# *A Register Analysis of Written Messages From the Papacy and the Universal House of Justice*

Dylane Yue Ting Ho, SEGi University, Malaysia

## Abstract

This study investigates selected features of the Catholic and Bahá'í registers through an analysis of written messages from the Papacy and the Universal House of Justice (UHJ) using a corpus-based approach. The research questions of this study pertain to the situational characteristics and language patterns associated with these registers, and the functional relationships that connect the situational context to the linguistic features. Since previous studies on the interface between language and religion have predominantly focussed on liturgical language in Christian texts, this research addresses a gap in the literature, primarily the description of religious registers of other faiths. Based on Biber and Conrad's (2019) framework of register analysis, a three-step process was applied to two corpora, each approximately 50,000 words in size. The results show that nearly all of the top 20 noun keywords in the Papal register belong to the animate and abstract/process categories, whereas the top 20 noun keywords in the UHJ register are almost evenly distributed across the animate, group/institution, and abstract/process categories. These findings illustrate how the Papacy and UHJ adopt distinctive registers to fulfil their social roles. As an ecclesiastical leader, the Pope guides the Catholic community with messages containing traditional Christian teachings, adapted to suit modern contexts. The UHJ, however, is the supreme administrative body of the Bahá'í community, and its messages outline systematic plans of action for the betterment of society. This study provides impetus for renewed interest in theolinguistics, presenting new opportunities for research on linguistic variation within the domain of religion.


Keywords: Register Analysis, Theolinguistics, Religious Language, Papacy, Universal House of Justice, Catholicism, Bahá'í Faith

**Introduction**

Theolinguistics, a term coined by van Noppen (1981) and subsequently adopted in works by Crystal, is "the study of the relationship between language and religious thought and practice" (Crystal, 2008, p. 484). Over the years, theolinguistics has diminished in prominence, having enjoyed a momentary peak in academic interest between 1960 and 1980. Theolinguistic studies in that period predominantly concerned Christian liturgical texts—the language of prayers and hymns, for example. Crystal (2018) claims that theolinguistic enquiry is still relevant today, arguing that critical questions of language diversity and change merit further investigation into language and religion. By diversity is meant the variation that not merely exists within Christianity and its various denominations, but also between other religious communities, such as Judaism, Islam, and the Bahá'í Faith. The other reason, language change, also calls for a renewed interest in theolinguistics. Given the complex and widespread changes in language over the past four decades, driven by globalisation, the development of the Internet, and religious practice, it is doubtful that descriptions of the language of religious texts and usage from the 1980s reflect contemporary usage.

Adopting Biber and Conrad's (2019) analytical framework of register analysis, this study is an exploration of linguistic features in the Catholic and Bahá'í registers. The focus of the analysis in this study is the specialised register of written correspondence from the heads of faith to the members of these religious communities. Hence, this study compares the registers of messages from the Papacy and the Universal House of Justice (UHJ), the institutions acting as the supreme governing bodies of the Catholic and Bahá'í communities, respectively.

This study aims to explicate some features of the Papal and UHJ registers by analysing the situational characteristics, identifying significant language patterns (with a focus on nouns), and interpreting their functional relationships. It seeks to answer the following research questions:

- What are the situational characteristics of written messages from the Papacy and the UHJ?

- What language patterns are observed in the top 20 noun keywords (and related collocations) associated with written messages from the Papacy and the UHJ?

- What are the functional relationships between the situational context and the linguistic features found in written messages from the Papacy and the UHJ?

**Literature Review**

One of the earliest theolinguistic studies is a description of liturgical language, defined as "a set of distinctive verbal forms used in official public worship on behalf of a religion" (Crystal, 1964, p. 149). Liturgical language is markedly formal in style, but the register is seen as a more pronounced feature. Accordingly, Crystal (1964) identified three "kinds of distinctiveness": archaisms, formulaic diction, and specialised vocabulary (p. 151). Archaisms are fossilised forms found in a language, such as grammatical words (e.g. pronouns *thou, thee* and *thy*), lexical words (e.g. *vouchsafe, whence*), and syntactic structures (e.g. vocatives with *O*, as in *O God*). Formulaic units are set expressions, which include idioms, proverbs and metaphors (e.g. *by the sweat of thy brow*). The specialised vocabulary encountered was roughly divided into seven categories:

1. Historically contextual vocabulary, with emotional overtones (e.g. *crucifix*, *disciple*)
2. Historically contextual vocabulary, with no emotional overtones (e.g. *centurion*, *synagogue*)
3. Vocabulary of personal qualities and activities (e.g. *charity*, *love*)
4. Religious concepts which can be given a Catholic definition (e.g. *heaven*, *salvation*)
5. Technical terms (e.g. *sermon*, *cardinal*)
6. Theological terms with precise definitions (e.g. *only-begotten*, *transubstantiation*)
7. General vocabulary in liturgical language that may appear in other registers (e.g. *deliverance*, *partake*)

(adapted from Crystal, 1964, pp. 154–5)

In a later development, a functional approach to the analysis of liturgical language established a connection between sociolinguistics and religious language (see Crystal, 1990).

Among the studies related to religious registers are those on the language of hymns and prayers. In a study on Wesleyan hymns, van Noppen (2001) found that the most prominent word in the first 40 keywords of all three different works by John Wesley (prose, sermons, and hymns) was *God*, followed by *Christ*, *Jesus*, *Lord*, and *Spirit*. After crossing the threshold of what he considers genre-conditioned vocabulary, he discovered content-related words common to the three genres: *all*, *faith*, *heart*, *grace*, *love*, *sin*, and *soul*. A study of collocations in the hymn corpus yielded more interesting findings. For instance, the keyword *blood* was found alongside words like *dip*, *flow*, and *gush*; all suggestive of the physical nature of blood as a liquid, rather than gory imagery. Hymns are considered a genre in van Noppen's terms, but their features consistently exhibit the hallmarks of a religious register. This is because the language patterns observed are characteristic of hymns, and their functions are associated with a specific situational context, namely the worship of God sung as though speaking to Him through Christ.

One part of a theolinguistic study on prayers involved the investigation of features of the language of intercessory prayers, which was of two kinds: collects and votive prayers. Lašťovičková (2013) described the structure of collects as formulaic, focussing on God's attributes and man's weaknesses, followed by the intentions for which the prayers are said. To achieve this, abstract words (e.g. *love*, *sin*, and *freedom*), and fixed expressions (e.g. *He is the salvation of mankind*) are typically used. Votive prayers, on the other hand, were found to be less formal, and due to the influence of spoken language, some were written using colloquialisms, contractions and abbreviated forms. These features were clearly incompatible with liturgical language, but interestingly, Lašťovičková (2013) noted that some contained archaisms, such as the pronouns *thou* and *thee,* as a form of politeness when addressing God. The findings suggest that even in votive prayers, where the expression of intentions is done extemporaneously, some semblance of a religious register is adhered to. It is likely that the church members relied on previous experiences with liturgical prayers to develop the most 'appropriate' ways to converse with God, leading to observable similarities in sentence structure and vocabulary.

**Methodology**

**Corpus Design**

Two corpora (henceforth referred to as the Papacy corpus and the UHJ corpus, respectively) were compiled. Each corpus is a collection of excerpts from written messages, totalling up to

approximately 50,000 words. As the full texts chosen were of varying lengths—some as short as 400 words, while others around 50,000 words—the corpora comprised differing numbers of texts. The length of each excerpt ranged between 1,000 and 1,100 words since there is generally a "high level of stability for… linguistic feature counts" across 1,000-word samples (Biber, 1990, p. 261). Smaller chunks of 250 to 600 words (depending on the length of the full text) were extracted from messages of fewer than 1,000 words.

To achieve representativeness, stratified sampling was applied, and the proportion of the number of documents for each message type was calculated. In the case of the Papacy corpus, messages (with official English translations) released between 1996 and 2020 - spanning the pontificates of Popes John Paul II, Benedict XVI, and Francis - were first identified, and accordingly, four message types were chosen. As for the UHJ corpus, messages from the same time period were identified, and five message types were selected. For both corpora, the selection of message types was based on the number of documents in the category - in the case of the Papacy corpus, whether English versions were available - and the average length of the messages. Therefore, the list of message types presented in Tables 1 and 2.

| Message types | Number of texts | Total number of words |
|---|---|---|
| Apostolic Exhortations | 8 | 8,314 |
| Apostolic Letters | 21 | 21,556 |
| Encyclicals | 6 | 6,255 |
| Messages for Lent | 14 | 14,524 |
| **TOTAL** | **49** | **50,649** |

Table 1: Composition of Papacy corpus

| Message types | Number of texts | Total number of words |
|---|---|---|
| Messages to the Conference of the Continental Boards of Counsellors | 5 | 4,696 |
| Messages to all National Spiritual Assemblies | 7 | 4,515 |
| Messages to the Bahá'ís of the World | 19 | 14,627 |
| Messages to the Believers in the Cradle of the Faith [Iran] | 17 | 12,025 |
| Ridván Messages | 14 | 14,793 |
| **TOTAL** | **62** | **50,656** |

Table 2: Composition of UHJ corpus

**Register Analysis**

Biber and Conrad's (2019) analytical framework of register analysis is a three-step process, beginning with an analysis of the situational characteristics of each register, followed by a quantitative analysis of selected linguistic features found in each corpus, and, finally, a functional interpretation of the associations between the situational characteristics and the linguistic patterns.

Table 3 shows the seven major situational characteristics in Biber and Conrad's (2019) framework for the situational analysis:

| Situational Characteristic | Aspects to consider |
|---|---|
| Participants | • *Addressor* (producer of the text)<br>• *Addressee* (intended receiver of the text)<br>• *On-lookers* (observers, not direct addressees) |
| Relations among participants | • *Interactiveness* (the extent to which participants interact directly with each other)<br>• *Social roles* (how participants function in the interaction)<br>• *Personal relationships* (how participants relate to one another)<br>• *Shared background knowledge* (familiarity with topic or specialist background knowledge) |
| Channel | • *Mode* (speech or writing)<br>• *Specific medium of communication* (e.g. radio, print) |
| Production circumstances | • *Production* (how the addressor produces the text)<br>• *Comprehension* (how the addressee processes the text) |
| Setting | • *Time and place* (shared by participants or not)<br>• *Place of communication* (private or public setting)<br>• *Time of communication* (contemporary or historical time period) |
| Communicative purposes | • *General purposes* (e.g. inform, persuade, entertain)<br>• *Specific purposes*<br>• *Factuality* (e.g. factual, opinionated, speculative)<br>• *Expression of stance* (personal attitudes and epistemic stance) |
| Topic | • *General topical domains* (e.g. science, religion, politics)<br>• *Specific topics* |

Table 3: Situational characteristics of registers (adapted from Biber & Conrad, 2019)

Information on the situational context of the Papal register was derived from entries in the Catholic Encyclopedia and the Modern Catholic Dictionary. As for the UHJ register, the situational characteristics were described based on the Bahá'í Holy Writings and credible, researched works.

The quantitative aspect of the research lies in the linguistic features analysis. This stage involves frequency and probability calculations to determine the language patterns that are more salient in one register compared with the other.

The linguistic feature of focus in this study is nouns—both common and proper—which denote specific meanings that differ from conventional dictionary definitions. This is because nouns normally make up a significant proportion of specialised vocabulary in a register. In the analysis, 20 nouns with the highest likelihood values for each corpus were grouped according to semantic categories (see Table 4), and content-word collocates for each noun were analysed to compile specialised vocabulary lists for both registers.

| Semantic category | Examples |
|---|---|
| animate | *teacher, child, person* |
| cognitive | *fact, knowledge, understanding* |
| concrete | *rain, sediment, modem* |
| technical/concrete | *cell, wave, electron* |
| quantity | *date, energy, minute* |
| place | *habitat, room, ocean* |
| group/institution | *committee, bank, congress* |
| abstract/process | *application, meeting, balance* |

Table 4: Semantic categories of nouns (adapted from Biber, 2006, p. 244)

To calculate the frequencies of the linguistic features in the corpus and compute the relevant statistical information, AntConc was used in this study. AntConc contains functions, such as a concordancer, a collocation search tool, and word and keyword list generators. The Keyword tool was crucial in generating keyword lists for the Papacy corpus in comparison with the UHJ one and vice versa to analyse the distribution of noun keywords according to the semantic categories in each register.

The Collocate tool of AntConc was utilised for the study of collocations. Taking into consideration the size of the corpora for this research and the amount of data they would yield, content-word collocates from a window span of two words on either side of each keyword were observed.

The functional analysis represents the qualitative aspect of the research, as connections between the situational characteristics and observed language patterns are made to describe the features of the registers examined. This step requires the researcher to provide explanations showing how the linguistic features are linked to the situational characteristics, thereby making this an interpretive process (Biber & Conrad, 2019).

**Results and Discussion**

**Situational Analysis**

*Participants.* The producer of papal messages is the Pope. The Papacy corpus used in this study contains messages from Popes John Paul II, Benedict XVI, and Francis. The Pope is a title conferred on the Bishop of Rome. In the Catholic community, he is seen as the "successor of St. Peter, the chief pastor of the whole Church, the Vicar of Christ upon earth" (Joyce, 1911), based on Christ's constitution of St. Peter as head of His Church in Matthew 16:18—"And I say also unto thee, That thou art Peter, and upon this rock I will build my church; and the gates of hell shall not prevail against it". In addition, the Pope has papal infallibility, which is freedom from error in "his teaching of faith and morals" (Hardon, 2000c). The intended recipients vary according to the message type, but, in general, the messages are addressed to members of the worldwide Catholic community: the bishops, clergy, consecrated persons and the lay faithful.

As for the messages from the UHJ, the addressor is the institution of the Universal House of Justice, which is composed of nine elected members. In the UHJ corpus, there were changes in the membership, as elections took place five times in 25 years. The UHJ is the supreme governing body of the international Bahá'í community, ordained in the Most Holy Book (Kitáb-i-Aqdas) by Bahá'u'lláh, the Prophet Founder of the Bahá'í Faith. In a Tablet,

Bahá'u'lláh (2005) writes that the members of the UHJ "have been charged with the affairs of the people", for they are "the Trustees of God among His servants and the daysprings of authority in His countries" (p. 128). As an institution, the UHJ has conferred infallibility, confirmed by 'Abdu'l-Bahá, the son of Bahá'u'lláh, in His Will and Testament (Vafai, 2005). The addressees of the messages are either institutions, such as the Continental Boards of Counsellors and National Spiritual Assemblies, or Bahá'í communities around the world.

***Relations among participants.*** In both registers, there is no direct interaction between the addressor and the addressee. The participants in both registers are not socially equal, in that the Papacy and the UHJ are institutions of authority that have jurisdiction over their respective religious communities. Interestingly, the participants in both communicative situations regard themselves as being in a close relationship, despite the profound differences in the power dynamic. This impression may be accounted for as the result of the participants' shared faith experience. Moreover, both the addressor and addressees share a great deal of background knowledge, especially with respect to the tenets and teachings of their faith.

***Channel***. Both are written registers, although in the case of the Papacy, the Pope delivered some messages as homilies for occasions like World Youth Day. All these messages are publicly available online for mass dissemination.

***Production circumstances.*** Because the messages are written, the Papacy and the UHJ would have planned and reviewed them before officially releasing them. The Papacy has its own press office, which manages the Pope's communications, whereas in the case of the UHJ, messages are sent to the Quality Control department at the Bahá'í World Centre (in Israel) for proofreading.

Concerning the comprehension circumstances, careful reading is often required, with spaces for reflection and collective study, whether formal or informal, made available to members of the community. This is because the messages are treated as having divine authority, and adherence to the exhortations enshrined in them is comparable to obedience to the laws of God.

***Setting.*** The time and place are not shared by participants during the interaction made in the messages by the Papacy and the UHJ, but the expectation is that the messages are read immediately once they are disseminated. The place of production for papal messages is the Vatican, while for UHJ messages, it is the Bahá'í World Centre in Haifa, Israel. As these two locations are the corresponding centres of administration for the Catholic and Bahá'í communities, they lend a greater sense of authority to the messages. The messages analysed in this study are set in a contemporary time period, from 1996 to 2020.

***Communicative purposes.*** The general purposes of the messages from the Papacy and the UHJ are to guide and educate the members of the Catholic and Bahá'í communities, respectively. The specific purposes vary depending on the message type. Encyclicals, for example, are written to "express the mind of the Pope to the people" on "doctrinal, moral, or disciplinary matters of universal significance" (Hardon, 2000b). Messages from the UHJ to the Conference of the Continental Boards of Counsellors normally present the goals of forthcoming Plans for the worldwide Bahá'í community; those addressed to the believers in Iran serve to encourage the community to remain resilient in the face of widespread persecution.

Regarding the aspect of factuality, both registers often contain references to passages from Holy Scripture in relation to current events. Although the messages can be considered factual, there is no conclusive way to assume that interpretations of Holy Scripture are speculative or otherwise, since religious experiences are subjective and different for each individual. An additional feature of messages from the UHJ is that they frequently incorporate information from reports sent to the Bahá'í World Centre by Bahá'í communities around the world.

On the aspect of stance, because the Papacy and the UHJ are institutions of authority in religious matters, they express an overt stance that is in alignment with the principles of the religion. They also express an epistemic stance in their messages, as their certitude stems from faith in God and interpretations of Holy Scripture.

*Topic.* The general topic area of both registers is clearly religion, but the specific topics vary. Some specific topics found in the Papacy corpus include interreligious dialogue, environmental issues, and commemorations of anniversaries. In the UHJ corpus, messages revolved around community building, involvement in the life of the society, and news of Bahá'í-related developments.

To conclude, both registers are similar in many aspects largely due to the similarities shared by the Papacy and the UHJ. However, it should not be assumed that the UHJ is the Bahá'í counterpart of the Papacy. The social characteristics discussed here do not present the entire scope of jurisdiction for each institution. It would suffice at this point to make one final distinction between the two institutions. The Pope is an ecclesiastical leader charged with the duty to guide the Catholic community on religious matters, often through dogmatic statements on the Christian doctrine. As the Bahá'í Faith does not have clergy, the UHJ serves as an administrative body that primarily serves "to exert a positive influence on the welfare of humankind, to promote education, peace and global prosperity, and to safeguard human honour and the position of religion" (Bahá'í International Community, 2022). While it also provides guidance on matters of faith, the UHJ strongly encourages the believers to deepen their own understanding of the teachings, in accordance with the Bahá'í principle of independent investigation of truth.

**Linguistic Features Analysis**

**The Papacy Corpus**

The 20 nouns with the highest likelihood values in the Papacy corpus were obtained from the first 35 words on the list (see Appendix A). The nouns were then sorted according to semantic category, as shown in Table 5.

The data show that a large proportion of nouns unique to the Papacy corpus belong to the animate and abstract/process categories. In the animate category, it is not surprising that there are references to *Jesus/Christ*. The Triune God in Christian doctrine—the Father, the Son, and the Holy Spirit—also appear as keywords in this category, although the classification of words associated with God is contentious, since semantically, words like *God* and *Spirit* do not refer to tangible or sentient beings akin to anything within the purview of human experience.

In the cognitive category, *Gospel* refers specifically to the teachings of Christ, as shown in the example below:

(1)     All of this has great relevance for the preaching of the *Gospel*, if we are really concerned to make its beauty more clearly recognized and accepted by all.

(PAP-APE-FRA-2013)

| Semantic category | Keyword |
|---|---|
| animate | Jesus / Christ |
|  | God[*] |
|  | Christian(s) |
|  | Lord |
|  | Spirit[*] |
|  | man |
|  | Father |
|  | Saint |
|  | Son |
| cognitive | Gospel |
| concrete | - |
| technical/concrete | - |
| quantity | - |
| place | - |
| group/institution | Church(es) |
| abstract/process | love |
|  | mercy |
|  | faith |
|  | Eucharist |
|  | mystery |
|  | death |
|  | Lent |
|  | gift |
|  | charity |

Table 5: Top 20 noun keywords in Papacy corpus, according to semantic category

* *God* and *Spirit* are considered animate nouns as opposed to abstract nouns, based on interpretations in Christian doctrine.

Only one group/institution noun appears among the top 35 keywords in the Papacy corpus: *Church*. This is to be expected, since the Church is an institution exclusive to Christianity, regardless of denomination. With the exception of *Eucharist* and *Lent*, the keywords in the abstract/process category are common nouns. The keywords *love*, *mercy*, *faith*, and *charity* are semantically related; in Crystal's (1964) categorisation of specialised vocabulary in Catholic liturgical language, they represent personal qualities and activities, which are "interpreted in the light of Christ's own usage and example" (p. 154). Moreover, they are consistent with van Noppen's (2001) observation of content-conditioned words found across all genres in Christianity. Another group of words is also related to Crystal's (1964) classification: *mystery*, *death*, and *gift* refer to religious concepts with Catholic definitions. These words relate to the theme of salvation through Christ's sacrifice, a mystery consisting of the passion, death, resurrection, and ascension of Jesus. Words from both groups are used in various combinations in the papal messages, as shown below:

(2)     *Mercy* and *love* for one's neighbour must therefore be the fruit of a living relationship with God and have God as their constant point of reference, since it is in closeness to Christ that we find our joy.

(PAP-LEN-JP2-2003)

(3)     By placing his *gift* in this context, Jesus shows the salvific meaning of his *death* and resurrection, a *mystery* which renews history and the whole cosmos.

(PAP-APE-BEN-2007)

(4)     In a similar way, *faith* precedes *charity*, but faith is genuine only if crowned by charity.

(PAP-LEN-BEN-2012)

A study of collocations thus aided in contextualising some of the specialised vocabulary found in the Papal register. The list of content-word collocates for the aforementioned 20 nouns is presented in Table 6.

*Face*, a collocate of *Christ*, refers to His physical appearance, but denotes a metaphorical meaning:

(5)     Lent is a favourable season for opening the doors to all those in need and recognizing in them the *face of Christ*. Each of us meets people like this every day.

(PAP-LEN-FRA-2016)

In (5), the *face of Christ* can be interpreted as the essence of Christ, considering the fact that 'goodliness' is an abstract concept.

Another collocate, *encounter*, has a specific meaning in Catholicism, as demonstrated here:

(6)     The *encounter with Christ*, constantly intensified and deepened in the *Eucharist*, issues in the Church and in every Christian an urgent summons to testimony and evangelization.

(PAP-APL-JP2-2004)

This phrase refers to "the relation with Christ" implied through the reception of a sacrament, mainly the *Eucharist*, which is another keyword (Hardon, 2000a).

These collocations are consistent with Crystal's (1964) findings, as phrases such as *works of mercy* and *paschal mystery* are exclusive to Catholic contexts. They make up the specialised vocabulary of the Papal register, and, thus, a recognisable Catholic style.

| Keyword | Left collocates | Right collocates |
|---|---|---|
| Jesus / Christ | *Jesus* Christ[*]<br>*face* of Christ<br>*love* of Christ<br>*encounter* with Christ<br>*Lord* Jesus (Christ) | Jesus *Christ*[*] |
| God | *Son* of God<br>*Word* of God<br>*love* of God<br>*children* of God<br>*People* of God | God's *love*<br>God's *mercy*<br>God's *word* |
| Christian(s) | - | - |
| Lord | the *Risen* Lord | Lord's *Day* |
| Spirit | *Holy* Spirit | - |
| man | - | man and *woman* |
| Father | *Heavenly* Father | Father's *will* |
| Saint | - | Saint *John*<br>Saint *Augustine*<br>Saint *Paul* |
| Son | - | Son of *God* |
| Gospel | - | - |
| gift | - | - |
| Church(es) | the *Catholic* Church(es)<br>the *Eastern* Church(es) | - |
| love | *God's* love | love of *God*<br>love of *Christ* |
| mercy | *works* of mercy<br>*God's* mercy | - |
| faith | *Christian* faith | - |
| Eucharist | - | - |
| mystery | *paschal* mystery | - |
| death | - | death and *resurrection* |
| Lent | *season* of Lent | - |
| charity | - | - |

Table 6: Collocates for keywords in Papacy corpus

[*] *Jesus* and *Christ* are mutual collocates, but also occur independently in texts in the Papacy corpus.

**The UHJ Corpus**

A similar process was applied to the UHJ corpus: from the first 35 words in the Keyword List (see Appendix B), the 20 nouns with the highest likelihood values were extracted. Table 7 presents the 20 noun keywords, sorted according to semantic category.

| Semantic category | Keyword |
|---|---|
| animate | Bahá'í(s) |
| | Bahá'u'lláh |
| | friends |
| | individual |
| | 'Abdu'l-Bahá |
| | youth |
| cognitive | - |
| concrete | - |
| technical/concrete | - |
| quantity | - |
| place | cluster(s) |
| | Iran |
| group/institution | community/ies |
| | Assembly/ies |
| | institution(s) |
| | institute(s) |
| abstract/process | Plan(s) |
| | Faith |
| | Cause |
| | activity/ies |
| | progress |
| | effort(s) |
| | capacity |
| | expansion |

Table 7: Top 20 noun keywords in UHJ corpus, according to semantic category

As with the Papacy corpus, the noun keywords in the UHJ corpus mainly belong to the animate, abstract/process and group/institution categories. Of note is the absence of *God* in the list. We might expect frequent occurrences of the token *God* as a defining feature of a religious register. However, in the 50,656-word UHJ corpus, there are only 87 instances of *God*, in comparison with 357 instances in the 50,649-word Papacy corpus.

In the animate category, there are notable similarities shared with the set of keywords in the Papacy corpus. *Bahá'í(s)* corresponds to *Christian(s)*, and distinguished personages in the Bahá'í Faith—*Bahá'u'lláh* and *'Abdu'l-Bahá*—parallel occurrences of central figures in Christian discourse, such as *Jesus Christ* and *Holy Spirit*. The use of the common noun *friends*, is noteworthy, as it refers to those who identify themselves as members of the Bahá'í community, as exemplified below:

(7)    Indeed, the *friends* should be on their guard, lest the development of capacity in the community not keep pace with the rise in receptivity of a disillusioned humanity.

(UHJ-RID-UHJ-2008)

(8)    Dear *friends*! You represent an army of able and highly motivated servants of the Cause throughout the world.

(UHJ-CCC-UHJ-1998)

In (7), *friends* appears as though it is separate from the addressee, but this reference is made clearer in (8), where the salutation, *Dear friends*, is followed by the pronoun *you*.

In the group/institution category, of note is *Assembly/ies*, which signifies established institutions in the Bahá'í Administrative Order:

(9) The evolution of local and national Bahá'í *Assemblies* at this time calls for a new state of mind on the part of their members as well as on the part of those who elect them, for the Bahá'í community is engaged in an immense historical process that is entering a critical stage.

(UHJ-RID-UHJ-1996)

Finally, in the abstract/process category, words such as *Faith* and *Cause* have specific meanings, distinct from those in the Papal register. *Faith*, for instance, differs from *faith* in the Papacy corpus, as the former refers to the Bahá'í Faith. A similar concept is associated with *Cause*, as is seen below:

(10) Suffice it to say that, four years into the current Plan, the tireless supporters of the *Cause* have brought the *Faith* of Bahá'u'lláh to the strongest position in which it has ever been in its history.

(UHJ-WRL-UHJ-2020-1)

The collocations for the 20 noun keywords provided more scope to the specialised vocabulary of the UHJ register. Table 8 presents the content-word collocates found in the UHJ corpus.

*Junior youth* is a technical term coined by the UHJ in 2000 to refer to young adolescents aged between 12 and 15. Along with another notable collocation, *training institute*, this demographic group receives significant attention in the Bahá'í community's educational process:

(11) In these places, the *training institute* is learning to ensure that sufficient human resources are being raised up to provide for the spiritual and moral edification of children and *junior youth* in ever-increasing numbers.

(UHJ-RID-UHJ-2018)

Interestingly, the collocation, *Plan's three protagonists* relates to three keywords—*individual*, *community*, and *institution*—suggesting that they play a pivotal role in the Plans:

(12) In more and more clusters, the programme of growth is increasing in scope and complexity, commensurate with the rising capacity of the *Plan's three protagonists*—the *individual*, the *community*, and the *institutions* of the Faith—to create a mutually supportive environment.

(UHJ-RID-UHJ-2014)

It can be observed that most of the collocations are common nouns with specialised meanings that pertain to processes (e.g. *community building*, *institute process*, and *expansion and consolidation*) rather than overtly theological terms.

| Keyword | Left collocates | Right collocates |
| --- | --- | --- |
| Bahá'í(s) | - | Bahá'ís of *Iran* |
| Bahá'u'lláh | *followers* of Bahá'u'lláh<br>*teachings* of Bahá'u'lláh<br>*Birth* of Bahá'u'lláh<br>*Cause* of Bahá'u'lláh<br>*Faith* of Bahá'u'lláh | Bahá'u'lláh's *Revelation*<br>Bahá'u'lláh's *teachings* |
| friends | *beloved* friends | - |
| individual | - | - |
| 'Abdu'l-Bahá | - | - |
| youth | *junior* youth<br>*children* and youth<br>*Bahá'í* youth | - |
| cluster(s) | - | - |
| Iran | *Bahá'ís* of Iran<br>*people* of Iran<br>*women* of Iran | - |
| community/ies | *Bahá'í* community<br>*worldwide* community<br>*national* communities<br>*local* communities | community *building* |
| Assembly/ies | *National (Spiritual)* Assembly/ies<br>*Local (Spiritual)* Assembly/ies<br>*Spiritual* Assembly/ies | - |
| institution(s) | *Bahá'í* institution(s) | - |
| institute(s) | *training* institute(s) | institute *process*<br>institute *courses*<br>institute *program(s)* |
| Plan(s) | *One/Four/Five/Ten Year* Plan<br>*Divine* Plan<br>*Minor* Plan<br>*Major* Plan<br>*series* of Plans<br>*global* Plan(s) | Plan's *three protagonists* |
| Faith | - | Faith of *Bahá'u'lláh*<br>Faith of *God* |
| Cause | - | Cause of *God*<br>Cause of *Bahá'u'lláh* |
| activity/ies | *core* activity/ies<br>*Bahá'í* activity/ies<br>*community-building* activity/ies | - |
| progress | *social* progress | - |
| effort(s) | *individual* effort(s)<br>*collective* effort(s) | - |
| capacity | *institutional* capacity<br>*administrative* capacity | capacity for *service* |
| expansion | *large-scale* expansion | expansion and *consolidation* |

Table 8: Collocates for keywords in UHJ corpus

**Functional Analysis**

In the Papal register, the noun keywords and their collocations belong mainly to the animate and abstract/process categories. This distribution can be attributed to the social characteristic of the Papacy. As head of the Catholic Church, the Pope fulfils his duty to advise on matters regarding the Christian faith in pronouncements that make numerous references to Christ, the Trinity, Mary, and other significant figures. Across the pontificates of the Popes studied, there is a consistent pattern in the usage of traditional Catholic terms that encourage the congregation to deepen their personal relationship with Christ. This can be achieved, for example, by practising *charity* and the *works of mercy*, and recognising the *face of Christ* in others.

On the other hand, the UHJ register contains keywords and collocations in the animate, group/institution, and abstract/process categories that reveal an organised process operating on a global scale. The UHJ, as the supreme governing body, directs the course of action for the worldwide Bahá'í community through a series of Plans. Thus, in the Plans spanning the 25-year period, the messages progressively describe how the establishment of the *training institute* enables the *Plan's three protagonists* to promote *community building*. The use of *friends* helps, in this context, to establish solidarity with the Bahá'ís; it denotes a horizontal power dynamic. In addition, this noun suggests that the UHJ and the Bahá'í community are engaged in a common purpose and are working shoulder to shoulder, rather than implying that the addressees are carrying out the directives of the addressor.

**Conclusion**

The distinction between the religious registers of messages from the Papacy and the UHJ is sufficiently significant to demonstrate that linguistic variation exists within the domain of religion. The linguistic features in this study—noun keywords and collocations with specialised meanings—not only prove that they are unique to a religious community, but also shed light on how these institutions, as religious leaders, operate in society.

As the findings indicate, the Papacy, though covering various topics in different message types, ultimately focusses on the sanctity of life and mankind's relationship with God. Emphasis on the Person of Christ and His teachings motivates the Catholic community to act in accordance with His example, that is, to do charitable deeds out of love for God and humanity. In these texts, the Papacy encourages certain lines of action, but does not make plans for the Catholic community at large.

The UHJ, on the other hand, has systematic plans of action in place, enabling it to chart the course for the Bahá'í community. Although the messages appear to be administrative in nature, constant references to the central figures of the Bahá'í Faith, such as Bahá'u'lláh and 'Abdu'l-Bahá, serve as a reminder that the Plans are part of a spiritual enterprise. The worldwide endeavour to build vibrant communities can be seen as the UHJ's fulfilment of its duty to promote the welfare of the generality of humanity. The efforts initiated to this end are not exclusive to the Bahá'ís alone, whether in their involvement or their outcomes.

**Limitations of the Study**

The scope and scale of this study were necessarily limited. Since only one type of linguistic feature was investigated, any attempt to make direct generalisations for both the Papal and

UHJ registers would be premature at this stage. Ideally, excerpts from all message types and spanning a longer period of time would have contributed to better representativeness in the corpora, thereby improving the validity of the findings. One barrier to achieving this was the limited number of papal messages with official English translations available, as some of these messages were only in Latin.

As for the data, while larger corpora may not have had a significant impact on the findings, part-of-speech (POS) tagging would have been useful in the corpus analysis. Furthermore, with POS-tagged corpora, analyses of different aspects of the various word classes could have been conducted to define the features of both registers more conclusively.

**Recommendations for Future Research**

A comparative analysis of language patterns across diverse religious registers can enable researchers to make generalisations about the unique characteristics of each. In Biber and Conrad's (2019) framework, this type of analysis is known as a multidimensional (MD) analysis, "a quantitative approach [comparing] many different registers, with respect to several different linguistic parameters", referred to as *dimensions* (p. 268). Thus, specialised corpora for registers in other religious communities would have to be developed. This opens exciting new prospects for theolinguists, with the potential to further uncover the relationship between language and religion. Although Christianity remains "the dominant religion of the English-speaking world" (Crystal & Davy, 2013, p. 148), other faiths are emerging in prominence, thanks to globalisation. Hence, theolinguistic analyses of registers found in other faiths would serve to foster better interreligious understanding, and even clarify some misconceptions about religion in this age.

**Appendix A**

**Keyword List for the Papacy corpus**

| Rank | Frequency | Likelihood | Keyword |
|:---:|:---:|:---:|:---:|
| 1 | 301 | 292.251 | us |
| 2 | 197 | 273.730 | Christ |
| 3 | 186 | 258.425 | Church |
| 4 | 143 | 198.621 | Jesus |
| 5 | 353 | 173.436 | God |
| 6 | 199 | 145.603 | love[*] |
| 7 | 358 | 139.307 | his |
| 8 | 122 | 134.193 | I |
| 9 | 93 | 129.127 | Christian[†] |
| 10 | 286 | 125.356 | our |
| 11 | 236 | 121.824 | he |
| 12 | 118 | 107.510 | Lord |
| 13 | 72 | 99.955 | mercy |
| 14 | 123 | 96.676 | faith |
| 15 | 65 | 90.232 | Christians[†] |
| 16 | 61 | 84.677 | Eucharist |
| 17 | 90 | 84.357 | himself |
| 18 | 66 | 82.611 | Spirit |
| 19 | 77 | 80.570 | man |
| 20 | 113 | 74.415 | him |
| 21 | 53 | 73.568 | Gospel |
| 22 | 52 | 72.179 | mystery |
| 23 | 50 | 69.402 | death |
| 24 | 251 | 68.680 | who |
| 25 | 63 | 62.687 | Father |
| 26 | 44 | 61.071 | Lent |
| - | 44 | 61.071 | Saint |
| 28 | 49 | 59.594 | gift |
| 29 | 48 | 58.246 | my |
| 30 | 407 | 57.212 | which |
| 31 | 38 | 52.741 | charity |
| 32 | 58 | 52.502 | poor |
| 33 | 42 | 51.522 | Son |
| 34 | 36 | 49.965 | Paul |
| 35 | 68 | 48.495 | truth |

Table A: Top 35 keywords in Papacy corpus

[*] Instances of *love* in the Papacy corpus comprised those belonging to the syntactic categories of noun (179 instances) and verb (20 instances).

[†] *Christian* and *Christians* are considered two forms of the same lexeme, *Christian*. As the Papacy corpus is not part-of-speech (POS)-tagged, instances of *Christian* found in the corpus comprised those belonging to the syntactic categories of noun (9 instances) and adjective (84 instances).

## Appendix B

### Keyword List for the UHJ corpus

| Rank | Frequency | Likelihood | Keyword |
|:---:|:---:|:---:|:---:|
| 1 | 204 | 282.957 | Bahá'í[*] |
| 2 | 125 | 173.283 | Plan |
| 3 | 224 | 170.687 | community |
| 4 | 122 | 169.121 | Bahá'u'lláh |
| 5 | 129 | 133.361 | Faith |
| 6 | 114 | 128.625 | friends |
| 7 | 65 | 90.069 | Cause |
| 8 | 61 | 84.524 | Bahá'ís[*] |
| 9 | 191 | 73.366 | your |
| 10 | 95 | 73.186 | individual[†] |
| 11 | 49 | 67.890 | Assemblies |
| 12 | 73 | 67.013 | institutions |
| 13 | 48 | 66.504 | 'Abdu'l-Bahá |
| 14 | 47 | 65.118 | clusters[‡] |
| 15 | 46 | 63.733 | Spiritual |
| 16 | 45 | 62.347 | National |
| 17 | 72 | 62.181 | progress[§] |
| 18 | 51 | 62.161 | activities |
| 19 | 44 | 60.961 | institute |
| 20 | 43 | 59.575 | cluster[‡] |
| 21 | 52 | 57.713 | youth |
| 22 | 72 | 55.739 | efforts |
| 23 | 364 | 55.665 | their |
| 24 | 38 | 52.646 | Iran |
| 25 | 264 | 49.918 | will |
| 26 | 63 | 48.393 | capacity |
| 27 | 38 | 44.731 | expansion |
| 28 | 67 | 44.627 | growth |
| 29 | 32 | 44.331 | Centre |
| 30 | 3329 | 43.847 | of |
| 31 | 77 | 43.491 | years |
| 32 | 31 | 42.946 | Councils |
| - | 31 | 42.946 | Order |
| 34 | 50 | 42.700 | members |
| 35 | 401 | 42.099 | are |

Table B: Top 35 keywords in UHJ corpus

[*] *Bahá'í* and *Bahá'ís* are considered two forms of the same lexeme, *Bahá'í*. As the UHJ corpus is not part-of-speech (POS)-tagged, instances of *Bahá'í* found in the corpus comprised those belonging to the syntactic categories of noun (6 instances) and adjective (198 instances).

[†] Instances of *individual* in the UHJ corpus comprised those belonging to the syntactic categories of noun (41 instances) and adjective (54 instances).

[‡] *Clusters* and *cluster* are considered two forms of the same lexeme, *cluster*.

[§] Instances of *progress* in the UHJ corpus comprised those belonging to the syntactic category of noun (71 instances) and verb (1 instance).

***Note:*** The likelihood values represent the probability of words appearing in a text of a particular register due to their unusually high frequencies found in the corpus in comparison with another. Hence, keywords with high likelihood values are more likely to belong to a register under examination. If these keywords were to be encountered in a random text not in the corpus, it can be assumed that the text employs the register to which the keywords are associated.

# References

Bahá'í International Community. (2022). *The Universal House of Justice: An official site of the Bahá'í Faith*, http://universalhouseofjustice.bahai.org/.

Bahá'u'lláh. (2005). *Tablets of Bahá'u'lláh revealed after the Kitáb-i-Aqdas*. Bahá'í Publishing Trust.

Biber, D. (1990). Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and Linguistic Computing*, 5(4), 257–269.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. John Benjamins Publishing Company.

Biber, D. & Conrad, S. (2019). *Register, genre and style* (2nd ed.). Cambridge University Press.

Crystal, D. (1964). A liturgical language in a linguistic perspective. *New Blackfriars 46*(534), 148–156.

Crystal, D. (1990). Liturgical language in a sociolinguistic perspective. In D. Jasper & R. C. D. Jasper (Eds.), *Language and the Worship of the Church*, 120–146. Macmillan.

Crystal, D. (2008). *A dictionary of linguistics and phonetics* (6th ed.). Blackwell.

Crystal, D. (2018). Whatever happened to theolinguistics? In P. Chilton & M. Kopytowska (Eds.), *Religion, language, and the human mind*. Oxford University Press.

Crystal, D. & Davy, D. (2013). *Investigating English style*. Routledge.

Hardon, J. A. (2000a). Encounter with Christ. In *Modern Catholic Dictionary*. Eternal Life.

Hardon, J. A. (2000b). Encyclical. In *Modern Catholic Dictionary*. Eternal Life.

Hardon, J. A. (2000c). Infallibility. In *Modern Catholic Dictionary*. Eternal Life.

Joyce, G. (1911). The Pope. In *The Catholic Encyclopedia*. Robert Appleton Company. Retrieved January 21, 2022 from New Advent: http://www.newadvent.org/cathen/12260a.htm

Lašťovičková, M. (2013). *The language of prayer: On the material of intercessory prayers*. Ph.D. dissertation, Palacký University Olomouc.

Noppen, J. P. van. (1981). Preface. In J. P. van Noppen (ed.), *Theolinguistics*, 1–4. Brussels: Studiereeks Tijdschrift Vrije Universiteit Brussel, Niuwe Serie 8.

Noppen, J. P. van. (2001). The English hymn: A discourse perspective. *Belgian Essays on Language and Literature (BELL): Papers of the Belgian Association of Anglicists in Higher Education (BAAHE)*, 11–25.

Vafai, S. (2005). *The essence of the Covenant: Features, history, and implications*. Palabra
 Publications.

**Contact email:** dylane@segi.edu.my