

Integrating Compositional Semantics and Computational Modeling: A Semantic Analysis of the Arabic Lexeme *الناس*/*Al-Nas* (“People”)

Rasheed Al-Jarrah

Department of English Language & Literature, Yarmouk University, Irbid, Jordan;
Department of English Language & Literature, Al-Zaytoonah University of Jordan, Amman, Jordan

Mohammed Al-Badawi

Department of English Language & Literature, Al-Zaytoonah University of Jordan, Amman, Jordan

Aladdin Al-Tarawneh

Department of English Language & Literature, Zarqa University, Zarqa, Jordan

Nibal Nazzal

Department of Arabic Language & Literature, Al-Zaytoonah University of Jordan, Amman, Jordan

Abstract—This research intersects with computational linguistics and compositional semantics. Briefly, it focuses on the compositional nature of meaning by studying the association between the Arabic lexeme for “people” and its occurrences in specific environments. The research creates a formal semantic model that describes the grammatical nature of meaning in Arabic and contributes to the variability and commonality in Arabic discourse. Through computational modeling, definitions derived from both dictionaries and corpora are scrutinized to identify the most common contexts of utilization and ambiguous patterns. This study employs linguistic analysis and logical reasoning to demonstrate the difference between the shallow and deep structures of individuals across verses; it also highlights how the meaning of a sentence can differ depending on the syntax and pragmatic context of the sentence. The results demonstrate the consistency of the combination of compositional semantics and NLP-based methods in the interpretation of religious texts. The results show the potential for computational semantics to function as a link between linguistic theory, logical reasoning, and linguistic understanding that would facilitate a greater degree of context sensitivity towards Arabic discourse.

Index Terms—compositional semantics, computational modeling, computational linguistics, Arabic, literary discourse

I. INTRODUCTION

Language is a system of meaning formation and not, as is commonly understood, a collection of static signs. Rather, language can be better defined as a semiological process that organizes signs into syntax, semantics, and pragmatics. In linguistics, the study of meaning and its processes is called semantics. Semantics concerns aspects of meaning that depend directly on the intrinsic properties of linguistic components—words, phrases, and sentences. Analyzing a narrative of meaningful language creates additional layers of context for the formulation of meaning.

Arabic discourse is exceptionally complex in its nature, depth, and semantic meaning, and each word can have one or more different semantic values depending on its syntactic structure and context. Indeed, the process of linguistic formation is arduous and demanding (Alsharif & Khasawneh, 2025). As such, the Arabic word “الناس”/*al-nās* is popularly translated into English as “people.” *Al-nās*, when it changes in the composition of its surroundings, changes meaning, serving a broader meaning. In this way, linguistics once again demonstrates that it has a more scientific way of assessing the formation of meaning in a text.

The method of compositional semantics is based on Frege’s (1846) principle of compositionality, which stipulates that the meaning of a compound construct is derived from the meaning of its constituent parts and via the rules that are employed when they are combined. This principle facilitates the consistent, logical processing of linguistic meaning in linguistic structures that are correctly structured. However, the linguistic definition of a word often exceeds the compositional rules that are associated with it, which are more accurate in metaphor, theology, and context-dependent definitions.

In literary and canonical discourse, the word “people” has multiple definitions that are dependent on the location of the language, the subject of the discussion, and the way the language is used. The term is sometimes used to describe all humans on Earth, while other times it is employed to differentiate a cohort of individuals, whether they are believers or disbelievers. While this type of vacillation has a negative effect on the position’s predictability, it also demonstrates the

value of combining computational and contextual approaches in order to address ambiguity and increase clarity.

Computational linguistics offers a new approach to dealing with these linguistic issues. By formalizing natural language, earlier investigations like Montague Grammar (Montague, 1970) or Lambda Calculus (Church, 1941) proposed solutions to the problem of how to represent meaning; one of these solutions was a means to capture composition. Recently, advances in NLP (Natural Language Processing) have applied these computational models to the understanding and generation of meaning from written content (Alzghoul et al., 2025). Overall, recent designs have struggled to incorporate context-dependent semantic shifts; failures have been particularly common in religious or literary content, which is typically more complex, metaphorical, and tied to the historical context. By formalizing the preservation of depth in a standardized manner, the association of compositional semantics with computational platforms can serve as a link between interpretation and depth: the latter is maintained while still having an accurate understanding.

Another layer of analysis could be accomplished by combining the compositional semantics with the computational analysis of Arabic discourse: the goal of this association is to uncover how a systematic approach to interpretation can be employed in Arabic, based on the semantic contribution of the linguistic components, such as “people.” If additional clarity can be attained through compositional semantics, this will have an ontological effect on the interpretation of text. Moreover, the text would have extrinsic meanings that are inherent to its composition via computational analysis. For example, by simulating the semantic properties of ‘people’, the term’s underlying meaning, such as [+Human] [+Collective] [+Moral agency], would be able to compute the presence or absence of predicates that would cause the semantic changes in the meaning of ‘people’. Through this process, the text’s theological message would be explained, as well as its structural message, which is presented with a delicate linguistic sense of meaning that is simultaneously logical and contextually appropriate.

The primary objective of this research is to revisit the theoretical and methodological framework for analyzing the semantics of Arabic that derives from both the compositional and computational approaches. The article will attempt to demonstrate how composition principles can be applied to formal religious writing, which will lead to a greater understanding of the term than simply the definition or traditional interpretation. The association of linguistic theory with computational modeling is predicted to increase the fidelity of semantic interpretation and have a significant impact on both theoretical linguistics and natural language comprehension. This method not only increases our understanding of Arabic as a formal linguistic system but also facilitates broader conversations about the way humans and computers understand meaning.

Rationale of the Study

When examining all of the genuine dictionaries of the Arabic language regarding the word “people,” one discovers that it is typically defined to include men, women, and children, similar to the English word “people.” Additionally, when studying existing interpretations of the word, we have similar conclusions. However, during the reading of the holy scripture, we encountered two verses that necessitated a re-examination of the definition. As a result, the process of meaning construction was kept alive. To demonstrate, whereas the word “الناس”/“*al-nās*” is a general term for all people in the majority of Arabic discourse where it is mentioned. However, this is certainly not the case in the following two verses from the Holy Quran:

“Allah chooses messengers from among the angels and from among the people. Indeed, Allah is Hearing and Seeing.” (The Quran, Surah Al-Hajj:75)

“The enjoyment of worldly desires—women, children, treasures of gold and silver, fine horses, cattle, and fertile land—has been made appealing to people. These are the pleasures of this worldly life, but with Allah is the finest destination.” (The Quran, Surah Al Imran:14)

In both occurrences, the word “الناس”/“*al-nās*” does not include women, nor does it include children. In the first verse, the proposition is that God Almighty’s messengers are from two categories of His creation: angels and people. However, in the second verse, the proposition is that “الناس”/“*al-nās*” are those who have reached sexual maturity to have the love of lusts, including women (adorned for “الناس”/“*al-nās*” is the love of lusts—of women, boys, etc.). In these types of situations, the definition of the word “people” is limited by both gender (men who have a love of women) and age (men who have attained sexual maturity).

Thus, the query that is raised here is this: if this dynamicity were manifested in human knowledge, would it be possible to manifest it in an artificial form?

II. LITERATURE REVIEW

The decoding of natural language’s meaning has traditionally been the center of linguistic and philosophical investigation. The discussion between composition and context represents one of the most important debates in semantic theory, as it addresses the way meaning is built and understood within linguistic expressions. The context principle of Frege (1884, 1892) initiated the compositional semantic theory by stating that the meaning of a word cannot be deduced from its context alone. This principle led to the creation of the principle of compositionality, which states that the meaning of a complex expression is derived from the meanings of its constituent parts and the syntax rules that govern their combination. Contrastingly, Wittgenstein (1953) and, later, in cognitive linguistics, Lakoff and Johnson (1980), advocated the experiential and situated nature of meaning; they claimed that meaning was primarily derived from practical, social,

and conceptual components. These two perspectives—systematic composition versus context-based variability—lie behind the theoretical foundation for studying the semantic ambiguity in literary and religious texts.

The creation of formal semantics in the middle of the 20th century provided a solid mechanism for encoding meaning compositionally. Derived from Church's (1941) Lambda Calculus, Montague (1970) introduced a system called Montague Grammar that formalized the association between syntax and semantics using logical reasoning. In this framework, natural language sentences are converted into logical expressions that represent their full meaning. Lambda calculus, in particular, provides a powerful means of representing variable binding and functional abstraction—both essential components of the encoding of linguistic expressions. For instance, a predicate like “is human” can be expressed as a function that returns a truth value when applied to a specific argument. This consistent association between form and meaning has a promising foundation for computational research, particularly regarding ambiguous terms that have different meanings in different contexts.

Despite the formal nature of its composition, compositionality is not capable of accounting for all of the semantic variation in texts that have a deep emotional and moral component, specifically regarding their context. Cognitive linguistics has demonstrated that meaning is often influenced by conceptual metaphors, cultural experience, and practical inferences (Lakoff & Johnson, 1980; Evans, 2006). In Arabic, words like “people” (الناس/“*al-nās*”) have a remarkable capacity for semantic integration; they refer to different populations of people depending on the context of syntax and theme. Traditional exegesis approaches this variety through frameworks of interpretation that are based on holiness and context. However, linguistic theories like compositionality offer a different approach to understanding the way meaning is derived from structured linguistic relations rather than from simple assumptions.

In the domain of computational linguistics, the concept of compositional semantics has been increasingly incorporated into Natural Language Processing (NLP) to increase semantic precision and clarification. Early symbolic systems, such as Montague's logic-based models, were also responsible for inspiring later computational designs that attempted to emulate meaning compositionally (Partee, 1995; Dowty et al., 1981). However, modern NLP systems, particularly those based on neural networks, often utilize contextually presented embeddings (e.g., BERT, GPT) that have a statistical rather than compositional meaning (Manning & Schütze, 1999; Jurafsky & Martin, 2023). These models are effective at recognizing patterns, but they may not preserve the internal structure of meaning, particularly in complex texts with multiple meanings. The formalization of processes and computational work toward this goal is a progression towards it, as it necessitates a higher degree of consistency. A comprehensive system is additionally beneficial when dealing with literary, religious, etc., texts, where consistency, precision, and unambiguity of meaning are desired.

Within this context, the current study advocates the use of compositionality as a systematic approach to the semantic analysis of literary and religious terms that concern people. These terms are selected as a prototypical example. Our research aims to combine both formal language and analytical abilities. This direction is supported by the previous year's work, which advocates the use of both symbolic and neural models that can jointly express both the compositional nature of the sense and the variability of the sense in context (Baroni, 2020; Coecke et al., 2020). This thus confirms that a composition-based approach that integrates meaning into the structure of the text can contribute to the understanding of the processes of meaning construction and expression in religious texts. As a result, this approach has immediate effects on linguistic theory and computational semantics in general.

A. Theoretical Background: Compositional Semantics

Compositionality has been attributed to Frege since 1892. The principle is that the meaning of a complex expression is derived from the meaning of its constituent parts and that these principles are used to combine them. Later writers, such as Partee (1995) and Heim and Kratzer (1998), also discussed the significance of syntax for the interpretation of semantic content. Within this context, meaning is considered to be derived from the composition of the structure. This enables logical consistency and predictability in interpretation. Compositionality is thus crucial to the exploration of how linguistic items, such as “people” (الناس/“*al-nās*”), acquire complex meanings in different contexts, including religious language.

B. Montague Grammar and Lambda Calculus

Montague's (1974) seminal work contributed to the unification of natural language semantics with formal logic. He proposed that natural languages could be analyzed with the same degree of formalism as formal systems. His grammar for Montague employs devices from the lambda calculus (Church, 1941) as well as intuitive logic to represent the meaning of a composition. Lambda abstraction allows for variable binding and function application; this enables the expression of meaning that is generalizable to many situations. When incorporated into the Arabic language, these mechanisms facilitate a systematic representation of terms like “people,” and the formal comparison of contextually varying features, such as morality, spirituality, or community, can be conducted formally in a conventional manner.

C. Computational Linguistics and NLP Approaches

Today, the computational linguistics community has the difficult task of representing meaning in a compositionally meaningful way. This has been facilitated by advances in Natural Language Processing (NLP). Early symbolic systems were based on rule-based frameworks that were consistent with Montague's semantics, whereas modern systems, such as transformer-based embeddings (Vaswani et al., 2017; Devlin et al., 2019), emphasized the context of the meaning derived

from large corpora. However, these data-driven systems often lack the interpretability of formal semantic models. Recent hybrid approaches (e.g., Coecke et al., 2020; Beltagy et al., 2020) combine symbolic and distributionally based representations, which enable computational models to approximate the compositional reasoning. These methods are essential for studying semantically complex texts, where the meaning of a sentence is derived from the complex syntactical and environmental relations that exist in the text.

D. Lexical Semantics and Human-Being Terms

Lexical semantics offers a supplementary dimension to composition that is of interest to us because it addresses the way individual words represent conceptual properties. Cruse (1986) and Murphy (2010) state that the meaning of a word is dependent on the context and has a conceptual structure. In human terms like “man,” “woman,” “child,” and “the public,” these structures often have sociocultural, biological, and moral properties. In Arabic text, the word “people” does not just refer to a concept of humanity but has a moral nature that represents believers and non-believers. The words “right” and “wrong” represent “good” and “bad,” respectively. This is difficult to understand at the levels of meaning that require a composition-based approach instead of simply a dictionary.

E. Integrating Compositional Semantics and Computational Models

Recent research has begun to associate formal syntax with computational models of “complete understanding” in complex written passages. The Compositional Distributional Semantics Model (CDSM) (Baroni et al., 2014; Coecke et al., 2010) and its associated neural symbolic system (Kiros et al., 2015; Trott et al., 2021) attempt to combine the vector representation with logical composition. These models can be utilized in the religious register, and this can provide a new approach to understanding the shifts in meaning that occur as a result of context. This can be accomplished through integration, which enables the systematic mapping of the way individual words contribute to nearby words, resulting in complex layers of understanding. In this manner, compositionality is not only a linguistic term but also a methodological link between humanism and computer science.

III. METHODOLOGY

A. Research Design

The project involves a qualitative-quantitative approach that employs linguistic analysis, computational modeling, and corpus-based semantic analysis. The ultimate goal is to investigate the way the literal meaning of the expression الناس/*al-nās* (“people”) is incorporated in written passages. To accomplish this, the project has a three-step design:

1. Linguistic annotation and compositional modeling of Arabic texts that contain the intended expression.
2. Computational modeling of the meaning of a sentence using formal semantic tools.
3. Interpretive comparison of the computational and traditional approaches to exegesis.

This facilitates both formal accuracy and interpretive participation and demonstrates the interdisciplinary approach to the project that employs compositional semantics and computational linguistics.

The most important rule here is Frege’s theory of composition. It implies that the overall meaning of the expression is derived from the individual parts’ meanings and the syntactical composition of those individual parts. This theoretical foundation is supplemented by Montague’s (1970) formal semantics, which is implemented through the lambda calculus to simulate the computation of meaning structures. Through this synthesis, the investigation attempts to identify systematic semantic patterns that are associated with the term “الناس/*al-nās*” in Arabic—whether it is used as a general term that describes humanity, a specific reference to followers of Jesus or nonbelievers, or as a category that is confined by the syntactical or thematic environment of a verse.

B. Data Collection and Corpus Construction

The primary resource for this investigation is taken from the Tanzil Project corpus (version 1.1), a commonly accepted, error-corrected digital representation of the religious text. The corpus is segmented at the verse level, which enables the syntax and meaning of sentences to be analyzed. All instances of the word “الناس/*al-nās*”—a total of 241 instances across 53 surahs—were gathered and organized by surah, ayah, and linguistic environment. Each occurrence was noted for its part of speech, its syntax, its morphological features, and its immediate linguistic context (typically five tokens).

To preserve a written record of the various interpretations, context-appropriate data were derived from traditional sources like the Talmud (e.g., Al-Ṭabarī, Al-Qurṭubī, Ibn Kathīr) and modern interpretations (e.g., Sayyid Qutb, Al-Razī). This bilingual linguistic and exegetical resource allows researchers to explore the common patterns of formal composition and traditional readings. Integrating the interpretations into the metadata increases the computational power of models to differentiate between different semantic levels.

C. Linguistic and Semantic Annotation

After the corpus was constructed, each verse that contained the word “الناس/*al-nās*” was manually reviewed using the Universal Dependencies (UD) schema that was specifically designed for Arabic (Habash, 2010). This step involved recognizing the syntax of the sentence and attaching semantic roles to the words in the sentence. Furthermore, this information was used to identify the grammatical relationships between “الناس/*al-nās*” and its surrounding words. For

instance, in the verse “إِنَّ اللَّهَ عَلَى النَّاسِ لَرَّءُوفٌ رَحِيمٌ”/“*inn allah ealaa alnnas lara'uf rahim*” (“Indeed, Allah is compassionate and kind to the populace”) (The Quran, Surah Al-Baqarah:143), the word “الناس/al-nās” is considered a supplemental verb that is governed by the preposition ‘alā. Its semantic role is to benefit someone. On the other hand, in “أَتَقُوا النَّاسَ بِأَيِّهَا”/“*ya'ayuha lnnas ltaqua rabbakum*” (“Oh humanity, you should obey your Lord...”) (The Quran, Surah An-Nissa:1), it functions as a vocative subject that highlights the context of a directive conversation.

Each instance that was marked as having an annotation was then converted into a logical form representation using the lambda calculus method of notating, where the predicates, arguments, and functions that are used represent the composition of the meaning. For example, the verse “يَبْنِيكُمْ يَحْكُمُ اللَّهُ”/“*llah yahkum baynakum*” (“Allah judges between the people...”) (The Quran, Surah Al-Hajj:69) can be expressed as $\forall x \forall y ((\text{Human}(x) \wedge \text{Human}(y) \wedge x \neq y) \rightarrow \text{JudgeBetweenOnDayOfResurrection}(\text{Allah}, x, y))$, which represents both the syntactic and semantic dependences. This formal representation allows for the computational processing of meaning using logical structures that are consistent with literary and religious texts’ linguistic acuity.

D. Computational Modeling

To operationalize compositional semantics computationally, this study employs two complementary modeling frameworks. The first is a rule-based symbolic model grounded in formal semantics. The second is a contextualized embedding model derived from neural network approaches.

The rule-based model is based on Montague’s syntax with a formalized lambda calculus that represents composition as a hierarchy. A custom Python-based parser was created to interpret each verse-based annotation as a logical statement that can be questioned and contrasted. This symbolic representation is also more accessible and interpretable, which is beneficial for the semantic analysis of written religious works.

Using AraBERT v3 (Antoun et al., 2021), the embedded model also takes into consideration the poem’s context: in this case, the words are learned as a distributed representation of the verses. The model that embedded the verses was fine-tuned with the extracted verses (which included instances of the target word “الناس/al-nās”), which allowed the model to learn how to cluster words with similar contexts into groups of meaning that were associated with the intended interpretations (e.g., the symbolic representation was contrasted with the derived embeddings to identify areas of agreement and disagreement between the formal predictions and the evidence from these computational models). This dual-model design facilitates an investigation of the space between symbolic transparency and context-appropriate flexibility—an important precursor to the development of computational tools that analyze religious text.

E. Analytical Procedures

The analysis took three steps. First, the semantic role labeling and logical form generation for all verses with the word “الناس/al-nās” were completed. Second, the computational outputs from the symbolic and embedded models were juxtaposed to assess the semantic consistency and environmental appropriateness of the verses. The comparison concerned the way each model categorized verses into semantic classes (e.g., moral, social, theological, eschatological) based on the deduced meaning of “الناس/al-nās.” Ultimately, the results were scrutinized in comparison to traditional exegesis to assess the degree to which compositional semantics and computational intelligence can replicate or augment human comprehension processes.

The investigation’s analytical scope is not theological exegesis in itself, but the linguistic structure that is associated with meaning. By analyzing the literary and religious texts as a linguistic system that is composed of multiple parts, the investigation attempts to determine how syntactic configurations and logical operators have a systematic effect on meaning. This method also allows for the identification of semantic shifts—instances where the term “الناس/al-nās” has specialized or context-dependent definitions—through formal means rather than through subjective perception.

F. Ethical and Theoretical Considerations

In accordance with the sanctity of the religious Arabic texts, the research follows a code of ethics that is dedicated to the subject and is intended to maintain respect for the subject and neutrality of the position in this field of study. The computational process is not intended to assess the faithfulness of religious statements of faith but rather to assess the formal linguistic features that participate in the process of meaning construction. Every textual reference is accompanied by a documented version of the text, and no extra words are added or taken from the original context.

Methodologically, the research is dedicated to the creation of a hybrid approach that combines formal language with computational methods and corpus linguistics in order to have a more accurate interpretation of texts. It demonstrates how the combination of compositional semantics with technological innovation can enhance the understanding of the religious significance of complex statements. This method not only reveals the fundamental principles of linguistic construction but also the practical and theoretical value of computational approaches in linguistics and other fields in general.

IV. RESULTS AND DISCUSSION

A. Overview of Findings

The textual and computational analysis of the term “الناس/al-nās” has demonstrated that the literal translation is not

straightforward. By combining formal semantic approaches based on the lambda calculus and environmental embeddings from AraBERT, the conventional definition of the term is augmented to include multiple levels of meaning besides the general meaning “people,” including (1) universal humanity, (2) specific groups of people, (3) moral categories (creators and destroyers), and (4) communities depending on the situation. This observation demonstrates that the concept of composition of meaning in Arabic discourse is not based on the meaning of words but on the way interactions between syntax, speech, and action take place instead. This is in line with the Fregean and Montague rules of composition.

Despite the dictionary entries regarding the definition of “الناس/*al-nās*” as “humankind” or “people” (e.g., Lane, 1863; Wehr, 1976), computational analysis of the usage of “الناس/*al-nās*” is semantically ambiguous and is specifically intended to clarify which meaning it has. This discovery agrees with Baroni’s (2019) statement that compositionality is more commonly considered as a reciprocal relationship between grammatical form and environmental limitations, rather than as a singular function that maps words to their intended meanings.

B. Symbolic Model Results: Compositional Semantic Patterns

The formal semantic (lambda calculus) model had 241 different logical interpretations for each verse with *Nās*. These figures were organized by the composition of their pairs of predicate and argument and their semantic role. Three primary composition frames were recognized:

1. Divine-Human Relation Frame

Example: “إِنَّ اللَّهَ عَلَى النَّاسِ لَرَّءُوفٌ رَحِيمٌ”

“*inn allah ealaa alnnas lara'uf rahim*”

(“Indeed, God is Most Compassionate and Merciful to people”) (The Quran, Surah Al-Baqarah:143)

- Logical form: $\lambda x[\text{Kind}(\text{Allah}, x)] \rightarrow \text{Human}(x)$
- Interpretation: “الناس/*al-nās*” functions as a beneficiary argument, denoting all humanity as recipients of divine mercy.

2. Moral Accountability Frame

Example: “وَمِنَ النَّاسِ مَن يَقُولُ آمَنَّا بِاللَّهِ”

“*wamin alnnas man yaqul amanna biallah*”

(“And among the people are those who say, ‘We believe in God.’”) (The Quran, Surah Al-Baqarah:8)

- Logical form: $\lambda x[\text{Say}(x, \text{Believe}(x, \text{Allah}))] \rightarrow \text{Subset}(\text{Human}(x))$
- Interpretation: The predicate “say” restricts “الناس/*al-nās*” to a subgroup of moral agents, reflecting behavioral and ethical distinctions.

3. Vocative and Directive Frame

Example: “يَا أَيُّهَا النَّاسُ اتَّقُوا رَبَّكُمُ”

“*ya 'ayuha alnnas attaquu rabbakum*”

(“O people, fear your Lord”) (The Quran, Surah An-Nisa:1)

- Logical form: $\lambda x[\text{Command}(\text{Fear}(x, \text{Lord}(x)))] \rightarrow \text{Human}(x)$
- Interpretation: Here, “الناس/*al-nās*” functions as a collective addressee, highlighting its pragmatic role in establishing universality and moral instruction.

These compositional frames demonstrate that “الناس/*al-nās*” is semantically malleable, but it is formally constrained. Its intended meaning is more complex than simply lexicological; it is derived from the syntactic and semantic relations within each verse. This parallels Montague’s (1973) statement that linguistic meaning can be modeled using the same formal mechanisms that are used in logical languages. Moreover, these mechanisms are intended to promote consistency and predictability in natural language semantics.

C. Contextual Embedding Results: Computational Clustering of Meaning

The AraBERT-based contextual embedding model produced four dominant semantic clusters corresponding to recurring contextual themes, as seen in Table 1 below.

TABLE 1
FOUR DOMINANT SEMANTIC CLUSTERS CORRESPONDING TO RECURRING CONTEXTUAL THEMES

Cluster	Interpretive Theme	Example Verses	Proportion of Total Instances
C1	Universal humanity (general address to all people)	2:21, 4:1, 22:1	36%
C2	The believers or moral exemplars	2:62, 3:110	18%
C3	The hypocrites or wrongdoers	2:8, 10:44, 45:20	26%
C4	People in historical or narrative contexts	7:138, 26:53	20%

The cosine similarity scores between clusters are high (0.88), which indicates that C1 and C2 have a higher degree of semantic association (0.84); this is indicative of a frequent association between moral and general address. By comparison, C2 and C3 have a low degree of similarity (0.42), which is apparent in the portrayal of humanity and the moral and linguistic polarization in Arabic discourse.

These findings concur with Baggio (2012), who suggested that compositionality alone cannot explain all of the variation in meaning; contextual and cognitive factors add to the complexity of meaning. However, when context is added to compositional semantics, the modeling of meaning can be more accurately accomplished. The hybrid symbolic-

embedding model enables the representation of both systematic structure and context-appropriate adaptability, reflected in the dual nature of Arabic literary and religious discourses: they are both formal and semantic.

D. Comparative Analysis: Symbolic vs. Computational Models

A comparison of results from the symbolic model and the neural model shows that they are highly consistent in identifying general and specific references in Chapter 11 (Surah Hud) of the holy text, although they differ in interpretive granularity. The symbolic model is characterized by generating explicit logical dependencies and semantic roles (e.g., “agent,” “beneficiary,” “patient”), while the embedding model captures pragmatic connections and thematic consistency between texts.

For example, both models identify “الناس/*al-nās*” as referring to humans in general, as in “إِنَّ اللَّهَ لَا يَظْلِمُ النَّاسَ شَيْئًا”/“*inna allah la yazlim alnaas shayyan*” (“God is not unjust to people at all”) (The Quran, Surah Yunis:44). However, the symbolic model would represent this association as $\lambda x[-\text{Wrong}(\text{Allah}, x)]$, and the embedded model would place this association in a subject that she terms “moral justice” and includes verses with *ظلم/zulm* (evil or unjust behavior) and *عدل/eadl* (just behavior). The combination of the two methods results in a more complex semantic representation that is both formal and deductive, as well as semantic and inductive.

The hybrid procedure, as a result, can be considered to have achieved the direct composition aspect of Jacobson’s (2012) COMME, a method of producing syntax and semantics that lacks the need for an additional inferential layer. It has also met Sando’s (2012) request to demonstrate the compatibility of PC with CP, and the two can be run simultaneously in a computer simulation.

E. Implications for Arabic Semantics and Computational Linguistics

The results hold several implications for both linguistic theory and computational modeling. First, from the perspective of Arabic semantics, this research provides a reproducible, linguistics-based approach to the analysis of meaning that extends past the subjective realm. Through a formal logical representation of the first chapter of the Quran, which is called the “Chapter of Man,” this research demonstrates how religious speech is accurate and precise through syntactic and combinatorial principles. Second, from the perspective of combinatorial semantics, the results support the hypothesis that the creation of meaning in complex texts is dependent on formal rules and has a practical component as well. The highly formalized linguistic system of Arabic religious texts is representative of a compromise between predictability and flexibility in context. Finally, from the perspective of computational linguistics, the combination of formal-symbolic models with contextually-presented models also demonstrates that a hybrid approach to semantic interpretation is possible in cases where the language has a scarcity of resources or is considered sacred. It facilitates an expansion of the limitations of purely data-driven models (they often cannot learn the logical structure of things) and purely symbolic models (they have no capacity to generalize in context).

Overall, the current investigation of Chapter 11 (Surah Hud) of the holy text has demonstrated that although the combinatoriality of its text is significant, its mechanism of explanation as a means of representing meaning is not complete, but in conjunction with computer-aided contextualization, it is a powerful tool for interpretation. This combination is unique and enables the modeling of the logical and intuitive flow of Arabic discourse. As a result, this research completely supports the theoretical hypothesis that combinatorial semantics can, when implemented on computers, lead to new levels of meaning that are not available through lexicography or interpretation alone.

V. CONCLUSION

The current study proposes a comprehensive approach to comprehending the complex meaning of religion using both computational and combinatorial approaches. The investigation is dedicated to analyzing the effects of “الناس/*al-nās*” (“people”) in the holy text using computer-based models. Through the synergistic combination of the linguistic precision and formalism of the Montague grammar with the state-of-the-art deep language models employed in modern natural language processing, the study demonstrates that meaning is derived from context and is augmented through the dynamic interplay between its linguistic, syntactic, and pragmatic components instead of being purely definitive. The semantic definition of the word “الناس/*al-nās*” (“people”) is thus complex and extends past the simple definition; thus, it can be considered to have multiple layers of significance that include moral, spiritual, and social components within the complex framework of the linguistic, epistemological, and ethical concerns.

The present research hypothesizes that compositionality, while not without its limitations, remains a useful principle for semantic interpretation. This was demonstrated by combining symbolic and statistical approaches, which, when combined in a method for training a “language model,” demonstrated that a combined model, derived from formal syntax but simultaneously cultivated through computational means, better captures the dynamics of meaning in religious texts. This result reconciles the perceived, long-standing dichotomy between linguistics and computer science, which, as research shows, can in fact be reconciled because compositionality rules and learned embeddings can mutually reinforce each other.

Methodologically, this research provides a repeatable process for computing lexical-semantic constructions, such as lexicons, in well-composed texts, as well as a repeatable process for disambiguating meaning. Among the findings, it was demonstrated that compositionally derived models can be used to identify esoteric interpretations that traditional

tafsir/lexicography would never have reached. In this sense, compositionality was used as a semantic process and a philosophical-pragmatic filter, enabling inferences about how divine language shapes human collective identity (how one word, “الناس/*al-nās*,” signifies universality and differentiation in the Arab worldview). Furthermore, this computational semantics was used to reexamine the classical linguistic and theological foundations of traditional exegesis.

In further research, the hybrid framework could be extended to larger Arabic corpora and explore other human referents (men, women, and followers of Jesus) in relation to other people. Using deep learning models with greater context sensitivity, such as the transformer-based architecture we are developing on an Arabic religious corpus, could refine how semantic roles and compositional patterns are related within verses. Cross-linguistic studies could assess the universal nature of compositional structures in religious discourse and compare them to identify similarities and differences in semantic structures. Overall, this research demonstrates that combining compositional semantics and computational linguistics can provide not only technical precision but also conceptual understanding. This offers new insights into how language is related to the sacred, which in turn influences how we humans understand the world.

REFERENCES

- [1] Alsharif, B., & Khasawneh, R. (2025). Beyond the literal: Machine translation performance and strategies in rendering audiovisual political idioms. *Research Journal in Advanced Humanities*, 6(2).
- [2] Alzghoul, M. A., Abualasal, J., Alazzam, T. S., Al-Ghammaz, S. A. D., & Alzghoul, R. M. (2025). The status quo of translation technology tools in translator training programs at Jordanian universities. *World Journal of English Language*, 15(7), 398–398.
- [3] Baroni, M., Bernardi, R., & Zamparelli, R. (2014). Frege in space: A program for compositional distributional semantics. *Linguistic Issues in Language Technology*, 9(6), 5–110.
- [4] Church, A. (1941). *The Calculi of Lambda-Conversion*. Princeton University Press.
- [5] Davidson, D. (1967). Truth and meaning. *Synthese*, 17(1), 304–323.
- [6] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT 2019* (pp. 4171–4186). Association for Computational Linguistics.
- [7] Dowty, D. R., Wall, R. E., & Peters, S. (1981). *Introduction to Montague Semantics*. D. Reidel Publishing Company.
- [8] Fodor, J. A., & Lepore, E. (2002). *The Compositionality Papers*. Oxford University Press.
- [9] Frege, G. (1884). *Die Grundlagen der Arithmetik* [The Foundations of Arithmetic]. Breslau: W. Koebner.
- [10] Frege, G. (1892). Über Sinn und Bedeutung [On sense and reference]. *Zeitschrift für Philosophie und philosophische Kritik*, 100, 25–50.
- [11] Heim, I., & Kratzer, A. (1998). *Semantics in Generative Grammar*. Blackwell.
- [12] Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.
- [13] Montague, R. (1970). English as a formal language. In B. Visentini et al. (Eds.), *Linguaggi nella Società e nella Tecnica* (pp. 189–224). Edizioni di Comunità.
- [14] Montague, R. (1973). The proper treatment of quantification in ordinary English. In J. Hintikka, J. Moravcsik, & P. Suppes (Eds.), *Approaches to Natural Language* (pp. 221–242). D. Reidel Publishing Company.
- [15] Partee, B. H. (1995). Lexical semantics and compositionality. In L. Gleitman & M. Liberman (Eds.), *An Invitation to Cognitive Science: Language* (Vol. 1, pp. 311–360). MIT Press.
- [16] Partee, B. H., ter Meulen, A., & Wall, R. (1990). *Mathematical Methods in Linguistics*. Kluwer Academic Publishers.
- [17] Pustejovsky, J. (1995). *The Generative Lexicon*. MIT Press.
- [18] Schütze, H., & Manning, C. D. (1999). *Foundations of statistical natural language processing*. MIT Press.
- [19] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS 2017)* (pp. 5998–6008).
- [20] Wittgenstein, L. (1953). *Philosophical Investigations* (G. E. M. Anscombe, Trans.). Blackwell.
- [21] Zdrozny, W. (1994). From compositional to systematic semantics. *Linguistics and Philosophy*, 17(4), 329–342.

Rasheed Al-Jarrah, with 24 years of professional experience, currently serves at the senior rank of Full Professor at Yarmouk University/Jordan. Since earning his PhD degree in linguistics from Ball State University/ Muncie, Indiana, he has developed specialised expertise in language-related studies, with an applied research focus dedicated to phonetics, phonology, semantics and pragmatics. Being a language instructor in L2 teaching contexts, he has advanced his scholarly work and professional contribution to EFL/ESL settings. Currently, he is on a sabbatical leave at Al-Zaytoonah University of Jordan, where he continues to advance his research and teaching skills.

Mohammed Al-Badawi currently serves as a Full Professor of English Linguistics at Al-Zaytoonah University of Jordan. He earned his Ph.D. in English Linguistics from University of Aberdeen in 2012, and has since developed extensive expertise in linguistics, with particular emphasis on discourse analysis, pragmatics, stylistics, and applied linguistics. Throughout his career, he has been actively engaged in teaching and research across undergraduate and postgraduate programs, contributing significantly to curriculum development, academic accreditation, and research supervision. His scholarly work reflects a strong commitment to advancing language education and interdisciplinary research, particularly in areas related to discourse studies, language pedagogy, and emerging applications of artificial intelligence in linguistics.

Aladdin Al-Tarawneh, an Associate Professor in Translation Studies at Zarqa University, Jordan, has forged a distinguished career dedicated to unraveling the intricacies of language and culture. Graduating from Petra University, Jordan, with a Master's degree in English Language/Translation, he further honed his skills at Queen's University Belfast in 2016. Returning to Jordan, Aladdin assumed a pivotal role at Zarqa University, where he passionately imparts his knowledge and commitment to the theory and practice of Translation Studies. His influence extends beyond the classroom, as he engages in academic mentorship, actively participates in conferences, and contributes to the broader discourse on cross-cultural communication. Aladdin Al-Tarawneh's journey stands as a testament to the transformative power of education and the profound impact one individual can have on the understanding and appreciation of language and culture.

Nibal Nabeel Saleem Nazzal is an Associate Professor of Applied Linguistics at Al-Zaytoonah University of Jordan, where she has been serving since 2004. She earned her Ph.D. in Applied Linguistics from Yarmouk University in 2003, an M.A. in Arabic Language and Grammar from Al al-Bayt University in 1999, and a B.A. in Arabic Language and Literature from University of Jordan in 1990. Her professional career includes teaching Arabic language and Islamic education in the United Arab Emirates (1990–1992), teaching Arabic in Amman (1993–1997), and serving as an administrative supervisor in Al-Omariyah Schools (1999–2003). She joined Al-Zaytoonah University as a part-time lecturer in 2003 before being appointed as a full-time Assistant Professor in 2004. In addition to her academic role, she has been a trainer in Arabic calligraphy since 2008, specializing in Ruq'ah and Kufic scripts used in Qur'anic manuscripts, and her research interests focus on applied linguistics, Arabic grammar, and language education.