

Vertiefungsseminar: Utopia Arabica

Modul: Arabische Literatur Vertiefung

Dozent: Ruben Schenzle, M.A.

Wintersemester 2020/2021

Hausarbeit

# **Data-Driven Study of Hope in Early Arabic Poetry**

27.03.2023

Mahmoud Kozae

Masterstudiengang Arabistik

17. Hochschulsemester

# Table of Contents:

Table of Contents:	1
1. Early Arabic Poetry and Utopianism	2
1.1 Social Context	2
1.2 Themes and Meanings	3
1.3 Studying Utopianism	4
1.4 Related Research	5
2. Designing a Data Scheme for Studying Hope in Arabic Poetry	6
2.1 Collecting and Structuring Text Data	6
2.2 Enhancing the Corpus with Annotations	8
2.2.1 Lemmatization and PoS Tagging	9
2.2.2 Building a Reverse Index of the Lemmas	9
2.2.3 GPT-3	10
2.2.3.1 Analysis of the Original Poems	11
2.2.3.2 Analysis of the Translated Poems	11
3. Lemma-Driven Verse Exploration and Interpretation	13
3.1 Overview of the Lemmas	13
3.1 Metaphysical Visions	14
3.2 Need for Peace	15
3.3 Hopes Related to Travel and Separation	17
3.4 Morals and Mental Attitudes	18
3.5 Psychology of Hope	20
4. Conclusion	22
5. Bibliography	23

# 1. Early Arabic Poetry and Utopianism

## 1.1 Social Context

In Pre-Islamic Arabia, numerous societies were dispersed over a vast geographical area in and outside the Arabian Peninsula, which alone is almost four million square kilometers; this landmass is equivalent to over two-thirds of the Indian subcontinent and nearly a third of Europe. The Arabian tribes inhabited many regions of Mesopotamia and Sinai, extending to the Nile's eastern bank. The vast dispersion of tribes resulted in a variety of lifestyles. Nomadic and urban societies were equally prevalent, with some tribes relying on hunting and gathering and others on complex farming systems. Commerce played an essential role in the lives of many tribes, with some engaging in it while others residing along caravan routes. Through trade, the Arabs frequently interacted with neighboring civilizations, particularly Persia and Byzantium, and ventured to the edges of their known world<sup>1</sup>.

The poetry transmitted from that period was not restricted to a single society, social class, profession, or gender. It was a shared cultural aspect composed by individuals from all walks of life, including the rich and the poor, kings and commoners, warriors and farmers<sup>2</sup>. In terms of language and style, the poetry of that time was not as diverse as the societies. It represented a uniquely shared cultural aspect that united people across a vast land. This poetic unity resulted from the presumed shared respect and acceptance of Quraish and Mecca as cultural centers<sup>3</sup>. Even though Arabs expressed themselves in various dialects distinct from Quraish's vernacular in daily life, they adopted the language of Quraish for poetic expression. Even non-poets expressed themselves in this language by reciting and chanting poetry on various occasions and identifying strongly with their tribespeople's poetry<sup>4</sup>.

Poetry in pre-Islamic Arabia reflected the cultural, social, and religious norms of the society. In addition to serving a creative function, it was applicable in various situations. Poets were held in high regard because it was believed that they were wise individuals who could offer sound advice to their communities. The historical and cultural traditions of their respective tribes were preserved through their poetry. Poetry served as a medium for social interaction, negotiation, and propaganda in this highly stratified and diverse society. Such a social role is manifested in the fact that poetry was frequently employed to commemorate significant events, such as battles, weddings, funerals, and other rites of passage. The religious beliefs of pre-Islamic Arabs were

---

<sup>1</sup> (Al-Asad 1988) pp. 1-5.

<sup>2</sup> (Daif 1984) pp. 16-17.

<sup>3</sup> (Daif 2003) p. 131.

<sup>4</sup> Ibid pp. 189-193.

reflected in their poetry, as hymns and praises for the tribes' various deities were composed in the same poetic style<sup>5</sup>.

From a contemporary perspective, life in pre-Islamic Arabia would be considered hard. This observation is not exclusive to Arabia but holds for any place on earth during that period. Life was devoid of the conveniences that later centuries and modern times offer. However, it is possible to assess the difficulties of pre-Islamic Arabia by comparing it with neighboring societies of the time. The society's geographical location was challenging, and living in the desert profoundly influenced the region's culture and lifestyle. The climate was highly livable, with moderate temperatures and a rare occurrence of natural disasters such as hurricanes, floods, and earthquakes. However, the landscape posed significant obstacles. Navigating through the rugged terrain was challenging, and the availability of food and water was limited<sup>6</sup>. As a result, the society was characterized by a spirit of thriftiness, and communal ties were strengthened by the need to support one another. This sense of solidarity even extended to strangers, as hospitality to travelers was highly valued. Nevertheless, the scarcity of resources often led to fierce competition and long-standing enmities among tribes. Consequently, war was a constant part of life, and raids on neighboring territories and ambushes against caravans were common occurrences<sup>7</sup>. These hardships were further compounded by the absence of a centralized authority or the rule of law, making resolving disputes challenging.

## 1.2 Themes and Meanings

The first classification of themes in Arabic poetry can be traced back to Abū -Tammām (d. 845 CE), who categorized all poetry until his time, not just pre-Islamic poetry. He identified ten themes, including *al-ḥamāsah*, *al-marāthī*, *al-adab*, *al-nasīb*, *al-hijā'*, *al-madīh*, *al-ṣifāt*, *al-siyar*, *al-nu'ās*, *al-milḥ*, and *maḍmatu al-nisā'*. These themes frequently intersected with one another, and a poem could be classified under multiple themes. In many cases, a poem could include sections that address different themes. Subsequently, different scholars compiled various enumerations of the themes, with Qudāma ibn Ja'far (d. 948 CE) listing six themes and ibn Raṣīq (d. 1064) listing nine. In his analysis of these categorizations, Šawqī Ḍaif identified the seven primary themes that are most applicable to pre-Islamic poetry, which include *al-madīh*, *al-hijā'*, *al-waṣf*, *at-tashbīh*, *al-marāthī*, *al-i'tidhār*, and *al-ḥamāsah*<sup>8</sup>.

The emergence of these themes in pre-Islamic poetry is unclear, but it is evident that themes and poem structure were prevalent and repetitive during this era. Poems often

---

<sup>5</sup> The elaborations of this passage is based on the discussion of themes in (Ḍaif 2003) pp. 195-219.

<sup>6</sup> (Al-Asad 1988) pp. 1-5.

<sup>7</sup> (Ḍaif 2003) pp. 62-66.

<sup>8</sup> Ibid. pp. 195-196.

began with *al-bukā` alā al-aṭlāl*, i.e., wailing over ruins, before moving onto *al-nasīb*, which is highly related to *al-bukā` alā al-aṭlāl* as it predominantly deals with lost or unrequited love<sup>9</sup>. The repetitiveness of themes makes it uncertain whether poets were genuinely expressing themselves or merely fulfilling expectations.

Themes were not the only constraints on poets, as there were many linguistic and aesthetic limitations in deciding what constitutes a proper poem. The poets had to adhere to a literary language different from their own, and meter and rhyme rules highly influenced word choice. This often meant that the musicality of utterances was more important than meaning<sup>10</sup>. Poetry was, therefore, a restrictive and unnatural means of expression that did not offer the possibility of free, spontaneous articulation of thoughts and feelings.

Early scholars and anthologists sometimes distinguished between *maṣnū`* crafted and *maṭbū`* natural poems, hinting that some poets spent extended time polishing and improving their poetry while others delivered theirs spontaneously. However, it is unlikely that any poetry was not a result of at least some polishing and revisions. The idea that some poets naturally speak in poetry is probably a glorification of their talents and not an accurate representation of how they composed their poetry<sup>11</sup>.

### 1.3 Studying Utopianism

Utopia is a literary genre that revolves around the depiction of an idealized society, often characterized by perfect living conditions and moral principles. The term “utopia” was first introduced by Thomas More (d. 1535 CE), though the concept can be traced back to earlier works, with Plato's (d. 347 BC) Republic being the prime example. Utopian works often provide a comprehensive, top-down reimagining of human life, within which the author's aspirations for a better world can be realized.

In relation to utopia, the concept of “utopianism” emerges as a philosophical and literary movement that seeks to transform the difficult aspects of life. While utopianism does not constitute a genre, it can be viewed as an analytical lens through which various texts and genres can be examined. The primary objective of such investigations is to discern the hopes and sentiments reflected in the literature of a particular culture, region, time, or author<sup>12</sup>.

It is important to note that the utopian genre, as traditionally understood, did not feature prominently in early Arabic literature. Most of this literature comprises poetry, which may not have been the most conducive medium for articulating utopian thoughts and sentiments. Nevertheless, it is plausible to assume that many hopes and related

---

<sup>9</sup> (El Tayib 1983) p.43.

<sup>10</sup> (Daif 1987) pp. 17-19.

<sup>11</sup> Ibid. pp. 13-14.

<sup>12</sup> The Information in this passages and the previous one are from the introduction of (Sargent 2010).

sentiments about living conditions and societal morals were either explicitly or implicitly expressed by the poets of the time.

The following chapters will delve into the methodology employed to explore the presence of hope and related themes in early Arabic poetry based on the assumption mentioned earlier. By applying the concept of utopianism as an analytical measure, the assessment will focus on how these hopes and sentiments were manifested in the works of various poets. This investigation will contribute to a more comprehensive understanding of the ideals and aspirations that shaped early Arabian society, as well as the extent to which the poets of the era sought to address and improve upon the challenges they faced.

## 1.4 Related Research

Two studies have been found that investigate hope in early Arabic poetry. One study by (Ḥassānīn 2020) is literary, and the other one by (Ḥiḍr 2014) is linguistic.

Ḥassānīn's literary study delves into both utopian and dystopian themes present in early Arabic poetry. The author conducts a systematic investigation of distinct groups of poets, analyzing their works within the context of their personal experiences and the historical events that shaped their lives. Particular emphasis is placed on the poetry of aṣ-Ṣa'ālīk, with the author arguing that their status as social outcasts played a significant role in generating insightful commentary on objections against the societal norms and the moral values of pre-Islamic Arabia.

Ḥiḍr's linguistic study focuses on the occurrences of specific words and syntactic patterns in both the Quran and early Arabic poetry. The poetry analysis is confined to the *diwāns* of the ten poets associated with the Mu'allaqāt. The primary objective of this research is to ascertain the most accurate grammatical rules governing the usage of particles, words, and syntactic structures related to the expression of hope.

Both studies contribute valuable insights to the understanding of hope in early Arabic poetry. Ḥassānīn's work provides a contextualized exploration of utopian and dystopian themes, shedding light on how poets navigated and critiqued the societies they lived in. Meanwhile, Ḥiḍr's study offers a linguistic perspective, enhancing comprehension of the grammatical intricacies associated with expressions of hope. The linguistic study also served as a guide in selecting lemmas for conducting the data-study.

## 2. Designing a Data Scheme for Studying Hope in Arabic Poetry

The term “data-driven research” is interchangeable with “inductive research.” Inductive research refers to drawing conclusions from specific observations and data rather than beginning with a general hypothesis. Data is a buzzword often associated with computer use and digital records. Nonetheless, data-driven, i.e., inductive research, can occur entirely independently of computers. Computers can optimize many aspects of a data-driven study, especially when it comes to searching through a massive corpus of data. It is noteworthy that not every computer-assisted literary research is necessarily inductive; there is a broad range of ways in which computer software can assist research. This study aims to identify some software-based methods that can help locate early Arabic poetry verses that discuss or express hope. Once these verses have been identified, they will be analyzed in more detail.

### 2.1 Collecting and Structuring Text Data

A wealth of digitized Arabic poetry corpora can be found on the internet. However, many of these corpora resulted from non-expert community efforts or have uncertain origins<sup>13</sup>. Consequently, locating digitized Arabic text corpora is an uncomplicated task, but discovering corpora compiled through extensive scholarly effort and with academic rigor can prove challenging. This study utilizes a corpus created by the “Ruins to Resurrection” project. This project took place concurrently with the Corpus Coranicum Project at the Berlin Brandenburg Academy of Sciences, led by Ghassan el-Masri and funded by the German Research Foundation<sup>14</sup>. The “Ruins to Resurrection” project produced digitizations of the Mu‘allaqāt, Mufaḍḍaliyāt, and Aṣma‘iyyāt based on the most reliable printed editions. The poems of the corpus are transcribed with full vowel marks *ḥarakāt* and writing signs *taškīl*. Scholarly translations for the Mu‘allaqāt and Mufaḍḍaliyāt are included in the corpus. This data is not publicly accessible online and was obtained through direct communication with the project lead.

It is necessary to restructure the corpus to facilitate the intended research and to conduct an effective analysis. Additional annotations can also be applied to enhance the corpus, which can help uncover new insights and extend possibilities for more profound discoveries. While manually annotating the corpus based on expert observation is the ideal scenario, it would require significant effort. Therefore, a more

---

<sup>13</sup> Most famous online corpora of non-scholarly origins are: [adab.com](http://adab.com) and [aldiwan.net](http://aldiwan.net) (last retrieved March 2023)

<sup>14</sup> See <https://gepris.dfg.de/gepris/projekt/248678153/ergebnisse> (last retrieved March 2023)

practical approach for the present study would be to explore the potential of using automated annotations.

The corpus was provided as Microsoft Word files. While it is computationally possible to work with this file format, it is more favorable to use database software that provides the possibility of performing sophisticated queries and analytics. For extended projects, it is necessary to compare such technologies, their nuances, and their tradeoffs. For this study, MongoDB<sup>15</sup> is chosen as it perfectly fits the purpose, and the author has extended experience working with it.

MongoDB is a database management system that stores data in collections, with each entry in the collection referred to as a document. In this particular case, a dedicated collection is used for storing Arabic poems, where each document in the collection corresponds to a verse. Meanwhile, a different collection is used for storing translations, with some auxiliary collections to be created, as explained later. It is not feasible to store the poems, and their translations in a single collection since the provided translations have different verse structures than the original poems. These structures make it challenging to programmatically align the translations with specific parts of the original poems due to the lack of consistent separators in the text. The following example represents a verse structured according to the schema designed for the Arabic poem MongoDB collection:

```
{
  "numerical_id": 0,
  "collection": "المعلقات",
  "poet": "Imru al-Qays (EI: Imru' al-Ḳays b. Ḥudj̄r)",
  "poem": "Arb_Mual_I",
  "order": 0,
  "tokens_r": [
    {"order": 0, "raw": "فِئَا"},
    {"order": 1, "raw": "نَبْكَ"},
    {"order": 2, "raw": "مِنْ"},
    {"order": 3, "raw": "ذِكْرَى"},
    {"order": 4, "raw": "حَيْبٍ"},
    {"order": 5, "raw": "وَمَنْزِلٍ"}
  ],
  "tokens_l": [
    {"order": 0, "raw": "بِسْفَطٍ"},
    {"order": 1, "raw": "الْوَى"},
    {"order": 2, "raw": "بَيْنٍ"},
    {"order": 3, "raw": "الدَّحُولِ"}
  ]
}
```

<sup>15</sup> See [mongodb.com/docs/atlas/](https://www.mongodb.com/docs/atlas/) (last retrieved March 2023)



```

    {"order": 4, "raw": "فَحْوْمَل"}
  ]
}

```

Having all verses in a single collection simplifies queries. The information about the collection, poet, poem, and order in the poem is preserved in corresponding fields in each entry. Thus, it is simple to extract a whole poem by using the MongoDB query:

```
{ $match: { $poem: "Arb_Mual_I" } }
```

Saving the words as spread-out arrays is useful when assigning annotations to individual words.

The scholarly translations collection is far simpler, as there is no need to add annotations to individual words. Following is an example entry:

```

{
  "numerical_id": 0,
  "collection": "المعلقات",
  "poet": "Imru al-Qays (EI: Imru' al-Ḳays b. Ḥudj̄r)",
  "poem": "Arb_Mual_I",
  "order": 0,
  "verse": "Halt, friends both! Let us weep, recalling a love
and a lodging by the rim of the twisted sands between Ed-Dakhool
and Haumal"
}

```

## 2.2 Enhancing the Corpus with Annotations

The present state of Arabic Natural Language Processing (NLP) software poses a significant challenge to the automated annotation of the poetry corpus currently under study<sup>16</sup>. Because of the software's limitations, automated annotations will likely contain several errors and inconsistencies. As a result, this study has opted for a qualitative instead of quantitative approach, relying on automated annotations only to assist in exploring the corpus and locating verses relevant to the research topic of hope and related sentiments. Despite the limitations of Arabic NLP software, it is essential to continue experimenting with the latest advancements in the field. This experimentation

---

<sup>16</sup> A survey on the status of Arabic NLP tools in (El Gayar and Suen 2019) pp. 32-36.

will help discover the potential of these tools in enhancing Arabic literary studies and should provide feedback that aids the further developments of these tools.

### 2.2.1 Lemmatization and PoS Tagging

NLP comprises numerous tasks that aim to extract meaning from human language. Two of these tasks, Lemmatization and Parts-of-Speech (PoS) Tagging, play a significant role in text analysis. By adding annotations such as lemmas and PoS tags, researchers can carry out more advanced searches and apply standard text-mining algorithms. Initially developed for English, these techniques have been extended to other languages. However, Arabic text processing presents a unique set of challenges. The absence of adequate lemmatizers and PoS taggers has long been a major obstacle.

The concept of a lemma, which refers to the base form of a word, does not apply well into Arabic lexicographical systems, as these systems organize words by their roots rather than their lemmas. Therefore, each developed lemmatizer independently decided on what a lemma should be, and this led to various definitions of a lemma among existing lemmatizers. Additionally, Arabic presents a similar problem with parts of speech. Unlike English, Arabic has no clear morphological distinctions between nouns, adjectives, and adverbs. As a result, it is always necessary to examine the syntax of a sentence to determine the correct PoS, which makes the task programmatically far more complex. Furthermore, there is no universal or commonly used tagset for Arabic, which presents an additional challenge.

Finally, the absence of crucial information hinders automatic lemmatization and PoS tagging of Arabic text. Vowel signs are often incomplete or missing in Arabic text, making it difficult to determine the correct word form.

The project CAMEL tools<sup>17</sup> provided the most promising results among the surveyed Arabic NLP software. The previously presented schema was extended with two fields for each token, containing the lemma and the PoS tag determined by the CAMEL disambiguator:

```
{ "raw": "نَبَكَ", "annotation": [ "verb", "بَكَى" ] }
{ "raw": "حَتَّى", "annotation": [ "prep", "حَتَّى" ] }
{ "raw": "دُمُوع", "annotation": [ "noun", "دُمُوع" ] }
```

### 2.2.2 Building a Reverse Index of the Lemmas

Queries for lemmas using the verse schema previously presented would be computationally demanding and time-consuming. Therefore, constructing a reverse index that organizes lemmas alphabetically and directly maps each lemma to its

---

<sup>17</sup> (Obeid et al. 2020)

occurrences in the corpus would be tremendously beneficial. Thus, a third MongoDB collection is created with individual documents for each lemma that adhere to the following schema:

```
{
  "lemma": "تَمَنَّى",
  "occurrences": [ [2865, 74, 0, 1],[ 3510, 12, 0, 3] ],
  "count": 2
}
```

An occurrence array contains in order: poem numerical id, line number, *šatr* (0 for first *šatr*, and 1 for the second *šatr*), and token's order in *šatr*.

### 2.2.3 GPT-3

This study is being conducted concurrently with the release of ChatGPT, a breakthrough technology that generates responses to text prompts resembling those of a human being. The remarkable performance of ChatGPT was a motivation to delve into the underlying technologies and their potential applications in literary research.

ChatGPT operates on a language model called Generative Pre-Trained Transformer 3 (GPT-3). This software uses probabilistic reasoning to produce text that fits the context of a given input prompt. ChatGPT, which is optimized for conversation, is built on GPT-3. OpenAI<sup>18</sup>, the developer of both GPT-3 and ChatGPT, provides a web API for programmatically interacting with the model. Consequently, the idea was conceived to utilize GPT-3 to examine each verse in the corpus programmatically, determining whether it expresses hope or a related sentiment and then annotating the corpus with GPT-3's responses.

The use of generative Artificial Intelligence (AI) for sentiment classification is a controversial decision, particularly given the existence of a dedicated field for such tasks: discriminative AI. However, all discriminative AI implementations require significant training data and multiple refinement iterations to produce satisfactory models. Even GPT-3 can be fine-tuned for classification tasks, but this requires providing at least 100 instances per class, which would be a substantial research undertaking. As a result, this study will be restricted to what is achievable with the general-purpose generative model.

GPT-3 generates different responses based on the prompt provided, and even the same prompt can produce diverse responses at different times. This unpredictability can pose challenges to the reproducibility of research findings, but it is still worthwhile to test the technique.

---

<sup>18</sup> As per time of writing, the latest papers by OpenAI are released under this link [openai.com/research](https://openai.com/research). The paper on GPT-3 is (Brown et al. 2020).

To prompt GPT-3, the original Arabic text and its English translations were used, with appropriate prompts determined through experimentation. Due to the financial costs associated with using GPT-3, which are calculated based on the number of tokens in each prompt and response, the experimentation process had to be concise. Additionally, applying multiple prompts for the entire corpus would be costly. Therefore, only one prompt was employed for the Arabic corpus and one for the English translations. It is important to note that processing Arabic text requires more Unicode tokens and is more expensive.

### 2.2.3.1 Analysis of the Original Poems

For Arabic poems, the following prompt gave the most systematic responses:

"Does this verse express hope? {verse text}"

Due to the financial constraints, no other prompts could be widely tested. Based on some brief tests, it was noticeable that longer and more elaborate prompts, like the one used for translations, gave highly chaotic responses. The prompt was applied for each verse in the corpus, and the responses were collected and added to the previously presented MongoDB collections:

```
{
  "numerical_id": 1,
  "collection": "المعلقات",
  "poet": "Imru al-Qays (EI: Imru' al-Ḳays b. Ḥudj̄r)",
  "poem": "Arb_Mual_I",
  "order": 1,
  "GPT": "No, this verse does not express hope. It is describing
a situation where something has been done, but the result is not
favorable.",
  "tokens_r": [...],
  "tokens_l": [...]}
}
```

### 2.2.3.2 Analysis of the Translated Poems

For English translations the following prompt was used:

"Does this verse express hope, wish, desire, or something missing from the speaker's life?

{verse text}

If none is expressed, answer with 'None'. Otherwise provide the expressed sentiment then elaborate."

Similarly, the MongoDB collection of the translated poems were extended:

```
{  
  "numerical_id": 1,  
  "collection": "المعلقات",  
  "poet": "Imru al-Qays (EI: Imru' al-Ḳays b. Ḥudjr)",  
  "poem": "Arb_Mual_I",  
  "order": 1,  
  "GPT": "None",  
  "verse": "...."  
}
```

### 3. Lemma-Driven Verse Exploration and Interpretation

The Arabic corpus comprises 158 poems authored by 102 poets, amounting to 4044 verses. On the other hand, in the English corpus, a total of 3280 verses were translated from 133 poems by 86 poets. The previous chapter presented a comprehensive account of the annotation process. The upcoming sections will delve into the analytical approaches used to identify the most relevant verses on the theme of hope.

The expression of meaning in natural language is primarily associated with semantics. This fundamental level is responsible for conveying the intended message through words. It is generally expected that words related to the semantics of hope and related emotions are used in contexts that express such sentiments. However, when it comes to poetry, there is no guarantee that identifying such words would lead to verses where the poet expresses utopian visions.

The process of annotating the corpus and analyzing the resulting annotations of famous poems prompted the research inquiry on the connection between the semantics of hope and the expression of a utopian vision. The choice of particular lemmas assisted in narrowing the focus of the close reading to a subset of poems. GPT analysis of English translations was far more accurate than the analysis of Arabic and helped identify verses worthy of deeper examination.

The following section briefly analyzes the occurrence of the selected lemmas. The subsequent sections discuss the themes and patterns discovered by closely reading the located verses. GPT responses will be provided but not individually discussed. Only a general evaluation of these responses is presented in the conclusion.

#### 3.1 Overview of the Lemmas

The following table shows the lemmas that were selected to drive the investigation of hope, and some count data related to their occurrence in the annotated corpus. All lemmas are included in the exact form, i.e. *rasm*, vowel marks, and additional writing signs, with which CAMEL tools produced them.

Lemma	Verse Count	Identified as Hope by GPT	Poets
لُو	78	24	43
وُد	13	3	10

بَعَى	8	2	8
لَيْتَ	8	0	8
نَزَعَ	7	2	7
شَوَّقَ	5	4	3
لَعَلَ	5	3	5
حَرَصَ	2	2	2
شَاقَ	2	1	2
تَمَنَّى	2	1	2
شَهَوَ	1	0	1

After a preliminary survey of the verses containing these lemmas, it is concluded that they occur in vastly diverse contexts. Additionally, there are some false identifications by CAMEL tools. Nonetheless, some relevant themes were discovered.

### 3.1 Metaphysical Visions

The poets express hopes beyond what is physically or biologically possible in the following two verses:

The first verse is by Ġuraiqa al-‘Absī:

بِمَا لَمْ تَكُنْ عَنْهُ النَّفْسُ تُطِيبُ

فَلَوْ كَانَ مَيِّتٌ يَفْتَدِي لَفَدَيْتَهُ

**Translation:** not included.

**Own Translation:** If a deceased can be ransomed (back into life), I would ransom him with what souls constantly desire.

**GPT response to the Arabic text:** “No, this verse does not express hope.”

Almost nothing is known about this poet, beyond his name and that few poems are attributed to him. The poem expresses many grievances against fate and the inevitability of death<sup>19</sup>. The wish that death would be avoidable and the dead could be brought back to life is a recurrent theme in utopian visions in literature.

The second verse is by Salāmah ibn Ġandal:

وَدُّ الْقُلُوبِ مِنَ الْبَيْضِ الرَّعَائِبِ

وَلِلشَّبَابِ إِذَا دَامَتْ بِشَاشَتُهُ

**Translation:** Not included.

**Own Translation:** And for the youth, if its beauty would last, the desires of the hearts of beautiful women.

**GPT response to the Arabic text:** “No, this verse does not express hope. It is talking about the beauty of youth and how it fades over time.”

In this beginning part of the poem, the poet mourns his fleeting youth<sup>20</sup>. The verse has an implicit hope for eternal youth, which is a typical utopian vision.

### 3.2 Need for Peace

War was famously a constant troublesome part of life in early Arabia. The following three verses contain expressions of hope concerning the constant threat of war.

The first verse is by Zabbān ibn Sayyār:

لَوْ كَانَ عَنْ حَرْبِ الصَّدِيقِ سَبِيلٌ

أَبْنِي مَثُولَةَ قَدْ أَطَعْتُ سَرَاتِكُمْ

**Translation:** Ye Sons of Manūlah! in sooth I obeyed your chiefs, hoping that there might be a way to avoid war with friends.

**GPT response to the Arabic text:** “No, this verse does not express hope. It is expressing a willingness to follow the instructions of the speaker, but does not express any hope for a particular outcome..”

**GPT response to the translation:** “Hope: The speaker expresses hope that there is a way to avoid war with friends.”

<sup>19</sup> (Hārūn and Šākir 1967) pp. 98-100.

<sup>20</sup> (Hārūn and Šākir 1979) pp. 119, 120.



In this opening verse, the poet promises his tribespeople that he would obey his chiefs if he would find a way to avoid war with friends. In the rest of the poem, the poet expresses his preparedness for war, and does not really further reflect on the thought of the first verse<sup>21</sup>. The thought in this verse is both situational and indicative of a recurrent social phenomenon of the time, which is the inevitability of wars due to competition over resources and status. The poem can be interpreted as containing a utopian vision of more abundant resources and a more peaceful society.

The second verse is by ‘Amr ibn al-Ahtam:

وَعَادَانِي شِوَاءٌ أَوْ قَدِيرٌ

وَلَوْ أَنِّي أَشَاءُ كُنْتُ جَسْمِي

**Translation:** And if I had chosen, I could have sheltered my body from hardship, and there had been meat, roast or seethed, for my breakfast.

**GPT response to the Arabic text:** “Yes, this verse expresses hope. It implies that if the speaker wished, they could be changed or given a different destiny.”

**GPT response to the translation:** “None.”

In the context of the poem, the poet says that he has the choice to save himself from the hardships of war, and that he only goes into war to honor the ancestors and to follow their example<sup>22</sup>. The theme of this verse is similar to the previous one by Zabbān. However, the counterfactual here is about the poet willfully choosing peace over war. Thus, what the poet is wishing for in this verse is that he would not choose war. The larger context of the poem indicates that his devotion to his ancestors is what is driving him to do so.

The third verse is by al-Mumazzig:

وَوَدَّ الَّذِينَ حَوْلَنَا لَوْ تَشْرِقُ

وَوَجَّهَهَا غَرْبِيَّةً عَنِ بِلَادِنَا

**Translation:** And he [the leader (?)] diverted them westwards away from our country – and those about us had been glad if he had gone eastwards.'

**GPT response to the Arabic text:** “No, this verse does not express hope. It expresses longing for a distant land and a wish that those around them could understand.”

<sup>21</sup> (Hārūn and Šākīr 1979) pp. 351-353.

<sup>22</sup> Ibid. pp. 409-412.

**GPT response to the translation:** “None.”

The verse is about a marauding army, and it is expressed that the people, who are the target of the attack, wish that this army would go in the opposite direction. In the poem, the poet is actually glorifying that army<sup>23</sup>. While hoping for peace is not the central theme of the poem, the verse nonetheless points out this sentiment.

### 3.3 Hopes Related to Travel and Separation

Life in the desert, with extensive distances between communities and tiresome means of travel, brought many challenging life conditions. The prominent theme of *nasīb* deals, to a large extent, with separation from a lover, precisely due to these conditions. The following two verses explicitly express hopes related to the travel conditions and ailments of separation.

The first verse is by Ṭarafa:

أَلَا لِيَتَنِي أَفْدِيكَ مِنْهَا وَأَفْتَدِي

عَلَى مِثْلِهَا أَمْضِي إِذَا قَالَ صَاحِبِي

**Translation:** Such is the beast I ride, when my companion cries 'Would I might ransom you, and be ransomed, from yonder waste!'

**GPT response to the Arabic text:** “No, this verse does not express hope. It expresses regret and a desire to give something in exchange for something else.”

**GPT response to the translation:** “Wish: The speaker wishes that they could be ransomed from the waste they are in.”

The exact meaning and context of the poem imply that the poet's companion wishes him that traveling through the desert on a camel is not as hard as it is. The companion explicitly wishes for the possibility to ransom the poet and himself from this hardship<sup>24</sup>. This occurrence has a connotation of a utopian aspiration and points out a genuine difficulty in the lives of people in the Arabian peninsula at the time.

The next verse is by al-Ḥārīṭ, which utilizes the famous exclamation expression *layta shi'rī*, meaning “I wish I know.” or “I wish I could know.”

<sup>23</sup> (Hārūn and Šākīr 1979) pp. 301, 302.

<sup>24</sup> (al-Zuzanī 1983) p.106.

لَيْتَ شَعْرِي مَتَى يَكُونُ اللَّقَاءُ<sup>25</sup>

أَذْنَتْنَا بَيْنَهَا ثُمَّ وُلَّتْ

**Translation:** not included in the provided corpus.

**Own Translation:** She hurt as with her staleness then departed, oh I wish to know when is the reunion!

**GPT response to the Arabic text:** “No, this verse does not express hope.”

The poet hopes to know when he can meet again with his beloved. The hope of the poet could be viewed as situational not discussing an element relating to society or the lifestyle. However, a broader look at the circumstances can regard the thought as expressing how hard it is to communicate with someone in a distant location and keep up with their news.

### 3.4 Morals and Mental Attitudes

Improving morals and mental attitudes is integral to any utopianistic hope to transform life and society. The following three verses express criticism and advice in this regard.

The first verse is from a poem by Ufnūn:

وَتَقْوَالَهُ لِلشَّيْءِ يَا لَيْتَ ذَا لِيَا

فَلَا خَيْرَ فِيمَا يَكْذِبُ المرءُ نَفْسَهُ

**Translation:** No profit is there to a man who strives to deceive his soul, who wails over Fate and cries: 'Ah, would that my lot were thus!

**GPT response to the Arabic text:** “No, this verse does not express hope. It is expressing regret and disappointment.”

**GPT response to the translation:** “None”

The general theme of this short poem is the inevitability of fate. The poet reportedly uttered the poem when he was bitten by a snake and was expecting an imminent death<sup>26</sup>. The poet promotes honesty and warns from self-deception. Additionally, he deems greed and wishing for the unattainable to be pointless.

The second verse is by Maraqqish the Elder:

<sup>25</sup> This seems to be an infrequently included verse. It is in the provided corpus, it is not in the (al-Zuzanī 1983) version or in the translation of A.J. Arberry, which the corpus used for this poem.

<sup>26</sup> (Hārūn and Šākīr 1979) pp. 260-261.

أَوْ يَسْبِقُ الْإِسْرَاعُ سَيِّبًا مَّقْبِلًا

فَلَعَلَّ بَطْأً كَمَا يَفْرُطُ سَيِّئًا

**Translation:** And perchance your delaying may send on ahead some evil thing [so that it will not affect us] : or it may be that, if ye hurry away, ye may miss some good that is coming to you.

**GPT response to the Arabic text:** “Yes, this verse expresses hope. It suggests that even if the two of you make a mistake, the speed of progress and success will eventually make up for it..”

**GPT response to the translation:** “None.”

The poet is advising his companions that haste is pointless, it does not bring any good or prevent any evil<sup>27</sup>. Such utterance is an important general life advice in regard to care and deliberation.

The final verse is the last verse in a poem by al-Jumaiḥ:

فِي سَجَبِلٍ مِّنْ مَّسُوكِ الضَّأْنِ مَنجُوبِ

فَأَقْنِي لَعَلَّكَ أَنْ تَحْظِي وَتَحْتَلِبِي

**Translation:** Then possess thyself in patience – it may be that yet shalt thou be happy, and milk thy herds into a great milk-skin – the skin of a sheep tanned with acacia-bark.

**GPT response to the Arabic text:** “Yes, this verse expresses hope.”

**GPT response to the translation:** “Hope: The speaker expresses hope that they will be able to milk their herds and have a great milk-skin.”

In the poem, the poet is addressing a woman named Umāmah. In this final verse, he is prompting her to be patient while wishing her that her patience would be rewarded with great wealth<sup>28</sup>. The verse shows an association between wealth and happiness, and expresses the hope for a final reward for specific actions and attitudes. The poem and the verse can be considered as expressing a utopian vision.

<sup>27</sup> (Hārūn and Šakir 1979) p. 222.

<sup>28</sup> Ibid. pp. 34-35.

### 3.5 Psychology of Hope

In the next two verses, poets express some self-reflection concerning their hopeful feelings:

The first verse is by Ta'abaṭa Sharra :

وَمَرَّ طَيْفٌ عَلَيَّ الْأَهْوَالِ طَرَأًا

يَا عَيْدُ مَالِكَ مِنْ شَوْقٍ وَإِيرَاقٍ

**Translation:** O return of remembrance! how with thee come longing and wakefulness, and the passing of a phantom darkling, spite of terrors by the way!

**GPT response to the Arabic text:** "Yes, this verse expresses hope. It is wishing someone a happy holiday and encouraging them to find a way to overcome their worries."

**GPT response to the translation:** "Wish: The speaker expresses a wish for the return of remembrance, and longing for something that has been lost"

This verse serves as an exclamation about the inescapable sentiment of yearning and its resultant effects, such as insomnia and other associated distresses<sup>29</sup>. Though it does not explicitly express hope, its relevance to the discussion lies in its recognition of this fundamental aspect of human psychology: the persistent inclination to desire what is better and what has been lost.

The second verse is by as-Samaw'al Aḥū Sa'yah:

قِيلَ أَقْرَأْ عُنْوَانَهَا وَقْرِئْتُ

لَيْتَ شِعْرِي وَأَشْعَرَنُ إِذَا مَا

**Translation:** not included in the provided corpus.

**Own Translation:** I wish to know that if I were prompted to read its meaning I would do so.

**GPT response to the Arabic text:** "No, this verse does not express hope."

The verse occurs in a highly philosophical and reflective poem in which the poet discusses the themes of life, death, aging, forgiveness, morality, and honest living<sup>30</sup>. It is unclear what the poet is referring to in this verse. The verse immediately before this one is about forgiveness, more precisely, ignoring insults. The one immediately after is a general statement

<sup>29</sup> (Hārūn and Šākīr 1979) p. 27.

<sup>30</sup> (Hārūn and Šākīr 1967) pp. 85, 86.

that he cannot judge himself on his own, whether he is virtuous or vicious. There is no plausible explanation for this verse that fits either context.

## 4. Conclusion

The data-driven approach adopted in this study concentrated the close reading efforts on 131 verses selected from a corpus of 4,044 verses. Upon examination, a mere 12 verses were found to contain themes pertinent to the investigation topic. It is conceivable, however, that poets expressed thoughts akin to the five themes identified in Chapter 3 without employing the lemmas specifically chosen for this inquiry. Future studies may benefit from utilizing the same method but with more comprehensive, well-categorized lemma sets that encompass not only hope but also lemmas related to peace, travel, morals, and other aspects commonly associated with utopianism.

It is essential to recognize that the corpus used in this study represents only the well-known anthologies of early Arabic poetry, leaving out numerous poems outside these compilations, particularly within the *diwāns* of various poets. The limitation arising from the incompleteness of the digitized corpus is an ongoing challenge encountered in computer-assisted Arabic literary research.

Regarding the use of GPT-3, it has been demonstrated that the model is not entirely reliable for the task at hand, particularly when analyzing individual verses. Often, even for a human reader, discerning the meaning of a verse necessitates reading the entire poem. Based on GPT-3's current limitations, it is feasible to include only short poems within a single prompt. As the model continues to be refined and improved, the capability to analyze entire long poems may be attainable in the near future. Another critical issue in employing GPT-3 is the lack of reproducibility in the results, which holds even for the English translations that the model is purportedly more adept at processing. Providing GPT-3 with the same prompt for the same verse multiple times yields inconsistent responses, which poses challenges to the reliability of the analysis.

Overall, a data-driven approach holds considerable potential for identifying themes of hope in poetry. However, improved outcomes are contingent upon the availability of a more accurate lemmatizer. Although the CAMEL tools lemmatizer represents a significant advancement over similar software from the past decade, it still generates numerous incorrect lemmatizations.

On the researchers' side, employing a more comprehensive lemma set and incorporating multi-word expressions could facilitate better identifying relevant verses. To fully harness the capabilities of artificial intelligence, it is crucial to conduct manual research and structure its results as training data for discriminative and generative AI models. This approach would be instrumental in ensuring that the outcomes derived from AI applications are reproducible and consistent.

By refining the data-driven approach and integrating advanced AI technologies, researchers can more effectively analyze poetry and uncover subtle themes of hope embedded within the texts. Investigating these themes will not only serve to deepen understanding of poetic works but also pave the way for more nuanced interpretations of historical and cultural contexts within which these works were created.

## 5. Bibliography

- Al-Asad, Nāsir ad-Dīn. 1988. *Maṣādir Aš-Ši'r al-Ġāhilī*. 5th ed. Cairo: Dār al-Ma'ārif.
- al-Zuzanī. 1983. *Šarḥ Al-Mu'allaqāt al-'Ašr*. Dār Maktabat al-Ḥayāh.
- Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. 2020. "Language Models Are Few-Shot Learners." arXiv. <https://doi.org/10.48550/arXiv.2005.14165>.
- Ḍaif, Šawqī. 1984. *Aš-Ši'r Wa Ṭawābi'u Aš-Ša'biyya 'alā Mar al-'Uṣūr*. 2nd ed. Cairo: Dār al-Ma'ārif.
- . 1987. *Al-Fann Wa-Maḍāhibuhu Fī al-Ši'r al-'Arabī*. 11th ed. Cairo: Dār al-Ma'ārif.
- . 2003. *Al-'Ašr al-Ġāhilī*. 11th ed. Cairo: Dār al-Ma'ārif.
- El Gayar, Neamat, and Ching Y. Suen, eds. 2019. *Computational Linguistics, Speech and Image Processing for Arabic Language*. Series on Language Processing, Pattern Recognition, and Intelligent Systems, vol. 4. Singapore ; Hackensack, NJ: World Scientific Publishing Co. Pte. Ltd.
- El Tayib, Abdulla. 1983. "Pre-Islamic Poetry." In *Arabic Literature to the End of the Umayyad Period*. New York: Cambridge University Press.
- Hārūn, 'Abd ul-Salām, and Aḥmad M. Šākir. 1967. *Šarḥ Al-Ašma'iyyāt*. 3rd ed. Cairo: Dār al-Ma'ārif.
- . 1979. *Šarḥ Al-Mufaddaliyyāt*. 6th ed. Cairo: Dār al-Ma'ārif.
- Ḥassānīn, 'Aṭīyya M. 2020. "Bawāi'ū Al-Yūtūbiyā Wa-d-Dīstūbiyā Fī al-'ašr al-Ġāhilī: Baina at-Tanzīr Wa-l-Mumārasa." Cairo.
- Ḥiḍr, Muṣṭafā M. A. 2014. *At-Tamannī Wa at-Taraġġī Fī al-Qur'ān al-Karīm Wa Aš-Ši'r al-Ġāhilī "Dirāsa Luġawiyya"*. Alexandria: Horus Publishing.
- Obeid, Ossama, Nasser Zalmout, Salam Khalifa, Dima Taji, Mai Oudah, Bashar Alhafni, Go Inoue, Fadhl Eryani, Alexander Erdmann, and Nizar Habash. 2020. "CAMEL Tools: An Open Source Python Toolkit for Arabic Natural Language Processing." In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, 7022–32. Marseille, France: European Language Resources Association. <https://aclanthology.org/2020.lrec-1.868>.
- Sargent, Lyman Tower. 2010. *Utopianism: A Very Short Introduction*. Oxford University Press. <https://doi.org/10.1093/actrade/9780199573400.001.0001>.