Words to Encourage Evangelization: A Comparison of Lexical Frequencies in the Writings of Pope Francis and Fr. Tomás Morales, SJ

Abstract

Studying language through naturally-occurring data is easily feasible nowadays thanks to the use of concordancers. Using the software package AntConc by Laurence Anthony, the author approaches the linguistic style of Pope Francis in four of his apostolic exhortations. His lexical frequencies are compared to those of Venerable Fr. Tomás Morales, a Spanish Jesuit (1908–1994) who is considered a prophet of our time for anticipating the teachings of the Second Vatican Council twenty years beforehand. The two have several aspects in common, mainly in making laypeople aware of the universal call to sanctity and the missionary responsibility acquired in baptism. The results of the comparison show a similarity in the lexical choices. The article’s conclusions and implications are considered in relation to the word in education and in evangelization.

Keywords: corpus linguistics, stylistic choices, Pope Francis
Introduction

The call for papers in this journal invited us to send proposals that would develop topics related to the Word in Education and several sub-themes. This paper relates to the subtheme of “encouraging and supportive ways of communication in teaching and parenting,” since evangelization could be somehow considered both teaching and parenting. The lexical frequencies in the writings of two pastors are compared: Pope Francis and the Spanish Jesuit Venerable Fr. Tomás Morales Pérez.

The paper begins with a description of corpus analysis and its applications for the study of literature. Then, the contextual part presents the two writers under comparison. The methodology section includes the corpus of writings being analysed and the software tool that performed the analysis, as well as the aim of the research. The results of the analysis are presented and discussed from a quantitative and qualitative perspective. Some conclusions are drawn in order to make a contribution to corpus analysis in literature.

Theoretical Framework

Corpus Analysis and Digital Humanities: The Concept and Its Applications

In linguistics, corpus analysis is a form of text analysis which permits comparisons between textual objects on a large scale. The potential of electronic corpora revolutionized linguistic research in the past few decades (Römer, 2011). Corpus analysis of electronic texts offers possibilities for applied linguistics, language pedagogy, lexicography, library and information science, and others. Corpus analysis is not only for linguists (Bowker, 2018); for instance, applications for pedagogical corpus analysis include syllabus or materials design, as well as direct applications for corpora in the second-language classroom (Sinclair, 2004; Reppen, 2010).

The label digital humanities refers to emerging methodologies that apply computational techniques to questions in the humanities. The
digital humanities offer what Berry (2011) called the “computational turn,” a fundamentally different way of engaging with digitized documents. Traditional humanities methods provide what has been called close reading, in contrast to distant reading, which electronic corpora enable. Close reading can only uncover a small portion of a given archive or corpus at once. “Distant reading refers to the processing of content in (subjects, themes, persons, places, etc.) or information about (publication date, place, author, title) a large number of textual items without engaging in the reading of the actual text” (Drucker, 2013, p. 62). This form of distant reading reveals facts that may not be necessarily noticed by a human reader. Corpus analysis provides empirical data, so it is especially useful for testing intuitions about texts (Szudarski, 2018). In addition, when distant reading is combined with traditional methods, this approach has the potential to uncover new patterns.

Corpus Analysis and Literature

In the context of digital humanities described above, corpus-based analytical techniques have also been applied to the study of literature. This application of corpus linguistics tools in literary texts is commonly referred to as corpus stylistics (Biber, 2011). Research in this field focuses on the distribution of words (analyzing keywords, lexical phrases, or collocations of words). One of the purposes is to identify textual features that may be especially distinctive of an author or a particular text. Biber (2011) clarified that this research approach blends the methods of corpus analysis and its qualitative and quantitative results with the traditional stylistic aims of stylistic research.

Computational and statistical research have dealt with authorship attribution and the literary style of writers such as Charles Dickens, William Shakespeare, Jane Austen, R. L. Stevenson, the Brothers Grimm, and Virginia Woolf, among others. A representative compilation of what can be achieved through corpus analysis in literature can be found in Keshabyan and Almela (2012) and Balossi (2015). The distant reading approach to the above-mentioned authors and their texts reveals facts which are not visible from a close reading approach and which make
a contribution to understanding an author’s style and achievement, for instance, the high number of words with violent connotations in the Grimms’ Cinderella, the differentiation of a novel’s characters through their vocabulary, or the different connotations in the use of diminutive forms.

This paper aims to make a contribution in proving the potential of this distant reading approach by comparing the lexical frequencies of the two writers described in the next section.

**The Writers**

Pope Francis and Fr. Tomás Morales present some common features. They are both Jesuits. They both speak Spanish as their mother tongue. The timespan of the periods in which they wrote is 60 years, as Fr. Morales wrote the works analyzed herein between the 1960s and the 1980s, and Pope Francis wrote the three selected documents between 2013 and 2019. Neither of them is an academic or a literary writer; writing is a pastoral and apostolic instrument for both of them. The two Jesuits concur in the ideas transmitted in these writings; they both had a large responsibility in the Church. Finally, the two authors both live what they preach. These coincidences and similarities bring meaning to the research question addressed in this paper.

After a brief comment of Pope Francis’s vocabulary, Fr. Morales will be presented since he is less known and some information about him and his written works will facilitate the comprehension of this paper’s aim and contribution.

**Pope Francis’s Vocabulary**

Pope Francis has shown himself to be “a great communicator.” His language shows the traces of his Latin American spirit. He speaks with the style of a pastor, in contrast with St. John Paul II, whose communicative style was connected with his knowledge of the word in poetry and theater. He is unpretentious and he makes his message accessible to
everyone. For this Pope, words are of fundamental importance due to their ability to build both bridges and walls, as reflected in a book which collects the key words in his ministry (McElwee & Wooden, 2020). His particular vocabulary has received the label of Bergoglismos (Nuñez, 2018). Expressions such as “primerear” [to take the first step], “Hagan lio” [Make a ruckus!], the sin of “habriaqueísmo” [wasting time talking about “what needs to be done”] or “No balconeen la vida” [Don’t stand aloof!] are a real challenge to convey in other languages. Translators try their best but suffer from the limitations of various languages to render the spontaneity and richness of the source expression.

This feature of the Pope’s linguistic style could be the object of study of a complete research paper. The above expressions are Ἰπαξ λεγόμενον (áπταξ λεγόμενον), meaning that they appear only once in a text. This paper rather focuses on the words most frequently used in Pope Francis’s writings, specifically the four apostolic exhortations under study.

Fr. Tomás Morales Pérez, SJ

Fr. Tomás Morales was born in 1908 in Macuto, Venezuela. After studying at a Jesuit school and studying law in Madrid and obtaining his doctorate in Bologna, he joined the Jesuits in 1932 and was ordained in 1941 in Spain. He dedicated his priesthood to the development of the laity. He founded secular institutes of the Crusaders of Mary for both men (1955) and women (1965); he set up the family movement Homes of Mary and the apostolic youth movement Militia of Mary. As a Jesuit, Fr. Morales drew on the formative influence of the Ignatian Spiritual Exercises to provide a spiritual backbone for these movements.

In life he was known for his holiness; the process of his beatification began in 2000, less than a decade after his death. Pope Francis confirmed that Fr. Morales had led a model life of heroic virtue and declared him Venerable in 2017.¹ His experience and educational approaches appear clearly in his books, of which there have been several editions and translations.

¹ For more information about his life and cause for beatification see https://padretomasmorales.weebly.com.
A summary of these books as well as the key aspect of each one is presented below.


Fr. Morales has been called the prophet of our time because before the Second Vatican Council he emphasized the essential and non-transferable role that laypeople have in the Church and their potential for evangelization. Thus, he anticipated the teachings of *Apostolicam Actuositatem* (1965), the Second Vatican Council’s Decree on the Apostolate of the Laity.

The *Forge of Men* is perhaps the most emblematic and popular of his writings. Requests for it arrive continually at the Postulation of his cause for canonization. In 2017, the book was translated into French, English, German, and Italian. The Polish translation was published in 2018, with a prologue by Sławomir Oder, Postulator of St. John Paul II’s cause for canonization. These translations allow speakers of other languages to have access to this unique book, presented outside the linguistic and cultural context in which it was developed. The originality of the work resides in its intimate connection with experience. Ignatian pedagogy permeates the book and offers the reader the experience of a new way to educate young people through virtues, education of heart, clear thinking, and strong will. In spite of these emerging translations of his works, Fr. Morales is still largely unknown outside the Spanish-speaking world.
Methodology

Study Corpus and Aims

This study compares two collections of texts: the four apostolic exhortations by Pope Francis and three books by Fr. Morales. The four apostolic exhortations are as follows:

1. Evangelii Gaudium [The Joy of the Gospel], 2013
3. Gaudete et exsultate [Rejoice and Be Glad], 2018
4. Christus vivit [Christ is Alive], 2019

An apostolic exhortation is a magisterial document written by the Pope which generally encourages the faithful to practice a particular virtue, attitude, or activity. An exhortation does not define Church doctrine and is not considered legislative. These four apostolic exhortations will be compared with Fr. Morales’s pedagogical trilogy.

1. Forja de Hombres [The Forge of Men]
2. Laicos en Marcha [Laity on the Move]
3. Hora de los laicos [The Hour of the Laity]

The purpose of these three books is also to encourage. Thus, The Forge of Men encourages the educator of youth to follow four cardinal points in the process of education: demanding spirit, fighting spirit, the cultivation of reflection, and the school of constancy. Laity on the Move presents encouraging principles to promote apostolic action among the laity; the reflections about Christianity in The Hour of the Laity encourage people to live the demands of our baptismal vocation in society.

As mentioned above, this research aims to make a contribution to prove the potential of distant reading. Distant reading, as defined in the theoretical framework, entails the processing of content or information without engaging in the reading of the actual text. The main research question concerns the most frequent vocabulary in the two sets of texts.
This approach has been applied to literature in the artistic sense of the word. This paper applies this distant reading to the domain of spiritual and religious literature, a field in which, to our knowledge, this type of analysis has not yet been performed.

**Analytical Tool**

The manual process of counting the words or the times each word occurs would be time-consuming, labor-intensive, and error-prone. However, this type of work is easily accomplished by a computer, which name carries the etymological meaning of counting. Specialized software tools can be used to calculate several different measures of frequency, statistical analysis of linguistic data, and other functionalities.

The software used in this research is Antconc Concordance, a free software program developed by Laurence Anthony of Waseda University in Tokyo, Japan. This software contains various tools for corpus analysis: a concordancer with numerous functions, a collocation searcher, a word lister, and a keyword tool. It also provides information about the number of words (tokens) and of different words (types) in the corpus. Thus, if the word “universal” occurs three times, it is counted as three tokens but only one type. The relationship between the number of types and the number of tokens is known as the type-token ratio (TTR). This study uses the Word List tool, which takes the words in the corpus and ranks them in order of frequency.

In the interface of the Word List tool (see Figure 1), the first column (Rank) shows the position of the word in the frequency list. The second column (Frequency) displays the number of times each word is repeated. The third column (Word) presents the word itself.

The processing of frequency lists in any running text will show the most frequently used words of any language on the top ranks: articles, prepositions, some connectors, and determiners. These words are called function words because they do not convey any lexical meaning. On the other hand, words with lexical meaning (nouns, verbs, adjectives, and adverbs) are called content words and they provide more valuable information in corpus analysis (Szudarski, 2018).
Function words will be ranked highest in frequency (for instance “the,” “and,” “for,” and “of”) regardless the type of texts. For this reason, the software includes “stop lists,” which can exclude the most common function words from being counted and displayed in the frequency list. A stop list was applied in this research so that the frequency list would start with only content words.

In summary, the steps in the research process were as follows:

1) Corpus compilation and adaptation to the format required by the software tool
2) Generation of a word list
3) Application of a stop list to ignore function words
4) Second generation of a word list
5) Comparison/contrast/interpretation

These steps are in accordance with the guidelines established for research in corpus linguistics (Biber et al., 1998). Research begins with quantitative findings and works toward qualitative ones. However, the procedure may have cyclical elements depending on the research question and on the data themselves.

Results

As stated above, one of the applications of corpus analysis is the testing of intuitions. At this stage of the paper, before communicating any results, we would like to interact with our readers by inviting them to hypothesize which content words appear most frequently in Pope Francis’s four apostolic exhortations. The answers to this question may depend on whether they have read the exhortations (or at least some of them) or not. However, any response will always be intuitive, because readers will rely either on thematic content, previous knowledge, familiarity with the author, or other similar variables. In contrast, the software tool will provide empirical evidence. The same applies to readers of Fr. Morales, who
are invited to conjecture about the content words which would be at the top of a frequency list.

The first step in the analysis was to calculate the word list for each of the four exhortations using the software tool. The highest ranked words are presented in Table 1.

**Table 1. Words most frequently found in Pope Francis’s apostolic exhortations**

<table>
<thead>
<tr>
<th>Exhortation</th>
<th>Evangeli Gaudium</th>
<th>Amoris Laetitia</th>
<th>Gaudete et exsultate</th>
<th>Christus vivit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three most frequently used words</td>
<td>Dios (God) Iglesia (Church) vida (life)</td>
<td>amor (love) familia (family) vida (life)</td>
<td>jóvenes (youths) vida (life) dios (God)</td>
<td>Dios (God) vida (life) señor (Lord)</td>
</tr>
</tbody>
</table>

It can be observed that the most common words in each exhortation are closely related to the central subject of the exhortations, for example, “love” in *Amoris Laetitia*. If the lexical frequency is calculated using the complete corpus of the four exhortations, “vida” (“life”) obviously becomes the word ranked first, as presented in Figure 1.

**Figure 1. Total frequencies in Pope Francis’s exhortations**
Regarding Fr. Morales, the same analysis was done with the Word List tool for the individual books. Table 2 summarizes the results.

<table>
<thead>
<tr>
<th>Book</th>
<th>The Forge of Men</th>
<th>Laity on the Move</th>
<th>The Hour of the Laity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three most frequently used words</td>
<td>vida [life]</td>
<td>amor [love]</td>
<td>jóvenes [youths]</td>
</tr>
</tbody>
</table>

For a second time, “vida” [“life”] is one of the most common words in all three books, so therefore it is the highest ranked word. Figure 2 shows the word list for the pedagogical trilogy.

**Figure 2. Total frequencies in Fr. Morales’s trilogy**

Those who have read these books would find that these top ranked words are significant and in close connection with the main content of the three books. Some clarifications will help to make this list more meaningful
for those who have not read the books. These clarifications, at the same
time, will illustrate how the lexical items of a text can provide a transpar-
ent view of the content.

Firstly, the words “hombre”[“man”] and “hombres”[“men”] could be
grouped together, as they share the same lemma. It is a common prac-
tice in corpus analysis to group and count together words that share the
same base form (Bowker, 2018). Thus, if adding the frequency of “man” to
the frequency of “men,” this lemma would rise to the second position in
the list. It has to be said that these words are used here in their inclusive,
generic meaning and should not be at all interpreted as exclusively mas-
culine. The Spanish language allows this grammatical use of the inclu-
sive masculine, which is highly recommended by the Spanish Royal
Academy of Language. Fr. Morales follows this recommendation in all
his writings.

Secondly, we find the similar case of the verbal forms “hacer”[“to do”]
and “hace”[“does”], keeping in mind that the latter could also mean the
homograph expression of time (hace = ago). They are both inflections of
the same verbal meaning. Fr. Morales requests actions and, in his appli-
cation of Ignatian pedagogy, he encourages actions. Therefore, a high
frequency of this verb is expected.

Finally, the proper names “Pablo”[“Paul”] and “Juan”[“John”] are
among the twenty most frequently used words in the corpus. The expla-
nation for this can easily be proven using the Concordance tool, another
utility within the AntConc software package. A concordance of the words
“Pablo” and “Juan” reveals that they are combined in an elevated number
of instances. These are the numerous references to Pope Saint John Paul
II. These references appear only in the latest edition of the trilogy, which
was updated by Fr. Morales with references to and quotations of the then
Pontifex. References to Pope Saint Paul VI and the Apostle Saint Paul are
also abundant. Regarding “Juan,” apart from the Pope mentioned above,
Fr. Morales makes reference to St. John of the Cross and St. John Chrysos-
tom. The high frequency of these two masculine proper names is a visible
indicator of Fr. Morales’s continuous reference to pontifical magisterium
and to the teachings of the saints.
In the case of these three examples, research began with quantitative findings and worked toward qualitative ones, as stated in the methodology (Biber et al., 1998).

Comparing and Contrasting Lexical Frequencies
The frequency lists of the two writers are displayed together in Figure 3. The coincidence the majority of the words most frequently used by the two Jesuits can be observed. With some variations in rank, they make use of the same content words or words within the same or related semantic scope. For example, the Pope does not use the generic “man/men,” but rather “youth” and “family.” Thus, both writers make profuse references to the person. This is very much related to Ignatian spirituality, in which the person is at the center.

Figure 3. Comparison and contrast of total frequencies

“Life,” “God,” and “Church” are ranked among the first five positions in both lists. “Faith,” “love,” and “heart” are also in both lists, though in different positions. Another similarity is the presence of names like “Christ” (in Fr. Morales) and “Jesús” (in Pope Francis). In this lexical choice, Fr. Morales
remains very close to the most frequent Ignatian vocabulary, while the Pope opts for a more popular form in consonance, perhaps, with a Latin American style. However, in both cases “Christ/Jesus” is also among the most common words.

The list shows the high frequency of the names “Paul” and “John” in Fr. Morales’s writings. As already mentioned, this fact signals his continuous references to Church teachings from the Popes or from the saints. In contrast, Pope Francis’s list shows a high frequency of the words “Lord” and “Spirit.” This could be interpreted as the Pope referring here to spiritual authorities other than himself, while Fr. Morales mentions and quotes the Pontifices and saints as authorities.

The verb “ser” is also ranked very high in both lists. In Spanish, the verb “to be” can be rendered in two different verbs: ser and estar. Estar makes more reference to temporary status, while ser brings the idea of permanence and of essence, a more ontological dimension. The presence of this infinitive among the ten most frequent words is surprising because nouns could be expected to be more frequent than verbs in the infinitive form. One possible interpretation could be the connection with the Ignatian features of perseverance, permanence, and going to the essential.

So far, the analyzed words are nouns and verbs. However, the list of top ranked items also includes adverbs and quantifiers: “siempre” [“always”], “todos” [“all”], and “mismo” [“the same”]. These words could have been included in a stop list and consequently removed from the frequency list. The decision to leave them was based on the fact that they are connected to Ignatian pedagogy and spirituality. The ideas of permanence, inclusiveness, and repetition are conveyed by these three words, respectively.

Conclusions

Comparability is one of the most important aspects of scientific research. Literary analysis was not able to produce comparability before the advent of electronic corpora and software tools for corpus exploration.
However, “computer tools don’t do the thinking,” as stated in one of the screens of Wordsmith Tools, another software package for corpus analysis (Scott, 1996). The frequency lists are the result of a rapid and precise programming algorithm. Nonetheless, the results could only be interpreted in light of the content and our knowledge of the writers.

The comparison of these two writers using software falls into the distant reading approach described in the theoretical framework. The issues discovered in this comparison could have never been found if the writings had been analyzed individually or without the Word List tool. Some subtle uses of language were appreciated by placing the texts next to each other. The comparison of lexical frequencies also revealed that the words which are ranked highest in the writings of both authors are everyday words. It was somehow surprising that these words were very similar in both.

These findings add to other already known features of their styles. Their language is simple and accessible to everyone. Their language is that of a speaker who is accustomed to be in contact with people. They use neologisms, vocabulary, and examples to reinforce a concept, sometimes to the point that the listener can hardly forget it. As another screen message reads in Wordsmith Tools, “much can be inferred from what is absent” (Scott, 1996). In Pope Francis and Fr. Morales, complex terms are absent though elegant and rich vocabulary are maintained. This absence of lexical complexity seems to reinforce their educational attempts.

An empirically derived statement about a tendency is obviously more accurate than any other statement derived from intuition. Thus, the empirical data from these frequency lists allow us to draw some conclusions. The conclusions are presented in three points to stay faithful to the three usual points around which Pope Francis articulates his sermons or speeches.

1. Fr. Tomás Morales, as a Jesuit, lived the Sentire cum Ecclesia. Those who have read his writings, those who met him, and those who know the movements he founded experience this Sentire cum Ecclesia. The frequency lists provide empirical evidence that he also used the same words as the Pontifex. It could be argued that
two different messages could be written with the same lexical items and nevertheless convey totally opposite meanings. This is not the case, since Fr. Morales’s writing are, in their content, in total communion with the Church’s teachings. The empirical results of this analysis may serve to support this notion.

2. This study has revealed the most frequent words which two Jesuits employed in a selection of their writings to encourage evangelization. But some of these words can serve in dialogue with those of different faiths or of none. Life, love, the world, young people, families, and men (the person is implicit in these last three), for example, are concepts of universal concern. The scope of the Church’s teachings reaches every person and all dimensions of humankind. These frequent words could provide a bridge towards those who do not consider the Church *Mater et Magistra* but who nonetheless search for the truth and work for the common good.

3. The high frequencies of some words recall the repetition exercise which St. Ignatius proposed in his Spiritual Exercises. A repetition exercise demands that a meditation be returned to with the purpose of extracting the best from it, the repeated ideas and words. Repetition should not be associated with boredom.

It could be affirmed that the distant reading approach applied with the Word List tool has proven the potential to uncover patterns which both authors share and that it could not have been discovered with the close reading approach, thus the research aim proposed in the paper has been accomplished. As a closing paragraph of the paper and after the three points of the conclusion, this quote recalls the importance of repetition in education in general and in Ignatian pedagogy:

The educator of the youth must be on guard against a temptation …: to think he has to be constantly saying new things …. Every day the same concept will have to be reiterated; each time in a different way, it is true, but always emphasizing the same basic theme. (Morales, 1987, p. 187)
References

https://doi.org/10.1007/s10579-019-09456-6

Balossi, G. (2015). A corpus linguistic approach to literary language and characteri-

Scientific Study of Literature, 1, 15–23.

structure and use. Cambridge: Cambridge University Press.

Bowker, L. (2018). Corpus linguistics is not just for linguists: Considering the poten-


Universidad de Murcia.


Studium.


Encuentro.

Pope Francis. (2013). Exhortación apostólica Evangelii Gaudium del Santo Padre 
Francisco sobre el anuncio del evangelio en el mundo actual. Retrieved from 
http://www.vatican.va/content/francesco/es/apost_exhortations/documents/papa-francesco_esortazione-ap_20131124_evangelii-gaudium.html

exhortations/documents/papa-francesco_esortazione-ap_20160319_ 
amoris-laetitia.html


