dominum
E8:	Et dixo samuel Plegat vos todo el pueblo de Israel en masphat & Rogare por vos adios
E4:	E dixo Samuel: ayuntad todo Ysrrael a Mizpa, & orare por vos otros al Señor
E19:	E dixo Samuel: Apañad a toda Ysrrael ala Mizpa, & fare oraçion por vos otros al Señor
HEB:	w:yamr smwal qb_w at-kl-ysral h:m_pt:h w:atpll bed:km al-yhwth
MSP:	Y Samuel dijo: --Reunid a todo Israel en Mizpa, y yo oraré por vosotros a Jehovah

LAT:	et postquam tacuerunt respondit Iacobus dicens viri fratres audite me
E6:	E depues que callaron, recudio Yague e dixo: Ermanos, escuchat a mi
MSP:	Cuando terminaron de hablar, Jacobo respondió diciendo: Hermanos, oídme

Once the parallel text was established for each one of the 14,600+ verses under consideration, we used the WordCruncher⁴ computer program to create an "every-word" index of the text. Using this index, we were then able to perform very quick yet complex searches of the data. In the case of clitic placement, for example, the program was able to locate all of the 1000+ cases of a preposition (*en, de, por, para (pora), a*) followed by an infinitive, which was in turn preceded by or followed by a clitic (*me, te, lo, la, etc.*). Examples of this construction are the following:

- (1a) yo saly a ty [por atorçerte] (E19, Numbers 22:32)
(1b) e son debdores [de le fazer] merced (E6, Romans
15:27)

Using the every-word index we were also able to extract the 500+ cases of one of three main verbs (*poder, querer, deuer*) adjacent to a clitic and an infinitive

- (2a) car [los quiso dios matar] (E8, 1 Samuel 2:25)
(2b) & [quiero lo complir] de bonament (E8, Ezra 6:12)
(2c) & non [quisieredes oyrme] (E7, Leviticus 26:21)

Needless to say, without the computer-based corpus and the computer-based searching capabilities it would have been much more difficult, if not impossible, to extract the more than 1,500 relevant examples of clitic placement with infinitives.

3. Clitic syntax

Before we look at the clitic syntax of the four versions of the medieval Spanish bible to see what evidence it gives for the date of composition and the influence of the source language, we will first provide "control" data from Old Spanish, against which we can compare the four bibles. We will first present an overview of the PREP+INF construction in Old Spanish (preposition + infinitive), followed by the V+INF construction (verb + infinitive).

3.1 PREP+INF

The first construction we will study is PREP+INF: *para fazerlo / para lo fazer*. A number of studies indicate that there were some peculiar shifts with this construction in Old Spanish (cf. Hanssen 1913, Sturcken 1953, Menéndez Pidal 1954, Gminder 1959, Ramsden 1963, Granberg 1989). These studies suggest that final placement (*para fazerlo*) was the norm previous to 1250, but from 1250 to 1500 it changed to medial placement (*para lo fazer*), after which it returned to final placement. Preliminary research that we have carried out on a separate 118 text / 5,300,000+ word computer corpus of historical Spanish prose adds some precision to these claims (for a description of this supplementary corpus, see Note 5). Table 2 summarizes the cases of the prepositions *de*, *por*, *para*, and *para* followed by an infinitive and one of the clitics *me*, *te*, *le*, and *les*. The data show that final placement is the norm until about 1280 (3a), when medial placement becomes the norm. It reaches its highest point between 1350 and 1400 (3b), when 87% of all cases have medial placement. It then decreases to about 50% by 1500, and after 1550 virtually all of the cases have reverted to the original final placement (3c).

- (3a) e nos as fecho *para fazerle* mucha merçed (Engaños 53:4 [1253])
 (3b) seran treinta dias *para lo fazer* (Caza 44v [1385-88])
 (3c) él promete ser contigo *para hacerlo* (Pecadores: 291 [1556])

Table 2. Clitic placement with "PREP+INF" in supplementary corpus

Period	PREP+CL+INF	PREP+INF+CL	% PREP+CL+INF
1200-1253	4	21	.16
1270-1280	23	80	.22
1280-1300	41	34	.55
1300-1350	79	46	.63
1350-1400	52	8	.87
1400-1482	80	40	.67
1482-1492	59	62	.49
1492-1508	121	144	.47
1526-1555	72	128	.36
1555-1700	4	634	.01

3.2 V+INF

Turning now to the V-INF construction, we find that this has also been the focus of a number of previous studies (cf. Spaulding 1927, Colburn 1928, Keniston 1937, Wanner 1982, Davies 1997). Although most are based on just a handful of texts, these studies suggest that cases of final placement (*quiero fazerlo*) were quite rare in Old Spanish, and that even by the early 1500s they accounted for only about 10% of all cases. As in the case of "PREP+INF" constructions, preliminary research that we have carried out on the 5,300,000+

word control corpus provides somewhat more precise data. Based on nearly 3000 cases of one of the three main verbs *poder*, *deuer*, or *querer* followed by an infinitive and containing one of the four clitics *me*, *te*, *le*, and *les*, we find the following distribution. In the 1200s nearly all cases have initial or medial placement (4a-b). Final placement gradually becomes more common through the 1300-1400s, but it is not until the 1500-1600s that it becomes the norm (4c).

- (4a) *si lo quiere fazer por su mesura* (Estoria de Espana: 72v [1270])
Initial
- (4b) *yo quierovos dezir el enxemplo* (Historia Troyana: 51 [1253])
Medial
- (4c) *Más quiero avergonzartolos* (Guzmán: 459 [1597])
Final

Table 3. Clitic placement with "V+INF" in supplementary corpus

Period	(CL)V(CL)-INF	V-INF-CL	% V-INF-CL
1200-1300	221	5	.02
1300-1400	163	23	.12
1400-1482	178	37	.17
1482-1555	218	95	.29
1555-1700	186	199	.51
1700-1800	101	540	.84
1800-1900	67	961	.93

4. The Latin-based bibles

Now that we have some idea of clitic placement in Old Spanish, we can examine the clitic syntax of the four medieval bibles. We first want to see whether the

bibles fit into the "syntactic milieu" of the period in which they are assumed to have been copied. If they do not, we will want to consider what other factors might account for their particular clitic syntax.

4.1 The E8 text

Littlefield (1983) suggests that the E8 text dates from the 1200s, and his judgment agrees with other scholars such as Berger (1899) and Solalinde (1929-31). Indeed Berger was of the opinion that it was the oldest of all of the Castilian bibles: "au moins en partie la plus ancienne version castillane de la Bible qui soit conservée." None of these researchers, however, have attempted to date the text with more precision than to say that it is from the thirteenth century. Let us see whether the clitic data can provide more information on a possible date. First, the PREP+INF data shows that only 16% of all cases have medial placement (10 medial, 52 final). Table 2 shows that this low 16% figure would place E8 early in the 1200s, at the same rate as other texts from 1200-1253. This suggests that E8 is one of the very early medieval bibles, which finds support in other linguistic data that Littlefield presents. Examples of the two competing clitic placements are presented below:

- (5a) fue bersabee al Rey salomon por Rogarle por adonias (1 Kings 2:19)
- (5b) yo vin ati por te dezir esta palaura (2 Samuel 14:15)

The data from the V-INF construction corroborates these findings. All 156 tokens have the older initial or medial placement (6a), which means there are no cases of the later final placement. Table 2 shows that in the control corpus of OSp there were occasional cases of the innovative final placement during the period 1200-1300

(6b), so the complete lack of this placement in E8 again suggests a very early date:

- (6a) car todo non lo podras ueer (Numbers 23:13)
 (6b) Et los que dotra guisa vsaren en esta Razon yerran & pueden reubiarlos por ello (Castilla 154r [1200-1284])

4.2 The E6 text

Montgomery, who provides the most complete study of the E6 text, suggests that E6 dates from about 1254-1270 (1962:9-11). Let us consider whether there is evidence for this from the clitic placement data. Looking first at the V-INF construction we find that E6 (like E8) has rather archaic syntax. All 58 of these constructions have the older initial or medial placement; there are no cases of the innovative final placement.

- (7) Et agora vn hombre quiere **me** tomar mis hijos
 (2 Kings 4:1)

When we look at the PREP+INF construction, however, we find something mildly unexpected. While there are 20 cases of final placement (8a), which is the norm for the early to mid-1200s, there are 13 cases of the later medial placement (8b).

- (8a) tornara Dios para agozar **se** sobre ty por bien
 (Deuteronomy 30:9)
 (8b) e debdores son de **les** fazer merced (Romans 15:27)

This relatively high 39% medial placement would place E6 in about 1280. While this is a decade or so later than Montgomery has proposed, it is still within the general range of 1254-1270 that he suggests. In summary, then, both the PREP+INF and V-INF data with both E6 and E8, the two Latin-based bibles, shows that they have clitic syntax that is representative of the 1200s. Let us now turn to the Hebrew-based bibles.

5. The Hebrew texts

As mentioned, the medieval bibles from Spain are of great interest to scholars because several of them were translated from the Hebrew text by Jewish scholars. As a result, this provides scholars with the unique opportunity to compare the language and style of the Hebrew and Latin-based bibles. In addition to the source language there is another factor, which is the group for which the bible was being translated. While some of the bibles that were translated from the Hebrew were destined for a Christian audience, others were meant to be kept within the Jewish community itself. In what follows, we will consider evidence from the corpora that suggest that this second factor did have some effect on the syntax of the text.

5.1 The E4 text

Littlefield suggests that the E4 text was originally translated from the Hebrew into Spanish between 1200 and 1400 (1987:xxxiii-xxvi). Basing his conclusions on a range of linguistic phenomena, he further suggests that the E4 manuscript itself was produced in about 1400-1440. Finally, he hypothesizes that the manuscript was made by Jewish translators, but that it was destined for a Christian audience. Let us consider whether the clitic syntax supports the rather late date of 1400-1440, as opposed to the earlier thirteenth century dates for E8 and E6. Turning first to the PREP-INF construction, we find that E4 has 90% medial placement, based on 90 cases of medial placement (9a) and just 10 cases of final placement (9b). Table 2 shows that medial placement was at about 87% during the period 1350-1400, which means that the high 90% medial placement in E4 fits in very well with the date that has been proposed by Littlefield.

- (9a) & salio Saul alo rresçebir por lo benedezir
(1 Samuel 13:10)
- (9b) cosa ligera es ala sombra de desuiarse diez grados (2
Kings 20:10)

The data from the V-INF construction also argues for a fifteenth-century text. Unlike E8 and E6, which have only initial and medial placement, we find that 4 of the 22 cases in E4 have the innovative final placement (10b). Table 3 shows that this 15% figure places the E4 text squarely in the early 1400s, which is what Littlefield had proposed, based on other linguistic criteria.

- (10a) non te podras leuantar delante tus enemigos
(Joshua 7:13)
- (10b) & non quiso Daudid beuerla (2 Samuel 23:16)

5.2 The E19 text

Littlefield proposes that the E19 text dates from approximately the same time as E4, the other Hebrew-based bible (1992: xx-xxi). He suggests that it was originally translated from Hebrew between 1200-1400, and that the E19 manuscript itself dates from about 1420. What distinguishes the E19 text from E4, E6, and E8, however, is that it is the only one of the three that was destined for a Christian audience. We will see that this is a significant difference, one that has important repercussions for its clitic syntax.

As soon as we turn to the PREP-INF construction, we realize that E19 is not a typical fifteenth-century text. While we would expect a high degree of medial placement (compare the 90% figure for E4), we find that only 22% of the 144 cases have medial placement (32 medial, 112 final):

- (11a) ouo voluntad de ty para te poner sobre la silla de
Ysrael (1 Kings 10:9)

(11b) & el deçendio a rresçebirme (1 Kings 2:8)

As Table 2 indicates, a figure of 22% medial placement would either place E19 in about 1270-1280 (as there began to be a shift towards medial placement) or it would place E19 in the late 1500s (when medial placement was in its declining stages). In either case, it would be about 150 years away from its assumed date of production.

The data for the V-INF construction tells much the same story. A full 9/28 cases have the innovative final placement (12a), and this 32% final placement would place E19 in the late 1500s.

- (12a) Ca non les podemos dar por mugeres a nuestras fijas
(Judges 21:18)
- (12b) & non quiso Dios destroyrte (Deuteronomy 10:10)

Hence it appears that in the case of E19, we must either abandon the idea that we can use syntactic data to date the texts, or we must come up with some other explanation for the unexpected facts. It would be unfortunate to discard syntactic evidence as a tool for dating, especially when the results were so favorable with E6, E8, and E4. Luckily, there is another explanation for the E19 results, which provides insight into the importance of both the source language and the intended audience.

We might assume that since E19 was destined for a Hebrew audience, as Littlefield claims, Hebrew syntax affected the translation more than with E4, which was destined for use by Christians. Let us therefore compare the Hebrew clitic syntax for both the PREP-INF and V-INF constructions, and see how it relates to the syntax of E19. First, we find that Hebrew had invariable final placement with the PREP-INF construction (13a). It is not difficult to see that the translator might adapt the Spanish syntax (which should only have about 10% final

placement in the early 1400s) towards the Hebrew syntax (13b), which would then explain the high 78% final placement in E19:

- (13a) yrwn rkb smym b:ez:k (Deuteronomy 33:26)
 (13b) el que causalga en los çielos para ayudarte

When we examine the V-INF construction, we again see influence of the Hebrew source text. As mentioned, the high 32% final placement with V-INF in E19 represents a degree of final placement that would not become common until the late 1500s. However, when we see that final placement was obligatory in Hebrew with this construction (14a), we can again suppose that there was an attempt to adapt the Spanish to that of the Hebrew original (14b).

- (14a) aykh awkl l:hwrys:m (Deuteronomy 7:17)
 (14b) como podre esterrarlos

An even clearer indication that there is a difference between E19 and E4 comes when we look at the cases where the clitic syntax of E19 and E4 differ with a given verse. For example, with the PREP+INF construction there are ten cases where clitic placement with E19 and E4 differ. In all ten cases, E19 has final placement (following the Hebrew), whereas E4 has medial placement (following the normal syntax of the early fifteenth century):

- (15a) E19: Rrodearon me el arca de Dios de Ysrrael [para matarme] (1 Samuel 5:10)
 (15b) E4: boluieron ami el arca de Ysrrael, [para me matar]
 (16a) E19: la mi mano sera contigo [para tomarte] toda Ysrrael (2 Samuel 3:12)
 (16b) E4: he mi mano, que es contigo [para te tomar] atodo Ysrrael

- (17a) E19: & feziste dioses otros, & fundiçiones [para ensañarme] (1 Kings 14:9)
 (17b) E4: & feziste para ti dioses agenos & cosas fundedizas [para me enojar]

We find the same thing with the V-INF construction. In all of the examples where E19 and E4 differ, E4 has the normal Spanish syntax while E19 follows the Hebrew. This is the case even when the clitic placement is awkward for early fifteenth century Spanish (such as after *non*, which nearly always produces initial placement):

- (18) HEB: w:la [ykl l:hsyb:h] aly:w (1 Kings 13:4)
 (18b) E19: & non [pudo tomarla] asi
 (18c) E4: & non [la pudo tomar asy]

6. The comparative effect of the Latin/Hebrew source text

We have just suggested that in the case of E19 the Jewish translators adapted the syntax of the text towards the Hebrew original, since the text was destined for use in the Jewish community. In the case of E4, however, this did not happen because the Jewish translators were creating a text that would be used in a Christian community, for whom the original Hebrew language would have had little meaning.

Let us now return to the Latin bibles. Could it be the case that in the E6 and E8 also there was an attempt to adapt the syntax to the original Latin source text? We have already seen that these two texts are quite representative of the 1200s. It may be the case, however, that like E19 they are simply following the source text, which in this case is the Late Latin syntax of the Vulgate. The data, however, shows no effect from the Latin source text. For example, with the V-

INF construction Latin had both initial and medial as well as final placement:

- (19a) si Dominus [nos vellet occidere] (Judges 13:23)
 (19b) quomodo [potero delere eas] (Deuteronomy 7:17)

When we begin to examine the relationship between the Latin source and E8 and E6, however, we see that there is no correspondence between the two. In none of the 32 cases of final placement in the Vulgate do the corresponding passages in the E8 or E6 text take final placement

- (20) LAT: sin autem ille qui voverat [voluerit redimere **eam**]
 E8: Mas si aquel que la prometio [la quisiere quitar] (Leviticus 27 15:4)
- (21) LAT: et non [potuit respondere **ei**] quicquam quia metuebat illum
 E8: & no[pudo Recudir] nada car lo temie (2 Samuel 3 11:4)
- (22) LAT: [volebamus tradere **vobis**] non solum evangelium Dei
 E8: [queriemos **uos** dar] non sola mientre ell euangelio de Dios (1 Thessalonians 2:8)

The same is true of the PREP-INF construction. In the Latin passages that correspond to the PREP+INF construction in E6 and E8, we find both medial and final placement:

- (23a) fecisti tibi deos alienos et conflaviles [ut **me** ad iracundiam provocares] (1 Kings 14:9)
 (23b) et omnis populus manicabat ad eum in templo [audire **eum**] (Luke 21:38)

But when we compare the Latin and Spanish verses we find that in at least half of the cases the clitic placement in the two languages is different:

- (24) E8: por que nos quesist passar Jordan [por **nos** meter] en mano delos amorreos (Joshua 7:7)
 LAT: quid voluisti transducere populum istum Iordanem fluvium [ut traderes **nos**] in manus Amorrei
- (25) E8: abner non veno por al ati sino [por te engaynnar] (2 Samuel 3:25)
 LAT: Abner . . . quoniam ad hoc venit [ut deciperet **te**]

Thus, whereas the clitic syntax of E19 was adapted to the Hebrew original to the point that it reflected a stage of Spanish quite foreign to the early 1400s, this does not happen with the two Latin texts. The only syntactic influence on these two texts is the Spanish of the 1200s, when they were produced.

Let us summarize the influence that the source text and the intended audience may have had on the translation of the four medieval Spanish bibles. All of the texts except E19 were destined for a Christian audience, and in the case of E4 it would have made little sense to follow the Hebrew source text. In the case of E6 and E8 also, the translator apparently felt little need to follow the syntax of the Latin Vulgate. Only in the case of E19, where the Hebrew text was destined for a religious group that felt an affinity for the culture of the source text, did the translator have an incentive to adapt the translation towards the source language. Indeed, as Mark Littlefield has noted (p.c.), the Hebrew text was one that they encountered on at least a weekly basis. They knew much of this text by heart, since sections of the Pentateuch were read in Hebrew each week in the synagogue. Thus in the case of E19, there is a clear rationale for the translators to be influenced by the source text as they translated the text to Spanish.

7. Conclusion

In this study we have used a nearly one million word parallel corpus of four medieval Spanish bibles to compare the syntax of the four bibles against each other, as well as against a computer-based corpora of the original Latin and Hebrew texts. We were also able to compare the data from the medieval bibles against a 'control' corpus of more than 5,300,000 words of historical Spanish prose. As can be readily appreciated, this type of comparison, which involved thousands of examples in the six different texts and many more in the control corpus, would have been very difficult without the aid of the computer-based corpus.

We hypothesized that with a large enough corpus and a syntactic construction that would yield a large number of examples, one might be able to create a "syntactic fingerprint" of the individual texts. One purpose of this would be to help provide a date for texts when previous dating was unclear. While the 1200s had already been established as the date of the two Latin-based bibles (E6 and E8), our clitic data independently confirmed this. It also suggested that E8 was the earlier of the two texts, which likewise had been hypothesized by other researchers. The data from the corpus also placed the E4 text in the early 1400s, precisely where it should have been. The data from the E19 text, on the other hand, was the only one that did not correlate well with other texts in the control corpus that came from its supposed date of production. This led us to suppose that there was another factor at play, which was the effect of the Hebrew source text. We hypothesized that in the case of E19 alone was there a rationale for having some effect from the source text, since E19 was the only text that was translated by and for the Jewish community, who would have been familiar with the

Hebrew source text. In summary, then, we hope to have shown that large computer-based corpora can serve as a useful tool to investigate a number of philological issues related to medieval texts.

Notes

¹ By far the most active researcher of the medieval Spanish bibles has been Margherita Morreale. General and introductory studies include Morreale (1960a, 1960b, 1963, 1974). In addition to these introductory studies, she has also carried out the most systematic survey of the comparative syntax of these bibles. See, for example, Morreale (1968, 1982a, 1992a, 1992b).

² Besides the medieval bibles listed in Table 2, which form the basis of this study, one might also wish to consult the other medieval bibles listed in the Works Cited section.

³ The URL (Web address) for medieval bibles with links to the Latin and Modern Spanish bibles online is <http://138.87.135.33/bibl-span/samuel.htm>.

⁴ For information about "WordCruncher" see <http://www.wordcruncher.com/software.html>. See also Mark Davies, "Omnipage and WordCruncher: Tools for Creating and Searching Digitalized Text Corpora," in *La Corónica* 23, 1995, pp 111-115.

⁵ The texts comprising the supplementary 5,300,000 word / 118 text corpus are the following: **1200 [14 texts / 776,700 words]** *Libro de los halcones, Tratado de la cetrería, Fueros de Castilla, Fueros de Aragón, Libro de los animales de caza, Poridat de poridades, Libro de los engaños, Estoria de España, Historia troyana en prosa y verso, General estoria, Libro de los cient capitulos, Castigos y documentos para bien vivir, Libro de consejo y de los consejeros, Gran conquista de Ultramar. 1300 [10 texts / 744,200 words] *Crónica de Sancho IV, Sumas de la historia troyana, Leyes del estilo, Crónica de veinte reyes,**

El Conde Lucanor, Libro de la montería, Historia troyana, Proverbios morales, Libro de la caza de las aves. **1400 [15 texts / 765,200 words]** Cuento de Tristán de Leonís, Menor daño de medicina, Retórica, De los oficios, Arte cisoría, El Corbacho, Libro de las doñas, Breve Invencionario, Espejo de medicina, Fábulas de Esopo, Breve confesionario, Claros varones de Castilla, Imitación de Cristo, Arnalte y Lucenda. **1500 [19 texts / 745,300 words]** Cárcel de amor, De las mujeres ilustres en romance, Grimalte y Gradissa, Historia de Grisel y Mirabella, Glosa sobre Lux bella, La Celestina, Amadís de Gaula; Sumario de la natural historia de las Indias, Diálogo de las cosas acaecidas en Roma, Diálogo de doctrina cristiana, Diálogo de la lengua, Relación o Naufragios, Relación de las comunidades de Castilla, Los corsarios Barbarroja, Brevisima relación de la destrucción de África, Los amores de Clareo y Floriseu, El crotalón, Lazarillo de Tormes. **1600 [16 texts / 701,100 words]** Guía de pecadores, Las moradas del castillo interior, De los nombres de Cristo, La perfecta casada, Guzmán de Alfarache, La vida del Buscón, Don Quijote de la Mancha, Comentaríos reales de los incas: dos Novelas ejemplares (Rinconete y Cortadillo, La ilustre fregona), Los trabajos de Persiles y Sigismunda, Vida de Marcos de Obregón, Novelas amorosas, Desengaños amorosos, Empresas políticas, Vida y hechos de Estebanillo González, El criticón. **1700 [17 texts / 669,800 words]** El día de fiesta por la mañana y por la tarde, El no importa de España, Historia de la conquista de México, Teatro crítico universal and Cartas eruditas, La poética, Vida, Historia del famoso predicador fray Gerundio de Campazas, El Pensador, Mayans y Siscar (Escritos económicos), Cartas marruecas, Gaspar Melchor de Jovellanos (Anthology), El Censor, La derrota de los pendantes, El Rodrigo, Diario de Madrid, Juan Meléndez Valdés (Cartas y discursos). **1800 España [13 texts / 425,500 words]** Las palabras, Escenas matritenses, El señor de Bembibre, Escenas andaluzas, Fernán Caballero (Selected prose), Pepita Jiménez, La pródiga, La regenta, Los pazos de Ulloa, Fortunata y Jacinta, La espuma, La puchera. **1800 LatAm [14 texts / 550,600 words]** Don Carrín de la Fachenda, Facundo, Guatimozin, último emperador de México, El ideal de un calavera, María, Clemencia, Enriqueillo, Carmen, Juvenilia, Amistad funesta, La noche buena, De sobremesa, El donador de almas.

Works Cited

- Américo Castro, et al. (1927) *Biblia Medieval Romanceada: I - Pentateuco*. Buenos Aires: U de Buenos Aires, Facultad de Filosofía y Letras.
- Berger, Samuel. (1899) "Les bibles castillanes." *Romania* 28: 360-408, 508-67.
- Colburn, Guy Blandin. (1928) "The complementary infinitive and its pronoun object." *Hispania* 11: 424-9.
- Gminder, Albert. (1959) "A study in fourteenth century Spanish syntax." Diss. North Carolina.
- Granberg, Robert A. (1989) "Object pronoun position in Medieval and Early Modern Spanish." Diss. UCLA.
- Hanssen, Friedrich. (1913) *Gramática histórica de la lengua española*. Halle A.S.: M. Niemeyer.
- Hauptmann, Oliver H. (1953) *Escorial Bible I.j.4. Vol. 1, The Pentateuch*. Philadelphia: U of Pennsylvania Press for Grinnell College Press.
- Keniston, Hayward. (1937) *The Syntax of Castilian Prose: The Sixteenth Century*. Chicago: U of Chicago P.
- Littlefield, Mark G. (1983) *Biblia Romanceada I.I.8 : the 13th-century Spanish Bible contained in Escorial MS.I.I.8*. Madison, WI: Hispanic Seminary of Medieval Studies.
- (1987) *Escorial Bible I.j.4. Vol. 2*. Madison, WI: Hispanic Seminary of Medieval Studies.
- (1992) *Escorial Bible: I.ii.19*. Madison, WI: Hispanic Seminary of Medieval Studies.
- Llamas, José P. (1950-55) *Biblia medieval romanceada judío-cristiana*. 2 vols. Madrid: Inst. Francisco Suárez.
- Menéndez Pidal, Ramón. (1954) *Cantar de Mio Cid: texto, gramática, y vocabulario*. Vol 1. 3rd ed. Madrid: Espasa Calpe.
- Montgomery, Thomas. (1962) El evangelio de San Mateo según el manuscrito escorialense I.I.6: texto, gramática, vocabulario. Madrid: Real Academia Española.
- (1970) El Nuevo Testamento según el manuscrito escorialense I-I-6 (Anejos del BRAE, 22). Madrid: Real Academia Española.
- Morreale, Margherita. (1960a) "Apuntes bibliográficos para la iniciación al estudio de las traducciones bíblicas medievales del castellano." *Sefarad* 20: 66-109.

- (1960b) "Vernacular Scriptures in Spain." *The Cambridge History of the Bible. Vol 2. The West from the Fathers to the Reformation*. Ed. G.W.H. Lampe. Cambridge: Cambridge UP. 465-491.
- (1963) "Las antiguas biblias hebreo-españolas comparadas en el pasaje del Cántico de Moisés." *Sefarad* 23: 3-21.
- (1968) "De la comparación bíblica en un romancamiento castellano del siglo XIII (MS Esc 1-1-6)." *Litterae Hispanae et Lusitanae: Festschrift zum fünfzigjährigen Bestehen des iberoromanischen Forschungsinstituts der Universität Hamburg*. Ed. Hans Flasche. Munich: Hueber. 241-98.
- (1974) "Al margen de la historia de la biblia latina y romance en España." *Iberoromania* 1:41-51.
- (1982) "Ca de Dios vos es dado el poder: la voz pasiva en tres romaneamientos bíblicos medievales." *Biblos* 58: 129-49.
- (1992a) "'Cosa': usos pronominales sintácticos y léxicos en tres romaneamientos bíblicos medievales." *Anuario de Letras* 29: 5-40.
- (1992b) "El participio y el participio adjetivado en tres romaneamientos bíblicos medievales (Libro de la Sabiduría)." *Verba* 19: 343-48.
- Paz, Antonio and Julián Paz. (1922) *Biblia (Antiguo Testamento) Traducida del Hebreo al Castellano por Rabi Mose Arragel de Guadalfajara (1422-1433) y publicada por el Duque de Berwick y de Alba*. 2 vols. Madrid.
- Ramsden, Herbert. (1963) *Weak Pronoun Position in the Early Romance Languages*. Manchester: Manchester UP.
- Solalinde, Antonio G. (1929-31) "Los nombres de animales puros e impuros en las traducciones medievales españolas de la Biblia." *Modern Philology* 27: 473-85, 28: 83-98.
- Spaulding, Robert K. (1927) "Puedo hacerlo versus lo puedo hacer and Similar Cases." *Hispania* 10:342-48.
- Sturcken, Henry Tracy. (1953) "Studies in thirteenth century Spanish syntax." Diss. North Carolina.
- Wanner, Dieter. (1982) "A History of Spanish Clitic Movement." *Proceedings of the Berkeley Linguistics Society* 8. Berkeley: BLS. 135-47.

HISTORIA QUE ILUSTR LA MANERA O VÍA POR LA CUAL ALFONSO FRAILE ESTÁ PRESENTE EN EL MUNDO A RAZÓN DE CORTAS TEMPORADAS Y NO SIEMPRE, SEGÚN TENÍAMOS ENTENDIDO ALGUNOS QUE ERA EL CASO

DIEGO DENI

ALFONSO FRAILE DESPERTÓ una mañana algo Azorado; entendió que había muerto durante la noche y que estaba en el Infierno. Lejos de quebrantar su buen ánimo, este contratiempo estimuló su habitual curiosidad, de modo que se apresuró a explorar el nuevo entorno. Su primera observación reveló que la geografía del mundo infernal coincidía punto por punto con la geografía del mundo terreno, si bien se echaba de ver que los objetos que en el mundo terreno se consumen en el aburrimiento de lo inánime, en el Infierno contraen la incertidumbre de lo que está ligeramente inclinado y supuran una extraña fascinación. Un armario común, por ejemplo, consiste en un relieve común de armario; el relieve de un armario infernal se diluye en el relieve de todas las formas posibles; su aspecto, si es que alguno tiene, prueba ser intolerable. . . . Los objetos infernales, concluyó don Alfonso, están agazapados en sí mismos, no con la rigidez de un artefacto inevitable, sino con las fluctuaciones de un

Maíces y camelias <i>José Luis Varela</i>	11
A Visão de Yosef e Maria <i>Glauco Ortolano</i>	45
Historia que ilustra la manera . . . <i>Diego Deni</i>	93

CRÍTICA

El juego es narrar y entramar historias: dos cuentos contemporáneos de la literatura mexicana <i>Alejandra Herrera Galván y Joaquina Rodríguez Plaza</i>	31
Using Large Computer-Based Corpora as a Philological Tool: An Analysis of Four Medieval Spanish Bibles <i>Mark Davies</i>	70
Searching for Singularity: Mother-Daughter Narration in Elena Poniatowska's <i>La "Flor de Lis"</i> <i>Kimberly Chisholm</i>	99

ENTREVISTAS

George D. Schade: Una breve historia de nuestros tiempos literarios <i>Kathryn Anne Everly</i>	15
Rosario Ferré entre el inglés y el español: " <i>Let oneself become the meeting place of both</i> " <i>Gloria Diaz Rinks y Elisabeth Sisson-Guerrero</i>	61

The *Dactylus* editorial board would like to thank the many people who generously contributed to the completion of our sixteenth edition. We first wish to recognize the efforts of those graduate students who took the time to carefully evaluate the works submitted. We also would like to extend our appreciation to the students and staff members, especially Antonio Valle de Antón, who helped raise money for publishing costs by selling fundraiser tickets. Thank you to those faculty members who made generous personal contributions. Special thanks go to Dean Teresa Sullivan and to the Texas Union CoSponsorship Review Board, who kindly supported the journal with institutional funds. Thank you also to the Huntington Art Gallery for granting us permission to display art from their collection on the cover. Furthermore, we are grateful to the staff members of the Department of Spanish and Portuguese and the Student Microcomputer Facility who graciously helped us to use their many resources. And finally, we would like to thank the writers and scholars who submitted their work for publication. The outstanding submissions which you contribute each year continue to make *Dactylus* an intellectual forum of which we all can be proud.