

LEXICAL VARIATION IN MEDIEVAL SPANISH: APPLYING QUANTITATIVE METHODS TO SPANISH BIBLICAL TEXTS¹

HIROTO UEDA
University of Tokyo

ABSTRACT:

In this paper I carry out a quantitative analysis of lexical coincidences between nine medieval Spanish versions of a section of the *Book of Judges* in the *Old Testament*. The purpose of this analysis is to obtain an overview of the lexical variation in Spanish medieval biblical translation through the application of co-occurrence methods, correlation coefficients, clusters, principal component analysis and my own method of integration. The results of these quantitative analyses are compared with the relationships proposed in previous studies concerning the degree of relationship between the different Old Spanish biblical translations.

KEY WORDS: lexical variation, Spanish medieval bibles, quantitative method

PALABRAS CLAVE: variación léxica, biblias medievales hispánicas, método cuantitativo

1. INTRODUCTION

In this paper I intend to carry out a quantitative analysis of lexical coincidences between medieval Spanish versions of a section of the *Book of Judges* in the *Old Testament*, for which all existing versions offer a complete translation. The purpose of this study is to obtain an overview of the lexical variation in Spanish medieval bibles through the application of various statistical methods.

First, in this section I explain the method for laying out comparative lexical data in form of a two-dimensional array of source texts and linguistic forms. I analyze Chapters 13 to 16 of the *Book of Judges* in the *Old Testament*, which

¹ The research developed for this paper was originally presented at the XIII *Oxford Forum of Iberian Studies on Variation and Change in Ibero-Romance*, held at the University of Oxford in 2008. The paper, which remained unpublished, has been expanded and translated into English for this occasion. I would like to thank Andrés Enrique-Arias for providing me with bibliographical references as well as for his suggestions concerning different stages of this research. Likewise I wish to thank Javier Muñoz-Basols for his help with proofreading the English version, and Javier Pueyo and Enrique Pato for their comments on my first draft of this paper. This study was supported by the *JSPS KAKENHI Grant Number 24520453* (Japan).

correspond to the story of Samson². Within the wide corpus of medieval Bibles, those chapters appear in the following nine texts (listed in chronological order³):

Fazienda. *Fazienda de ultramar*. University of Salamanca 1997. Date of composition: ca. 1200, date of copy: 1230. Translated from Hebrew.

E8. Escorial I.i.8. Date of composition: ca. 1250, date of copy: ca. 1400. Translated from Latin.

GE. *General estoria*. Madrid, National Library 816. Date of composition: ca. 1270-1280, date of copy: 1272-1275. Translated from Latin.

E4. Escorial I.i.4. Date of composition: the first third of the fifteenth century. 1400-1430, date of copy: 1400-1430. Translated from Hebrew.

E7. Escorial I.i.7. Date of composition: the first third of the fifteenth century, date of copy: ca. 1400-1430. Translated from Hebrew.

E19. Escorial I.ii.19. Date of composition: the first third of the fifteenth century, date of copy: ca. 1420. Translated from Hebrew.

Ajuda. Ajuda Library 52-xii-1. Date of composition: the first third of the fifteenth century, date of copy: ca. 1420-1430. Translated from Hebrew.

Alba. *Biblia de Alba*. Madrid, Palacio de Liria. Date of composition: 1420-33, date of copy: 1420-1433. Translated from Hebrew.

E3. Escorial I.i.3. Date of composition: the first third of the fifteenth century, date of copy: ca. 1425-1450. Translated from Hebrew.

Of these versions, E7 and E19 on the one hand as well as Ajuda and E3 on the other are testimonies of the same translation. Still I include these versions in the analysis to see their lexical proximity compared to other parallel texts.

As for the lexicon, for the purpose of investigating the lexical variation between the texts cited above, I have eliminated those cases in which all the translation use the same Spanish equivalente. For example: (13:04) *bever, vino*,

² The texts used in this analysis have been provided by Andrés Enrique-Arias, who heads the *Biblia Medieval* research project. The materials are currently available in the project's website at www.bibliamedieval.es.

³ The chronological data are from Enrique-Arias (2011). For bibliographical information on Medieval Spanish biblical texts, see Morreale (1969), Sánchez-Prieto Borja (2002) and Avenzoa and Enrique-Arias (2005).

comer; (13:06) *marido*, *decir*; (13:08) *hacer*, *nacer*; (14:03) *fija*, *hermano*; (14:18) *leon*, *arar*; (15:15) *asno*; (15:19) *agua*; (16:03) *puerta*, etc. These words are repeated in all the texts analyzed and thus do not show anything particular about the lexical character of any of the texts.

In order to focus exclusively on lexical aspects, graphophonemic, morphological and syntactic permutations have been disregarded. For this reason, I have changed the spelling of <çi>, <çe> to <ci>, <ce> and I have normalized the use of <i> and <u> as vowels, and <j> and <v> as consonants⁴. Variable spellings such as <np> - <mp>, <nn> - <ñ>, <pp> - <p>, etc. have all been converted into a common form (i.e., the second option in each pair). Some specific cases, such as translation equivalents *sizra*, *sidra*, *sisra* 'intoxicating drink' have been considered as variants of the same lexical unit *sidra*. In addition, in those cases in which variants reflect some change in progress (i.e. *omne* / *ombre*; *aparescer* / *aparecer*; *levar* / *llevar*; *ondrar* / *(h)onrar*, etc.), the variants in question have been treated as equivalent across all texts.

Morphosyntactic variation does not affect the discussion undertaken in this paper either, so that (13:20) *cielos* / *cielo*; (13:07) *aver un fijo* / *aver fijo*; *su nombre* / *nombre* (13:06); *fablar con* / *a* / *por* / *fablar* (14:07); *guardarse*, *ser guardado* / *guardar* (13:13), etc. have all been changed to the latter form listed in each grouping.

In contrast to inflectional variants, derivational ones are treated as cases of lexical variation. These are the result of word formation, and when a word is derived from another, the two words are considered as different forms from the lexical point of view. For example: *levarse* / *levantarse* (13:11); *mostrar* / *amostrar* / *demostrar* (13:23); *septimo* / *seteno* (14:17); *cabra* / *cabrito* (13:15); *carreillada* / *carrillo* (15:17); *aborrir* / *aborrescer* (14:16); *advinanciella* / *advinança* (14:18); *raposa* / *raposo* (15:04); *llegar* / *allegar* (16:09); *quebrar* / *quebrantar* (16:21); *crescer* / *acrescer* (16:22); *estruir* / *destruir* (16:24).

With the purpose of determining the distance and similarity between the texts, in this study lexical variation does not only refer to a single word, but can also refer to a number of words that constitute an expression with the same or similar meaning. What I am interested in analyzing is how a given semantic unit is expressed in varying forms, for example: (15:01) *después de tiempo* / *después de esto a pocos días* / *acabo de días* / *a cabo de un año*. While these expressions are not precisely synonymous, these variants are translation equivalents inserted

⁴ For graphemic questions see Fernández López (1996) and Sánchez-Prieto Borja (1998). I discussed a particular aspect of the spellings of <i>, <j> and <y> in Ueda (2010).

in the same context of occurrence in different texts. Another example is: (14:09) *por el camino / por la carrera / andando*.

Within the same context, I also find non-synonymous expressions: *ver / echar con / aver pleito con / entrar con / ir a / estar con* (15:01); *fermoso / mejor* (15:02); *casa / camara / cillero* (16:09). These variants should be considered as lexico-pragmatic variants.

On the other hand, there are also cases of mismatched parts of speech, for example (13:2), in which no grammatical correspondence is found:

E8: Fue vn hombre de saraa & del linage de dan que *ouo nombre manue* (...)

GenEs: En aquella sazón era en Saraa vno del Linage de Dan & *llamauanle Manue* (...)

E3: E auja vn omne de çora de linaje de dan & *su nonbre manoah* (...)

The above example demonstrates a case of synonymous expressions: *llamarse* ('to be called') and *su nombre es* ('his name is'). Such examples are included in this analysis as cases of lexical variation.

In total I have found 520 cases of lexical variation based on the definitions explained above. I begin the analysis by means of a two-dimensional table contrasting all the examples, with the corresponding line number (chapter and verse) on the vertical axis and the identifiers of the source texts on the horizontal axis⁵:

(c,v)	Fazienda	E8	GE	E3	Ajuda	E19	E7	E4	Alba
(13:02)	generacion	linaje	linaje	linaje	linaje	linaje	linaje	linaje	generacion
(13:02)	varon	omne	omne	omne	omne	omne	omne	omne	varon
(13:02)	aver nombre	aver nombre	llamar	su nombre	su nombre	aver por nombre	aver por nombre	su nombre	llamavan
(13:02)	emprenarse	-	aver fijo	parir	parir	parir	parir	parir	parir
(13:02)	mañera	mañera	-	mañera	mañera	mañera	mañera	mañera	mañera
(13:03)	muger	-	-	muger	muger	muger	muger	muger	muger
(13:03)	emprenar	concebir	-	empreñar	empreñar	concebir	concebir	empreñar	encintar
(13:03)	aver fijo	aver fijo	aver fijo	parir fijo	parir fijo	parir fijo	parir fijo	parir fijo	parir fijo

Table 1. Contrastive summary

To calculate a correlation coefficient for each pair of source texts, it is necessary to convert this contrastive table into a table listing the occurrences of

⁵ To consult the analysis programs in Excel VBA which I have developed for treatment of digital texts see <http://lecture.ecc.u-tokyo.ac.jp/~cueda/gengo/index.html>.

each linguistic form:

(cv)	Fazienda	E8	GenEs	E3	Ajuda	E19	E7	E4	Alba
a (14:16)	+								
a cabo de dias (14:08)	+			+	+			+	+
a cabo de dias (15:01)				+	+			+	+
a cabo de un año (14:08)						+	+		
a cabo de un año (15:01)						+	+		
a la luz de la mañana (16:02)				+	+				
a pocos de dias (14:08)			+						
a pocos dias (14:08)		+							
abezar (13:08)				+					

Table 2. Distribution of occurrences

2. CORRELATION MATRIX AND CLUSTER

To observe the relationships between each pair of texts, the correlation coefficient Phi in its modified version ('Phi) is calculated from four data points:

X / Y	Y (+)	Y (-)
X (+)	a (x+, y+)	b (x+, y-)
X (-)	c (x-, y+)	d (x-, y-)

Table 3. Four points: a, b, c, d

where the point (a) corresponds to the number of cases in which both X and Y are positive (+) (i.e. a given semantic unit is observed in both texts), (b) corresponds to the number of incidences of X and not Y (i.e. the semantic unit appears in text X but not in text Y), (c) is the inverse of (b); and in (d) both X and Y are negative (-). The equation for the modified correlation coefficient Phi is:

$$\text{Phi}' = \frac{a}{\sqrt{(a+b)(a+c)}}$$

This is a formula modified from the well-known Phi coefficient:

$$\text{Phi} = \frac{ad - bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$$

It has been changed in order to exclude cases in point (d) indicating common

negative observations (i.e. a given expression does not appear in either of the two texts being compared), which has been proved to be irrelevant to the study of the variable lexicon.

The result of the calculation is as follows:

[PHIX]	Fazienda	E8	GE	E4	E7	E19	Ajuda	Alba	E3
Fazienda	1.000								
E8	0.254	1.000							
GE	0.172	0.468	1.000						
E4	0.231	0.272	0.181	1.000					
E7	0.229	0.249	0.181	0.474	1.000				
E19	0.221	0.212	0.168	0.466	0.865	1.000			
Ajuda	0.241	0.258	0.197	0.615	0.482	0.474	1.000		
Alba	0.249	0.260	0.210	0.466	0.399	0.395	0.457	1.000	
E3	0.252	0.274	0.198	0.613	0.464	0.451	0.882	0.437	1.000

Table 4. Modified correlation coefficients, phi

Observing this table, it is evident that some pairs of texts show a high level of correlation, while others show a relatively low one. For example the correlation coefficient of the pair of Ajuda and E3 is 0.882, the highest of the group. The correlation coefficient of the pair E7 - E19 is also high, 0.865. Should one want to assess the proximities relative to a fixed variable, these figures can also be listed in ascending or descending order. For example, with regard to the text of Ajuda, the figures can be presented as follows: E3: 0.882, E4: 0.615, E7: 0.482, E19: 0.474, Alba: 0.457, E8: 0.258, Fazienda: 0.241, GE: 0.197, indicating a descending level of correlation.

From the correlation matrix, a dendrogram can be produced via cluster analysis (method of mean values between groups)⁶:

⁶ For further information on cluster analysis, see Romesburg's monograph (1989), recommended for its completeness and applicability for research in the social sciences.

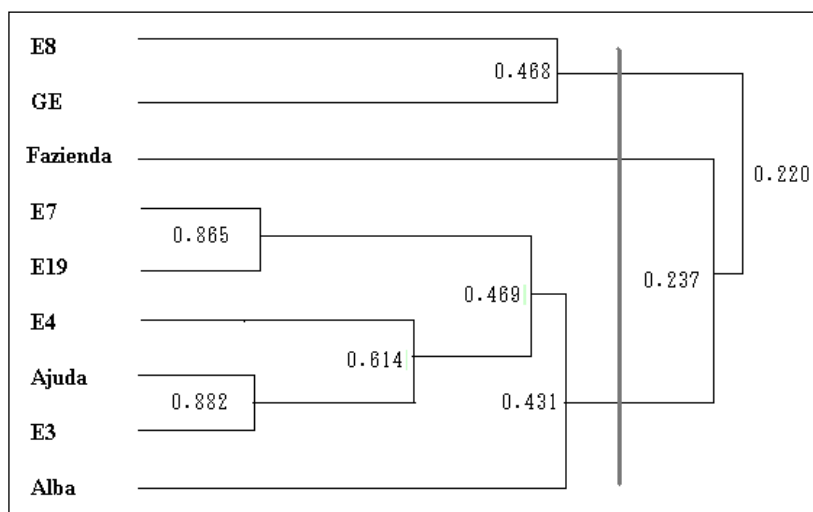


Figure 1. Dendrogram

The graph can be interpreted either from right to left or from left to right. Starting from the right side, it can be observed that the figure is basically split into two groups {E8, GE} and {Fazienda, E7, E19, E4, Ajuda, E3, Alba}, with Fazienda the first to deviate from the latter group, and then Alba. The remaining texts can be grouped as follows: {E7, E19} and {E4, Ajuda, E3}. By drawing a vertical line at any point, it is possible to establish provisional groupings. For instance, the vertical line in Figure 2 enables a classification into three main groups: {E8, GE}, {Fazienda}, {E7, E19, E4, Ajuda, E3, Alba}. For each group and subgroup, the correlation coefficient that occurs at the point of connection is indicated. As this method uses the average value of coefficients between the groups, the figure is calculated as the sum of all correlation coefficients between relevant pairs divided by the total number of pairs.

In previous studies various groupings of biblical texts have been proposed⁷. According to the latest schemes derived by Lazar (1994) and Pueyo (1996), and modified by Pueyo (2008) as indicated by Avenzoa (2008), the texts discussed here can be grouped by their line of descent as follows: {Fazienda}, {E3, Ajuda}, {Alba}, {E4}, {E7, E19} in the group of translations of the Hebrew original, separately from the texts translated from the Latin Vulgate: {E8, GE}.

My lexicostatistic taxonomy basically coincides with this genealogical group. The position of Fazienda, of Hebrew origin, is noteworthy, as it maintains its independence up to the point marked as 0.237, where it joins the rest of the family of texts translated from Hebrew. Alba is also independent to some extent

⁷ See Berger (1899), Llamas (1950-1955), Morreale (1960), Amigo Espada (1983), Avenzoa (2001).

from the rest of the Hebrew group. Within the family of Hebrew Bibles the closeness between E7 and E19, on the one hand, and between Ajuda and E3 on the other, is due to each pair being two testimonies of the same translation.

The advantage of a dendrogram in the study of comparable texts is its ability to offer a multistratic classification based on accurate quantitative measurements. Ideally, the grouping based on qualitative philological analysis should coincide with the quantitative taxonomy offered by multiple correlation analysis. This is precisely the case with the lexical taxonomy described in this study.

3. PRINCIPAL COMPONENT ANALYSIS

The principal component analysis offers a new interpretation of the variables, which in our case are the biblical texts. It is intended to draw new component lines to explain the variation in the data matrix⁸. The result of this analysis is as follows:

Texto	Componente 1	Componente 2	Componente 3	Componente 4	Componente 5
Fazienda	-0.112	0.340	-0.313	0.763	0.310
E8	-0.004	-0.569	0.218	0.469	0.265
GenEs	-0.094	-0.591	0.290	0.096	-0.065
E3	0.455	-0.189	-0.352	0.113	-0.241
Ajuda	0.466	-0.171	-0.332	0.087	-0.223
E19	0.400	0.273	0.498	0.121	-0.028
E7	0.405	0.238	0.499	0.164	-0.009
E4	0.398	-0.103	-0.180	-0.038	0.077
Alba	0.269	-0.070	-0.080	-0.355	0.846

Table 5. Principal components

According to the table, the component 1 divides texts of the 13th century {Fazienda, E8, GE} from those of the 15th {E4, E7, E19, Ajuda, Alba, E3}. Component 2 in turn separates {Fazienda} from {E8, GE} by placing them at opposite ends of the distribution of values of that component, reflecting the independence of translation between the extreme two groups⁹. The chart below

⁸ See Woods et al. (1986: 273-295). The Principal Component Analysis is one of the multivariate methods seeking internal structure of numerical information. It presents the main components explaining the greatest variance in distribution, instead of the initial coordinate axes. By interpreting the first few components it is possible to find new classification criteria of data variables in a reasonable way.

⁹ I owe much to Francisco J. Pueyo in these interpretations: «Fazienda's biblical fragments tend to be a very literal translation of the Hebrew text; the translator

summarizes the strength of the various components, with the first two explaining 32.8% and 15.5% of the variability in the data, respectively:

	Component 1	Component 2	Component 3	Component 4	Component 5
Eigen value	2.956	1.395	1.284	0.911	0.851
%	32.846	15.503	14.266	10.121	9.454
% (accumulated)	32.846	48.349	62.615	72.736	82.190

Table 6. Percentage of explanatory value

Then data in table 3.2 shows that the first division of the texts into those translated in the 13th century and those translated in the 15th century is twice as important as the division based on source language: {Fazienda}, which was translated from Hebrew as opposed to {E8, GE}, both translated from Latin.

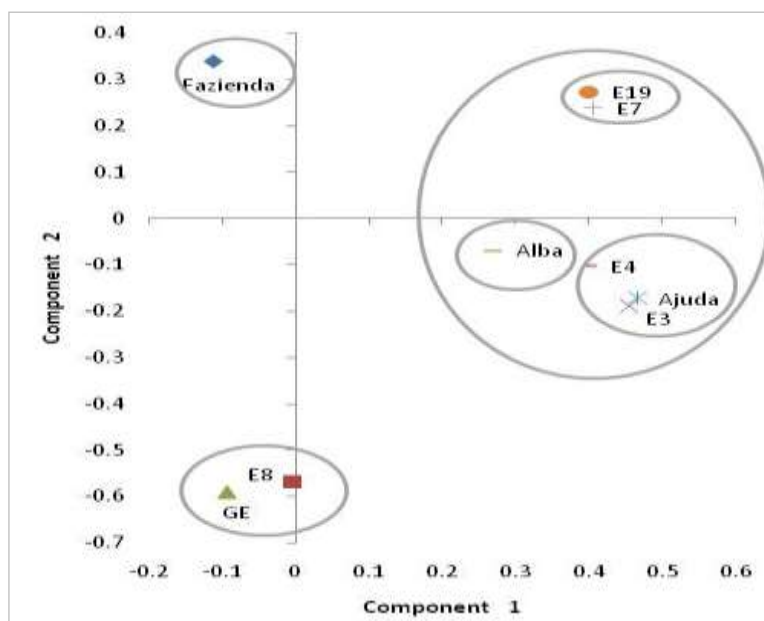


Figure 2. Tree groups

The lexical similarity between E8 and GE is remarkable, which is also verified in Table 4 and Figure 1. Another highlight is the fact that Alba is closer to the

probably had little to do with the design of the book and therefore did not have his own literary intention. Precisely the most literary text of these versions would be that of *General Estoria*, which inserts the biblical story into a historical narrative and sometimes paraphrases the biblical translated text. It is actually easier to explain Component 2 considering that the *Fazienda* is a translation from Hebrew and not from Latin, which can clearly influence the lexical choice of both groups. Similarly, it is worth pointing out that in this Component 2 positive values of *Fazienda* and E19/E7, and relatively low negative values of Alba would indicate different degrees of independence on the part of each one of the translators» [my translation].

group {E4, Ajuda, E3} than to the group {E7, E19} within the family of texts translated from Hebrew.

4. CONCENTRATION

In Ueda (2006) I have developed a method for rearranging the two axes of the two-dimensional distribution table. The purpose of this method is to look at the distribution of occurrences (Table 2), through a procedure called "diagonalization" of a continuous distribution pattern, transforming the order of variables (in our case, texts) in a horizontal line and/or the order of linguistic forms in a vertical line. The distribution thus diagonalized allows a unified interpretation of both the vertical axis and the horizontal one. In the calculation, for the merit of accuracy, I use the Minkowski's formula instead of Euclidean distance, usually used to search the resemblance of distribution. The points marked in Figure 3 represent the cohesion of texts and linguistic forms at the same time:

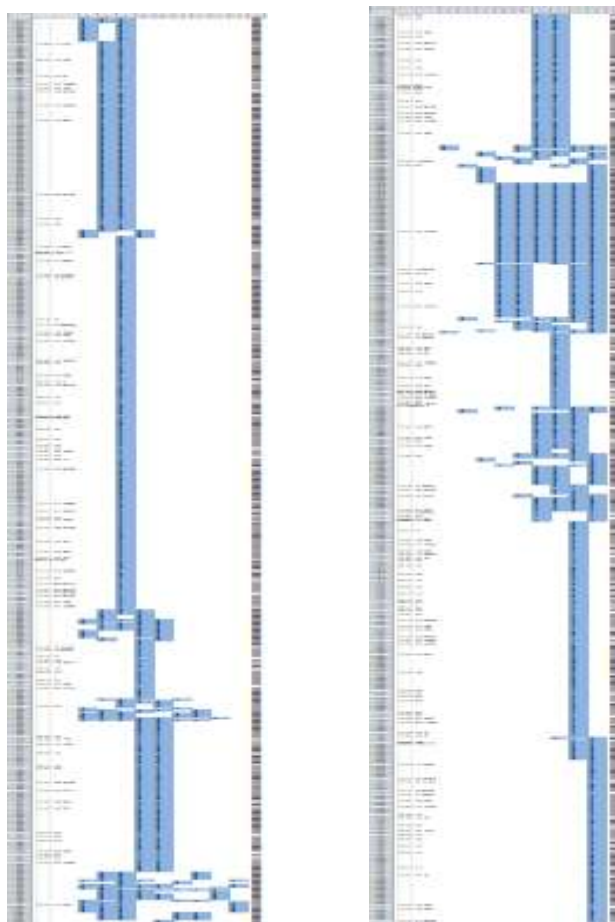


Figure 3. Concentration

I have named this method *concentration*. By applying it I have found linguistic forms that are exclusive to each group: {1: Fazienda}, {2: E8, GE}, {3: E4, Ajuda, E3}, {4: E7, E19} and {5: Alba}¹⁰ (See Appendix).

This list consists of unique linguistic forms, i.e. those that are not present in more than one group, to the exclusion of other parallel texts in the same verse. In order to provide a quantitative analysis for this phenomenon of co-occurrence I have prepared the following table:

¹⁰ Llamas (1944: 233) explains the originality of Alba: «neither Arragel [translator of Alba] really depends on the Escorial version [E3], nor the latter on the version of the Rabbi» [my translation]. And then he emphasizes the Hebrew character of E3 adding that «(...) after I-I-3 [E3] Arragel's version is the most Jewish of all the Castilian bibles» [*id.*].

	Fazienda	E8	GE	E4	E7	E19	Ajuda	Alba	E3
Fazienda	240								
E8	82	440							
GE	46	188	400						
E4	81	135	89	490					
E7	79	115	83	233	485				
E19	76	101	77	226	424	486			
Ajuda	83	127	93	301	238	233	488		
Alba	90	127	102	228	195	192	220	489	
E3	86	131	92	299	229	223	433	213	480

Table 7. Co-occurrence

I would call attention here to the similarities between E8 and GE on the one hand, and E4, E3 and Ajuda on the other, which has been demonstrated also in 2.

Pueyo (1996: LVII) presents an equation of the *Book of Judges* as follows: [E3 / Ajuda ~ E19 = E7 ~ E4] ≠ Alba. My lexicostatic taxonomy matches this equation except for the position of E4. According to the lexical data in Chapters 13 to 16 of the *Judges*, E4 is closer to {E3 Ajuda} than to E7 (see Figures 1 and 2). Indeed, according to the latest study of Pueyo (2008), presented in Avenozza (2008: 40), in regards to the book of *Judges*, along with the *Pentateuch* and *Joshua*, «[E4] does not reproduce any other known romance translations, which makes it unique in the panorama of medieval Bibles (...). Thus, E4 was isolated in the scheme designed by Pueyo (...)» [my translation].

5. CONCLUSION

To obtain a more complete description and classification of medieval texts and an explanation of the causes of the similarities and differences between them, several linguistic and extralinguistic analyses should be carried out¹¹. Among the linguistic factors to be analysed, graphophonemic, morphosyntactic¹², and lexical aspects should be considered in order to complement this study with special attention to quantitative aspects¹³. Some of the extralinguistic work, as

¹¹ See, for example, Colón Domènech (2002).

¹² Enrique-Arias (2008), Vincis (2009).

¹³ Francisco J. Pueyo in evaluating an earlier version of this paper suggests the following two points which deserve further study: «(1) Since we have taken into account different points copies of the same text (E7/E19 and E3/Ajuda) it would have been interesting to detail and explain the type of variation found between each pair of copies (i.e. which categories exhibit more or less variation). This analysis would have contributed significantly to a better understanding of the mechanisms of lexical adaptation that occurred in the copying process in the Middle Ages. (2) The high degree of lexical correlation between E8 and GE requires some form of qualitative discussion, considering whether it could be caused by a direct textual relationship or

characteristics of graphs, paleographical allographs, abbreviations, etc., should also use statistical methods. Each one of these perspectives would offer interesting and complementary contributions, which should be assessed jointly in order to improve our understanding of the world of medieval biblical texts.

The aim of this paper has been to propose the use of quantitative methods exclusively in lexical analysis. The sample text used in this analysis has been limited to four chapters of the *Book of Judges*. However, it could serve as a good example of how the quantitative methods described here could be applied to the linguistic analysis of other medieval Spanish texts.

REFERENCES

- ANDERBERG, Michel R. (1973): *Cluster analysis for applications*. New York: Academic Press.
- AVENOZA, Gemma (2008a): «Las traducciones de la Biblia en castellano y sus comentarios», in G. del Olmo (coord.), *La Biblia en España*, vol. 1.2. Barcelona: Trotta, pp. 13-75.
- AVENOZA, Gemma (2008b): «Manuscritos bíblicos ibéricos», *Congreso Internacional Códices Literarios Españoles (Edad Media) VIII Centenario del códice del Cantar de Mio Cid*. San Millán de la Cogolla: Fundación San Millán de la Cogolla.
- AVENOZA, Gemma y Andrés ENRIQUE-ARIAS (2005): «Bibliografía sobre las Biblias romanceadas castellanas medievales», *Boletín Bibliográfico de la Asociación Hispánica de Literatura Medieval*, 19, pp. 409-454.
- CASTRO, Américo, Agustín MILLARES CARLO y Ángel J. BATTISTESSA (1927): *Biblia medieval romanceada. I. Pentateuco*. Buenos Aires: Universidad de Buenos Aires, Facultad de Filosofía y Letras.
- COLÓN DOMÈNECH, Germán (2002): «Sobre versiones del *Cantar de los cantares* en castellano medieval y renacentista», in Pedro Álvarez de Miranda y José

simply by the influence of underlying Latin words of the Vulgate. For example, in *Judges* 13:9 we have E8 y *aparesció* 'he appeared' and also GE y *paresció*; in contrast, all Hebrew versions agree in translating y *vino* 'and he came'. This coincidence does not necessarily presuppose the direct textual relationship between E8 and GE (highly debated issue, on the other hand, in the field of Bible translations); rather it may result from the presence of Vulgate *apparuit* in the same position (while in Hebrew we have *vayabó* 'and he came'). The result is that two different translations may have lexical coincidences owing to the influence of a lexical item in the underlying text. It would be necessary to control this type of factor to determine if it occurs in a consistent way» [my translation].

- Polo (eds.), *Lengua y diccionarios, Estudios ofrecidos a Manuel Seco*. Madrid: Arco/Libros, pp. 95-116.
- ENRIQUE-ARIAS, Andrés (2004): «Texto subyacente hebreo e influencia latinizante en la traducción de la *Biblia de Alba* de Moisés Arragel», en Victòria Alsina, Jenny Brumme, Cecilio Garriga, Carsten Sinner (eds.), *Traducción y estandarización. La incidencia de la traducción en la historia de los lenguajes especializados*. Madrid-Frankfurt: Iberoamericana, pp. 99-122.
- ENRIQUE-ARIAS, Andrés (2006): «Sobre el parentesco entre la *Biblia de Alba* y la *Biblia* de la Real Academia de la Historia ms. 87», *Romance Philology*, 59, pp. 21-43.
- ENRIQUE-ARIAS, Andrés (2008a): «Apuntes para una caracterización de la morfosintaxis de los textos bíblicos medievales en castellano», in Johannes Kabatek (ed.), *Sintaxis histórica del español y cambio lingüístico. Nuevas perspectivas desde las tradiciones discursivas*. Madrid-Frankfurt: Iberoamericana, pp. 109-125.
- ENRIQUE-ARIAS, Andrés (2008b): «Biblias romanceadas e historia de la lengua», in Concepción Company, José Moreno de Alba (eds.), *Actas del VII Congreso Internacional de Historia de la Lengua Española*. Mérida (Yucatán), 4-8 de septiembre de 2006. Madrid: Arco/Libros, pp. 1781-1794.
- ENRIQUE-ARIAS, Andrés (2011): «Traduciendo la Biblia en la Castilla medieval. Nuevas perspectivas a la luz de la edición electrónica integral de los romanceamientos bíblicos medievales en castellano», *Letras de Deusto*, 133, 41, pp. 13-41.
- FERNÁNDEZ LÓPEZ, María del Carmen (1996): «Una distinción fonética inadvertida en el sistema gráfico medieval», in Alegría Alonso González (ed.), *Actas del III Congreso Internacional de Historia de la Lengua Española: Salamanca, 22-27 de noviembre de 1993*, vol. 1. Madrid: Arco/Libros, pp. 112-123.
- LAZAR, Moshe (1995): *Biblia ladinada, Escorial 1.1.3. Critical edition, notes and commentaries*, vol. 1, Genesis-Kings. Madison: The Hispanic Seminary of Medieval Studies.
- LITTLEFIELD, Mark G. (1983): *Biblia romanceada 1.1.8*. Madison: The Hispanic Seminary of Medieval Studies.
- LITTLEFIELD, Mark G. (1992): *Escorial Bible 1.11.19. Edition, study and notes*. Madison: The Hispanic Seminary of Medieval Studies.
- LITTLEFIELD, Mark G. (1996): *Escorial Bible 1.1.7. Edition, introduction, notes and glossary*. Madison: The Hispanic Seminary of Medieval Studies.

- LLAMAS, José (1944): «La antigua Biblia castellana de los judíos españoles», *Sefarad*, 4, pp. 219-244.
- MORREALE, Margherita (1960): «Apuntes bibliográficos para la iniciación al estudio de las traducciones bíblicas medievales en castellano», *Sefarad*, 20, pp. 66-109.
- MORREALE, Margherita (1969): «Vernacular scriptures in Spain», in *The Cambridge History of the Bible*, vol. 2. Cambridge: Cambridge University Press, pp. 465-491 y 533-535.
- PUEYO MENA, Francisco Javier (1996): *Biblia romanceada. Biblioteca Nacional de Madrid. Ms. 10.288. Estudio, edición y notas*. Madison: The Hispanic Seminary of Medieval Studies.
- PUEYO MENA, Francisco Javier (2008): «Biblias romanceadas y en ladino», in Jacob M. Hassán, Ricardo Izquierdo Benito y Elena Romero (coords.), *Sefardíes: Literatura y lengua de una nación dispersa*. Cuenca: Universidad de Castilla-La Mancha, pp. 193-263.
- ROMESBURG, H. Charles (1989): *Cluster analysis for researchers*. Malabar, Florida: Robert E. Krieger Publishing.
- SÁNCHEZ PRIETO BORJA, Pedro (1998): *Cómo editar los textos medievales*. Madrid: Arco/Libros.
- SÁNCHEZ PRIETO BORJA, Pedro (2002): «Biblias romanceadas», in Carlos Alvar y José Manuel Lucía Megías (eds.), *Diccionario filológico de literatura medieval*. Madrid: Castalia, pp. 212-223.
- SCHONFIELD, Jeremy (ed.) (1992): *Companion volume to the facsimile edition. La Biblia de Alba. An illustrated manuscript Bible in Castilian, by Rabbi Moses Arragel*. Madrid: Fundación Amigos de Sefarad.
- UEDA, Hiroto (2006): «Patternized configuration of geolinguistic data: Application to Spanish lexical variation», in Markku Filppula, Juhani Klemola, Marjatta Palander, Esa Penttilä (eds.), *Topics in Dialectal Variation*. Joensuu, Finland: University of Joensuu, Faculty of Humanities, pp. 163-187.
- UEDA, Hiroto (2009): «Palatal graphemes in a medieval Spanish biblical text: a corpus analysis of “i, j, y” in Genesis, Biblia de Alba», in Yuji Kawaguchi, Makoto Minegishi and Jacques Durand (eds.), *Corpus analysis and variation in linguistics*. Amsterdam: John Benjamins Publishing Company, pp. 239-257.
- VERD, Gabriel María (1971): «Las biblias romanceadas. Criterios de traducción», *Sefarad*, 31, pp. 319-351.

VINCIS, Valentina (2009): «Describir textos a partir de elementos de cohesión: un corpus de romanceamientos bíblicos medievales», in Andrés Enrique-Arias (ed.), *Diacronía de las lenguas iberorrománicas: nuevas aportaciones desde la lingüística de corpus*. Madrid-Frankfurt: Iberoamericana, pp. 285-304.

WOODS, Anthony, Paul FLECHER and Arthur HUGHES (1986): *Statistics in language studies*. Cambridge: Cambridge University Press.

APPENDIX

1: Fazienda: *a* (14:16), *acordarse* (16:23), *adivinanciella* (14:18), *adivinar la palabra* (14:14), *adivinar la razon* (14:19), *adobar* (13:15), *adormir* (16:19), *alegrarse* (16:23), *aplegar* (14:08), *aplegarse* (14:09), *aprovescer* (14:19), *aptarse de* (16:01), *aver nombre* (16:04), *bever* (13:14), *brusco de las cabras* (13:15), *como era usage* (14:10), *complir palabras* (13:12), *contar* (14:06), *convidar* (13:16), *cortarle los cabellos* (16:19), *cosa suzia* (13:04), *crebantar con sus manos* (14:09), *cremar* (15:05), *cuello* (16:03), *cumplir las tus palabras* (13:17), *dar* (16:18), *dar* (16:23), *de noche* (16:03), *des que naciere* (13:05), *descubrir su coraçon* (16:17), *dezir* (14:02), *dezir la adivinanza* (14:17), *dezir so cuer* (16:18), *echar celada* (15:09), *echar celada* (16:02), *echar la mano en* (16:29), *echarse sobre sus fazes a tierra* (13:20), *emprenar* (13:03), *emprenarse* (13:02), *emprenarse* (13:07), *ensañe* (14:08), *estas cosas* (13:23), *fazer entender* (13:23), *fazer entender* (13:23), *fer sacrificio a* (16:23), *filisteo* (14:19), *fontana dagua* (15:19), *fuessa* (16:31), *gran comer* (14:10), *guiar de la mano* (16:26), *huessos* (15:15), *jazer* (16:03), *jogar* (16:25), *leon muerto* (14:08), *levar consigo* (14:11), *levarse* (13:11), *limbral* (16:03), *mancebo* (13:05), *mancebo* (13:07), *mancebo* (13:08), *mancebo* (13:12), *mancebo* (13:24), *mancebo* (16:26), *menuzar* (14:06), *miel* (14:14), *mil cien siclos de plata* (16:05), *mostrar* (13:25), *non tener nulla cosa en su mano* (14:06), *orar* (13:08), *pariente* (16:31), *placer* (13:23), *podestad* (16:05), *pregar* (13:08), *prender* (13:19), *prender* (13:23), *prender* (14:08), *prender* (16:03), *prender* (16:03), *prender muger* (14:03), *provescer sobre* (14:06), *que naciere* (13:07), *que no son de tu lei* (14:03), *rico omne* (16:23), *saltar* (13:20), *seder* (13:09), *septimo* (14:15), *soltar* (14:18), *suzio* (13:07), *tentar* (16:26), *trabajar en el molino* (16:21), *traer* (16:25), *trameter* (13:08), *troal dia desu muert* (13:07), *vaqujella* (14:18), *vengar* (16:28), *venir* (14:05), *venir* (16:31), *vestidura* (14:12), *villa* (14:15).

2: E8, GE: *aver fijo (13:03), aver fijo (13:07), dominar (16:05), donde venia (13:06), falagar (14:15), venir (16:05), falagar (16:05), mañana (16:02), ser señores (14:04), varon (13:08), venir (14:07), alegrarse (16:25), atar (15:04), atender (16:27), aver (16:06), barra (16:03), cabdieillo (16:30), capdieillo (16:27), capdieillos (16:08), cara (13:06), casa (16:09), clavo (16:14), cosa que no sea limpia (13:04), cuidar (15:02), dar voces (15:14), de nervios (16:07), demandar (13:18), despojar (14:15), despues (16:04), dexar ir (15:05), dezir (16:04), durante su vida (16:30), echar (16:03), empuxar (16:19), entender (13:21), estar (16:31), facha (15:04), fazer dormir (16:19), fazer juegos (16:25), fazer parar (16:25), fierro (16:17), fincar (16:13), ir (14:19), jugar (16:25), levar (16:03), ligadura (16:09), ligar (16:05), ligar (16:12), ligar (16:13), ligar (16:21), lizo (16:13), llamar (16:09), llamar (16:12), mandar venir (16:25), matar (15:12), mentir (16:15), meter en mano de (15:12), meter en nuestras manos (16:24), mill & cient dineros de plata (16:05), morar (15:08), morar (16:04), no ser mi culpa (15:03), nunca (16:17), parentesco (16:31), piedra (13:19), poner guardas (16:02), poner nombre (15:17), por la carrera (14:09), querer (13:16), raer (16:17), raer (16:19), recudir (16:07), rogar (16:28), romper (16:12), salir (14:05), segudir muy fuerte (16:30), ser derramado (15:09), somo (16:27), tañer (16:26), tejer (16:13), tomar (16:03), tomar (16:29), tomar (16:31), tomar vengança (15:07), trabajar (16:21), valentia (16:09), valentia (16:17), varon (16:27), vedar (15:01), veer (15:01), vencer (16:13), venir (15:14), venir (16:18),*

3: E3, E4, Ajuda: *arder en fuego (15:14), casa delos presos (16:21), descender fuego (15:05), detener (13:15), empreñar (13:03), encima (16:27), entrar con (15:01), esta vez (15:03), fablar mentiras (16:13), fazer (14:10), fazer oir (13:23), fender (14:06), fender (15:19), ferida grande (15:08), heredar (14:15), moço (13:24), muerte (16:30), peña (15:08), pico de la peña (15:11), posar (15:14), preñada (13:07), prender (16:06), pueblo (16:30), que non fue fecho con ellas obra (16:11), rescebimiento (15:14), rodilla (16:19), sacrificar (16:23), señor (16:27), sepultura (16:31), ser (16:06), ser libre (15:03), ser su nombre (16:04), su nombre (13:02), subir (14:19), sus hermanos & todos los de casa de su padre (16:31), telar (16:14), tocar en (15:12), trebejar (16:27), villa (14:18), villa (16:03)*

4: E7, E19: *a cabo de un año (14:08), a cabo de un año (15:01), aborrrir (15:02), alcançar (14:18), alçar navaja (13:05), almaizares (14:12), almaizares y paños (14:19), andar (16:18), aparejar (13:15), apodestar en (14:04), apretar (16:29), ara (13:20), arriba (16:27), atahona (16:21), atar (15:04), atar (15:13), atar*

(16:13), *aver por nombre* (13:02), *aver por nombre* (16:04), *bezarse* (13:25), *caçar* (15:04), *castillo* (15:08), *castillo de la peña* (15:11), *cata* (13:10), *catulo* (14:05), *celada tener puesta* (16:09), *costumbre* (13:12), *criatura* (13:08), *criatura* (13:24), *daño* (15:03), *dar alabaças* (16:24), *dar salvacion* (15:18), *de cada dia* (16:16), *del vientre* (13:07), *demostrar* (13:23), *denunciar la apuesta* (14:14), *departir toda su voluntad* (16:17), *descubrir la poridad* (14:17), *desgarrar* (14:06), *desque viniere lo que dixiste* (13:12), *desta vez* (15:03), *dezir las mentiras* (16:13), *duque* (16:27), *echarle prisiones* (16:21), *echarse de fazes en tierra* (13:20), *empobrescer* (14:15), *encuentro* (15:14), *entender* (14:04), *entrar en* (15:08), *entregar en poder de* (15:12), *escabollir* (16:20), *escarnescer* (16:25), *escarnio* (16:27), *estar* (16:29), *estar en el mejor solaz* (16:25), *fablar* (13:23), *fasta el dia clarao* (16:02), *fazer gran sacrificio* (16:23), *fazer moledor* (16:21), *fortaleza* (15:08), *foyo* (15:19), *fresco* (15:15), *generacion* (14:03), *influir* (15:14), *ir* (15:14), *ir a* (15:01), *ir con* (14:11), *ir para* (16:05), *lleno de ira* (14:19), *llevar* (14:10), *logar* (16:03), *manejo* (15:04), *marido* (15:02), *matança grande* (15:08), *mentir* (16:13), *mimbre* (16:07), *mimbre* (16:09), *mirar* (13:20), *muchedumbre* (14:08), *non añadir mas a aparescer* (13:21), *pagarse de* (16:04), *palabra* (16:16), *panes* (15:05), *par mientes* (16:05), *parecer bien en ojos* (14:07), *perder la mi fuerça* (16:11), *placer en los ojos* (14:03), *por alegria* (16:23), *poridad* (14:18), *preguntar* (16:05), *quebrar* (16:09), *quebrar* (16:12), *querer ofrescer su anima ala muerte* (16:16), *quiça* (14:15), *reir* (16:25), *renuncio* (14:15), *rescibir* (13:23), *ruego* (13:09), *ser atadura* (16:11), *ser del señor* (14:04), *ser llegados* (16:23), *soler fazer* (14:10), *su derecha* (16:29), *su izquierda* (16:29), *tan sola esta vegada* (16:28), *ternera* (14:18), *tesoro* (16:18), *texedura* (16:14), *tirarse* (15:14), *tirarse la fuerça* (16:17), *tomar preso* (16:05), *tornarse el alma* (15:19), *traer* (16:25), *traer* (16:31), *varon* (15:11), *varon* (15:15), *varones de mi pueblo* (14:16), *vegada* (16:15), *velar* (16:02), *venir* (14:19), *venir a* (16:01)

5: Alba: *abraçar* (16:29), *absolver* (14:19), *adextrar* (16:26), *adversario* (16:23), *aflacar* (16:07), *aflacar* (16:11), *afligir* (16:05), *afligir* (16:06), *afligir* (16:19), *aparcearse con* (14:11), *arancar* (16:12), *asentar* (15:09), *asir* (16:21), *asir en* (16:03), *aun agora* (16:13), *aver nombre* (15:19), *avolverse a* (16:01), *burlar* (16:13), *buscar ocasion* (14:04), *cavillo* (14:05), *celda* (15:01), *ceradura* (16:03), *cerca de* (14:16), *cibdat* (14:18), *cognoscer* (16:09), *con gran gozo* (16:23), *confortar* (13:25), *congregarse* (16:23), *contar toda su voluntad* (16:18), *convenir* (14:02), *cordel* (16:12), *cosa no tener en su mano* (14:06), *cosa polluda* (13:14), *costumbrar* (14:10), *cuerpo muerto* (14:08), *dar alaridos et gozos* (15:14), *dar en*

poder nuestro (16:24), dar en su poder (15:13), derrotar (15:08), descruir (16:24), dezir (15:18), dezir mentiras (16:13), dezir todo su coraçon (16:17), echar espia (16:02), echar mientes en (16:05), el tiempo del segar trigos (15:01), empuxar con muy grand fuerça (16:30), en aquel tiempo (14:04), encintar (13:03), enduzir (14:15), enduzir (16:05), enseñar (13:23), enseñar (13:23), escabulirse (16:20), espalda (16:03), esprimir con sus palmas (14:09), estar con (15:01), estar en cinta (13:05), estar ledos y callados (16:02), estenderse (15:09), examen (14:08), facer oracion (13:08), fasta cabo de (16:03), fasta este presente dia (15:19), fasta ya que alumbre la mañana (16:02), fazar grand salvacion (15:18), fazer gran matança (16:24), fazer honra (13:17), fazer juegos (16:27), fazer molinero (16:21), fazer ofrenda de (13:15), fierro (16:21), forçar (13:15), garote (16:14), garote (16:14), gente (16:30), grand ferida (15:08), habitar (15:08), haredar (14:15), holocasta (13:23), ir (15:11), ir (16:31), ir muy corriendo (13:10), joven (14:10), juizio (13:12), la otra (16:29), lançar (15:17), legar a tentacion de muerte (16:16), ligamientos (15:14), llamar (16:04), llamavan (13:02), llegar (14:05), llegar (14:08), lo que dixeste (13:12), maginar (15:02), maillar contra el (14:05), mal querer (14:16), mal querer (15:02), miesses (15:05), mill ciento pesos de plata (16:05), mojado (16:07), muela (15:19), muger meretris (16:01), niño (13:05), niño (13:07), non circuncisos (15:18), non tornar mas a aparecer (13:21), nueva mente (13:09), oler (16:09), passar navaja (13:05), pegar (16:14), peña (15:11), pesar mucho (14:19), plañer (14:17), poner cerca (16:26), poner en fierros (16:21), poner sobre (13:19), principe (16:27), principe (16:30), principes (16:08), proponer (14:12), proposicion (14:15), proposicion (14:18), provescer sobre (14:19), provezir (15:14), pujar sobre (14:06), quanto (13:13), que non ayan conellos ninguna obra fecho (16:11), que secas no fuesen (16:07), quedar preso (16:06), querer (16:15), rapar (16:19), rapar (16:22), responder (15:02), responder (15:06), sacrificar grandes hostias (16:23), señorar (14:04), sif (15:08), soltar (15:05), solver la proposicion (14:14), sufrir (16:26), tan sola esta vez (16:28), tañer (15:12), tejado (16:27), tela (16:14), tener conmigo (16:15), tener coraçones alegres (16:25), tirarsele la fuerça (16:19), todas aquestas cosas (13:23), tomar amor con (16:04), tornar (13:08), tornarse a (14:19), una mano (16:29), varon (16:19), veer (15:14), venir en tu nombre (13:17)